

*Anais da*  
**JORNADA  
CIENTÍFICA E DE  
EXTENSÃO  
2018**

Nesta edição

---

Resumos expandidos

Trabalhos completos

## **Analisando a disponibilidade de um sistema e-Health integrado com infraestruturas de edge, fog e cloud**

**Matheus Felipe Ferreira da Silva Lisboa Tigre<sup>1</sup>, Guto Leoni<sup>2</sup>, Patricia Takako Endo<sup>1</sup>**

<sup>1</sup>Universidade de Pernambuco (UPE) – Campus Caruaru

<sup>2</sup>Universidade Federal de Pernambuco (UFPE)

matheus0906.mhcai@gmail.com, patricia.endo@upe.br, guto.leoni@gprt.ufpe.br

**Resumo.** *Internet das coisas e cloud computing são duas áreas que, cada vez mais, estão sendo integradas, visando a melhor utilização dos recursos de ambas as partes para alcançar um objetivo comum. Esta união oferece diversas vantagens, como maior disponibilidade, e maior capacidade de processamento e armazenamento, por exemplo. As aplicações e-Health tem aproveitado as vantagens desta integração, especialmente porque, por tratarem da saúde de pacientes, as mesmas são críticas e precisam estar em funcionamento o maior tempo possível. Neste contexto, o presente artigo propõe modelos matemáticos pra analisar como as falhas no sistema integrado de edge, fog e cloud podem impactar na disponibilidade de um sistema de e-Health.*

### **1. Introdução**

A Internet das Coisas (do inglês *Internet of Things* - IoT) tem encontrado um ecossistema cada vez mais favorável para se desenvolver, podendo beneficiar tecnologicamente os mais variados tipos de aplicações. Em [Atzori et al. 2010], os autores sugerem que o conceito básico de IoT é a presença pervasiva de uma variedade de dispositivos conectados que podem interagir uns com os outros para alcançar algum objetivo em comum. Estas coisas podem ser sensores, atuadores, *smart phones*, computadores, utensílios domésticos, carros, enfim, qualquer objeto que possa ser conectado ou monitorado [Biswas e Giaffreda 2014]. Aplicações de IoT tem o potencial de impactar significativamente o dia-a-dia dos usuários em diversas áreas, como: cidades inteligentes, fábricas inteligentes, controle de tráfego, *assisted living*. O setor de saúde, por exemplo, possui grande interesse em pesquisas na área de IoT, pois este possui um grande potencial para reduzir custos, melhorar a qualidade de vida dos pacientes, tal como a sua experiência [Islam et al. 2015].

Apesar de tudo, a IoT ainda possui diversos desafios a serem ultrapassados, que vão desde o gerenciamento da grande quantidade de dispositivos heterogêneos até a capacidade limitada de armazenamento e processamento de dados dos mesmos [Botta et al. 2014]. Essas limitações se tornam ainda mais evidentes quando levamos em consideração aplicações IoT que demandam alta quantidade de armazenamento e processamento e redes banda larga de alta velocidade para que possam ser feitas tomadas de decisões em tempo real [Lee e Lee 2015]. Por exemplo, no caso de sistemas de *e-Health*, enquanto os dados dos sensores podem ser utilizados para diagnósticos, desde fraturas na coluna até detecção de câncer, análises de imagens médicas requerem uma grande capacidade de armazenamento, que os dispositivos de IoT não possuem.

Alguns desses desafios de IoT podem ser amenizados integrando *fog* e *cloud computing*. *Cloud computing* é hoje uma tecnologia bem desenvolvida que pode prover uma melhoria em relação a capacidade de processamento, armazenamento, escalabilidade e disponibilidade. Já a *fog computing*, um paradigma que é relativamente novo, reconhece as limitações da comunicação das atuais redes de telecomunicações e provê conectividade para estender o paradigma da *cloud* para a borda da rede, permitindo aplicações verticalmente isoladas e sensíveis à latência [Vaquero e Rodero-Merino 2014, Bonomi et al. 2012]. Muitas aplicações necessitam de ambos, a proximidade da *fog* e a disponibilidade e geo-localização da *cloud*, particularmente para casos de análise de big data. Assim, a *fog*, geralmente, fica responsável pela análise e transmissão de dados em tempo real, enquanto a *cloud* pelo armazenamento e processamento mais complexo [Mell e Grance 2011].

Apesar das vantagens que *a fog* e a *cloud computing* trazem para aplicações de IoT, eles também trazem uma nova camada de complexidade no gerenciamento dos recursos. Enquanto melhoram a disponibilidade, eles também acrescentam novos pontos de falhas; além do dispositivo de IoT, também podem ocorrer falhas em nós da *fog* e nos componentes da infraestrutura da *cloud*, ou no sistema como um todo. Disponibilidade é um fator crítico para aplicações de *e-Health*, em especial aquelas que monitoram ativamente a situação de saúde de pacientes e que podem tomar decisões em tempo real. Qualquer que seja o tempo que o sistema fique indisponível, esta situação pode colocar a vida dos pacientes em risco.

Alguns trabalhos da literatura já analisaram soluções para identificar os principais pontos de falhas de sistemas IoT ou *e-Health*, como [Araujo et al. 2014], [Zeng et al. 2015], [Li et al. 2017]. Porém, nenhum deles analisa um ambiente onde são integradas as infraestruturas de *edge*, *fog* e *cloud*.

Neste contexto, esse trabalho apresenta um modelo estocástico para entender como falhas em dispositivos de IoT e *fog*, e/ou na infraestrutura de *cloud* impactam na disponibilidade de sistemas de *e-Health*. Para isso, foram feitas análises de disponibilidade e sensibilidade através de modelos *Stochastic Petri Nets* (SPN) [Murata 1989] e *Reliability Block Diagrams* (RBD), [Čepin 2011].

## 2. Sistema de monitoramento de e-Health

A Figura 1 apresenta a arquitetura do sistema de monitoramento de *e-Health* que utiliza dispositivos de *fog* (como a Raspberry Pi<sup>1</sup>) e infraestrutura de *cloud* para processar e armazenar dados vitais do paciente. É considerado que o pacientes possuem sensores (como sensores de glicose, batimento cardíaco, e de detecção de ataques epiléticos) que coletam os dados referentes à saúde física do paciente e estão conectados a um microcontrolador (como o Arduino<sup>2</sup>). Também foram consideradas duas aplicações que consomem os dados coletados: (a) uma aplicação na *fog*, and (b) uma aplicação na *cloud*.

A aplicação da *fog* verifica a normalidade dos dados e, em caso de alguma anomalia, a aplicação irá executar uma ação, como por exemplo, uma chamada para um serviço de emergência. Além disso, dados fisiológicos também podem ser enviados para

---

<sup>1</sup><https://www.raspberrypi.org/>

<sup>2</sup><https://www.arduino.cc/>

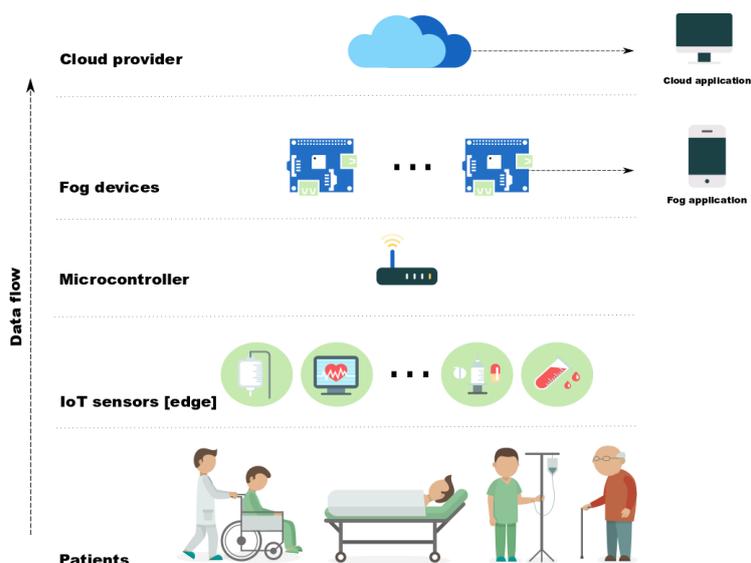


Figura 1. Arquitetura do sistema de monitoramento de e-Health

a aplicação da *cloud* para um processamento posterior, como por exemplo, para treinar um algoritmo de *machine learning* para examinar as condições do paciente com o passar do tempo e para comparar com dados de outros pacientes e prover um diagnóstico médico melhor <sup>3</sup>.

### 2.1. Modelo estocástico

Para analisar como falhas impactam na disponibilidade do sistema de *e-Health*, propomos um modelo analítico utilizando Redes de Petri Estocásticas (SPN) e *Reliability Block Diagrams* (RBD). A Figura 2 mostra o modelo SPN que representa o sistema de monitoramento de *e-Health*. Os seguintes componentes são considerados: sensor, microcontrolador, dispositivo da *fog* e infraestrutura da *cloud*. Para representar esses componentes, utilizamos *building blocks* compostos por dois *places* (um para representar quando o componente está funcionando (*ON*) e outro para representar quando ele falha (*DOWN*)), e duas transições (que representam a falha e o reparo de um componente específico).

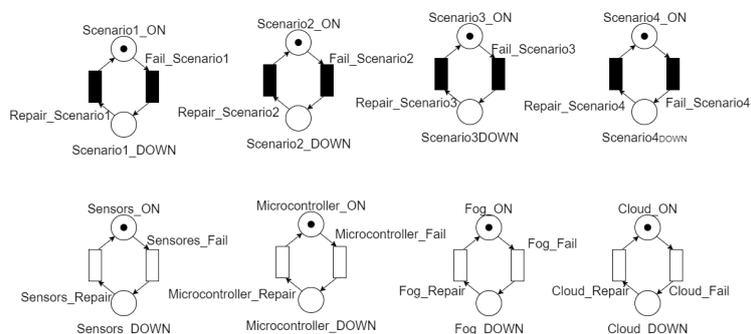


Figura 2. Modelo do sistema de monitoramento de e-Health

No *building block* referente à *cloud*, o *place Cloud\_ON* representa que a *cloud* está funcionando, e o *place Cloud\_OFF* que a *cloud* falhou ou está indisponível. A transição

<sup>3</sup><http://www.nvidia.com/object/deep-learning-in-medicine.html>

*Cloud\_Fail* representa o evento de falha da *cloud* (tempo médio de falha (do inglês, *mean time to fail*, MTTF)), enquanto *Cloud\_Repair* representa o tempo necessário para reparar a *cloud* (tempo médio de reparo (do inglês, *mean time to repair*, MTTR)). A modelagem dos outros três componentes (sensores, microcontrolador, e fog) seguem o mesmo raciocínio.

Também existem outros quatro *building blocks* com transições imediatas (no topo da figura) que representam o status do sistema em diferentes cenários. Da esquerda para direita, esses *building blocks* representam, respectivamente:

- **Cenário 1:** A disponibilidade do sistema de *e-Health* depende de todos os componentes do sistema, ou seja, uma falha em qualquer um deles irá resultar na indisponibilidade do sistema. Esse cenário não apresenta redundância, e as duas aplicações são complementares. Neste caso, a *cloud* necessita obrigatoriamente dos dados da *fog*;
- **Cenário 2:** O sistema de *e-Health* utiliza unicamente a aplicação da *cloud* para enviar dados vitais dos pacientes, sem utilizar a aplicação e os componentes da *fog*;
- **Cenário 3:** Semelhante ao cenário 2, o sistema de *e-Health* utiliza somente uma aplicação (neste caso, da *fog*) para enviar dados vitais dos pacientes, dessa vez não utilizando a aplicação e os componentes da *cloud*; e
- **Cenário 4:** Todos os componentes são levados em consideração mas as duas aplicações (*fog* e *cloud*) são redundantes, ou seja, o sistema ficará indisponível somente quando ambas não estiverem funcionando ao mesmo tempo.

Cada cenário possui um *building block* composto pelos places *ON* e *DOWN*, e duas transições imediatas. Esses *places* tem o mesmo comportamento dos citados anteriormente, bem como as suas transições. A diferença é que cada transição é ativada através de uma função de guarda ao invés de utilizar valores de MTTF ou MTTR. Essas funções de guarda são mostradas na Tabela 1.

Tabela 1. Funções de Guarda

Transição	Função de Guarda
<i>Fail_Scenario_1</i>	(#Sensors_ON=0)OR(#Microcontroller_ON=0)OR (#Fog_ON=0)OR(#Cloud_ON=0)
<i>Repair_Scenario_1</i>	(#Sensors_ON=1)AND(#Microcontroller_ON=1)AND (#Fog_ON=1)AND(#Cloud_ON=1)
<i>Fail_Scenario_2</i>	(#Sensors_ON=0)OR(#Microcontroller_ON=0)OR (#Cloud_ON=0)
<i>Repair_Scenario_2</i>	(#Sensors_ON=1)AND(#Microcontroller_ON=1)AND (#Cloud_ON=1)
<i>Fail_Scenario_3</i>	(#Sensors_ON=0)OR(#Microcontroller_ON=0)OR (#Fog_ON=0)
<i>Repair_Scenario_3</i>	(#Sensors_ON=1)AND(#Microcontroller_ON=1)AND (#Fog_ON=1)
<i>Fail_Scenario_4</i>	(#Sensors_ON=0)OR(#Microcontroller_ON=0)OR ((#Fog_ON=0)AND(#Cloud_ON=0))
<i>Repair_Scenario_4</i>	(#Sensors_ON=1)AND(#Microcontroller_ON=1)AND ((#Fog_ON=1)OR(#Cloud_ON=1))

Para detalhar mais o dispositivo de *fog* e a infraestrutura da *cloud*, ambos foram modelados utilizando RDB. Para representar o dispositivo de *fog* (Figura 3), considerou-se que ele é composto por *hardware* (HW), Sistema Operacional (OS), e a aplicação (APP) que consome os dados vitais dos pacientes. Se qualquer um desses componentes falhar, o dispositivo da *fog* é considerado indisponível; para isso, o RBD está organizado em série.

Por outro lado, o provedor da *cloud* (Figura 4) é composto por *hardware* (HW), Sistema Operacional (OS), máquina virtual (VM) e uma aplicação (APP). Nesse caso,

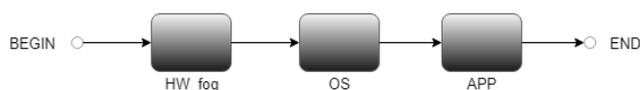


Figura 3. Modelo do dispositivo da fog

também foi considerado que a falha em qualquer um dos componentes resulta na falha da *cloud* como um todo (e portanto, o RDB também está em série).

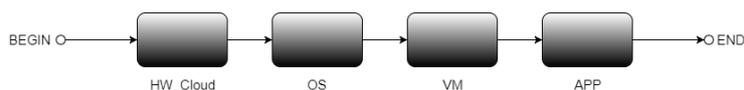


Figura 4. Modelo do servidor de cloud

A disponibilidade de cada um dos quatro cenários é a probabilidade de ter um *token* no *place* que representa o estado *ON* do respectivo *building block* do cenário (veja Figura 2). Para modelar a métrica referente à disponibilidade na SPN, foi utilizada a seguinte expressão:

$$P\{scenario_x\_ON\} > 0 \quad (1)$$

onde  $x$  é o número do cenário (de 1 to 4), como descrito anteriormente.

### 3. Resultados

Tanto a modelagem, como a análise dos dados foram feitas utilizando a ferramenta Mercury. Para efetuar a análise dos modelos, utilizamos valores de MTTF e MTTR dos componentes existentes na literatura. Com exceção dos valores de MTTF e MTTR do *hardware* do dispositivo referente à *fog* (*HW\_Fog*), que foram estimados através de uma média dos valores de HW encontrados em [Araujo et al. ], [Matos et al. 2017], [Silva et al. 2013], [Tang et al. 2004], [Kim et al. 2009], porque esses valores não estão disponíveis e provavelmente são secretos. Todos os valores utilizados para alimentar os modelos são apresentados na Tabela 2.

Tabela 2. Valores utilizados nos modelos de SPN e RBD, obtidos em [Araujo et al. ], [Matos et al. 2017], [Silva et al. 2013], [Tang et al. 2004], [Kim et al. 2009], [Novacek ], [Balc et al. ]

Component	MTTF (h)	MTTR (h)
HW_Cloud	8760	1.667
HW_Fog	4765.793684	3.4702988846
OS	2893	0.25
VM	1880	0.167
Microcontrolador	44957	5
Sensor	28011	5

### 3.1. Análise de disponibilidade

Os resultados da análise de disponibilidade de cada cenário estão presentes na Tabela 3. O cenário 4 apresenta a melhor disponibilidade devido a sua característica de possuir duas aplicações rodando de maneira redundante; por outro lado, o cenário 1 possui a pior disponibilidade porque uma falha em qualquer uma das duas aplicações torna o sistema indisponível. O cenário 2 apresenta uma disponibilidade maior em comparação com o cenário 3 devido ao alto valor de MTTF (leva mais tempo para falhar) e menor valor de MTTR (leva menos tempo para falar) da infraestrutura da *cloud*.

Tabela 3. Resultados de disponibilidade e *downtime*

Cenários	Disponibilidade (%)	downtime (h/ano)
Cenário 1	99.8482	13.29768
Cenário 2	99.9320	5.9568
Cenário 3	99.8872	9.88128
Cenário 4	99.9710	2.5404

### 3.2. Análise de sensibilidade

Análise de sensibilidade é um método que visa identificar quais fatores mais influenciam um sistema ([Frank 1978, Hamby 1994]). Variando os parâmetros de entrada e verificado como as variáveis de saída se comportam, pode-se descobrir parâmetros cruciais que impactam no processo geral.

Para a análise de sensibilidade, foram utilizados valores de MTTF e MTTR dos componentes do sistema como parâmetros, variando-os em cinco valores em uma variação limitada pelo máximo e mínimo valores (10% mais e 10% menos o valor inicial).

Tabela 4 apresenta os três componentes (e seus respectivos índices de sensibilidade) que mais impactam na disponibilidade do sistema. As Figuras 5, 6 e 7 representam a variação da disponibilidade dos três parâmetros que mais impactam na disponibilidade do sistema no cenário 1. Como esperado, quando o valor de MTTF aumenta (Figuras 5 e 7), a disponibilidade também aumenta, mas nesse caso, o MTTF do dispositivo da *fog* (Figura 5) acaba impactando mais na disponibilidade do que o MTTF da *cloud* (Figura 7), e isso pode ser visto claramente na inclinação da curva do gráfico. Um aumento de 20% no MTTF do dispositivo da *fog* implica em uma redução de cerca de 1.2 horas no *downtime* anual, enquanto o mesmo aumento no MTTF da *cloud* resulta em uma menor redução no *downtime*, somente cerca de 0.56 horas.

Tabela 4. Índices dos três parâmetros que mais afetam as métricas de cada cenário

Cenário 1		Cenário 2		Cenário 3		Cenário 4	
Parameter	Índice	Parâmetro	Índice	Parâmetro	Índice	Parâmetro	Índice
MTTF_Fog	1.69E-4	MTTF_Cloud	7.87E-5	MTTF_Fog	1.69E-4	MTTF_Sensores	3.60E-5
MTTR_Fog	1.67E-4	MTTR_Cloud	5.84E-5	MTTR_Fog	1.52E-4	MTTR_Sensores	3.56E-5
MTTF_Cloud	7.87E-5	MTTF_Sensors	3.60E-5	MTTF_Sensores	3.60E-5	MTTF_Microcontrolador	2.24E-5

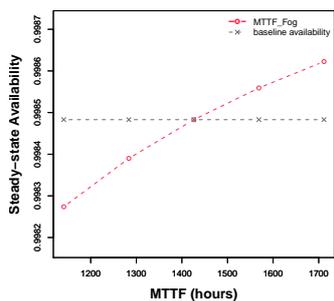


Figura 5. Resultados de disponibilidade variando somente o valor de MTTF do dispositivo da *fog* no Cenário 1

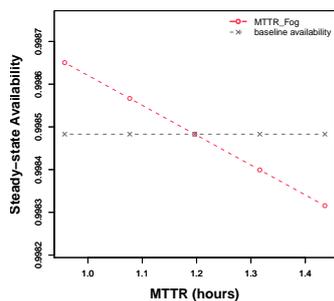


Figura 6. Resultados de disponibilidade variando somente o valor de MTTR do dispositivo da *fog* no Cenário 1

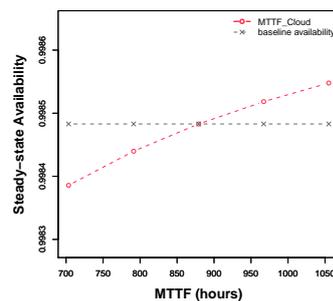


Figura 7. Resultados de disponibilidade variando somente o valor de MTTF do dispositivo da *cloud* no Cenário 1

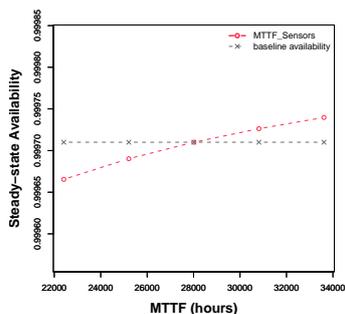


Figura 8. Resultados de disponibilidade variando somente o valor de MTTF do sensor no Cenário 4

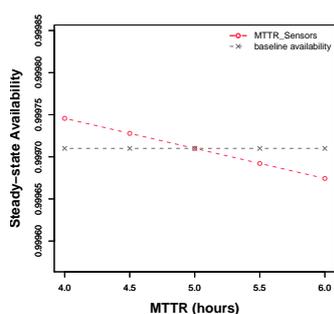


Figura 9. Resultados de disponibilidade variando somente o valor de MTTR do sensor no Cenário 4

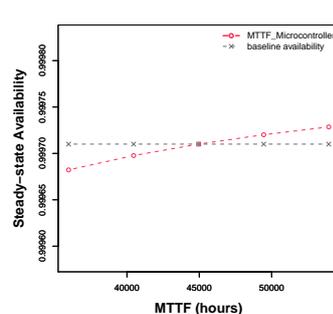


Figura 10. Resultados de disponibilidade variando somente o valor de MTTF do microcontrolador no Cenário 4

O cenário 4 (com aplicações redundantes) apresenta uma interessante diferença em comparação com o cenário 1 (Figuras 8, 9, e 10). O sensor e o microcontrolador são os componentes que mais afetam a disponibilidade do sistema, mas ainda assim impactam menos que a infraestrutura de *cloud* ou dispositivo *fog* no cenário 1.

### 3.3. Análise de escalabilidade

Para verificar a capacidade de escalabilidade do sistema, foi feita uma análise com diferentes quantidades de sensores para encontrar até que ponto um aumento na quantidade deles pode ou não impactar na disponibilidade do sistema. Utilizando o cenário 1, foi feita uma análise de sensibilidade, considerando que, no cenário que está sendo observado exista mais de um sensor e que para que o fluxo do sistema continue funcionando corretamente, todos estejam funcionando, ou seja, ocorrendo a falha de um ou mais sensores, todos que falharam devem ser reparados para que o sistema volte a funcionar correta-

mente. Esta análise foi feita para observar o quanto o aumento na quantidade de sensores impacta na disponibilidade do sistema como um todo.

Apesar do aumento de sensores, o impacto de seu MTTF e MTTR não sofreu grande alteração ou seja, é possível notar que, mesmo que o sistema utilize uma quantidade relativamente grande de sensores, a *cloud* e a *fog* ainda assim continuam sendo os componentes que mais impactam na disponibilidade do sistema neste cenário. Os sensores continuaram ficando em terceiro na ordem de impacto da disponibilidade sendo que, à medida que sua quantidade era aumentada, os índices do MTTF e do MTTR dos sensores sofreram um aumento, indicando que, como esperado, à partir do momento que se replica um componente, isto irá alterar de alguma forma a disponibilidade. Os índices referentes aos sensores obtidos na análise de sensibilidade podem ser vistos na Tabela 5.

Tabela 5. Índices referentes aos sensores obtidos nas análises de sensibilidade

	1 sensor	10 sensores	100 sensores	1000 sensores
Parâmetro	Índice	Índice	Índice	Índice
MTTF_Sensors	3.3049E-5	3.3061E-5	3.3062E-5	3.3067E-5
MTTR_Sensors	3.2100E-5	3.2151E-5	3.2132E-5	3.2129E-5

Os valores de disponibilidade obtidos nesta análise, tal como o *downtime*, podem ser observados na Tabela 6. Como pode ser visto, a diferença do *downtime* do sistema com um para o mesmo com mil sensores é de apenas 0,0049 horas o que equivale a menos de um minuto; demonstrando que apesar do aumento na quantidade de sensores, a disponibilidade do sistema não foi tão prejudicada.

Tabela 6. Resultados de disponibilidade e *downtime* para o cenário 1

	1 sensores	10 sensores	100 sensores	1000 sensores
Disponibilidade (%)	99.730069	99.730064	99.730055	99.730013
Downtime (horas/ano)	23.6459	23.6463	23.6471	23.6508

A mesma análise foi repetida para o cenário 4, devido ao fato de que os valores de MTTF e MTTR do sensor foram os que mais impactaram na disponibilidade do sistema. Ela foi feita para verificar se neste cenário, onde os sensores, comparado com os demais componentes, apresentam um impacto maior na disponibilidade, e se ao aumentar o número de sensores, a disponibilidade é impactada de alguma forma.

Feita a análise, foi visto que mesmo com um aumento de 1 para 1000 sensores, o impacto observado, apesar de maior que no cenário 1, não é tão significativo. Com apenas um sensor, a disponibilidade do sistema neste cenário foi de aproximadamente 0,9997089, e com 1000 ficou em 0,9997083; e o *downtime* subiu de aproximadamente 2,54 horas para 2,55 horas, ocorrendo um aumento de apenas 0,05 hora (cerca de 3 minutos). A Tabela 7 apresenta os resultados de disponibilidade e *downtime*.

#### 4. Conclusão e próximos passos

Nesse trabalho, foram apresentados modelos SPN e RBD com o intuito de analisar a disponibilidade de aplicações de *e-Health* utilizando infraestrutura integrada com *fog* e

Tabela 7. Resultados de disponibilidade e *downtime* para o cenário 4

	1 sensor	10 sensores	100 sensores	1000 sensores
<b>Disponibilidade (%)</b>	99.97089	99.97088	99.97087	99.97083
<b>Downtime (horas/ano)</b>	2.5499	2.5503	2.5511	2.5548

*cloud*. Foram apresentados diferentes cenários, e as análises mostraram que o cenário onde as duas aplicações (*cloud* e *fog*) funcionam de maneira independente apresenta uma maior disponibilidade. A análise de sensibilidade mostra os componentes que mais impactam na disponibilidade dos cenários apresentados. Os resultados comprovaram que os sensores e os dispositivos da *fog* são os componentes que mais afetam a disponibilidade de aplicações de *e-Health*. Através da análise de escalabilidade também pôde ser observado como um aumento no número de sensores do sistema acaba por não impactar tanto em sua disponibilidade.

Como trabalhos futuros, pretende-se prototipar o sistema de monitoramento de *e-Health* proposto, para efetuar experimentos e obter valores experimentais para alimentar os modelos matemáticos, bem como ampliar os cenários até então apresentados.

### Agradecimentos

Este trabalho recebeu apoio financeiro do CNPq e da UPE, através do Edital PIBIC-IC/UPE/CNPq – 2017/2018.

### Referências

- Araujo, J., Maciel, P., Torquato, M., Callou, G., e Andrade, E. Availability evaluation of digital library cloud services. In *IEEE/IFIP Dependable Systems and Networks (DSN)*.
- Araujo, J., Silva, B., Oliveira, D., e Maciel, P. (2014). Dependability evaluation of a mhealth system using a mobile cloud infrastructure. In *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on*, pages 1348–1353. IEEE.
- Atzori, L., Iera, A., e Morabito, G. (2010). The internet of things: A survey. *Computer networks*, 54(15):2787–2805.
- Balc, C., Cretu, A., Munteanu, R., Iudean, D., Balan, H., e Karaisas, P. Reliability modeling for an automatic level control system. In *International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*.
- Biswas, A. R. e Giaffreda, R. (2014). Iot and cloud convergence: Opportunities and challenges. In *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, pages 375–376. IEEE.
- Bonomi, F., Milito, R., Zhu, J., e Addepalli, S. (2012). Fog computing and its role in the internet of things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pages 13–16. ACM.
- Botta, A., De Donato, W., Persico, V., e Pescapé, A. (2014). On the integration of cloud computing and internet of things. In *Future Internet of Things and Cloud (FiCloud), 2014 International Conference on*, pages 23–30. IEEE.

- Čepin, M. (2011). Reliability block diagram. In *Assessment of Power System Reliability*, pages 119–123. Springer.
- Frank, P. M. (1978). *Introduction to system sensitivity theory*, volume 1. Academic press New York.
- Hamby, D. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32(2):135–154.
- Islam, S. R., Kwak, D., Kabir, M. H., Hossain, M., e Kwak, K.-S. (2015). The internet of things for health care: a comprehensive survey. *IEEE Access*, 3:678–708.
- Kim, D. S., Machida, F., e Trivedi, K. S. (2009). Availability modeling and analysis of a virtualized system. In *Dependable Computing, 2009. PRDC'09. 15th IEEE Pacific Rim International Symposium on*, pages 365–371. IEEE.
- Lee, I. e Lee, K. (2015). The internet of things (iot): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4):431–440.
- Li, Y., Orgerie, A.-C., Rodero, I., Parashar, M., e Menaud, J.-M. (2017). Leveraging renewable energy in edge clouds for data stream analysis in iot. In *Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pages 186–195. IEEE Press.
- Matos, R., Dantas, J., Araujo, J., Trivedi, K. S., e Maciel, P. (2017). Redundant eucalyptus private clouds: Availability modeling and sensitivity analysis. *Journal of Grid Computing*, 15(1):1–22.
- Mell, P. e Grance, T. (2011). The nist definition of cloud computing.
- Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580.
- Novacek, G. Tips for predicting product reliability. <http://circuitcellar.com/cc-blog/tips-for-predicting-product-reliability/>. Accessed: 2018-01-05.
- Silva, B., Maciel, P., Tavares, E., e Zimmermann, A. (2013). Dependability models for designing disaster tolerant cloud computing systems. In *Dependable Systems and Networks (DSN), 2013 43rd Annual IEEE/IFIP International Conference on*, pages 1–6. IEEE.
- Tang, D., Kumar, D., Duvur, S., e Torbjornsen, O. (2004). Availability measurement and modeling for an application server. In *Dependable Systems and Networks, 2004 International Conference on*, pages 669–678. IEEE.
- Vaquero, L. M. e Rodero-Merino, L. (2014). Finding your way in the fog: Towards a comprehensive definition of fog computing. *ACM SIGCOMM Computer Communication Review*, 44(5):27–32.
- Zeng, W., Koutny, M., e Watson, P. (2015). Opacity in internet of things with cloud computing (short paper). In *Service-Oriented Computing and Applications (SOCA), 2015 IEEE 8th International Conference on*, pages 201–207. IEEE.

## Uma Abordagem Híbrida Para Predizer Pontuações Motoras da Doença de Parkinson no Contexto de Mineração de Dados

Marília Lima<sup>2</sup>, Roberta Fagundes<sup>1,2</sup>

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

<sup>2</sup>Universidade de Pernambuco (UPE)  
Recife– PE – Brasil

mncal@ecomp.poli.br, roberta.fagundes@upe.br

**Abstract. Objective:** To predict Parkinson's disease scores based on a hybrid approach using a combination of unsupervised and supervised learning. **Methodology:** CRISP-DM (Cross Industry Standard Process for Data Mining) was used. **Results:** Mean errors 0.03922218, 0.03592167, and standard deviation are 0.00082, 0.00081, for linear regression and the hybrid approach, respectively. It was obtained evidences that the approach presented a prediction error smaller than the Linear Regression with the value of the p-value of  $2.2e-16$  to the level of significance of 5%. **Final Considerations:** the hybrid approach presented satisfactory results and could potentiate the result of the estimation in databases.

**Resumo. Objetivo:** prever as pontuações motoras da doença de Parkinson baseado em uma abordagem híbrida que utilize uma combinação de aprendizado não-supervisionado e supervisionado. **Metodologia:** utilizou-se a CRISP-DM (Cross Industry Standard Process for Data Mining). **Resultados:** a média dos erros 0.03922218, 0.03592167, e desvio padrão são 0.00082, 0.00081, para regressão linear e a abordagem híbrida, respectivamente. Obteve-se evidências que a abordagem apresentou um erro de predição menor que a Regressão Linear com o valor do p-value de  $2.2e-16$  ao nível de significância de 5%. **Considerações Finais:** a abordagem híbrida apresentou resultados satisfatórios e pode potencializar o resultado da estimação em bases de dados.

### 1. Introdução

A utilização de técnicas de Inteligência Computacional para resolução de problemas está se tornando frequente em diferentes áreas, tais como medicina, agricultura, educação (SHAKOOR *et al.*, 2017). Neste contexto, a área médica apresenta como um problema a identificação do impacto das alterações motoras causadas pela doença de parkinson.

A doença de parkinson é uma doença crônica e que afeta o sistema nervoso central (LANA, 2007), sendo seus sintomas motores os mais característicos (FERRAZ; BORGES, 2002). Pessoas que têm a doença de parkinson costumam apresentar movimentos involuntários, posturas anormais entre outras características (COELHO *et al.*, 2006). De acordo com Coelho *et al.* (2006) o impacto das alterações motoras da

doença de parkinson pode ser medido para exploração motora da escala unificada de avaliação da doença de parkinson (UPDRS).

Neste contexto, prever essa alteração motora é de extrema importância, pois se os sintomas desta doença forem diagnosticados no início e iniciado o tratamento, isso pode possibilitar ao paciente uma melhor qualidade de vida. Uma vez que a medicação se tomada no estágio inicial da doença oferece um maior alívio dos sintomas. Podendo ser um dos primeiros sintomas da doença de parkinson a deficiência vocal ( Little *et al.*, 2009).

Desta forma, a extração de características em bases de dados pode ser um excelente auxiliar no processo de tomada de decisão nesta área de atuação, sendo esta extração de informação realizada por técnicas de inteligência computacional. Entre estas técnicas pode ser realizado um aprendizado supervisionado ou um aprendizado não-supervisionado para conseguir informações sobre estes problemas. O aprendizado supervisionado apresenta a característica de o processo de aprendizado ser realizado com auxílio de um supervisor, logo é necessário um conjunto de dados de pares de entrada e saída, onde o conjunto de saída (variável dependente, por exemplo) deve estar associado ao conjunto de entrada (BARRETO, 1999). Já o aprendizado não-supervisionado tem como característica não necessitar de um agente externo que reconheça os padrões dos dados (BARRETO, 1999).

Neste cenário, entre as técnicas de Inteligência computação não-supervisionadas existe o agrupamento que tem como objetivo formar grupos com uma maior similaridade e os diferentes grupos formados devem apresentar uma maior dissimilaridade (JAIN and MURTY, 1999). Dentre os algoritmos de agrupamento existem diferentes tipos desses algoritmos como os particionais, hierárquicos. Nos algoritmos particionais pode-se destacar o K-means por ser simples e amplamente utilizado (PELLEG *et al.*, 2000).

Já entre as técnicas supervisionadas existe a regressão. Segundo Montgomery (2012) modelos de regressão têm por objetivo analisar as relações entre as variáveis. As técnicas de regressão podem ser classificadas em paramétricas e não paramétricas. Dentro as técnicas paramétricas existem a regressão linear simples que faz a relação de uma variável (X) com a variável dependente (Y) utilizando o método de estimação dos mínimos quadrados (MONTGOMERY, 2012).

Porém algumas destas técnicas em sua versão básica apresentam certa fragilidade, por isso abordagens híbridas podem vir a melhorar o resultado da estimação em bases de dados, uma vez que combinam as vantagens dos métodos envolvidos (POOJA, 2015). Neste contexto, este artigo tem por objetivo desenvolver uma abordagem híbrida que utilize uma combinação de aprendizado supervisionado e não-supervisionado no contexto de mineração de dados para prever as pontuações motoras da doença de Parkinson, essa predição é realizada utilizando UPDRS. Para isso foram utilizados o algoritmo de agrupamento K-means e a regressão linear simples.

Este trabalho está organizado em cinco seções. A seção 2 apresentará a fundamentação teórica, na seção 3 será apresentada a metodologia que será utilizada neste trabalho, na seção 4 os resultados e a discussão do trabalho, por fim, na seção 5 serão apresentadas as considerações finais do mesmo.

## 2. Fundamentação Teórica

Esta seção apresentará conceitos fundamentais para entendimento deste trabalho, deste modo os conceitos apresentados serão sobre regressão linear e o algoritmo de agrupamento K-means. As subseções a seguir apresentarão estes conceitos.

### 2.1. Regressão Linear

A regressão linear é uma técnica de regressão paramétrica que tem como objetivo ver a relação de uma variável dependente (Y) com variáveis independentes (X) (MONTGOMERY, 2012). O modelo de regressão linear simples é apresentado na equação (1)

$$y_i = \alpha + \beta x_i + \epsilon_i \quad (i = 1, 2, \dots, n) \quad (1)$$

Onde o  $\alpha$  representa o intercepto da reta com o eixo dos y,  $\beta$  o coeficiente de inclinação da reta e  $\epsilon$  o erro. A equação (2) apresenta a o cálculo do y estimado.

$$\hat{y}_i = \alpha + \beta x_i \quad (2)$$

O erro é a diferença entre o verdadeiro y e o y estimado, é apresentado na equação (3)

$$\epsilon_i = y_i - \hat{y}_i \quad (3)$$

O método de estimação utilizado por essa regressão é o método de mínimos quadrados. O método de mínimos quadrados busca fazer que a soma dos erros quadrados seja a menor possível. E é apresentado na equação (4)

$$\sum_{i=1}^n \epsilon^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

### 2.2. Algoritmo de Agrupamento K-means

No algoritmo de agrupamento k-means para fazer o agrupamento na base de dados é necessário a definição do número de grupos (k) que deseja-se formar, para que sejam gerados de forma aleatório os centróides do algoritmo e formados os grupos em torno desses centróides (JAIN and MURTY, 1999).

Para formar os grupos em volta de cada centróide pode-se utilizar a distância euclidiana, como apresenta a equação (5)

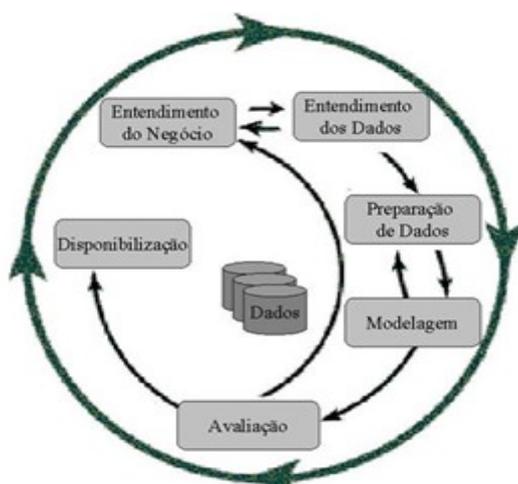
$$\arg \min_{c_i \in C} \text{dist}(c_i, x)^2 \quad (5)$$

A atualização dos centróides, isso é feita levando em consideração a média de todos os pontos de cada cluster, como apresenta a equação (6)

$$c_i = \frac{1}{|s_i|} \sum_{x_i \in s_i} x_i \quad (6)$$

## 3. Metodologia

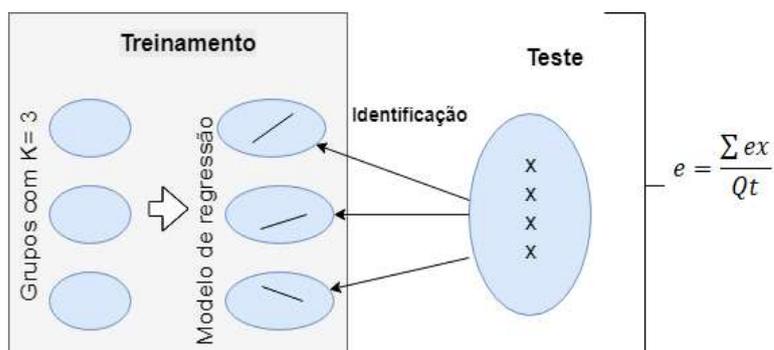
A metodologia que será utilizada neste trabalho é o CRISP-DM (Cross Industry Standard Process for Data Mining). A CRISP-DM é uma das metodologias mais utilizadas no contexto de mineração de dados. A metodologia consiste em seis etapas, a saber: entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação e, por fim, disponibilização (Chapman et al, 2000). Assim como apresenta a Figura 1.



**Figura 1. Metodologia CRISP-DM Adaptado de (Chapman et al, 2000)**

A descrição das etapas da metodologia será apresentada a seguir:

- Entendimento do negócio: essa etapa consiste em entender qual o problema que deve ser estudado e o que busca resolver na base. Neste trabalho, o problema refere-se à predição de pontuações motoras da doença de parkinson. A base de dados utilizada está disponível publicamente, e é formada por dados reais e composto de medições biomédicas de voz de 42 pessoas com doença de parkinson em estágio inicial. A base contém 5875 instâncias de gravações de voz desses indivíduos dentre um período de seis meses e não contém dados faltosos. Utilizou-se dados de vozes, pois segundo Little *et al.* (2009) um dos primeiros sintomas que identificam a doença de parkinson é a deficiência vocal.
- Entendimento dos Dados: essa etapa tem como objetivo entender quais variáveis devem ser utilizadas, fazendo isso através da descrição das variáveis da base de dados. No trabalho, este processo foi realizado através da correlação entre a variável depende e as variáveis explicativas.
- Preparação dos Dados: nesta etapa deve-se preparar toda a base de dados, realizando procedimentos como padronização dos valores da base, tratamento de dados faltosos. No trabalho foi realizada a padronização dos dados entre os valores zero e um.
- Modelagem: essa etapa tem como objetivo selecionar as técnicas de mineração de dados que irão ser utilizadas na resolução do problema. No presente trabalho, será utilizada a técnica de regressão linear simples, e foi desenvolvida uma abordagem híbrida que utiliza a combinação de uma técnica de aprendizado não-supervisionado e com uma técnica de aprendizado supervisionado. A técnica de aprendizado não-supervisionado utilizada foi o K-means e a técnica de aprendizado supervisionado foi a regressão linear simples. A Figura 2 apresenta a abordagem híbrida desenvolvida no presente trabalho.



**Figura 2. Abordagem Híbrida desenvolvida**

Nesta abordagem os dados foram divididos de forma aleatória em um conjunto de treinamento e outro de teste. No conjunto de treinamento foi realizada a execução do algoritmo K-means e os modelos de regressão também foram treinados. Para utilizar o algoritmo K-means é necessário a definição do número de grupos. Neste contexto, segundo Linden (2009) o número de grupos do algoritmo K-means é definido de forma *ad hoc* pelo usuário que irá usar o algoritmo.

Por esse motivo, valor da quantidade de grupos desenvolvida na abordagem foi de três, então foram gerados três grupos utilizando o algoritmo K-means e dentro de cada grupo foi treinado um modelo de regressão. Esse modelo de regressão na Figura 2 é representado pela a reta que estar dentro de cada grupo. Dessa forma, foram treinados três modelos de regressão, um para cada grupo gerado pelo algoritmo K-means.

Posteriormente, para testar a abordagem desenvolvida foi necessário realizar uma identificação dos dados do conjunto de teste. Essa identificação foi feita utilizando a menor distância de cada dado do conjunto de teste em relação aos centróides dos grupos. Logo, os dados do conjunto de teste foram agrupados em relação ao centróide de menor distância.

Desse modo, os dados do conjunto de teste eram testados com o modelo de regressão que apresentassem uma maior similaridade com os dados do conjunto de treinamento. E assim, cada dado do conjunto de teste será testado no modelo de regressão que seja mais adequado. Desse modo, para obter o erro final desta abordagem o cálculo do erro foi feito com seguinte equação:

$$e = \frac{\sum e_x}{Q_t} \quad (7)$$

- Onde,  $e$  representa o erro geral da abordagem;
  - $e_x$  representa o erro de cada  $x$  do conjunto de teste de cada grupo; e
  - $Q_t$  representa a quantidade de grupos do teste.
- Avaliação: nesta etapa são analisados os resultados dos experimentos realizados. No trabalho, a avaliação foi feita através média e desvio padrão, assim como teste de hipóteses dos erros de predição obtidos (erro médio quadrado).
  - Disponibilização: nesta etapa é realiza-se a disponibilização para indústria de todas as informações obtidas nas etapas de modelagem e avaliação. No trabalho,

esta etapa não foi realizada, uma vez que essa metodologia é uma metodologia desenvolvida para indústria e que quando utilizada para trabalhos na área de mineração de dados é bem aceita até etapa de avaliação.

#### 4. Resultados e Discussão

Esta seção apresentará os resultados e discussão deste trabalho. As subseções a seguir apresentarão sobre a análise descritiva da base de dados, a realização dos experimentos do presente trabalho e, por fim, os testes estatísticos realizados.

##### 4.1 Análise Descritiva das variáveis do estudo

A base de dados contém 5875 instâncias, 23 variáveis, sendo duas variáveis dependentes, a saber: pontuação motoras da doença de parkinson e pontuações total da doença de Parkinson. Deste modo, a variável pontuações total da doença de parkinson foi excluída da base, uma vez que esse estudo tem por objetivo analisar a pontuação motoras da doença de parkinson. Foi realizada uma padronização com os valores da base, a padronização realizada foi entre zero e um. A correlação entre as variáveis da base de dados foi realizada em todas as variáveis, desta forma, a variável que obteve o maior grau de correlação com a variável dependente foi a variável idade, que obteve uma correlação de 0.274, sendo assim uma correlação positiva. E foi a variável utilizada tanto no modelo de regressão linear simples, quanto na abordagem híbrida. A Figura 3 apresenta a correlação entre a variável idade e a variável pontuação motora.

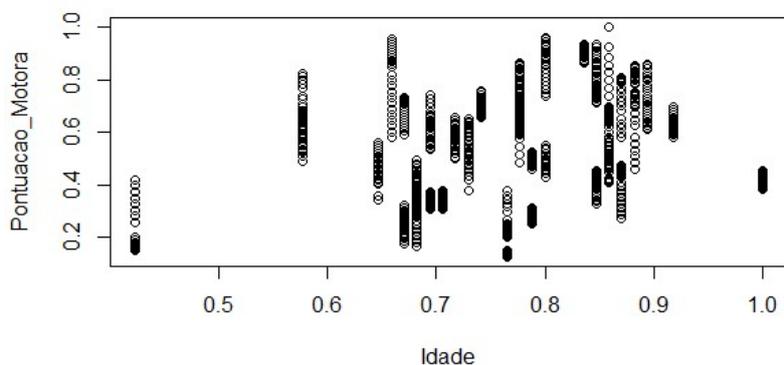


Figura 3. Correlação entre a variável idade e pontuações motoras da doença de Parkinson

##### 4.2 Realização dos Experimentos

Os experimentos deste trabalho foram realizados seguindo o método de simulação de Monte Carlo, desta forma, foram realizadas mil iterações para se obter os erros de predição. Os dados foram particionados de forma aleatória, e foram divididos em um conjunto de treinamento (contém 70% dos dados) e em um conjunto de teste (contém 30% dos dados). A partir do treinamento do modelo de regressão linear simples e da abordagem híbrida desenvolvida foram obtidos os erros de predição das 1000 iterações. O erro de predição utilizado no trabalho é o erro médio quadrado. Por fim, foi calculado a média e desvio padrão dos erros. Assim como é apresentado na Figura 4 o pseudocódigo utilizado no presente trabalho.

- 1: **Utilize** a base de dados
- 2: **Requerer** MC = 1000
- 3: **Para** todo  $i$  1 até MC faça:
  - 3.1 **Particione** aleatoriamente o conjunto de dados simulados em conjunto de Treinamento (70% dos dados) e de teste (30% dos dados)
  - 3.2 **Aplique** o modelo de Regressão Linear para o conjunto de treinamento
  - 3.3 **Aplique** a abordagem proposta para o conjunto de treinamento
  - 3.4 **Calcule** o erro médio quadrado da Regressão Linear e da abordagem
- 4: **Fim** para
- 5: **Calcule** a média e desvio padrão dos erros (erro médio quadrado).

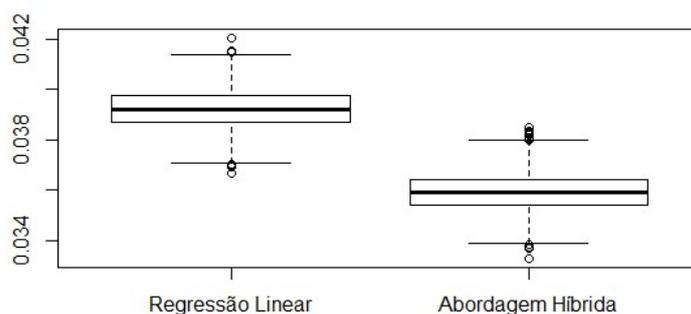
**Figura 4. Pseudocódigo do algoritmo utilizado no trabalho**

A Tabela 1 apresenta as médias e os desvios padrões dos erros obtidos pelos experimentos realizados. Como pode-se observar os desvios padrões tanto do modelo de regressão linear quanto da abordagem híbrida são similares demonstrando que existe uma similaridade na variabilidade de ambos os algoritmos. Porém a média dos erros da abordagem é menor que da regressão linear simples.

**Tabela 1. Média e desvio padrão do modelo de regressão linear e da abordagem híbrida**

Algoritmo	Média	Desvio Padrão
Regressão Linear	0.03922218	0.000819834
Abordagem Híbrida K=3	0.03592167	0.000808044

A Figura 5 apresenta o boxplot destes erros de predição da simulação de monte carlo realizada, e pode-se verificar que a abordagem híbrida obtém menor mediana, assim como similaridade na variabilidade amostral com a regressão linear. Pode-se então perceber que abordagem apresenta um modelo consistente em relação ao erro médio quadrado.



**Figura 5. Boxplot do modelo de Regressão Linear de Abordagem Híbrida**

As Figuras 6 e 7 apresentam os histogramas da distribuição dos erros do modelo de regressão linear e da abordagem, respectivamente. Podendo verificar que os erros obtidos com a simulação de Monte Carlo, tanto da regressão linear quanto da abordagem híbrida, não demonstram evidências que seguem uma distribuição normal. Desse modo foram realizados testes estáticos para conseguir melhores evidências sobre a normalidade desses dados. Estes testes serão apresentados na próxima subseção.

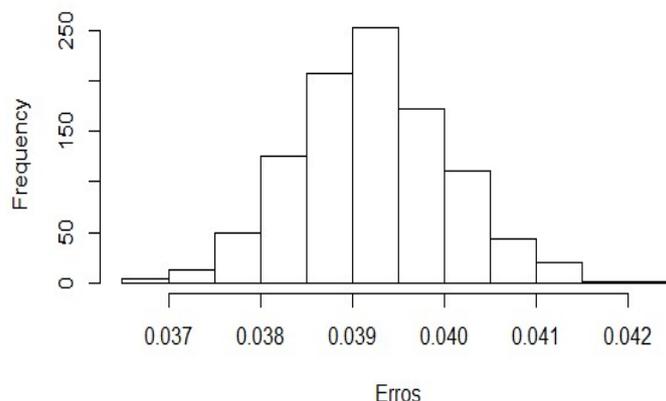


Figura 6. Histograma da distribuição dos Erros da Regressão Linear

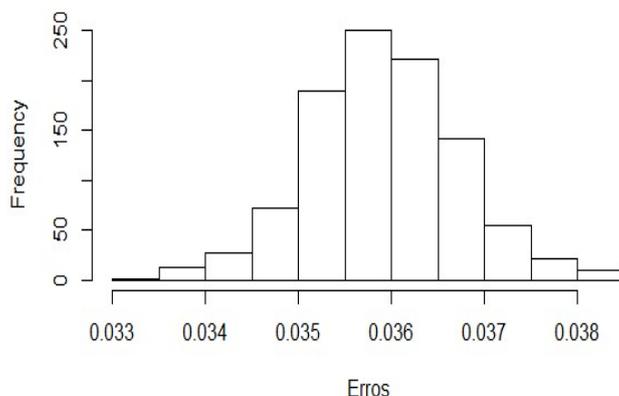


Figura 7. Histograma da distribuição dos erros da Abordagem Híbrida

### 4.3 Testes Estatísticos Realizados

Os teste estatísticos realizados foram o de kolmogorov-simirnov e teste hipótese. O teste de kolmogorov-simirnov foi utilizado para verificar se os erros seguem uma distribuição normal, visto que não houve como concluir pelos os histogramas já apresentados. A Tabela 2 apresenta o teste de normalidade realizado ao nível de significância de 5% para a regressão linear e a abordagem híbrida.

Tabela 2. Teste de Normalidade (Kolmogorov-simirnov)

Algoritmo	Valor do p-value
Regressão Linear	2.2e-16
Abordagem Híbrida	2.2e-16

Pelos valores de p-value obtidos nesta estatística e considerando o nível de significância de 5% conclui-se que os dados não seguem uma distribuição normal. Desta forma, como foi comprovado que erros de predição não seguem uma distribuição normal, foi utilizado o teste de wilcoxon no teste de hipótese.

O teste de hipótese foi elaborado da seguinte forma:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

Onde,  $\mu_1$  representa a média da abordagem híbrida e  $\mu_2$  a média da regressão linear, desta forma, hipótese nula ( $H_0$ ) definida é que a regressão linear apresenta os erros de predição maior ou igual a Abordagem Híbrida, e a hipótese alternativa ( $H_1$ ) é definida como a abordagem híbrida apresenta erros de predição menores que a regressão linear.

Pelas hipóteses de pesquisa elaboradas foi realizado um teste de hipótese unilateral à esquerda com nível de significância de 5%. O valor do p-value obtido no teste de hipótese foi de  $2.2e-16$ . Pode-se concluir que a hipótese nula ( $H_0$ ) foi refutada, logo há evidências que a abordagem híbrida apresenta erros de predição menores que a regressão linear, uma vez que pode-se obter evidência sobre a hipótese alternativa ( $H_1$ ).

Desta forma, pode-se verificar através de dados simulados que a abordagem híbrida desenvolvida apresenta menores erros de predições que a regressão linear. Demonstrando um modelo mais preciso e consistente. Neste cenário, a abordagem híbrida demonstra maior precisão na predição de pontuações motoras com relação a idade do indivíduo que a regressão linear.

## 5. Considerações Finais

Diante do que foi exposto a abordagem híbrida obteve resultados satisfatórios para a predição de pontuações motoras da doença de Parkinson. Já que combinando as vantagens dos métodos utilizados apresentou erros de predições menores, estando esse fato comprovado estatisticamente através do teste de hipótese.

E como na área médica a precisão no diagnóstico é de extrema importância, logo a abordagem híbrida pode prever de forma mais precisa, se comparado com a regressão linear. Deste modo, a predição das pontuações motoras da doença de parkinson utilizando a abordagem híbrida desenvolvida é mais eficiente para os dados do estudo. A utilidade da predição de pontuações motoras da doença de parkinson se dar no auxílio ao tratamento das pessoas que tem essa doença.

Adicionalmente, compreende-se que a construção de uma abordagem híbrida utilizando aprendizado supervisionado e não-supervisionado pode ser um excelente auxiliar no processo de extração de informação em bases de dados em problemas reais. No presente trabalho, essa extração de informação refere-se a prever as pontuações motoras da doença de parkinson, ou seja, depois do modelo treinado pode ser estimado essas pontuações de outros indivíduos.

Pelos fatos apresentados, pretende-se fazer o aperfeiçoamento desta abordagem através de métodos mais rebuscados, como utilizar uma regressão linear múltipla, ao usar uma regressão simples, além de testar uma maior variação de quantidade de grupos, formando diferentes cenários de testes. E assim, conseguir maiores evidências sobre esta abordagem, como também a precisão da predição de pontuações motoras da doença de Parkinson.

## Referências

BARRETO, Jorge Muniz. Inteligência Artificial no limiar do século XXI. Florianópolis: PPP edições, v. 97, 1999.

- COELHO, M. S.; PATRIZZI, L. J.; OLIVEIRA, A. P. R. Impacto das alterações motoras nas atividades de vida diária na Doença de Parkinson. *Neurociências*, v. 14, n. 4, p. 178-81, 2006.
- FERRAZ, Henrique Ballalai; BORGES, Vanderci. Doença de parkinson. *Revista brasileira de medicina*, v. 59, n. 4, p. 207-219, 2002.
- JAIN, Anil K.; MURTY, M. Narasimha; FLYNN, Patrick J. Data clustering: a review. *ACM computing surveys (CSUR)*, v. 31, n. 3, p. 264-323, 1999.
- LANA, R. C. et al. Percepção da qualidade de vida de indivíduos com doença de Parkinson através do PDQ-39. *Revista Brasileira de Fisioterapia*, v. 11, n. 5, 2007.
- LINDEN, Ricardo. Técnicas de agrupamento. *Revista de Sistemas de Informação da FSMA*, v. 4, p. 18-36, 2009.
- Little MA, McSharry PE, Hunter EJ, Ramig LO (2009), 'Suitability of dysphonia measurements for telemonitoring of Parkinson's disease', *IEEE Transactions on Biomedical Engineering*, 56(4):1015-1022
- MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, G.Geoffrey. *Introduction to linear regression analysis*. John Wiley & Sons, 2012
- SHAKOOR, Md Tahmid et al. Agricultural production output prediction using Supervised Machine Learning techniques. In: *Next Generation Computing Applications (NextComp)*, 2017 1st International Conference on. IEEE, 2017. p.182-187.
- PELLEG, Dan et al. X-means: Extending k-means with efficient estimation of the number of clusters. In: *Icml*. 2000. p. 727-734.
- Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rudiger Wirth. *Crisp-dm 1.0 step-by-step data mining guide*. CRISP-DM Consortium, 2000.
- POOJA, M. R.; PUSHPALATHA, M. P. A hybrid decision support system for the identification of asthmatic subjects in a cross-sectional study. In: *Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, 2015 International Conference on. IEEE, 2015. p. 288-293.

<https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring>, último acesso em 02 de outubro de 2018

## Uma Abordagem Híbrida Para Predizer Pontuações Motoras da Doença de Parkinson no Contexto de Mineração de Dados

Marília Lima<sup>2</sup>, Roberta Fagundes<sup>1,2</sup>

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

<sup>2</sup>Universidade de Pernambuco (UPE)  
Recife– PE – Brasil

mncal@ecomp.poli.br, roberta.fagundes@upe.br

**Abstract. Objective:** To predict Parkinson's disease scores based on a hybrid approach using a combination of unsupervised and supervised learning. **Methodology:** CRISP-DM (Cross Industry Standard Process for Data Mining) was used. **Results:** Mean errors 0.03922218, 0.03592167, and standard deviation are 0.00082, 0.00081, for linear regression and the hybrid approach, respectively. It was obtained evidences that the approach presented a prediction error smaller than the Linear Regression with the value of the p-value of  $2.2e-16$  to the level of significance of 5%. **Final Considerations:** the hybrid approach presented satisfactory results and could potentiate the result of the estimation in databases.

**Resumo. Objetivo:** prever as pontuações motoras da doença de Parkinson baseado em uma abordagem híbrida que utilize uma combinação de aprendizado não-supervisionado e supervisionado. **Metodologia:** utilizou-se a CRISP-DM (Cross Industry Standard Process for Data Mining). **Resultados:** a média dos erros 0.03922218, 0.03592167, e desvio padrão são 0.00082, 0.00081, para regressão linear e a abordagem híbrida, respectivamente. Obteve-se evidências que a abordagem apresentou um erro de predição menor que a Regressão Linear com o valor do p-value de  $2.2e-16$  ao nível de significância de 5%. **Considerações Finais:** a abordagem híbrida apresentou resultados satisfatórios e pode potencializar o resultado da estimação em bases de dados.

### 1. Introdução

A utilização de técnicas de Inteligência Computacional para resolução de problemas está se tornando frequente em diferentes áreas, tais como medicina, agricultura, educação (SHAKOOR *et al.*, 2017). Neste contexto, a área médica apresenta como um problema a identificação do impacto das alterações motoras causadas pela doença de parkinson.

A doença de parkinson é uma doença crônica e que afeta o sistema nervoso central (LANA, 2007), sendo seus sintomas motores os mais característicos (FERRAZ; BORGES, 2002). Pessoas que têm a doença de parkinson costumam apresentar movimentos involuntários, posturas anormais entre outras características (COELHO *et al.*, 2006). De acordo com Coelho *et al.* (2006) o impacto das alterações motoras da

doença de parkinson pode ser medido para exploração motora da escala unificada de avaliação da doença de parkinson (UPDRS).

Neste contexto, prever essa alteração motora é de extrema importância, pois se os sintomas desta doença forem diagnosticados no início e iniciado o tratamento, isso pode possibilitar ao paciente uma melhor qualidade de vida. Uma vez que a medicação se tomada no estágio inicial da doença oferece um maior alívio dos sintomas. Podendo ser um dos primeiros sintomas da doença de parkinson a deficiência vocal ( Little *et al.*, 2009).

Desta forma, a extração de características em bases de dados pode ser um excelente auxiliar no processo de tomada de decisão nesta área de atuação, sendo esta extração de informação realizada por técnicas de inteligência computacional. Entre estas técnicas pode ser realizado um aprendizado supervisionado ou um aprendizado não-supervisionado para conseguir informações sobre estes problemas. O aprendizado supervisionado apresenta a característica de o processo de aprendizado ser realizado com auxílio de um supervisor, logo é necessário um conjunto de dados de pares de entrada e saída, onde o conjunto de saída (variável dependente, por exemplo) deve estar associado ao conjunto de entrada (BARRETO, 1999). Já o aprendizado não-supervisionado tem como característica não necessitar de um agente externo que reconheça os padrões dos dados (BARRETO, 1999).

Neste cenário, entre as técnicas de Inteligência computação não-supervisionadas existe o agrupamento que tem como objetivo formar grupos com uma maior similaridade e os diferentes grupos formados devem apresentar uma maior dissimilaridade (JAIN and MURTY, 1999). Dentre os algoritmos de agrupamento existem diferentes tipos desses algoritmos como os particionais, hierárquicos. Nos algoritmos particionais pode-se destacar o K-means por ser simples e amplamente utilizado (PELLEG *et al.*, 2000).

Já entre as técnicas supervisionadas existe a regressão. Segundo Montgomery (2012) modelos de regressão têm por objetivo analisar as relações entre as variáveis. As técnicas de regressão podem ser classificadas em paramétricas e não paramétricas. Dentro as técnicas paramétricas existem a regressão linear simples que faz a relação de uma variável (X) com a variável dependente (Y) utilizando o método de estimação dos mínimos quadrados (MONTGOMERY, 2012).

Porém algumas destas técnicas em sua versão básica apresentam certa fragilidade, por isso abordagens híbridas podem vir a melhorar o resultado da estimação em bases de dados, uma vez que combinam as vantagens dos métodos envolvidos (POOJA, 2015). Neste contexto, este artigo tem por objetivo desenvolver uma abordagem híbrida que utilize uma combinação de aprendizado supervisionado e não-supervisionado no contexto de mineração de dados para prever as pontuações motoras da doença de Parkinson, essa predição é realizada utilizando UPDRS. Para isso foram utilizados o algoritmo de agrupamento K-means e a regressão linear simples.

Este trabalho está organizado em cinco seções. A seção 2 apresentará a fundamentação teórica, na seção 3 será apresentada a metodologia que será utilizada neste trabalho, na seção 4 os resultados e a discussão do trabalho, por fim, na seção 5 serão apresentadas as considerações finais do mesmo.

## 2. Fundamentação Teórica

Esta seção apresentará conceitos fundamentais para entendimento deste trabalho, deste modo os conceitos apresentados serão sobre regressão linear e o algoritmo de agrupamento K-means. As subseções a seguir apresentarão estes conceitos.

### 2.1. Regressão Linear

A regressão linear é uma técnica de regressão paramétrica que tem como objetivo ver a relação de uma variável dependente (Y) com variáveis independentes (X) (MONTGOMERY, 2012). O modelo de regressão linear simples é apresentado na equação (1)

$$y_i = \alpha + \beta x_i + \epsilon_i \quad (i = 1, 2, \dots, n) \quad (1)$$

Onde o  $\alpha$  representa o intercepto da reta com o eixo dos y,  $\beta$  o coeficiente de inclinação da reta e  $\epsilon$  o erro. A equação (2) apresenta a o cálculo do y estimado.

$$\hat{y}_i = \alpha + \beta x_i \quad (2)$$

O erro é a diferença entre o verdadeiro y e o y estimado, é apresentado na equação (3)

$$\epsilon_i = y_i - \hat{y}_i \quad (3)$$

O método de estimação utilizado por essa regressão é o método de mínimos quadrados. O método de mínimos quadrados busca fazer que a soma dos erros quadrados seja a menor possível. E é apresentado na equação (4)

$$\sum_{i=1}^n \epsilon^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

### 2.2. Algoritmo de Agrupamento K-means

No algoritmo de agrupamento k-means para fazer o agrupamento na base de dados é necessário a definição do número de grupos (k) que deseja-se formar, para que sejam gerados de forma aleatório os centróides do algoritmo e formados os grupos em torno desses centróides (JAIN and MURTY, 1999).

Para formar os grupos em volta de cada centróide pode-se utilizar a distância euclidiana, como apresenta a equação (5)

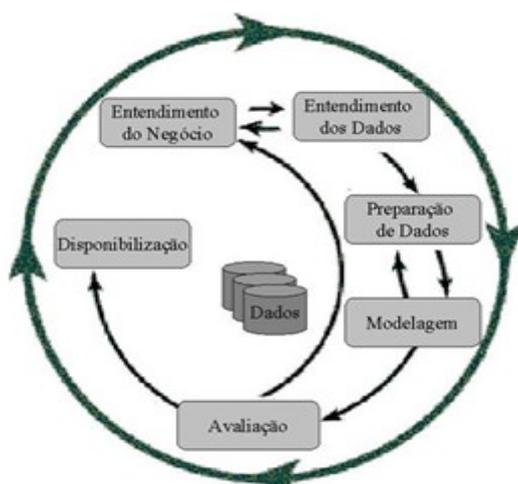
$$\arg \min_{c_i \in C} \text{dist}(c_i, x)^2 \quad (5)$$

A atualização dos centróides, isso é feita levando em consideração a média de todos os pontos de cada cluster, como apresenta a equação (6)

$$c_i = \frac{1}{|s_i|} \sum_{x_i \in s_i} x_i \quad (6)$$

## 3. Metodologia

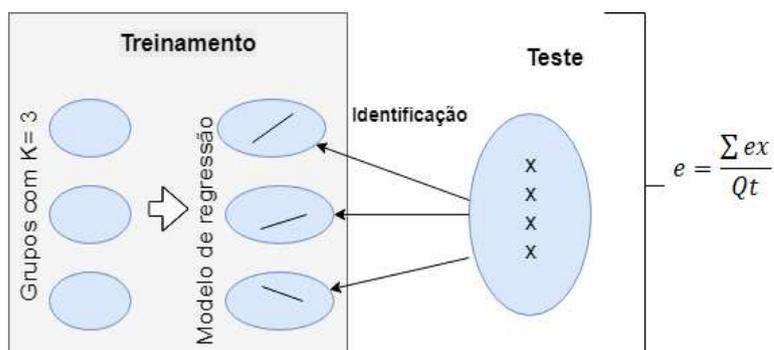
A metodologia que será utilizada neste trabalho é o CRISP-DM (Cross Industry Standard Process for Data Mining). A CRISP-DM é uma das metodologias mais utilizadas no contexto de mineração de dados. A metodologia consiste em seis etapas, a saber: entendimento do negócio, entendimento dos dados, preparação dos dados, modelagem, avaliação e, por fim, disponibilização (Chapman et al, 2000). Assim como apresenta a Figura 1.



**Figura 1. Metodologia CRISP-DM Adaptado de (Chapman et al, 2000)**

A descrição das etapas da metodologia será apresentada a seguir:

- Entendimento do negócio: essa etapa consiste em entender qual o problema que deve ser estudado e o que busca resolver na base. Neste trabalho, o problema refere-se à predição de pontuações motoras da doença de parkinson. A base de dados utilizada está disponível publicamente, e é formada por dados reais e composto de medições biomédicas de voz de 42 pessoas com doença de parkinson em estágio inicial. A base contém 5875 instâncias de gravações de voz desses indivíduos dentre um período de seis meses e não contém dados faltosos. Utilizou-se dados de vozes, pois segundo Little *et al.* (2009) um dos primeiros sintomas que identificam a doença de parkinson é a deficiência vocal.
- Entendimento dos Dados: essa etapa tem como objetivo entender quais variáveis devem ser utilizadas, fazendo isso através da descrição das variáveis da base de dados. No trabalho, este processo foi realizado através da correlação entre a variável depende e as variáveis explicativas.
- Preparação dos Dados: nesta etapa deve-se preparar toda a base de dados, realizando procedimentos como padronização dos valores da base, tratamento de dados faltosos. No trabalho foi realizada a padronização dos dados entre os valores zero e um.
- Modelagem: essa etapa tem como objetivo selecionar as técnicas de mineração de dados que irão ser utilizadas na resolução do problema. No presente trabalho, será utilizada a técnica de regressão linear simples, e foi desenvolvida uma abordagem híbrida que utiliza a combinação de uma técnica de aprendizado não-supervisionado e com uma técnica de aprendizado supervisionado. A técnica de aprendizado não-supervisionado utilizada foi o K-means e a técnica de aprendizado supervisionado foi a regressão linear simples. A Figura 2 apresenta a abordagem híbrida desenvolvida no presente trabalho.



**Figura 2. Abordagem Híbrida desenvolvida**

Nesta abordagem os dados foram divididos de forma aleatória em um conjunto de treinamento e outro de teste. No conjunto de treinamento foi realizada a execução do algoritmo K-means e os modelos de regressão também foram treinados. Para utilizar o algoritmo K-means é necessário a definição do número de grupos. Neste contexto, segundo Linden (2009) o número de grupos do algoritmo K-means é definido de forma *ad hoc* pelo usuário que irá usar o algoritmo.

Por esse motivo, valor da quantidade de grupos desenvolvida na abordagem foi de três, então foram gerados três grupos utilizando o algoritmo K-means e dentro de cada grupo foi treinado um modelo de regressão. Esse modelo de regressão na Figura 2 é representado pela a reta que estar dentro de cada grupo. Dessa forma, foram treinados três modelos de regressão, um para cada grupo gerado pelo algoritmo K-means.

Posteriormente, para testar a abordagem desenvolvida foi necessário realizar uma identificação dos dados do conjunto de teste. Essa identificação foi feita utilizando a menor distância de cada dado do conjunto de teste em relação aos centróides dos grupos. Logo, os dados do conjunto de teste foram agrupados em relação ao centróide de menor distância.

Desse modo, os dados do conjunto de teste eram testados com o modelo de regressão que apresentassem uma maior similaridade com os dados do conjunto de treinamento. E assim, cada dado do conjunto de teste será testado no modelo de regressão que seja mais adequado. Desse modo, para obter o erro final desta abordagem o cálculo do erro foi feito com seguinte equação:

$$e = \frac{\sum e_x}{Q_t} \quad (7)$$

- Onde,  $e$  representa o erro geral da abordagem;
  - $e_x$  representa o erro de cada  $x$  do conjunto de teste de cada grupo; e
  - $Q_t$  representa a quantidade de grupos do teste.
- Avaliação: nesta etapa são analisados os resultados dos experimentos realizados. No trabalho, a avaliação foi feita através média e desvio padrão, assim como teste de hipóteses dos erros de predição obtidos (erro médio quadrado).
  - Disponibilização: nesta etapa é realiza-se a disponibilização para indústria de todas as informações obtidas nas etapas de modelagem e avaliação. No trabalho,

esta etapa não foi realizada, uma vez que essa metodologia é uma metodologia desenvolvida para indústria e que quando utilizada para trabalhos na área de mineração de dados é bem aceita até etapa de avaliação.

#### 4. Resultados e Discussão

Esta seção apresentará os resultados e discussão deste trabalho. As subseções a seguir apresentarão sobre a análise descritiva da base de dados, a realização dos experimentos do presente trabalho e, por fim, os testes estatísticos realizados.

##### 4.1 Análise Descritiva das variáveis do estudo

A base de dados contém 5875 instâncias, 23 variáveis, sendo duas variáveis dependentes, a saber: pontuação motoras da doença de parkinson e pontuações total da doença de Parkinson. Deste modo, a variável pontuações total da doença de parkinson foi excluída da base, uma vez que esse estudo tem por objetivo analisar a pontuação motoras da doença de parkinson. Foi realizada uma padronização com os valores da base, a padronização realizada foi entre zero e um. A correlação entre as variáveis da base de dados foi realizada em todas as variáveis, desta forma, a variável que obteve o maior grau de correlação com a variável dependente foi a variável idade, que obteve uma correlação de 0.274, sendo assim uma correlação positiva. E foi a variável utilizada tanto no modelo de regressão linear simples, quanto na abordagem híbrida. A Figura 3 apresenta a correlação entre a variável idade e a variável pontuação motora.

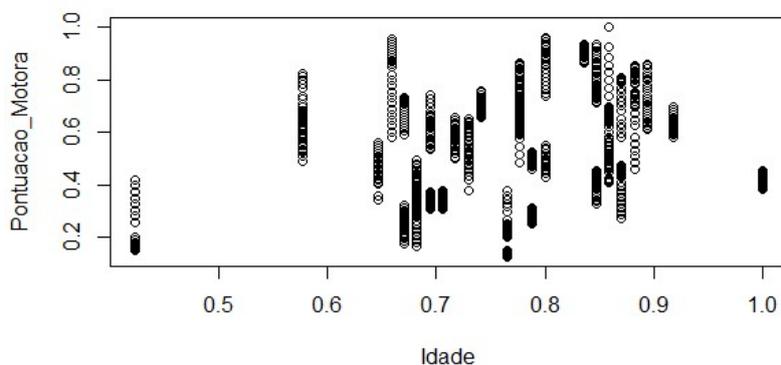


Figura 3. Correlação entre a variável idade e pontuações motoras da doença de Parkinson

##### 4.2 Realização dos Experimentos

Os experimentos deste trabalho foram realizados seguindo o método de simulação de Monte Carlo, desta forma, foram realizadas mil iterações para se obter os erros de predição. Os dados foram particionados de forma aleatória, e foram divididos em um conjunto de treinamento (contém 70% dos dados) e em um conjunto de teste (contém 30% dos dados). A partir do treinamento do modelo de regressão linear simples e da abordagem híbrida desenvolvida foram obtidos os erros de predição das 1000 iterações. O erro de predição utilizado no trabalho é o erro médio quadrado. Por fim, foi calculado a média e desvio padrão dos erros. Assim como é apresentado na Figura 4 o pseudocódigo utilizado no presente trabalho.

- 1: **Utilize** a base de dados
- 2: **Requerer** MC = 1000
- 3: **Para** todo  $i$  1 até MC faça:
  - 3.1 **Particione** aleatoriamente o conjunto de dados simulados em conjunto de Treinamento (70% dos dados) e de teste (30% dos dados)
  - 3.2 **Aplique** o modelo de Regressão Linear para o conjunto de treinamento
  - 3.3 **Aplique** a abordagem proposta para o conjunto de treinamento
  - 3.4 **Calcule** o erro médio quadrado da Regressão Linear e da abordagem
- 4: **Fim** para
- 5: **Calcule** a média e desvio padrão dos erros (erro médio quadrado).

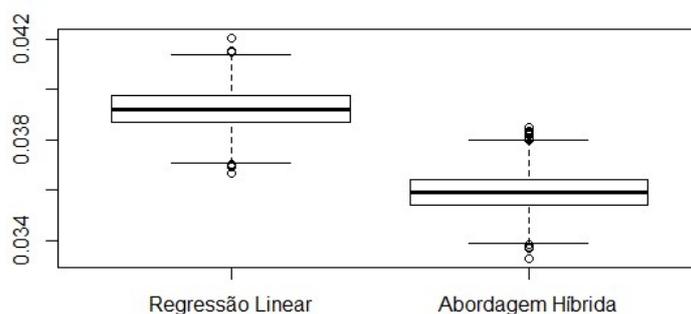
**Figura 4. Pseudocódigo do algoritmo utilizado no trabalho**

A Tabela 1 apresenta as médias e os desvios padrões dos erros obtidos pelos experimentos realizados. Como pode-se observar os desvios padrões tanto do modelo de regressão linear quanto da abordagem híbrida são similares demonstrando que existe uma similaridade na variabilidade de ambos os algoritmos. Porém a média dos erros da abordagem é menor que da regressão linear simples.

**Tabela 1. Média e desvio padrão do modelo de regressão linear e da abordagem híbrida**

Algoritmo	Média	Desvio Padrão
Regressão Linear	0.03922218	0.000819834
Abordagem Híbrida K=3	0.03592167	0.000808044

A Figura 5 apresenta o boxplot destes erros de predição da simulação de monte carlo realizada, e pode-se verificar que a abordagem híbrida obtém menor mediana, assim como similaridade na variabilidade amostral com a regressão linear. Pode-se então perceber que abordagem apresenta um modelo consistente em relação ao erro médio quadrado.



**Figura 5. Boxplot do modelo de Regressão Linear de Abordagem Híbrida**

As Figuras 6 e 7 apresentam os histogramas da distribuição dos erros do modelo de regressão linear e da abordagem, respectivamente. Podendo verificar que os erros obtidos com a simulação de Monte Carlo, tanto da regressão linear quanto da abordagem híbrida, não demonstram evidências que seguem uma distribuição normal. Desse modo foram realizados testes estáticos para conseguir melhores evidências sobre a normalidade desses dados. Estes testes serão apresentados na próxima subseção.

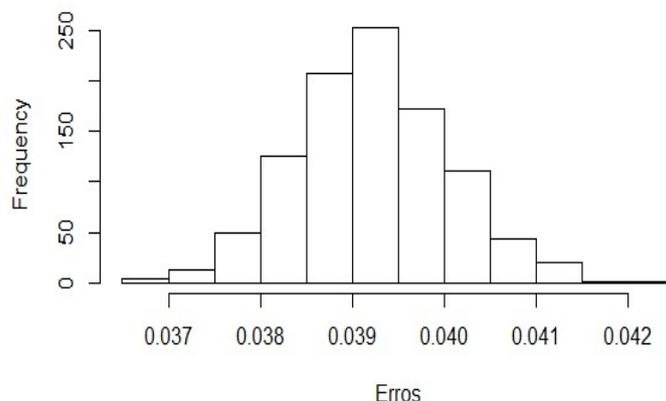


Figura 6. Histograma da distribuição dos Erros da Regressão Linear

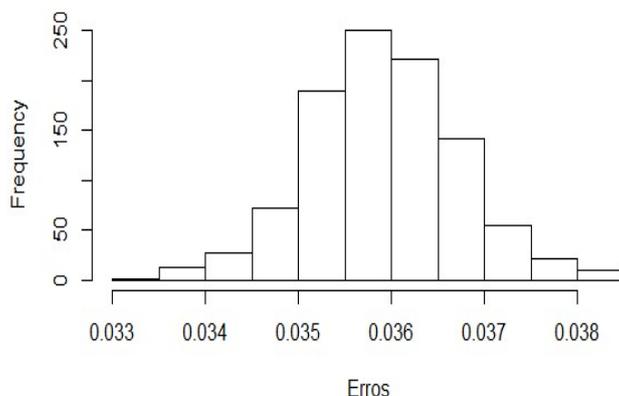


Figura 7. Histograma da distribuição dos erros da Abordagem Híbrida

### 4.3 Testes Estatísticos Realizados

Os teste estatísticos realizados foram o de kolmogorov-simirnov e teste hipótese. O teste de kolmogorov-simirnov foi utilizado para verificar se os erros seguem uma distribuição normal, visto que não houve como concluir pelos os histogramas já apresentados. A Tabela 2 apresenta o teste de normalidade realizado ao nível de significância de 5% para a regressão linear e a abordagem híbrida.

Tabela 2. Teste de Normalidade (Kolmogorov-simirnov)

Algoritmo	Valor do p-value
Regressão Linear	2.2e-16
Abordagem Híbrida	2.2e-16

Pelos valores de p-value obtidos nesta estatística e considerando o nível de significância de 5% conclui-se que os dados não seguem uma distribuição normal. Desta forma, como foi comprovado que erros de predição não seguem uma distribuição normal, foi utilizado o teste de wilcoxon no teste de hipótese.

O teste de hipótese foi elaborado da seguinte forma:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

Onde,  $\mu_1$  representa a média da abordagem híbrida e  $\mu_2$  a média da regressão linear, desta forma, hipótese nula ( $H_0$ ) definida é que a regressão linear apresenta os erros de predição maior ou igual a Abordagem Híbrida, e a hipótese alternativa ( $H_1$ ) é definida como a abordagem híbrida apresenta erros de predição menores que a regressão linear.

Pelas hipóteses de pesquisa elaboradas foi realizado um teste de hipótese unilateral à esquerda com nível de significância de 5%. O valor do p-value obtido no teste de hipótese foi de  $2.2e-16$ . Pode-se concluir que a hipótese nula ( $H_0$ ) foi refutada, logo há evidências que a abordagem híbrida apresenta erros de predição menores que a regressão linear, uma vez que pode-se obter evidência sobre a hipótese alternativa ( $H_1$ ).

Desta forma, pode-se verificar através de dados simulados que a abordagem híbrida desenvolvida apresenta menores erros de predições que a regressão linear. Demonstrando um modelo mais preciso e consistente. Neste cenário, a abordagem híbrida demonstra maior precisão na predição de pontuações motoras com relação a idade do indivíduo que a regressão linear.

## 5. Considerações Finais

Diante do que foi exposto a abordagem híbrida obteve resultados satisfatórios para a predição de pontuações motoras da doença de Parkinson. Já que combinando as vantagens dos métodos utilizados apresentou erros de predições menores, estando esse fato comprovado estatisticamente através do teste de hipótese.

E como na área médica a precisão no diagnóstico é de extrema importância, logo a abordagem híbrida pode prever de forma mais precisa, se comparado com a regressão linear. Deste modo, a predição das pontuações motoras da doença de parkinson utilizando a abordagem híbrida desenvolvida é mais eficiente para os dados do estudo. A utilidade da predição de pontuações motoras da doença de parkinson se dar no auxílio ao tratamento das pessoas que tem essa doença.

Adicionalmente, compreende-se que a construção de uma abordagem híbrida utilizando aprendizado supervisionado e não-supervisionado pode ser um excelente auxiliar no processo de extração de informação em bases de dados em problemas reais. No presente trabalho, essa extração de informação refere-se a prever as pontuações motoras da doença de parkinson, ou seja, depois do modelo treinado pode ser estimado essas pontuações de outros indivíduos.

Pelos fatos apresentados, pretende-se fazer o aperfeiçoamento desta abordagem através de métodos mais rebuscados, como utilizar uma regressão linear múltipla, ao usar uma regressão simples, além de testar uma maior variação de quantidade de grupos, formando diferentes cenários de testes. E assim, conseguir maiores evidências sobre esta abordagem, como também a precisão da predição de pontuações motoras da doença de Parkinson.

## Referências

BARRETO, Jorge Muniz. Inteligência Artificial no limiar do século XXI. Florianópolis: PPP edições, v. 97, 1999.

- COELHO, M. S.; PATRIZZI, L. J.; OLIVEIRA, A. P. R. Impacto das alterações motoras nas atividades de vida diária na Doença de Parkinson. *Neurociências*, v. 14, n. 4, p. 178-81, 2006.
- FERRAZ, Henrique Ballalai; BORGES, Vanderci. Doença de parkinson. *Revista brasileira de medicina*, v. 59, n. 4, p. 207-219, 2002.
- JAIN, Anil K.; MURTY, M. Narasimha; FLYNN, Patrick J. Data clustering: a review. *ACM computing surveys (CSUR)*, v. 31, n. 3, p. 264-323, 1999.
- LANA, R. C. et al. Percepção da qualidade de vida de indivíduos com doença de Parkinson através do PDQ-39. *Revista Brasileira de Fisioterapia*, v. 11, n. 5, 2007.
- LINDEN, Ricardo. Técnicas de agrupamento. *Revista de Sistemas de Informação da FSMA*, v. 4, p. 18-36, 2009.
- Little MA, McSharry PE, Hunter EJ, Ramig LO (2009), 'Suitability of dysphonia measurements for telemonitoring of Parkinson's disease', *IEEE Transactions on Biomedical Engineering*, 56(4):1015-1022
- MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, G.Geoffrey. *Introduction to linear regression analysis*. John Wiley & Sons, 2012
- SHAKOOR, Md Tahmid et al. Agricultural production output prediction using Supervised Machine Learning techniques. In: *Next Generation Computing Applications (NextComp)*, 2017 1st International Conference on. IEEE, 2017. p.182-187.
- PELLEG, Dan et al. X-means: Extending k-means with efficient estimation of the number of clusters. In: *Icml*. 2000. p. 727-734.
- Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rudiger Wirth. *Crisp-dm 1.0 step-by-step data mining guide*. CRISP-DM Consortium, 2000.
- POOJA, M. R.; PUSHPALATHA, M. P. A hybrid decision support system for the identification of asthmatic subjects in a cross-sectional study. In: *Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, 2015 International Conference on. IEEE, 2015. p. 288-293.

<https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring>, último acesso em 02 de outubro de 2018

## **Estudo Experimental com as Ferramentas CMSR *Data Miner*, *KnowledgeSTUDIO* e *RapidMiner***

**Breno A. Campos<sup>1</sup>, Paulemir G. Campos<sup>1</sup>**

<sup>1</sup> Universidade de Pernambuco (UPE) – Campus Caruaru  
Caruaru – PE – Brasil

brenoaraujocampos10@hotmail.com, pgc@poli.br

**Resumo.** *Este artigo aborda um estudo experimental com as ferramentas de Mineração de Dados (MD) CMSR Data Miner, KnowledgeSTUDIO e RapidMiner. Objetivando verificar em que problemas poderão ser aplicadas, incluindo se podem lidar com Big Data, essas ferramentas foram usadas com três bases de dados. Adicionalmente, almeja-se também com esse trabalho contribuir com gestores organizacionais em seus processos de tomada de decisão, ajudando-os a aplicar ferramentas de MD mais adequadas no seu trabalho cotidiano.*

### **1. Introdução**

Em busca de criar programas de computador que solucionem satisfatoriamente problemas do mundo real, estudiosos e pesquisadores em Inteligência Artificial (IA) demonstraram desde os anos 70 do século XX a necessidade de incluir nesses sistemas muito conhecimento específico sobre o problema [Russel e Norvig 2013].

Porém, o emprego de ferramentas automáticas para obtenção de dados associado à maturidade dos sistemas de bancos de dados com grandes quantidades de dados originários de bases de dados e de vários outros mecanismos de armazenamento provocaram uma verdadeira explosão de dados. Por outro lado, essa vastidão de dados contém muito pouco conhecimento útil, incentivando, assim, a origem da área de pesquisa conhecida por Mineração de Dados (MD), integrante da IA [Han et al. 2012].

Além disso, a partir do início do século XXI, o enorme interesse em desenvolvimento e difusão de aplicações para *Cloud Computing* (Computação em Nuvem), favoreceu e muito a geração da era do *Big Data*, constituído principalmente por dados não estruturados provenientes de redes sociais [Silva e Campos 2015].

Ainda é válido ressaltar que no setor público ou privado, é evidente o anseio por tecnologia apropriada, conduzindo a obter-se uma gestão empresarial eficiente. Logo, num cenário competitivo mundial entre as organizações, é notório o acréscimo de empresas que adotam Tecnologia da Informação (TI), *Data Mining* (DM), *Business Intelligence* (BI) e inovação tecnológica em suas tomadas de decisões e processos

organizacionais [Laudon e Laudon 2010], [Witten et al. 2011], [Han et al. 2012], [Salgado et al. 2015], [Sousa Neto et al. 2015], [Silva e Campos 2015], [Mendonça 2015].

Desta forma, este artigo descreve um estudo experimental com três ferramentas de MD que contribuam na tomada de decisão nas mais variadas áreas do conhecimento humano. Adicionalmente, objetiva-se destacar características dessas ferramentas, como em que problemas podem ser utilizadas, em que sistema operacional, licença de *software* proprietária ou pública, que algoritmos de Mineração de Dados, formas de Visualização dos Dados de saída, e até se é possível usá-las com *Big Data*.

Ademais, este artigo contém: A segunda seção, tratando de ferramentas de MD; A terceira seção especificando os experimentos com três ferramentas de MD; Na quarta seção ilustram-se os resultados obtidos; E, na quinta e última seção, são efetuadas as conclusões desta pesquisa.

## 2. Ferramentas de Mineração de Dados

Extrair conhecimento útil de forma eficiente de grandes bases de dados não é uma tarefa fácil. Assim, muitos centros de pesquisa e organizações somam esforços no desenvolvimento de ferramentas de Mineração de Dados [Witten et al. 2011], [Han et al. 2012], [Sousa Neto et al. 2015].

Hoje em dia, há diversas ferramentas para apoiar a descoberta de conhecimento de enormes bases de dados em várias áreas do conhecimento humano, desejando, portanto, ajudar gestores de organizações científicas e/ou empresariais em seus processos de tomada de decisão. Evidenciam-se as ferramentas de Mineração de Dados a seguir: Oracle *Data Mining*<sup>1</sup>, IBM SPSS<sup>2</sup>, SAS *Enterprise Miner*<sup>3</sup>, TANAGRA [Rakotomalala 2005], WEKA [Hall et al. 2009], Orange [Demsar et al. 2013], KNIME<sup>4</sup> e Miner3D<sup>5</sup>.

Destaca-se que as ferramentas WEKA, Orange e KNIME, todas com licença pública de *software*, já foram analisadas por Melo et al. (2016). Quanto às ferramentas Oracle *Data Mining*, IBM SPSS e SAS *Enterprise Miner* (estas com licença de *software* proprietária) também já foram investigadas por Luna et al. (2017). Logo, objetivando continuar esse estudo, serão consideradas para verificação detalhada as ferramentas de Mineração de Dados CMSR *Data Miner*, *KnowledgeSTUDIO* e *RapidMiner*.

---

<sup>1</sup> Disponível em: <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/overview/index.html>, acessada em setembro de 2018.

<sup>2</sup> Disponível em: <http://www.ibm.com/analytics/us/en/technology/spss/>, acessada em setembro de 2018.

<sup>3</sup> Disponível em: [http://www.sas.com/en\\_us/software/analytics/enterprise-miner.html](http://www.sas.com/en_us/software/analytics/enterprise-miner.html), acessada em setembro de 2018.

<sup>4</sup> Disponível em: <https://www.knime.org/>, acessada em setembro de 2018.

<sup>5</sup> Disponível em: <http://www.miner3d.com/>, acessada em setembro de 2018.

## 2.1. CMSR Data Miner

CMSR *Data Miner*<sup>6</sup> é uma ferramenta de Mineração de Dados (MD), que possui mecanismo de regras, de *shell* especializado com modelagem preditiva. Foi criado para aplicativos de negócios com ênfase em banco de dados, inclusive pode lidar com *Big Data*. Esse minerador suporta algoritmos necessários para banco de dados corporativos, inclusive contém métodos de preparação e formatação de dados, e técnicas e algoritmos de MD, tais como: Redes Neurais Artificiais (RNA), Árvores de Decisão (usam o coeficiente de Cramer como critério padrão de divisão dos nodos), Modelos de Regressão, de Agrupamentos (*clustering*) e de Associação (algoritmo Apriori).

CMSR também apresenta recursos de matriz de confusão, análise detalhada de pontos de acesso (*hotspot*) e gráficos de visualização de dados em 2D e 3D (duas e três dimensões). Além disso, a modelagem preditiva desse minerador possui como base o aprendizado de máquinas (*machine learning*), fazendo com que esse programa apresente características de eficiência e de inteligência.

Por fim, CMSR utiliza um subconjunto da linguagem de banco de dados (SQL99 como é conhecida) que é derivada do SQL (*Structured Query Language*), facilitando o aprendizado dessa ferramenta de MD para os usuários com experiência em SGBD (Sistemas de Gerenciamento de Banco de Dados).

## 2.2. KnowledgeSTUDIO

*KnowledgeSTUDIO* é um de vários pacotes de *software* desenvolvidos pela empresa Angoss<sup>7</sup>, que permite modelagem avançada para apoio à tomada de decisão de alto desempenho. Em outras palavras, *KnowledgeSTUDIO* provê modelagem preditiva avançada incluindo desenvolvimento e análise de *scorecard*, bem como, módulo de preparação e formatação de dados e o emprego de algoritmos de Mineração de Dados (MD) do paradigma supervisionado, tais como, Redes Neurais Artificiais, Árvores de Decisão, Regressão Linear e Regressão Logística; e do paradigma não-supervisionado, a exemplo de Análise de *Clusters* (Agrupamentos) e Análise Fatorial.

Adicionalmente, apresenta um ambiente de trabalho denominado *workflow* para extração de conhecimento num formato Canvas (estilo de visualização em tela única, como se fosse uma pintura) e semelhante ao fluxo de descoberta de conhecimento de bases de dados (KDD – *Knowledge Discovery in Databases*).

Além disso, possibilita o uso das técnicas de análise de algoritmos *Bagging* e *Boosting*, permite integração com arquivos textos (*flatfiles*), planilhas Excel, SPSS, SAS, *framework* Hadoop [Sousa Neto e Lins Filho 2014] e banco de dados via ODBC (*Open Data Base Connection*), suporta *Big Data* (versão *KnowledgeENTERPRISE*<sup>8</sup>), possibilita ainda integração com as linguagens de programação R e Python, bem como

<sup>6</sup> Disponível em: <https://www.roselladb.com/starprobe.htm>, acessada em agosto de 2018.

<sup>7</sup> Disponível em: <http://www.angoss.com/>, acessada em agosto de 2018.

<sup>8</sup> Disponível em: <http://www.angoss.com/predictive-analytics-software/software/knowledgeenterprise/>, acessada em agosto de 2018.

várias formas de visualização dos dados de saída, tais como, gráficos, tabelas e relatórios de fácil interpretação.

### 2.3. *RapidMiner*

*RapidMiner*<sup>9</sup> é uma plataforma de *software* que permite análise de dados, unificando preparação de dados, algoritmos de *machine learning*, e emprego de modelagem preditiva.

O *RapidMiner* é constituído de três versões: *RapidMiner Studio (stand-alone)*, *RapidMiner Server* (arquitetura cliente-servidor) e *RapidMiner Hadoop* (apropriada para *Big Data*). Apresenta um ambiente de trabalho visual baseado em *workflow*. Possibilita acesso a dados estruturados e não-estruturados, tais como texto, imagens e arquivos multimídia. Suporta arquivos CSV (*Comma Separated Values*), planilha Excel, SAS, Stata<sup>10</sup>, ARFF (*Attribute Relation Format File*, formato de arquivo padrão do WEKA), e acesso a banco de dados Access, SQL Server, Oracle, MySQL, Postgres, entre outros. Também pode-se acessar bancos de dados NoSQL, tais como MongoDB e Cassandra.

Possui vários métodos para preparação dos dados (incluindo tratamento de *missing values*, seleção de atributos e transformação de dados), pode ser aplicado a problemas de classificação, regressão, agrupamentos (*clustering*) e associação. Tem disponíveis muitos algoritmos de *machine learning* prontos para uso, tais como, Redes Neurais Artificiais (por exemplo, MLP – *Multi Layer Perceptron* e SOM – *Self Organizing Maps*), Árvores de Decisão (como ID3 e C4.5), *Clustering* (por exemplo, K-Means) e Regressão (como, Linear e Logística). Com o *RapidMiner* é possível executar *scripts* em linguagens de programação R e Python. Por fim, pode-se visualizar os dados de saída por meio de diversos tipos de gráficos e tabelas.

### 3. Experimentos com CMSR Data Miner, KnowledgeSTUDIO e RapidMiner

Com as bases de dados de Câncer de Mama<sup>11</sup>, de Íris<sup>12</sup> e de Vinho<sup>13</sup>, obtidas do repositório de Aprendizado de Máquina da Universidade da Califórnia [Lichman 2013], foram efetuados experimentos com as ferramentas CMSR Data Miner (versão 12.1), KnowledgeSTUDIO (versão 10.7) e RapidMiner Studio (versão 9.0.002) no ambiente do Sistema Operacional Windows 8 de 64 bits.

Focando problemas de Classificação e de Agrupamento (*clustering*), devido às características das bases de dados (atributo alvo ou classe serem categóricos ou

---

<sup>9</sup> Disponível em: <https://rapidminer.com/>, acessada em setembro de 2018.

<sup>10</sup> Disponível em: <https://www.stata.com/>, acessada em outubro de 2018.

<sup>11</sup> Disponível em: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>, acessada em agosto de 2018.

<sup>12</sup> Disponível em: <http://archive.ics.uci.edu/ml/machine-learning-databases/iris/>, acessada em agosto de 2018.

<sup>13</sup> Disponível em: <http://archive.ics.uci.edu/ml/machine-learning-databases/wine/>, acessada em agosto de 2018.

nominais), nesse trabalho experimental verificaram-se funcionalidades de pré-processamento dos dados, algoritmos de Mineração de Dados (MD) aplicados aos tipos de problemas já citados, bem como recursos de visualização de dados, métodos de validação do conhecimento obtido, e se poderiam manipular *Big Data*.

Nos experimentos usando a ferramenta CMSR *Data Miner* (Figura 1) empregaram-se os algoritmos *Decision Tree* (Figura 1a) e *Clustering* (Figura 1b, Rede Neural SOM). Já com *KnowledgeSTUDIO* (Figura 2) verificaram-se os algoritmos *Deep Learning* (método baseado em Redes Neurais Multi Camadas ou Níveis com aprendizado não linear), *Logistic Regression* e *Clustering* (Figura 2b, *K-Means*), bem como o método WOE (*Weight of Evidence Optimizer*, cria automaticamente intervalos de valores para as variáveis numéricas de entrada da base de dados, Figura 2a). E, utilizando *RapidMiner* (Figura 3), aplicaram-se os algoritmos *Decision Tree* (Figura 3a), e, *Naive Bayes* e *Deep Learning* (Figura 3b). No mais, é importante destacar que os parâmetros dos algoritmos de MD utilizados foram os valores padrões (*default*) predefinidos em cada uma das ferramentas de MD avaliadas, exceto em eventual menção explícita na divulgação dos resultados expostos na próxima seção desse artigo.

Por fim, pode-se constatar experimentalmente o desempenho desses algoritmos de MD aplicados em três bases de dados diferentes com essas três ferramentas de MD.

#### 4. Resultados

Após os experimentos executados com as ferramentas CMSR *Data Miner*, *KnowledgeSTUDIO* e *RapidMiner Studio* obtiveram-se os resultados mostrados nas Figuras 1, 2 e 3, e na Tabela 1 mais adiante.

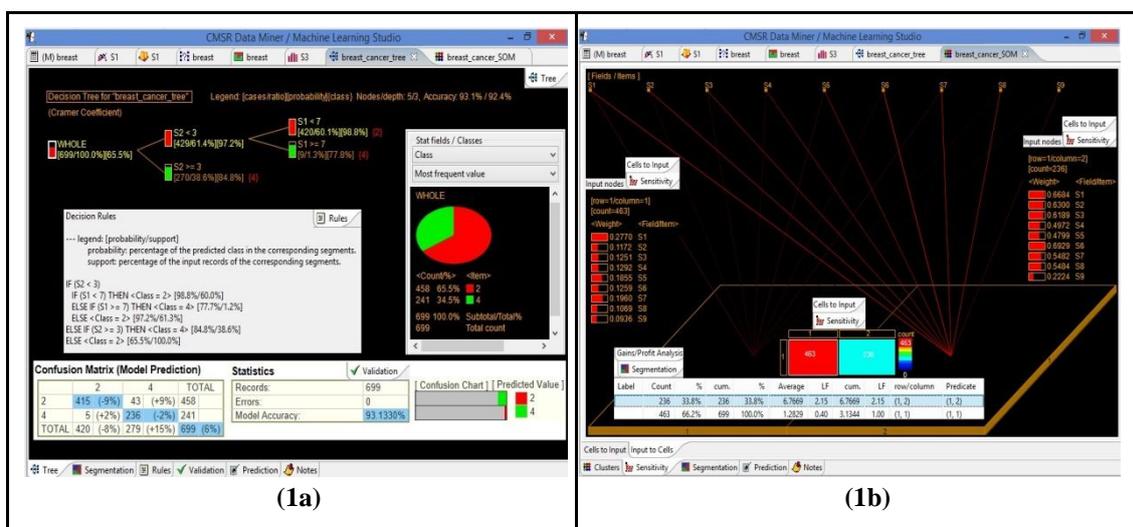
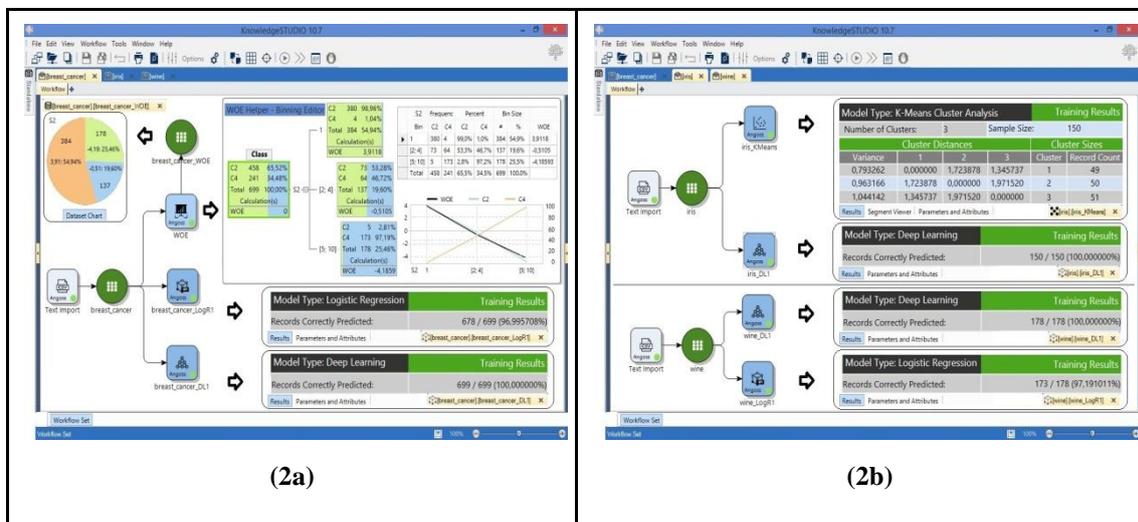


Figura 1. Ferramenta CMSR *Data Miner*. (1a) *Decision Tree*. (1b) *Clustering* (Rede SOM).

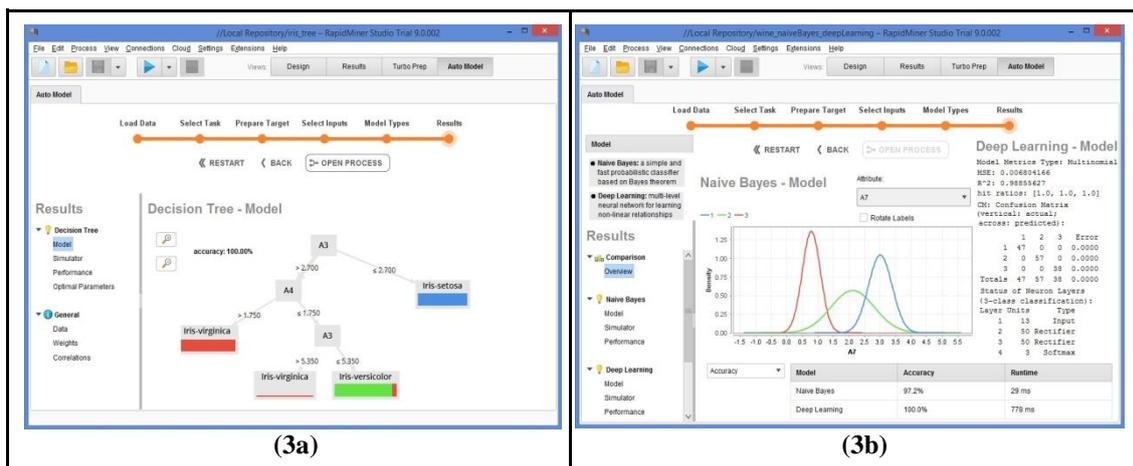
A Figura 1 ilustra o uso da Ferramenta CMSR *Data Miner* com a base de dados de Câncer de Mama (possuindo 699 padrões, 458 da Classe ‘2’ e 241 da Classe ‘4’) aplicando-se os algoritmos *Decision Tree* (Figura 1a, parâmetros: *Node split criteria: Cramer Coefficient*, e *Pruning: Maximum Depth 5*) e *Clustering* (Figura 1b, Rede

Neural SOM, parâmetro *Network: Cols: 2 e Rows: 1*). Essa Árvore de Decisão obteve 93.1% de classificação correta. Já com a Rede Neural SOM foram criados dois *clusters*, o '*cluster (1,1)*' agrupando 463 padrões (Classe '2') e o '*cluster (1,2)*' contendo 236 padrões (Classe '4').



**Figura 2. Ferramenta KnowledgeSTUDIO. (2a) Usando Logistic Regression e Deep Learning (Base de Dados: Câncer de Mama). (2b) Usando K-Means e Deep Learning (Base de Dados: Íris), e Deep Learning e Logistic Regression (Base de Dados: Vinho).**

Já na Figura 2 são apresentados alguns dos recursos gráficos da ferramenta KnowledgeSTUDIO, com o ambiente *workflow* (Figuras 2a e 2b) que facilita a adoção de vários modelos de aprendizado de máquina para extração de conhecimento útil de bases de dados, gerando vários tipos de gráficos e relatórios. Por exemplo, na Figura 2a pode-se visualizar os resultados da aplicação dos métodos WOE (*Weight of Evidence Optimizer*), gerando-se automaticamente para a variável 'S2' os intervalos: [1],[2;4]e[5;10], considerando as duas classes de saída 'C2' e 'C4'), *Logistic Regression* (com 96.9% de classificação correta) e *Deep Learning* (atingindo-se 100.0% de acertos) à base de dados de Câncer de Mama. E na Figura 2b, em relação à base de dados Íris (contendo 150 padrões distribuídos igualmente em três classes), foram empregados os algoritmos de *Clustering* (K-Means, parâmetro *Number of clusters: 3*; obtendo-se '*Cluster 1*': 49 padrões, '*Cluster 2*': 50 padrões, e '*Cluster 3*': 51 padrões) e *Deep Learning* (classificação correta em 100.0%); e considerando à base de dados de Vinho (com 178 padrões, Classe-1: 59 padrões; Classe-2: 71 padrões; Classe-3: 48 padrões), adotaram-se os algoritmos *Deep Learning* (100.0% de classificação correta) e *Logistic Regression* (obtendo-se 97.2% de acertos).



**Figura 3. Ferramenta RapidMiner. (3a) Usando Decision Tree (Base de Dados: Íris). (3b) Usando Naive Bayes e Deep Learning (Base de Dados: Vinho).**

Por sua vez, a Figura 3 exibe os resultados com a ferramenta *RapidMiner*. Empregando o algoritmo *Decision Tree* (Figura 3a) na base de dados de Íris, obteve-se um desempenho de classificação de 100.0% de acertos. Em seguida foram aplicados os algoritmos *Naive Bayes* (modelo Estatístico) e *Deep Learning* à base de dados de Vinho, atingindo-se os índices de acertos de classificação de 97.2% e 100.0% (Figura 3b), respectivamente.

**Tabela 1. Propriedades das Ferramentas de Mineração de Dados CMSR Data Miner, KnowledgeSTUDIO e RapidMiner**

Propriedades	Ferramentas de Mineração de Dados		
	<i>CMSR Data Miner</i>	<i>KnowledgeSTUDIO</i>	<i>RapidMiner</i>
Sistema Operacional	Windows, Linux e Mac OS X.	Windows e Linux.	Windows, Linux e Mac OS X.
Licença	Proprietária*	Proprietária*	Proprietária*
Suporte a <i>Big Data</i>	Sim.	Sim.	Sim.
Tipos de Problemas	Classificação, Regressão, Agrupamentos, Associação.	Classificação, Regressão, Agrupamentos.	Classificação, Regressão, Agrupamentos, Associação.
Pré-Processamento dos Dados	Transformação de Dados.	Missing Values, Extração, Manipulação, Transformação de Dados.	Missing Values, Seleção de Atributos, Transformação de Dados.

\*Disponibiliza licença grátis para fins educacionais.

Tabela 1 (Continuação). Propriedades das Ferramentas de Mineração de Dados  
CMSR Data Miner, KnowledgeSTUDIO e RapidMiner

Propriedades	Ferramentas de Mineração de Dados		
	CMSR Data Miner	KnowledgeSTUDIO	RapidMiner
Árvores de Decisão	<i>Decision Tree (Default: Coeficiente de Cramer)</i>	<i>Decision Tree, Strategy Tree, Random Forest.</i>	ID3, C4.5, <i>Random Forest, Gradient Boosted Trees.</i>
Regras de Decisão	Sim.	Não.	Não.
Redes Neurais	MLP, SOM.	<i>Deep Learning.</i>	<i>Perceptron, MLP, Deep Learning, SOM.</i>
Regressão	Linear, Exponencial, Logística.	Linear, Logística, <i>Partial Least Squares (PLS), Constrained Logistic, ARIMA (Auto Regressive Integrated Moving Average).</i>	Linear, Logística, <i>Kernel Logística, Vetor Linear, Polinomial Local, e Polinomial.</i>
Meta-Aprendizado	<i>Stacking.</i>	<i>Bagging, Boosting.</i>	<i>Bagging, Boost e Stacking.</i>
Agrupamentos	Redes SOM.	<i>Cluster Analysis (K-Means, Expectation-Maximization).</i>	<i>Kernel K-Means, K-Means, Support Vector, X-Means e Redes SOM.</i>
Regras de Associação	Apriori.	Não.	Apriori.
Estatísticos (Bayes)	Não.	<i>Naive Bayes.</i>	<i>Naive Bayes, Kernel Naive Bayes.</i>
Visualização dos Dados	Diversos Gráficos e Tabelas.	Diversos Gráficos, Tabelas e Relatórios, incluindo <i>Tableau</i> <sup>14</sup> .	Diversos Gráficos, Tabelas e Diagramas.
Avaliação dos Dados	<i>Training-and-Test.</i>	<i>Training-and-Test, Cross-Validation.</i>	<i>Holdout, Cross-Validation, Bootstrapping.</i>

<sup>14</sup> Disponível em: <https://www.tableau.com/>, acessada em outubro de 2018.

E, de acordo com os resultados detalhados na Tabela 1, nota-se que:

- Essas ferramentas possuem interface gráfica icônica (vide Figuras 1, 2 e 3) e amigável de fácil manuseio através de seleção de itens de *menu* e/ou botões icônicos (usabilidade muito superior a *softwares* com interface modo texto ou *prompt* de comandos), executam no Windows e Linux, têm licença proprietária (mas, dispõem de licença gratuita para uso educacional), podem ser aplicadas a problemas de Classificação, Regressão e Agrupamentos, contém métodos para transformação dos dados, possuem vários algoritmos de Mineração de Dados, têm métodos de Avaliação dos Dados, diversas formas para Visualização dos Dados de saída, principalmente gráficos e tabelas, e podem lidar com *Big Data*;
- Apenas CMSR *Data Miner* e *RapidMiner* podem ser usadas em problemas de Associação, executadas no Mac OS X, e possuem o algoritmo de meta-aprendizado *Stacking*;
- Já *KnowledgeSTUDIO* e *RapidMiner* manipulam *Missing Values*, têm os algoritmos de meta-aprendizado *Bagging* e *Boosting*, dispõem de algoritmos estatísticos baseados no Teorema de *Bayes* (por exemplo, *Naive Bayes*), e o modelo de Rede Neural *Deep Learning*;
- E, só CMSR *Data Miner* possui algoritmo de Regras de Decisão.

## 5. Conclusões

Segundo os experimentos e resultados alcançados com as ferramentas CMSR *Data Miner*, *KnowledgeSTUDIO* e *RapidMiner*, conclui-se que os objetivos dessa pesquisa foram atingidos, proporcionando a gestores organizacionais subsídios que apoiam suas decisões, com a escolha e utilização de alguma(s) dessas ferramentas de Mineração de Dados (MD) que seja(m) mais adequada(s) ao contexto de suas empresas.

No entanto, em trabalhos futuros, almeja-se ampliar esses resultados usando outras ferramentas de MD ainda não investigadas por Melo et al. (2016) e Luna et al. (2017), bem como, evidentemente, nem as ferramentas já analisadas neste artigo.

## Referências

- Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M. e Zupan, B. (2013) “Orange: Data Mining Toolbox in Python”, *Journal of Machine Learning Research*, Volume 14, p. 2349-2353.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. e Witten, I. H. (2009) “The WEKA Data Mining Software: An Update”, *SIGKDD Explorations*, Volume 11, Issue 1.
- Han, J., Kamber, M. e Pei, J. (2012) *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 3rd ed., San Francisco.

- Laudon, K. C. e Laudon, J. P. (2010) *Sistemas de Informação Gerenciais*, Pearson Prentice Hall, 9a. ed., São Paulo.
- Lichman, M. (2013) UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>], School of Information and Computer Science, University of California, Irvine.
- Luna, B. G., Barboza, M. F. X., e Campos, P. G. (2017) Estudo Experimental com as Ferramentas Oracle Data Mining, IBM SPSS e SAS Enterprise Miner. In *Anais da Jornada Científica e de Extensão - JCE 2017 da UPE Caruaru*, p. 20-29, Caruaru-PE.
- Melo, A. B., Silva, G. A., Morais, P. L. e Campos, P. G. (2016) Estudo Experimental com as Ferramentas de Mineração de Dados WEKA, Orange e KNIME. In *Anais da Jornada Científica e de Extensão - JCE 2016 da UPE Caruaru*, p. 55-62, Caruaru-PE.
- Mendonça, C. M. C. (2015) *A Computação em Nuvem e as suas Relações com os Arranjos de Governança de TI e o Alinhamento Estratégico*, Tese de Doutorado em Administração, Universidade Federal do Rio Grande do Norte, Natal-RN.
- Rakotomalala, R. (2005) “TANAGRA: Un Logiciel Gratuit pour L’Enseignement et la Recherche”, Actes de EGC’2005, RNTI-E-3, Volume 2, p. 697-702.
- Russel, S. e Norvig, P. (2013) *Inteligência Artificial*, Tradução de: “Artificial Intelligence: A Modern Approach”, Editora Campus, 3a. ed.
- Salgado, C. C. R., Araújo, A. G. e Sousa Neto, M. V. (2015) *Tecnologia da Informação e Vantagem Competitiva: Exploração dos Artigos sob a Ótica da Bibliometria*. In *XVIII SEMEAD’2015, Seminários em Administração*, 16p., Natal-RN.
- Silva, I. M. e Campos, F. C. (2015) *Big Data e Cloud Computing Perspectivas de Uso em Engenharia de Produção: Uma Análise Bibliométrica de 2004-2014*. In *XXXV ENEGEP’2015, Encontro Nacional de Engenharia de Produção*, 19p., Fortaleza-CE.
- Sousa Neto, M. V. de e Lins Filho, M. L. (2014) *Desafios na Era do Big Data: Uma Revisão*. In *XXI SIMPEP’2014, Simpósio de Engenharia de Produção*, 14p., Bauru-SP.
- Sousa Neto, M. V. de, Medeiros Junior, J. V., Anez, M. E. M. e Moraes, E. A. (2015) *Building Information Technology Capabilities: A Case Study of the Development of an Integrated Management System*. In *International Conference of the System Dynamics Society’2015, Proceedings of the 33rd International Conference of the System Dynamics Society*, 12p., Cambridge-USA.
- Witten, I. H., Frank, E. e Hall, M. A. (2011) *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann, 3rd ed.

## Política Urbana em Tela: O Olhar do Espectador Especializado

Douglas Henrique Albuquerque da Costa<sup>1</sup>; Paula Gonçalves da Silva<sup>2</sup>

<sup>1</sup>Universidade de Pernambuco [UPE]

Caruaru– PE – Brasil

[Douglas-albuquerque@hotmail.com](mailto:Douglas-albuquerque@hotmail.com)

<sup>2</sup>Universidade de Pernambuco [UPE]

Caruaru– PE – Brasil

[paula.goncalves@upe.br](mailto:paula.goncalves@upe.br), [paulaemcena@gmail.com](mailto:paulaemcena@gmail.com)

**Abstract:** Cinema has opened space for reflection on social reality. The article analyzed the speeches produced by emancipated spectators [Ranciere, 2012], being journalists, academics and the like, about the representation of the city of Recife in the films of Kleber Mendonça Filho and shows how the dialogue with the past has been suffocated by a vertical and hostile growth.

**Resumo:** O cinema tem aberto espaço para a reflexão sobre a realidade social. O artigo analisou os discursos produzidos por espectadores emancipados [Ranciere, 2012], sendo eles jornalistas, acadêmicos e afins, sobre a representação da cidade do Recife nos filmes de Kleber Mendonça Filho e apresenta como o diálogo com o passado vem sendo sufocado por um crescimento verticalizado e hostil.

### 1 – Introdução

O cinema, também conhecido como sétima arte, historicamente tem agregado mais do que histórias em suas telas, mas também críticas políticas e sociais. Com a chegada das novas tecnologias, o uso do audiovisual tem trazido muito mais do que “apenas imagens e sons” servem de suporte para reflexões sobre aspectos da vida social que são influenciadas tanto pela gestão pública como a privada.

A história do cinema de Pernambuco passou por vários momentos desde o Ciclo do Recife na década de 1920 até o atual Novo Cinema Pernambucano, os quais tiveram a cidade do Recife representada ao longo desse período de diferentes maneiras. Atualmente pode-se notar nas obras de alguns cineastas contemporâneos pernambucanos, uma crítica as recentes transformações na cidade ou ao próprio modelo de gestão urbana da cidade que aparentemente segue uma lógica de progresso tendo o ideal de modernidade como padrão social de referência.

Lino e Nascimento [2017], destacam exemplos de como essas questões urbanísticas foram parar nas telas dos festivais de cinema recifense:

Curiosamente, realizadores de curtas, médias ou longas-metragens desenvolviam um número crescente de documentários, ficções [ou documentários ficcionais ou ficções documentais], nos festivais de cinema e na internet com a temática da cidade e de suas potenciais mazelas, muitos deles iniciando suas carreiras justamente em obras

experimentais e curtas-metragens que tinham a cidade e suas questões emergentes como mote. [LINO e NASCIMENTO, 2007, p. 51]

Dentre as diversas obras do cinema produzido em Pernambuco recentemente, os filmes de Kleber Mendonça Filho chamam atenção pela repercussão nacional e internacional e pela potência política das obras. Citando dois de seus trabalhos: *O Som ao Redor* [presente na lista de melhores filmes pela *New York Times*, 2012], que narra o cotidiano de alguns moradores dentro de seus apartamentos em um bairro composto por prédios; e *Aquarius* [presente no Festival do Cannes e melhores filmes pela *New York Times*, ambos em 2016] que traz Clara, uma jornalista aposentada moradora de um antigo prédio que é alvo de uma construtora que pretende demoli-lo e construir um novo.

As duas obras retratam a nova imagem que a cidade do Recife vêm ganhando, os seus moradores em meio a isso e também em comum um cineasta que utiliza a arte como posicionamento político. Em reportagens podemos encontrar exemplos de fatos reais que se tornam problemáticos e coincidem com os filmes, como o Edifício Caiçara (construído na década de 1930 e totalmente demolido em 2016) e as Torres Gêmeas (construções que destoam da arquitetura do Recife Antigo, bairro onde se localizam).

Destaca-se que o entendimento de política e cinema segue o pensamento do filósofo Rancière [2012], para o qual a política da arte está na possibilidade de interpretação e novas traduções da realidade realizada pelo espectador.

A cidade representada nos filmes e as críticas relacionadas ao desenvolvimento urbano repercutem para além das telas através das diferentes interpretações dos espectadores, dentre eles o que Rancière [2012] chama de espectador especializado, ou seja, aqueles que produzem novos conteúdos a partir da obra original [jornalistas, acadêmicos e afins]. Assim, face ao grande poder que potencializa a mídia e à repercussão das matérias jornalistas de cunho crítico e de artigos acadêmicos buscaremos os discursos produzidos por esses espectadores sobre a representação fílmica da capital pernambucana nas obras de Kleber Mendonça Filho.

Nesse sentido, o presente artigo tem como objetivo geral realizar um estudo crítico sobre a cidade representada nos trabalhos do cineasta Kleber Mendonça Filho, a partir do olhar do espectador especializado em relação a alguns filmes que agregam de forma direta ou indireta questões sobre as políticas urbanas no Recife. Ressalta-se que não pretendemos dar soluções para administração pública da cidade, mas sim refletir sobre problemas sociais identificados no discurso dos espectadores especializados de modo a ampliar o debate sobre a cidade. Destacamos que não será discutido como e quando esses espectadores especializados chegaram às suas conclusões e opiniões, mas sim como essas percepções foram ordenadas em texto. É a partir da análise dos discursos [GILL, 2015] desses espectadores que esse artigo será desenvolvido.

Para uma melhor compreensão do artigo dividimos o texto em quatro partes além dessa introdução: o referencial teórico dedicado ao debate sobre os pressupostos do planejamento urbano moderno, a metodologia de pesquisa, análise dos dados e considerações finais.

## **2- Referencial Teórico: Pressupostos do planejamento urbano moderno**

Quando traçamos uma linha do tempo sobre o desenvolvimento das cidades, nota-se que as primeiras civilizações se estabeleceram perto dos rios para melhor desenvolvimento da agricultura, passando pela idade média, período no qual as cidades concentram-se no comércio até chegar à modernidade, quando a vida nas cidades passa a ser rodeada por indústrias, moradias e vias automotivas. Embora as cidades sejam diferentes, em

cada período histórico, há coincidentes patamares demográficos definidores desse conjunto de cidades nas mais variadas regiões do mundo [FILHO, SERRA]. Nele [o espaço urbano], as sociedades deixam impressas suas marcas, suas obras e as contradições que permeiam a existência humana [RESENDE, 2013].

Com o crescimento das cidades após a revolução industrial uma série de medidas [um exemplo sendo a necessidade de uma malha de transportes] foram tomadas para ordenamento da população, assim a partir desse ponto nota-se que o espaço urbano é formado por várias áreas: o centro da cidade, local de concentração de atividades comerciais, de serviço e de gestão; áreas industriais e áreas residenciais, distintas em termos de forma e conteúdo social; áreas de lazer; e, entre outras, aquelas de reserva para futura expansão. Somando esses fatos anteriormente citados com a descentralização da cidade, resultado principalmente no crescimento da cidade tendo em vista expansão. [CORREA, 1995].

Na cidade industrial capitalista a preocupação com a necessidade e a relação da sociedade com o espaço geográfico, fica mais visível quando nota-se a preocupação com a mobilidade urbana e projetos habitacionais. Porém não foram questões que surgiram repentinamente, o urbanismo moderno desde o final do século XIX já prometia melhorias para a ocupação do espaço. No início século XX, a insalubridade das cidades ampliou o debate sobre a necessidade de ordenamento urbano e segundo Gegner [2016] cinco soluções [ou promessas] foram apontadas pelo urbanismo moderno: solução da questão habitacional; ordem e progresso; separação das funções; maior mobilidade; e dissolução da cidade em uma paisagem urbana.

Ao observar a cidade contemporânea podemos considerar que algumas dessas questões, como mobilidade e a necessidade de construção de moradias suficientes e de boa qualidade, ainda carecem de solução nos dias atuais. Optamos então por nos concentrar, neste artigo, em uma das promessas do urbanismo moderno, identificadas por Gegner [2016], que está mais distantes da observação do dia a dia, mas que ainda interfere nas decisões do planejamento urbano contemporâneo: ordem e progresso. A escolha de apenas desse tema deve-se a ideia de que o progresso das cidades está relacionada a questão do crescimento urbano. A cidade do Recife servirá de exemplos ao discorrer dos pontos propostos.

Como uma forma de concertar a situação deplorável da urbanização das cidades industriais, os modernistas trouxeram como solução o ordenamento voltado para o progresso econômico com a criação de grandes avenidas, bulevares e conjuntos habitacionais ortogonais para facilitar a circulação das pessoas e mercadorias [Gegner,2016]. As moradias construídas nas novas vias passaram a ser disputadas pela elite econômica o que levou a uma grande especulação imobiliária.

Atualmente esse fenômeno de especulação pode ocorrer de mais de uma forma. Saboya [2008], apresenta exemplos como: construção ou reforma de pavimentos, o que oferece posteriormente venda de apartamentos e a adesão de terrenos que podem se valorizar posteriormente. A seguir, podemos ler o que pode levar a valorização de determinadas áreas de uma cidade.

Nas grandes cidades do Brasil e do mundo, este fenômeno se efetiva a partir da criação de novos eixos de circulação, implantação de linhas facilitadoras do deslocamento de pessoas, disponibilização de serviços mais sofisticados e diversificados,

implantação de equipamentos urbanos como parques e promoção de verticalização e de adensamento [RESENDE, 2013, p. 81]

Essa ideia de ordem e progresso pode estar relacionada a um modo de crescimento urbano organizado em que para a construção de empreendimentos é analisado o impacto sobre seu entorno, como vias, calçadas e circulação e também sobre a identidade histórica e cultural do lugar. Porém, Gegner [2016] ressalta que para a construção das grandes avenidas foi necessário destruir marcos históricos das antigas cidades, acabar com a história foi parte do programa de modernização. No Brasil, o progresso vinha com uma ideologia de “desenvolvimentismo” que pregava a necessidade de apagamento do passado escravocrata e rural levando a uma destruição do patrimônio colonial em cidades como Rio de Janeiro, São Paulo e Recife.

Esse processo de modernização, permitiu com que novos prédios surgissem, substituindo os antigos ou sendo vizinhos dos mesmos. Ainda hoje no Recife discute-se a criação de edificações no bairro do Recife [marco inicial da cidade] que vão contra toda uma arquitetura histórica. Momento em que o governo e/ou a iniciativa privada visam o crescimento local e econômico, e os arquitetos não se contrapõem a esse modelo de progresso e fazem seus trabalhos mesmo que representem uma perda (histórica ou cultural) para a cidade. O que resulta em edificações que descaracterizam o conjunto histórico e arquitetônico, representando uma perda irreversível para o patrimônio nacional [LACERDA, 2007]. Cita-se como exemplos contemporâneos da cidade do Recife: Livraria Cultura [construída como extensão do Shopping Paço Alfandega] e os “piers” Mauricio de Nassau e Duarte Coelho [também conhecidos como Torres Gêmeas].

Harvey [2001, p.173], no livro *A Produção Capitalista do Espaço*, diz que “o empreendedorismo enfoca muito mais na economia política do lugar do que o território.” O que pode levar a pensar, que o bem maior da construção é apenas o retorno financeiro deixando de lado o bem estar social, que se refere ao social englobando pessoas, meio ambiente e afins. O pensamento se completa com o apoio que o empreendedorismo recebe de uma parceria “público-privada” [governo e empreiteiras] que visa o retorno econômico mesmo que não seja de imediato, no qual um “disponibiliza” o pedaço de território e o outro faz seu uso.

Essa visão de empreendedorismo com ênfase na economia política do território nos remete ao Projeto Novo Recife, o qual prever a construção de 12 torres de 40 andares, próximo ao Bairro do Recife, no Cais José Estelita. O projeto foi alvo de protestos capitaneados pelo Movimento Ocupe Estelita em relação a modificação da paisagem, utilização do espaço restrito ao público alvo dos incorporadores e sem retorno social para a cidade. O apoio de profissionais de áreas diversas ao Movimento Ocupe Estelita [BUENO, 2018] pode significar como o uso do território é uma preocupação de grande parte dos cidadãos.

Além disso, com esses novos empreendimentos, a cidade restringe sua articulação ambiental e natural. A construção das 12 torres de 40 andares [apartamentos de R\$ 400.000 a R\$ 1.000.000] e estacionamento com porte para 5 mil carros, ficariam de frente ao rio Capibaribe, além de agredir a paisagem do Recife traria problemas ambientais, tal como o impedimento da passagem do vento [tendo em vista a proximidade dos prédios], desconforto térmico nas proximidades e poluição devido a grande circulação de carros.

O incomodo causados pelos altos prédios não se restringem aos bairros mais históricos e avançam pela cidade. As grandes construções que vem sendo realizadas, provocam medo e incomodo [pela presença de congestionamentos e degradação ambiental] até mesmo para os moradores das áreas já acostumados com os edifícios [NUNES, 2008]. Fatos que contribuíram para a criação da Lei 16.719/01 [também conhecida como Lei dos Doze Bairros], a qual estabelece: condições e controles no uso, ocupação e verticalização do solo em 12 bairros da zona norte do Recife.

Se a ideia inicial do urbanismo moderno era que apenas a ordem poderia contribuir para o progresso da cidade, vemos hoje que o progresso está relacionado a construções de pavimentos altos, sem diálogo com a rua e cobrindo a arquitetura e ambientes históricos. Segundo informações tiradas do JCOOnline, 2016, um indicativo [dos fatos anteriormente citados] é o caso do Edifício Caiçara, uma construção de 1930 totalmente demolido em 2016, para dar a espaço a um novo empreendimento. De acordo com a FUNDARPE, 2013 [Fundação do Patrimônio Histórico e Artístico de Pernambuco], o imóvel não era relevante para tombamento por não possuir valor histórico para o Estado. Entretanto, quando analisamos o seu redor e vemos tantos arranha céus, percebe-se que havia um registro de momento histórico naquele local, onde era o Edifício Caiçara.

### **3 – Metodologia**

O presente artigo é de natureza qualitativa e utilizou fontes de dados secundários, tais como a pesquisa bibliográfica [artigos, livros e sites], para a construção do referencial teórico e metodológico, e documental [críticas e resenhas dos espectadores especializados] como fonte empírica discursiva e objetivos explicativos, descritivos e exploratórios. Para a análise dos dados utilizamos análise de discurso conforme proposto por Gill [2015].

O discurso pode ser identificado como manifestações verbais, seja escrito ou em modo oral, referentes a qualquer coisa em que o emissor esteja empenhado a transmitir. Analisar esses discursos nos permite entender construções sociais, fenômenos ou problemas, principalmente por pontos de vistas diferentes. A análise de discurso é o nome dado a uma variedade de diferentes enfoques no estudo de textos [sejam escritos ou falados], desenvolvida a partir de distintas tradições teóricas e diversos tratamentos em diferentes disciplinas [GILL, 2015].

Os discursos utilizados para o desenvolvimento desse trabalho foram os emitidos pelos espectadores especializados. Críticos, jornalistas e acadêmicos que usam os veículos de comunicação para expressar suas opiniões e suas interpretações, que nesse trabalho se concentram nas obras cinematográficas de Kleber Mendonça Filho. Ressalta-se que a diferença entre os textos jornalísticos e acadêmicos se adere pelo fato desse segundo grupo de textos terem mais bases teóricas, referentes às ciências sociais. Assim permitindo uma produção variada de estilos e expressões. Em contrapartida, não podemos dizer que um tipo de texto é mais concentrado na produção fílmica enquanto o outro se concentra na relação “filme-cidade”, esses conceitos foram relativos na pesquisa, ambos percorrendo as matérias de jornais ou os argumentos dos artigos.

Para Aarts e Bauer [2015], toda pesquisa social empírica seleciona evidência para argumentar e ser base de investigação e demonstração. Nesse caso, os documentos foram selecionados, inicialmente foram classificados 142 arquivos, porém apenas 39 foram utilizados. O critério para eliminação dos demais foram: 1 – Fichas Técnicas, justamente por serem informações dos métodos técnicos; 2 – Discursos do próprio Kleber Mendonça Filho, já que o ponto da análise são das perspectivas do público especializado e não do próprio diretor; 3 – Textos que não traziam a cidade de forma que se encaixasse nas questões analisadas.

Depois de filtrados, em um segundo momento, para análise foi utilizado o software NVivo11 [um banco de dados] para analisar as informações. O material em vídeo e texto foi importado para a plataforma e com a utilização de nós [grupos] para codificação do conteúdo em acordo com as categorias analíticas definidas no decorrer da pesquisa e agrupamento de matérias com temas em comum. Ao final chegamos a 2 discursos que mais se evidenciaram nos documentos analisados, os quais comentaremos no próximo tópico.

#### **4 – Análises dos Resultados**

Na análise dos 39 documentos coletados emergiram 2 discursos principais: “passado ou futuro da cidade” e “crescimento da cidade” os quais serão comentados a seguir intercalados por citações de alguns textos analisados como ilustração.

##### **4.1 – Passado ou Futuro da Cidade**

Presente principalmente nos documentos relacionados aos filmes *O Som ao Redor* e *Aquarius* que estão relacionados aos fatos e características históricas da cidade do Recife. Com ênfase no passado do Recife, também bem como faz a relação entre a memória e o futuro da cidade.

Aqui somos deparados em um momento nostálgico de como o Recife era em décadas passadas. Quais características dessas décadas foram deixadas para trás?

Quando a câmera desvenda o que é observado pelo personagem, vê-se a rua do passado, quatro carros antigos estacionados, muitas árvores, casas e chão batido. Nenhum prédio, apenas residências com muros baixos e algumas sem portão. Sem câmeras de segurança, cercas ou grades. [TAVARES, 2016, p.51]

Os filmes *O Som ao Redor* e *Aquarius* têm como locação a zona sul da cidade, em *O Som ao Redor* o foco são as ruas de Setúbal e *Aquarius* a orla da praia [ambos no bairro de Boa Viagem]. A urbanização do bairro tem início em meados da década de 1920, com a abertura de uma avenida margeando a orla. Essa interferência em uma área até então considerada rural e “desconhecida” pelo resto do estado, se dá pelo Plano de Melhoramentos de 1917 do engenheiro e sanitarista Saturnino de Brito, contratado pelo governo do Estado de Pernambuco, que percebeu as condições de salubridade do local [Alves, 2018].

Tendo como base a matéria de Alves [2018] para o JC Online, até o ano de 1936, Boa Viagem era considerado um bairro rural e que em meados dos anos 70 a predominância era de casas térreas [correspondente a 22%], e prédios de no máximo quatro pavimentos [referentes a 40,12%] tal como o edifício *Oceania* presente no filme *Aquarius*. Essa configuração paisagística converge com a representação destacada por Tavares [2016] sobre o passado da cidade.

Porém essas características são alteradas com a aprovação da lei de Uso e Ocupação do Solo nº 16.176 em 1996, que possibilitou a substituição de casas e prédios baixos, que permitiam maior interação com a rua, por muros e prédios altos, com até duas linhas de muralhas, fachadas de vidro e guaritas. Em 80 anos, o bairro que era rural, isolado e sem urbanização foi transformado no mais verticalizado do estado, mesmo assim a busca para a criação por novos pavimentos continua.

A rápida transformação do bairro ainda deixa marcas, no território e na memória, o que nos remete ao artigo de Souza [2013] o qual analisa a locação do filme *O Som ao Redor* como a representação de um bairro em acelerada modificação urbana, mas com resquícios do passado recente:

O cenário escolhido para contar as histórias é um bairro da Cidade do Recife [Pernambuco], que embora esteja em plena transformação urbana e social, ainda conserva algumas características do passado. [SOUZA, 2013, p.2]

O passado que vai sendo destruído, substituído pelo novo empreendimento, se materializa tanto na luta da personagem Clara pelo direito de permanecer no edifício em que mora, no filme *Aquarius*, quanto na memória da personagem Sofia, de *O Som ao Redor* que visita sua antiga casa que será demolida. Tavares [2016] identifica novamente a relação entre passado e futuro ao considerar que não apenas a memória de Sofia [ou da própria cidade], mas a vida daquela localidade será alterada:

A casa demolida não irá alterar somente a memória de Sofia, também afetará, em diferentes níveis, todos os moradores ao redor. No mesmo lugar habitado apenas por uma família, em breve se erguerá um prédio de vinte e um andares, que abrigará um número muito maior de famílias e mudará o fluxo de pessoas e veículos na região. Com isso a paisagem sonora igualmente se altera. O prédio alto também vai modificar a vista e o sol que chega aos vizinhos [TAVARES, 2016, p.52].

A locação principal do filme *Aquarius*, moradia da personagem Clara, na realidade se chama Oceania e trata-se de “um dos poucos prédios da orla da Zona Sul da Cidade que não segue ao modo de morar “vertical”, com muros altos e portaria” [TRIGUEIRO, 2016]. O prédio que tinha tido o pedido de tombamento na Fundação do Patrimônio Artístico e Cultural de Pernambuco negado há cerca de 18 anos atrás teve o processo reaberto após o sucesso do filme.

Se na vida “real” o filme contribuiu para a preservação da memória arquitetônica, nas ficções a memória da cidade, seja simbolizada por Sofia ou pelo *Aquarius*, luta para não ser submergida pelo crescimento vertical do Recife, o que nos remete ao segundo discurso que emergiu nos documentos sobre os filmes *Recife Frio*, *O Som ao Redor* e *Aquarius*: crescimento da cidade.

#### **4.2 – Crescimento da Cidade**

Para citar o crescimento, precisa se analisar tanto seu desenvolvimento horizontal quanto vertical. Em questões horizontais, o crescimento estaria ligado a descentralização tanto em termos demográficos como espaciais [Côrrea, 1995]. Mas o crescimento vertical estaria ligado a concentrações, tomando como exemplos edificações com a maior quantidade de pavimentos possível. Ainda para Côrrea [1995], o limitado espaço e crescimento horizontal, faz, sobretudo, acentuar a verticalização, demolindo os prédios mais antigos dando espaço para os mais elevados.

No Recife, seu crescimento vertical se intensificou bastante depois da aprovação da lei de Uso e Ocupação do Solo [nº 16.176/1996] e mudou sua paisagem,

principalmente a de Boa Viagem, como podemos observar na crítica de Vigneron [2016] sobre a abertura do filme Aquarius:

Essa transformação aparente remete as mudanças em larga escala pelas quais a cidade do Recife passou desde os anos 1980, época em que o início do filme se detém, através de imagens da orla da cidade que dão conta de seu crescimento vertical nas últimas décadas. [VIGNERON, 2016, p.3].

Tanto o edifício Oceania [usado como locação] quanto o edifício Caiçara, o qual inspirou o prédio/filme “Aquarius”, são exemplos de arquiteturas que predominavam na localidade com o tipo de prédio que até mesmo o andar mais alto ainda podia ter contato direto com a rua. O Edifício Caiçara foi demolido, o Oceania teve seu processo de tombamento reaberto, conforme explicado anteriormente, e no filme Clara luta para que o Aquarius permaneça a existir não só pela memória da cidade, mas como um tipo de construção ainda possível no contemporaneidade.

No filme, a imobiliária pretende construir o edifício New Aquarius, um condomínio de luxo com segurança e área de lazer, e tenta convencer Clara de que é um excelente negócio para todos os envolvidos, o retorno financeiro do empreendimento se sobressai conforme identificado por Vaz [2018]:

O mercado imobiliário, que não pensa no espaço como lugar de morada de sujeitos com vidas e histórias, pensa em como lucrar mais, em como transformar o espaço urbano em produtor financeiro, não pertence a um sistema que humaniza os espaços, mas que constrói arquiteturas que garantem o caos, apequenando o sujeito em relação às dimensões físicas dos edifícios e ao sentimento de [des]pertencimento. [VAZ, 2018, p.207].

O New Aquarius simboliza o tipo de construção vertical, com muros altos, áreas de lazer e segurança privada comuns na paisagem urbana de alguns bairros da cidade. A altura dos prédios e também de seus muros fazem com que as ruas fiquem sem movimentos e entregues a “esquisitice”, já que seus moradores tem direito a lazer e segurança privada, restringindo ainda mais sua locomoção ao próprio prédio. Essa arquitetura pode ser um dos motivos pelo qual Boa Viagem é o bairro de Recife com maior registro de roubos de carro, de acordo com o site PorAqui, 2017, do dia 1º de janeiro até a data da reportagem [22 de fevereiro] 96 casos já haviam sido registrados pela Secretaria de Defesa Social. O que nos remete a reportagem de Guerra [2016] e artigo de Queiroz e Nóbrega [2017]:

O que Kléber parece criticar em seus filmes é o falso conceito de desenvolvimento e crescimento urbano que desumaniza as cidades e cria ambientes hostis. [GUERRA, 2016, p.35]

Recife, onde estaria ocorrendo uma metamorfose urbanística tanto abordam este descompasso urbanístico que realmente ocorre na cidade, quanto apresentam suas consequências e tensões sociais derivadas desse descompasso. [QUEIROZ e NÓBREGA, 2017, p.79].

Percebemos que o crescimento da cidade vertical da cidade é percebido não como uma melhoria da vida urbana, mas sim um espaço de tensões sociais e desumanização da cidade. Aparentemente o empreendedorismo territorial parece ser a força propulsora desse crescimento.

## **5-Considerações Finais**

Para a melhoria da cidade industrial, em crescimento demográfico e espacial, o urbanismo moderno proclamou a ordem e o progresso segundo o qual era preciso

ordenar o espaço de modo a garantir a circulação dos automóveis, mercadorias e das pessoas e construir edificações que atendessem a um maior número de pessoas mesmo que para isso fosse necessário destruir marcos históricos da cidade.

Ao analisarmos os discursos dos espectadores especializados sobre as obras do cineasta Kleber Mendonça Filho, percebemos que a destruição da história [memória] e o crescimento vertical estão entre as críticas presentes na representação da cidade do Recife. Como se a cidade ainda utilizasse o lema ordem e progresso para direcionar as suas transformações urbanas mas sem os benefícios prometidos. Em contraponto a uma melhor circulação de pessoas e automóveis, o empreendedorismo territorial baseado na verticalização da cidade produziu ruas desertas e ambientes hostis.

Com os resultados [ou discursos] colhidos, sendo centralizados em “Crescimento da Cidade” e “Passado e Futuro da Cidade” foi permitido traçar uma linha do tempo de transformações na cidade do Recife. Apesar de terem sido divididos em dois pontos os discursos se cruzam, quando vemos a história da capital pernambucana em transformação: um bairro com casas térreas que com novas leis e projetos deu lugar ao mais verticalizado da cidade.

Foi possível realizar um estudo crítico sobre a cidade representada nos trabalhos do cineasta Kleber Mendonça Filho, a partir do olhar do espectador especializado, seus discursos serviram para aprofundar mais a base teórica/científica e entender pontos de vistas variados que convergiam em determinados pontos. Embora os filmes sejam ficções, as traduções dos espectadores especializados se apresentam como um material rico de conteúdo para análise reflexiva sobre o futuro da cidade e as linhas que norteiam a administração pública sobre o ordenamento do espaço urbano. O trabalho demonstra sua importância em variados pontos: na pesquisa de espectadores; na análise de discursos; na identificação de questionamentos a políticas direcionadas a cidade e concepção do desenvolvimento urbano.

É colocado como sugestão para continuidade desses temas a análise fílmica dos filmes de Gabriel Mascaro que trazem questões sociais parecidas com as de Kleber Mendonça Filho; e artigos direcionados a Especulação Imobiliária e a divisão do Espaço Urbano.

## 6 – Referencias

- AARTS, Bas, BAUER, Martin W. **A Construção do Corpus: um principio para a coleta de dados qualitativos.** 2015
- ALVES, Cleide. **Boa Viagem: de área rural ao bairro mais verticalizado do Recife.** Disponível em: <  
<http://especiais.jconline.ne10.uol.com.br/recifeemtransformacao/boa-viagem-de-area-rural-ao-bairro-mais-verticalizado-do-recife/> > Acesso em: 09/09/2018
- BUENO, Chris. **Ocupe Estelita: movimento social e cultural defende marco histórico de Recife.** Disponível em: <  
[http://cienciaecultura.bvs.br/scielo.php?script=sci\\_arttext&pid=S0009-67252014000400003](http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252014000400003)> Acesso em: 27/08/2018
- CÔRREA, Roberto Lobato. **O Espaço Urbano.** Editora Ática, Série Princípios, 3a. edição, n. 174, 1995. p.1-16.
- FILHO, Oswaldo Amorim, SERRA, Rodrigo Valente. Evolução e perspectivas do papel das cidades médias no planejamento urbano e regional. **Cidades Médias Brasileiras.**

- GEGNER, Martin. A cidade brasileira: moderna demais para o século XXI? **Modernização ambivalentes: perspectivas e transnacionais**. Recife, Editora UFPE. 2016
- GILL, Rosalind. **Análise de discurso**. Pesquisa qualitativa com imagem, texto e som. Editora Vozes, São Paulo, 2015.
- GUERRA, Rafael. Aquarius e a cidade, ou como lutar e permanecer. **Cine**. 2016.
- HARVEY, David. **A Produção do Espaço Capitalista**. São Paulo. Annablume, 2005. Coleção Geografia e Adjacências.
- JOnline. Edifício Caiçara é demolido na Avenida Boa Viagem. 2016 Disponível em: < <https://jonline.ne10.uol.com.br/canal/cidades/geral/noticia/2016/04/07/edificio-caicara-e-demolido-na-avenida-boa-viagem-229935.php> > Acesso: 13/08/2018
- LACERDA, Norma. **INTERVENÇÕES NO BAIRRO DO RECIFE E NO SEU ENTORNO: indagações sobre a sua legitimidade**. Sociedade e Estado, Brasília, v. 22, n. 3, p. 621-646. set./dez. 2007
- LINO, Bárbara, NASCIMENTO, Cristiano. De esculpir e materializar os tempos. A representação cinematográfica da paisagem vertical da cidade do Recife. **Revista de Estudos Brasileños**. Número Especial. Salamanca, Espanha. 49-62. Segundo Semestre, 2017.
- NUNES, Ferreira. **O Impacto da “Lei dos Doze Bairros” sobre a construção civil n cidade do recife**. Recife, 2018.
- PORAQUI. Boa Viagem é o bairro do Recife com maior número de roubos de veículos, 2017. Disponível em: < <https://poraqui.com/boa-viagem/boa-viagem-e-o-bairro-do-recife-com-maior-numero-de-roubos-de-veiculos/> > Acesso em: 19/10/18
- QUEIROZ, Carlos Eduardo Japiassú de; NÓBREGA, Maria de Lourdes Carneiro da Cunha. Registros híbridos: uma análise da estrutura dos filmes “O som ao redor” e “Recife frio”, do diretor Kleber Mendonça Filho. **Revista Tempos e Espaços em Educação**. São Cristóvão, Sergipe, Brasil, v. 10, n. 21, p. 77-98, jan./abr. 2017.
- RANCIÈRE, Jacques. **O Espectador Emancipado**. ArtForum. Março, 2007.
- RESENDE, Ubiratan Pereira. **Especulação imobiliária e verticalização urbana: um estudo a partir do Parque Municipal Cascavel em Goiânia**. Geografia [Londrina], v.22, n.2. p. 79-102, maio/ago. 2013
- SABOYA, Renato. **O que é especulação imobiliária?** Disponível em: <<http://urbanidades.arq.br/2008/09/o-que-e-especulacao-imobiliaria/>> Acesso em: 27/08/2018.
- SOUZA, Rita de Cássia Alves de. A subalternidade impingida e a r[existência] dos serviços no filme O som ao redor de Kleber Mendonça Filho: reflexões sobre a poética da cotidianidade no cinema brasileiro contemporâneo. **Revista Temática**. Ano IX, n. 10 – Outubro/2013
- TAVARES, Cássia Jorge. A cidade construída no cinema: uma análise do filme o som ao redor de Kleber Mendonça Filho. **Trabalho de conclusão de curso**. Porto Alegre: Universidade Federal do Rio Grande do Sul, Curso de Jornalismo, 2016.
- VAZ, Aline Aparecida de Souza. A Recife como Mise en Scène na obra de Kleber Mendonça Filho. **Vozes e diálogos**. Itajaí, v. 17, n. 01, jan./jun. 2018
- VIGNERON, Victor. Aquarius: o filme-edifício e a experiência de Habitar. **Blog Junho**, 2016.

## Tratamento de Erros em Jogos Educacionais: Um Estudo Qualitativo aplicado ao jogo Pizzaria Dividendo

Pedro Macêdo, Wylliams Santos, Mirele Lima, Alexandre Maciel

REACT Research Labs, Caruaru – PE – Brazil  
Universidade de Pernambuco (UPE), Caruaru – PE – Brazil

pedroh.programacao@gmail.com, wbs@upe.com,  
mirele.moutinho@upe.com, amam@ecomp.poli.br

**Abstract.** *Educational games have the ability to engage and motivate students. In this sense, the game Dividendo Pizzaria, potentialize the understanding of mathematics, but presents limitations when offering new explanations to the students. This work proposes a model for constructive error messages applied to the Pizzeria Dividendo game, considering the pedagogical aspects. A focus group session was conducted as a qualitative method for data collection and model evaluation. Based on the data collected, it was possible to detect that error messages can be incorporated into the Pizzeria Dividendo, offering directed support to the student's difficulties.*

**Resumo.** *Os jogos educativos têm a capacidade de engajar e motivar os estudantes. Nesse sentido, o jogo Pizzaria Dividendo, potencializa o entendimento da matemática, mas apresenta limitações ao oferecer novas explicações aos alunos. Este trabalho propõe um modelo para mensagens de erro construtivas aplicadas ao jogo Pizzaria Dividendo, considerando os aspectos pedagógicos. Uma sessão de grupo focal foi utilizada como método qualitativo para coleta de dados e avaliação do modelo. Com base nos dados coletados, foi possível detectar que as mensagens de tratamento de erro podem ser incorporadas no Pizzaria Dividendo, oferecendo suporte direcionado às dificuldades do estudante.*

### 1. Introdução

Com a tecnologia sempre presente na vida dos jovens estudantes, o modelo de ensino atual, onde o aluno é apenas um telespectador, torna-se menos atrativo para as crianças e adolescentes. Neste contexto, Prensky (2001), afirma que os professores precisam se

adequar a essa nova realidade buscando usar a tecnologia como uma ponte entre o aluno e o professor.

Segundo Brito (2005), uma forma de empregar a tecnologia no ensino é através da utilização de *softwares* educacionais que auxiliam no processo de aprendizagem, sendo uma importante ferramenta tanto para alunos, possibilitando o acesso a conteúdo didático, quanto para os professores, auxiliando no processo de ensino.

Sob o mesmo ponto de vista, Macêdo, et. al (2017) propuseram um jogo aplicado à educação, intitulado Pizzaria Dividendo. O jogo que tem como objetivo apoiar o aprendizado da matemática, mais precisamente no aprendizado de números fracionários para estudantes no quinto ano do ensino fundamental.

Porém, apesar de oferecer auxílio na disciplina de matemática, o jogo citado anteriormente apresenta limitações ao oferecer novas explicações ao aluno. Ao errar o aluno não é informado sobre o porquê errou, ou simplesmente quando excedeu o tempo disponível na fase do jogo. As mensagens que aparecem no jogo podem ser expressadas de maneira construtiva. Segundo Bugs (2011), as mensagens construtivas auxiliam o jogador a diagnosticar porque aquele problema aconteceu e como corrigir.

Este trabalho tem como objetivo examinar e propor as melhores formas de apresentar mensagens de erros construtivas no jogo Pizzaria Dividendo, visando aperfeiçoar o conteúdo passado pela aplicação de forma a auxiliar os alunos com a ferramenta.

## **2. Mensagens de Erro em Softwares Educacionais**

Aceitar o erro como parte do processo de aprendizagem e refletir sobre os motivos desses erros são características que estão naturalmente presentes nos jogos, afirma Fardo (2013). Considerando, que os alunos do ensino fundamental dessa geração já nasceram na era da informática e, portanto, estão mais familiarizados com a tecnologia, os jogos podem ser utilizados em caráter educativo, tanto para o ensino das matérias, quanto para a conscientização sobre a importância dos erros nesse contexto.

Durante a prática do jogo, quando um educando falha, conclui-se que as possibilidades de ele não conseguir assimilar o que era requisitado pelo jogo pode ter sido por diversos motivos, como a falta de compreensão ou então pelo nível de

dificuldade da fase. É necessário entender como o fracasso pode ser depurado sem que o educando perca o estímulo pelo jogo. Segundo Paula (2015) a depuração ocorre quando o jogador avalia qual o motivo do seu erro, construindo conhecimento a partir dessa experiência. Dessa forma, pode-se mostrar ao educando que o seu insucesso é capaz de se tornar uma ferramenta de impulso para o mesmo poder seguir adiante.

Os erros em jogos educacionais devem levar o jogador a aprender com eles, sem que haja receio de errar ou encará-los como uma falta de capacidade por parte do jogador. Para isso, os erros devem ser tratados de forma pedagógica, eles devem ser construtivos, indica Papert (1994).

Sendo assim, um *software* educacional, deve ser um ambiente interativo, proporcionando aos alunos adquirir conhecimentos através de levantamento de hipóteses, aperfeiçoando suas ideias e testando cada uma delas. Aplicações educativas podem oferecer benefícios aos seus usuários, principalmente quando essas conseguem aproveitar o erro do aluno para enviar informações úteis, motivando o usuário a continuar tentando, como é apontado por Bugs et al. (2011).

A maioria dos sistemas educacionais fornecem mensagens negativas ou positivas, de acordo com as informações que transmitem ao usuário. Mensagens negativas podem intimidar o usuário, trazendo aspectos desfavoráveis ao jogo, por isso, estas devem ser evitadas. Já as mensagens positivas podem estimular o usuário a tentar novamente, não o fazendo desistir após o acontecimento dos erros. No entanto, mesmo que sejam positivas, estas podem não ser consideradas construtivas. Para que sejam consideradas mensagens construtivas elas precisam oferecer informações que auxiliem no processo de aprendizado, possibilitando uma reflexão de onde ocorreu a falha, viabilizando uma oportunidade para identificar esse erro e tentar corrigi-lo, afirma Bugs et al. (2011).

Nielsen (1993) fornece um conjunto de diretrizes a serem utilizadas como apoio na construção de mensagens de erros construtivas. Sendo assim, é preciso que as mensagens possam: Informar sobre o motivo do erro, fornecer dicas úteis para construir conhecimento, não exibir a resposta para que o estudante possa tentar novamente, evitar o uso de figuras negativas ou termos com conotação negativa e dar opções de retorno a questão.

### **3. Trabalhos Relacionados**

Berti (2007) trata da avaliação da aprendizagem, tal como, os elementos de contribuição para a avaliação diagnóstica das estratégias utilizadas pelos alunos na resolução de problemas e nos algoritmos e principalmente descrever possíveis formas de ações didático-pedagógicas para o trabalho com os erros dos alunos ao longo do desenvolvimento da aprendizagem.

O presente trabalho destaca-se dos demais por propor uma atualização ao jogo já utilizado em sala de aula, avaliando se as mensagens de tratamento de erro presentes no projeto Pizzaria Dividendo, são capazes de gerar conhecimento aos estudantes que a utilizam. Oferecendo, caso necessário, novas formas de explorar esse assunto dentro da aplicação.

#### 4. Grupo Focal

O presente estudo utilizou o grupo focal como método qualitativo para coleta de dados. O grupo focal é caracterizado como uma entrevista, conduzida com um grupo de pessoas com conhecimento na área de estudo (Kontio et al. 2014). Esse trabalho adotou o *guideline proposto* Kontio et al. (2014) para condução e discussão da sessão de grupo focal (Figura 1).

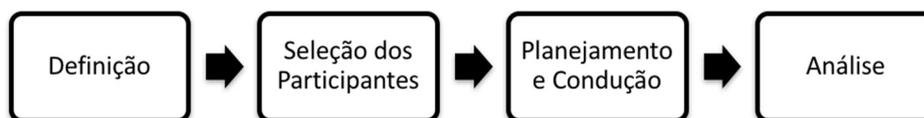


Figura 1. Processo para condução de grupo focal. Adaptada de Kontio et. al (2014)

##### 4.1 Definição do problema de pesquisa

O objetivo deste grupo focal é promover uma discussão para encontrar alternativas em relação a tratamento de erros do jogo educativo proposto no presente trabalho. Discutir como as mensagens de erro do jogo podem ser apresentadas de uma maneira que os alunos aprendam através de mensagens construtivas.

##### 4.2 Seleção dos participantes

De acordo com Kontio et al. (2014), a experiência e o nível de engajamento dos participantes podem influenciar diretamente na qualidade do grupo focal aplicado. Portanto, foram selecionados quatro discentes da Universidade de Pernambuco<sup>1</sup> para condução do grupo focal. Todos os participantes são alunos do sétimo período e

---

<sup>1</sup> www.upe.br

estavam cursando o componente curricular de "Tecnologias de Informação e Ensino a Distância". O assunto em questão, tratamento de erros, fora amplamente discutido durante as aulas da disciplina.

#### **4.3. Planejamento e Condução**

Todos os participantes concordaram formalmente em participar da sessão de grupo focal, de acordo com os termos de livre consentimento previamente assinados. A sessão teve duração de 50 minutos e foi gravada com o auxílio de um *smartphone*, facilitando uma reprodução fidedigna da sessão. Após a sessão, o moderador (investigador principal) realizou a transcrição completa do áudio para possibilitar a análise dos dados.

A sessão do grupo focal teve início com uma visão geral acerca dos objetivos da presente pesquisa. O moderador esclareceu o assunto abordado, assim como os pontos específicos para engajar os participantes em relação ao objeto em discussão. Em seguida, o moderador realizou a leitura das perguntas, guiando e garantindo a condução da sessão, mantendo o grupo focado, para não se afastarem do tópico. O moderador também apresentou as perguntas em um documento de fácil visualização e análise por parte de todos os participantes.

De forma a promover o engajamento dos participantes, algumas perguntas simples foram administradas. Essas perguntas estavam relacionadas com a temática de tecnologias na sala de aula e como os professores se relacionam com estas ferramentas.

- P1. Vocês aprovam a ideia de utilizar jogos educativos em sala de aula?
- P2. Vocês já utilizaram algum jogo educativo?
- P3. Os professores estão aptos para utilizar a tecnologia em sala de aula?
- P4. Os alunos sabem utilizar a tecnologia como ferramenta de estudo?

Com o objetivo de responder o problema de pesquisa (seção 4.1), foram definidas perguntas, de acordo com o *guideline* proposto por Kontio et. al (2014).

- P5. Quais das alternativas de tratamento de erros existentes neste trabalho podem ser modificadas para gerar um melhor desempenho do aluno?
- P6. Como as modificações sugeridas podem afetar o desempenho do aluno?
- P7. O modelo está suficientemente claro?
- P8. Quais são os principais problemas/desafios na adoção do jogo (modelo atual)?

#### **4.4. Análise**

As perguntas iniciais (P1 à P4) revelaram que todos os participantes defendem a utilização da tecnologia em sala de aula, pois tira o professor como centro de todo conhecimento e os alunos são direcionados à descoberta do conhecimento a partir dos seus estudos, de forma que o professor auxiliasse a detectar as principais dificuldades para direcionamento dos estudos daquele estudante.

Todos os participantes reforçam que os jogos podem ser utilizados para facilitar o uso da tecnologia em sala de aula. Porém, mesmo com os benefícios da tecnologia, os respondentes do estudo alegaram que a maioria dos professores não estão aptos a utilizá-la como ferramenta de estudos. As opiniões se dividem quando o assunto é treinamento, pois alguns participantes acreditam que o enfoque em treinamentos especializados pode promover a utilização das ferramentas. Outros participantes pontuaram que mesmo com os treinamentos, a ampla utilização das tecnologias nas salas de aula ainda é desafiadora.

Em relação às perguntas P5 à P8, os participantes apresentaram que o *Software* Pizzaria Dividendo precisa de algumas atualizações no tocante ao gerenciamento de mensagens de erro. Os participantes sugeriram uma modificação, utilizando mensagens com cunho explicativo para motivar o aluno. As sugestões fornecidas envolvem explicações com imagens e gráficos, apresentadas na próxima seção.

De acordo com os participantes, as mensagens atuais do jogo detinham-se apenas a informar que o aluno errou naquela operação, não explicando ao aluno qual o seu erro. Foi sugerido que as mensagens oferecessem de forma detalhada uma pequena animação indicando qual foi o erro cometido, assim como o caminho que ele deveria fazer para acertar.

Ao final da sessão, os respondentes apontaram que mudanças comportamentais e de gestão são necessárias para possibilitar ampla adoção das tecnologias educacionais nas escolas, como instrumentos de pesquisa e estudo.

### **5. Proposta de Mensagens de Erro para o jogo Pizzaria Dividendo**

O jogo Pizzaria Dividendo, proposto por Macêdo et al. (2017) tem como objetivo a análise do emprego de jogos educativos como ferramenta auxiliar no desempenho dos alunos no estudo da matemática, mais precisamente no estudo de números fracionários.

Nele, um jogador precisará atender clientes de uma pizzaria, mas todos os pedidos são realizados em formato de fração (Figura 2).



Figura 2 . Pedido do cliente no jogo (Macêdo, et al., 2017)

A versão atual do jogo não possui mensagens de erros construtivas que descreva onde o aluno errou. Com base na análise do grupo focal, as opções de mensagens de tratamento de erro são incorporadas para melhorar a mecânica do jogo Pizzaria Dividendo.

Nesse jogo, a mensagem de erro aparece de duas formas, são elas: (i) quando o jogador erra na escolha de um dos ingredientes da pizza que atenderia o cliente, ou (ii) quando o jogador ultrapassar o limite de tempo para realizar as ações e completar o pedido. Apesar de serem informações diferentes, a versão atual da aplicação não trata as alternativas de forma distintas, sendo assim, o jogador não tem como saber se errou ou se o seu tempo acabou.

Para diferenciar o erro do jogador e a demora para realização da ação, dois tipos de mensagens podem ser aplicados. Em um dos tipos de mensagem será indicado que o jogador não conseguiu cumprir a tempo, estimulando-o a fazer cálculos mais rápidos. A outra abordagem, que foi amplamente debatida no grupo focal, requer uma explicação de acordo com o erro do jogador.

A primeira proposta de mensagem apresenta o erro do jogador, demonstrando tudo que o jogador realizou em forma de fração e indicar a resposta correta ao problema, mostrando uma imagem da previsão do pedido ao final. Entretanto, essa

proposta foi descartada por ferir um dos princípios proposto por Nielsen (1993), de não mostrar a resposta para que o aluno possa tentar novamente, e por poder gerar soluções duvidosas, visto a possibilidade de poluição visual. A Figura 3 apresenta a proposta no Pizzaria Dividendo.



Figura 3. Proposta para correção e aviso de erro

A segunda sugestão de mensagem de erro e a mais aceita entre os participantes, foi que o jogo saísse da tela já conhecida, e desse uma pequena explicação de frações, utilizando-se de imagens e poucos textos. Indicando como montar a pizza de acordo com alguns valores de frações. Mostrando também, parte do conteúdo presente no jogo, porém, sem apresentar a resposta, para que a criança possa pensar como resolver os próximos problemas. A Figura 4 demonstra como essas informações podem ser apresentadas no jogo.



Figura 4. Ao clicar em avançar, o chefe menciona dicas sobre frações

Em relação aos aspectos pedagógicos, foram aplicados os padrões citados por Paula (2015), permitindo que através das mensagens de erro o educando possa avaliar os seus erros, e a partir das explicações recebidas, construir novos conhecimentos. As mensagens de erro propostas incentivam o aluno a continuar tentando, de modo a não perceber o erro como algo definitivo e sim como uma situação que pode ser melhorada na próxima tentativa, seguindo o que foi descrito por Paula e Valente (2015).

As propostas de mensagens de erro cumprem com os aspectos pedagógicos, oferecendo mais conteúdo ao seu jogador, não afetando a mecânica do jogo proposto pelo Pizzaria Dividendo.

## **6. Conclusões e Trabalhos Futuros**

Neste trabalho foi possível analisar que a tecnologia é um fator diferencial para a sala de aula, mas que é preciso estar bem adaptada ao conteúdo que será estudado. O *software* Pizzaria Dividendo, apesar de conseguir atender à necessidade de uma nova forma de apresentar o conteúdo, precisa de adequações para que possa gerar conhecimento de uma forma eficiente.

As modificações são necessárias quando uma mensagem de erro for emitida. É importante que o aluno entenda porque errou e como poderá prosseguir, de forma a obter resultados mais positivos. Dessa forma, a aplicação poderá fornecer um melhor conteúdo, divertido e educativo, mas sem perder a essência lúdica do jogo.

Como próximos passos, pretende-se desenvolver a nova versão do jogo Pizzaria Dividendo, considerando e incorporando as sugestões fornecidas ao longo da sessão do grupo focal. Além disso, mais estudos empíricos serão executados na escola para validação da ferramenta por parte dos alunos e professores.

## **Referências**

- Berti, N., Carvalho, M. (2010) “Erro e estratégias do aluno na Matemática: contribuições para o processo avaliativo”.
- Brito, G., Moura, A. (2005) “Sistema ROSA-P2P: uma arquitetura distribuída para Integração de objetos de aprendizagem”. Simpósio Brasileiro de Informática na Educação, p. 190-200.

- Bugs, D., Veiga, M., Silveira M. (2011) “Diretrizes para Elaboração de Mensagens de Erro Construtivas”. Simpósio Brasileiro de Informática na Educação.
- Fardo, M. (2013) “A gamificação aplicada em ambientes de aprendizagem”. RENOTE.
- Kontio, J., Lehtola, L., Bragge, J. (2004) “Using the focus group method in software engineering: obtaining practitioner and user experiences”. Empirical Software Engineering.
- Macêdo, P., Moutinho, M., Santos, W. (2017) “Jogo Digital como Auxílio no Estudo da Matemática: Um Estudo de Caso com Estudantes do Ensino Fundamental I”. Anais do Workshop de Informática na Escola. p. 548.
- Nielsen, J. (1994) “Usability engineering. Elsevier”.
- Papert, S. (1994) “A máquina das crianças. Porto Alegre: Artmed”.
- Paula, B., Valente, J. (2015) “Errando para aprender: a importância dos desafios e dos fracassos para os jogos digitais na Educação”. RENOTE.
- Prensky, M. (2001) “Digital natives, digital immigrants part 1”. On the horizon.

## **Uso de Softwares Logísticos como Vantagem Competitiva: Um estudo Bibliométrico do software WMS (Warehouse Management System)**

**Alex Gustavo da Silva Moura, João Mário Leão, José Augusto de Lira Lima, José  
Guilherme de Souza Silva, Shirley Danauey da Silva, Gabriela de Vasconcelos <sup>1</sup>**

alex.gus@hotmail.com, j.mario.leao14@gmail.com,  
augustolima95@hotmail.com, guilherme.silva1@hotmail.com,  
shirley.danauey@gmail.com devasconcelos.gabriela@gmail.com

### **Resumo**

As empresas necessitam adaptar-se as mudanças ocorridas no mercado, visando seus processos produtivos, integrando a sua rotina produtiva mecanismos que auxiliam desde da logística de suprimento até a de distribuição. A partir disso surge as ferramentas de gerenciamento integrada aos sistemas de informações, como o caso do sistema de gerenciamento de armazéns, Warehouse Management System (WMS). O uso desse mecanismo permite que as empresas integrem o ambiente competitivo do mercado, com capacidade de atender grandes demandas e ofertar qualidade no produto e na entrega do mesmo. Diante disso, o artigo promoveu um estudo bibliométrico, através da análise, discussão e mapeamento dos resultados evidenciando o objeto da pesquisa o software de auxílio na armazenagem de produtos em armazéns, o WMS. Para tanto, adotou-se uma abordagem qualitativa, com vistas mapear a produção científica em torno da temática, através do Portal de Periódicos da Capes a partir da base de dados Web of Science, que evidenciou uma amostra reduzida, de apenas 15 artigos enquadrados no eixo da pesquisa. Sendo, através da análise identificado a relevância do tema e a abordagem sistêmica sobre o WMS nas organizações que utiliza o software em seus processos administrativos.

### **Abstract**

As companies have adapted to the changes occurring in the market, their production processes, integrating their own generation of mechanisms to help supply logistics to distribution. From this it emerges as tools of management integrated to the information systems, like the system of management of warehouses, Warehouse Management System (WMS). The use of this mechanism allows companies to integrate the competitive environment of the market, with capacity to meet great demands and quality of project without product and delivery. Therefore, the article promoted a bibliometric study, through the analysis, discussion and mapping of the results evidencing the object of research, the aid software in the storage of products in warehouses, the WMS. In this paper, we present a sampling of only 15 sections in the research axis, which is the most important of all the studies. . Using the analysis name in the analysis and the method on WMS in organizations about the use of software in their administrative actions.

---

<sup>1</sup> Orientadora; Professora das disciplina de Administração de Vendas, Logística e Projeto Experimental 6 na Universidade de Pernambuco, *Campus Caruaru*.

## **1.Introdução**

As mudanças nos hábitos de consumo da sociedade exigem das empresas alterações em sua produção, entretanto a percepção dos gestores diante dessas transformações são equivocadas, pois em sua visão o crescimento de consumidores no momento consequentemente exige o aumento da produção instantaneamente. Há realização das atividades produtivas de forma a operar com o uso de um determinado tempo e ainda mantendo a cadeia de suprimentos, requer controle, informação e planejamento, impossibilitando ações como atrasos, desperdícios e a singularidade dos produtos. Como Ballou (2006) destaca definir estratégias em uma empresa como forma de orientá-la a alcançar seus objetivos é de fundamental importância para os gestores, a fim de evitar as causalidades do mercado. “As demandas dos clientes obrigam as organizações a se adequarem e a procurarem formas para aprimorar e incrementar sua competitividade por meio de inovação nas operações.” (Junior e Morais, 2009).

Através das novas tecnologias as empresas constituíram mecanismos que agilizam, organizam e otimizam os processos da produção, como “a aplicação de tecnologia de informação, proporcionando um grande aumento de eficiência” (Bezerra e Monteiro, 2009). Há implantação dos sistemas de informação contribuem para o controle e gerenciamento do fluxo de informação de uma organização (Ballou,1995), ou seja, contribuindo para a entrada e saída de informações, favorecendo os desafios de se gerir uma cadeia de suprimento. Essa implementação dos sistemas informatizados em ambientes produtivos capacitaram os processos tornando-os mais ágeis e eficientes garantindo resultados que vão rente aos objetivos almejados no início da implementação. Bezerra e Monteiro (2009) expõem os sistemas integrados de informação como vantagem competitiva para as empresas, então se subentende que as contribuições do uso dessa ferramenta também exprimem benefícios para o desempenho do empreendimento.

Diante desse cenário de produção, a logística das empresas necessitam realizar avanços que priorizem a eficiência de seus resultados, adaptando do manual para o informatizado, contribuindo para que os gastos reduzam e o desperdício consequentemente seja eliminado. Em vista disso o artigo propõe verificar o uso software logístico como vantagem competitiva para as empresas, evidenciando o software de auxílio na armazenagem de produtos em armazéns, o Warehouse Management System (WMS), mapeando através de um estudo bibliométrico acerca da análise e discussão dos resultados evidenciando o produto da pesquisa em questão e a conclusão do estudo a partir da exploração dos resultados obtidos.

## **2.Referencial Teórico**

### **2.1 Logística de Suprimentos**

A atividade de logística de suprimentos em uma organização é de grande importância para um bom funcionamento da empresa, pois é ela que faz o diferencial na produtividade, contribuindo também para a diminuição de seus custos, através de negociações, busca de fornecedores e materiais alternativos. De acordo com Platt e Nunes (2007) a logística é encarregada pela administração, planejamento, implementação e controle do ciclo de armazenagem de produtos, serviços e troca de mensagens entre fornecedores e cliente. Com o tempo a cadeia de abastecimento integra

responsáveis pela matéria- prima, distribuidores, indústrias, prestadores de serviços e o consumidor final.

Na cadeia de suprimentos, podemos analisar alguns fluxos de informações onde neles são identificados necessidades, desejos, pedidos entre outros, um fluxo entre

**Fornecedor**  $\longleftrightarrow$  **Empresa**  $\longleftrightarrow$  **Cliente**

Como mostra a imagem:



**Figura 1. Fluxo de Informações**

**Fonte: Adaptada de Novaes (2001, p.38)**

O fluxo de informações parte dos clientes, tanto internos quanto externos, apresentando um processo que deve disponibilizar um fluxo de produtos e/ou serviços. Como a figura apresenta, o cliente leva a informação para empresa, onde a mesma filtra a informação que pode ser uma necessidade do cliente, caso não tenha recursos na organização para sanar a demanda do cliente, essa informação é passada para o fornecedor que se encarrega de providenciar o produto ou serviço solicitado. O fluxo de informações parte do cliente e acaba no cliente, o fluxo de material sai a partir da informação do cliente, convertendo a informação em produto ou serviço que será disponibilizado para o cliente. A empresa fica responsável pela mediação da informação. Como apresenta a figura 2

**Fornecedor**  $\rightarrow$  **Empresa**  $\rightarrow$  **cliente**  $\rightarrow$  **empresa**  $\rightarrow$  **fornecedor**

**Figura 2. Fluxo de informações**

**Fonte: Autoria Própria**

Para Platt & Nunes (2007), o fluxo de informações geralmente deve ser disparado pelo cliente, através da identificação de suas necessidades, com o conseqüente fluxo de material e/ou serviços vindo à direção oposta para atender essa demanda.

## 2.2 A Tecnologia de Informação como Vantagem Competitiva

No cenário atual, cada vez mais surgem competições nos ambientes empresariais. Isso ocorre quando outras instituições realizam o mesmo trabalho que sua empresa, só que melhor (WAGNER, 1952). Neste contexto surge em paralelo a busca por vantagens competitivas. Essas devem ser observadas desde a fase inicial de um produto, seja procurando o menor custo em sua produção, redução de desperdícios, uma gestão de estocagem eficiente, melhor logística de transporte e decorre até o momento de sua venda, onde segundo Las Casas (2007, p.104) “os preços são determinados para cobrir custos operacionais e também proporcionar lucros, além de afetar as expectativas dos clientes”, ou seja, o preço também pode ser um diferencial atrativo aos clientes.

Profissionais e Acadêmicos da área convergem em suas pesquisas, ao concordar que o desafio principal no tocante a economia empresarial é em como conquistar e manter uma vantagem competitiva (BREZNIK, 2012). Pesquisas passadas mostram que o período médio em que as corporações conseguem manter uma vantagem competitiva vem diminuindo ao longo do tempo (WIGGINS e RUEFLI, 2005), muito disso proveniente a conscientização de empresários e busca contínua por modernização, gerando o aumento na aplicabilidade e uso de TI como fonte de melhoria. (BOWERSOX e CLOSS, 2010).

Porter (1985) reforça que Vantagem Competitiva remete a capacidade de manter-se em uma posição no mercado, onde a empresa obtém retornos acima da média em relação a seus concorrentes. Sendo assim, ele nos apresenta dois tipos básicos de vantagem competitiva: liderança de custos e diferenciação. Com isso, fica tangível visualizar a TI como um meio excelente para a obtenção dessa vantagem.

Como exemplo prático, o uso do Sistema de Gerenciamento de Armazéns (WMS) como um recurso de Tecnologia da Informação pode ser destacado. Possuindo um papel fundamental no auxílio a tornar a empresa mais interligada de forma a potencializar a sua competitividade e atingir as metas esperadas e tornar todo o processo mais rentável (LEITE, 2009), permite que o gerente responsável pela logística possa controlar as operações à distância, observando apenas se o funcionamento do sistema está adequado às operações logísticas, aumentando assim a possibilidade de atividades a serem realizadas em um mesmo momento.

### **2.3 Importância dos Sistemas de Informação para o Gerenciamento Logístico**

A produção de produtos ou serviços, necessitam de processos colaborativos e eficazes, que contribuam para que a empresas alcancem seus objetivos de produção. Para a execução desse processo, as indústrias começaram a adotar métodos que auxiliam e otimizam o processo de funcionamento e gerenciamento da produção da empresa. Integrando meios que a tornem eficiente, contribuindo para a mecanização da logística implementada nesse setor operacional da empresa. A partir disso começaram a se utilizar Sistemas de Informação (SI) que possibilita reunir, armazenar, retomar e disseminar determinadas informações (Turban et al, 2007). Esse processo auxilia principalmente nas funções operacionais, administrativas e na tomada de decisões nas organizações.

A redução do tempo gasto para executar um serviço ou o aumento da qualidade de produtos elaborados, com a manutenção dos níveis de qualidade são apenas algumas

das vantagens ao utilizar os sistemas informatizados. Sua importância também está associada aos métodos desenvolvidos para melhorar os resultados da organização como custo operacional, processos logísticos inteligentes e a inclusão de fornecedores e clientes através dessas ferramentas (Alves, Carneiro, Fonseca, Nunes & Santos, 2008). Interligar essa ferramenta ao processo logístico permite a comunicação não só dos ambientes funcionais da empresa como também a área de suprimentos, abordando vendedores e clientes. Ampliando as informações sobre “vendas, embarques, programas de produção, disponibilidade de estoques, situação dos pedidos” (Alves, Carneiro, Fonseca, Nunes & Santos, 2008, p.5), ou seja, ao relacionar os sistemas com as operações logísticas da empresa auxiliamos na obtenção de resultados positivos e também na mecanização e otimização das práticas operacionais.

Neste ambiente organizacional os sistemas de informação são dispostos como pacotes de *software* de computador, permitindo acesso fácil e rápido através desse meio, auxiliando na transmissão de informações entre os envolvidos em todos os setores. Entretanto aplicar um sistema a rede logística de uma empresa depende do objetivo almejado e das necessidades percebidas nessa operacionalização. Por isso, os principais sistemas logísticos se subdividem em (Alves, Carneiro, Fonseca, Nunes & Santos, 2008):

- Sistema de Gerenciamento de Pedidos (SGP);
- Sistema de Gerenciamento de Armazéns (SGA);
- Sistema de Gerenciamento de Transportes (SGT);

Cada um transmite informações para atividades e objetivos específicos, influenciando no processamento padrão das organizações, entretanto trazendo benefícios satisfatórios para cada especificação exigida. Esse formato ágil de troca de informações em cada gerenciamento auxilia no desempenho dentro do mercado, destacando-se a qualidade interna que transparece nos resultados externos.

#### **2.4 Sistema WMS (Warehouse Management System)**

O sistema Warehouse Management System (WMS) trata-se de um software voltado para a área de gerenciamento do armazenamento de uma organização. Estão entre suas funções o recebimento, inspeção, endereçamento, estocagem, separação, embalagem, carregamento, expedição, emissão de documentos, inventário, administração de contêineres entre outras, que, agindo de forma integrada, atendem às necessidades logísticas, evitando falhas e maximizando os recursos da empresa (Guarnieri et al., 2006). Ou seja, o sistema WMS faz um controle de estoque aprofundado, criando endereços para armazenar produtos casados ou similares afim de que todos estejam perto um do outro, evitando a circulação desnecessária dentro do estoque, o sistema beneficia também os colaboradores, pois eles apenas executam o que foi gerado pelo sistema.



**Figura 3. Ferramenta do WMS**  
**Fonte: MGITECH | Tecnologia móvel**

O software, numa visão ecológica, consegue eliminar o uso de folhas de papel que antes eram utilizadas para controlar todo o processo, pelo fato de que sistema consegue armazenar todas as informações inseridas e elaboradas por ele. Fabrimatel (2014) destaca que o sistema tem como seus principais objetivos a redução de tempo em movimentações no armazém, planejar recursos, extinguir ruídos, fazer a redução de custos operacionais, evitar perdas e desvios de materiais no decorrer de operações e ofertar ao cliente um serviço mais prático, eficaz e eficiente. Ou seja, o sistema otimiza e beneficia a empresa com sua objetividade, pelo fato de ser bastante preciso, além disso consegue reduzir a mão de obra humana, que em muitos casos acabam gerando desperdícios de materiais que poderiam ser evitados, e também geram um passo à frente de seus concorrentes pela agilidade que tem em operar e entregar os produtos solicitados, dado o pedido, que em outras palavras entendemos como uma vantagem competitiva.

### 3. Metodologia

A pesquisa desenvolvida utiliza-se do método qualitativo com objetivo exploratório. Foi utilizada uma análise exploratória com o intuito de captar dados relevantes e de opinião de partes interessadas, como mostra Gil (2008, p. 27) “Pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar visão geral, de tipo aproximativo, acerca de determinado fato”. Esse tipo de pesquisa é realizado especialmente quando o tema escolhido é pouco explorado e torna-se difícil sobre ele formular hipóteses precisas e operacionalizáveis. A partir disso, escolheu-se o uso dos métodos bibliométricos, pois seu uso possibilita a análise detalhada acerca da produção científica a respeito de uma determinada temática, afim de compreender as características da produção científica (Lima & Vasconcelos, 2017).

O progresso tecnológico, segundo Araújo (2006), incentivou os estudos bibliométricos, de modo que, atualmente pode-se observar softwares específicos para realização dos mesmos. Nesse cenário, os estudos bibliométricos estão crescendo nos últimos tempos, servindo de base para o levantamento de informações sobre a produção científica (Souza, 2013). Efetuando-se, assim, um diagnóstico da produtividade científica em torno do tema em questão, com a utilização de técnicas quantitativas e/ou

qualitativas (Vanti, 2002). O estudo bibliométrico preocupa-se com a produtividade dos cientistas e com a análise de citações. Sendo assim, adotou-se a pesquisa bibliométrica acerca da temática deste artigo, utilizou-se como instrumento de coleta de dados o Portal de Periódicos Capes adotando a base de dados *Web of Science* e as seguintes palavras-chave: Warehouse Management System e Logistic. Com uma amostra 16 artigos dos quais foram aproveitados 15 deles, pois 1 não tinha como objeto central as às palavras-chave utilizadas nesse estudo, pela a plataforma Web of Science para captação do material pesquisado e a análise foi feita na plataforma VOSWiewer, software que disponibiliza uma série de análises gráficas baseadas na co-ocorrência dos itens analisados, mostrando a conexão entre termos e autores relacionados, proporcionando a divisão em grupos denominados *clusters*. Cada *cluster* é representado por uma cor e agrega todos os itens considerados similares. O tamanho dos círculos dos mapas demonstra o número de ocorrência do item e a proximidade entre dois itens revela seu grau de relação, quanto mais próximos, tanto mais relacionados (VAN ECK et al., 2010).

#### 4. Análise dos Resultados

Sendo assim, chegou-se a amostra da pesquisa o equivalente a 15 artigos, conforme a quadro a seguir:

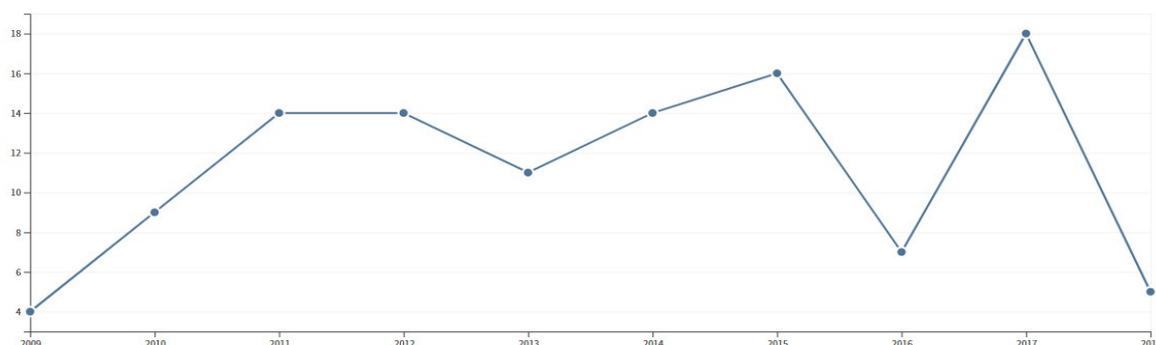
Título da Publicação	Eixos temático	Objetivo	Resultados
<i>Design and application of Internet of things-based warehouse management system for smart logistics</i>	Indústria 4.0; Internet das coisas; Logística Inteligente; Warehouse Management System (WMS).	Propor um sistema de gerenciamento de armazém baseado na Internet das coisas com uma abordagem avançada de análise de dados.	O WMS baseado na Internet das coisas, mostra que a produção do armazém e a eficiência da coleta podem ser melhorada.
<i>Implementation of Warehouse Management System Technology in a Logistics Operator (3PL)</i>	Gerenciamento de Armazenagem; Operação de Logística; Tecnologia.	Descrever o processo de implantação da tecnologia <i>Warehouse Management System</i> .	Através do WMS o operador logístico realiza suas atividades diárias como recebimento e armazenamento de forma eficiente, pois o sistema fornece dados precisos e confiáveis.
<i>Research on and Implementation of the Logistics Warehouse Management System</i>	Implementação; Logística de Armazenagem; Sistema de Gestão.	Compreender o uso do WMS, para fazer com que a produção da empresa possa, de maneira adequada e eficiente, suprir as demandas e alcançar os objetivos.	O sistema WMS foi implementado, alinhado a análise teórica para, verificando a melhoria o desempenho dos tradicionais.
<i>Construction of Intelligent Logistics Warehouse Management</i>	Identificação por radiofrequência (RFID); Gerenciamento de	Analisar as características e funções da tecnologia RFID, discutindo o	Desenvolveu-se um sistema inteligente de informações para gerenciamento de

<i>Information System based on RFID Technology</i>	Armazenagem; Gestão de inventário; Logística Inteligente;	método de construção do sistema inteligente de informações gerenciais do armazém logístico.	armazéns logísticos baseado na tecnologia RFID.
<i>JAVA-based Logistics Warehouse Management System Design</i>	Java; Gerenciamento de armazenagem; Logística.	A partir do Java, espera-se poder fornecer alternativas de práticas de engenharia de estrutura mais viáveis .	O Java é um projeto que trouxe uma melhoria adicional para a construção de informações do sistema de gerenciamento de armazém de logística moderna.
<i>Development and Application of a Warehouse Management Information System in Undergraduate Practice Course of Logistics Engineering</i>	Gerenciamento de armazéns; Logística de informações tecnológicas; RFID; WSN.	Propor um sistema de gerenciamento de armazéns projetado com tecnologia de informação logística.	O WMS cobriu para as aplicações técnicas de vários cursos de plataforma de educação profissional de Engenharia Logística na Beijing Union University.
<i>Application of RFID technology in development of warehouse logistics management system by Internet of things</i>	Armazéns; Gestão Logística; Internet das Coisas; RFID; Sistema de Gestão.	Analizar a aplicação da tecnologia RFID na gestão de armazéns logísticos de cada processo e de cada elo.	Mostram que a tecnologia RFID pode minimizar o custo de armazenamento e melhorar a eficiência do sistema de logística
<i>The Design and Development of Warehouse Management Information System on Hongxing Logistics</i>	Gerenciamento de Armazenagem; Planejamento de processos de negócio; Sistema de Informações de gerenciamento de armazém.	Fazer uma pesquisa sobre o sistema de informações de gerenciamento de armazém da empresa de logística Hongxing.	O Sistema de informações de gerenciamento de armazém torna-se inteligente, reduz o custo da empresa de logística de forma eficaz e atende a demanda de serviços.
<i>The logistics warehouse management and control integrated system based on information fusion</i>	Armazém; Integração de gestão e controle; Fusão de informação; Sistema ciberfísico;	Analisar as características tecnológicas e potencial tecnológico do CPS.	Um método que descreve a função de feedback básica da unidade CPS e proporcionou uma arquitetura de sistema comum do sistema de acordo com a característica de gestão de armazéns.
<i>Design of intelligent Logistics Warehouse Environment Monitoring and Management System Based on Zigbee</i>	Rede de sensores sem fio (WSN); Sistema de medida e controle.	Projetar um sistema inteligente de medição e controle de armazém com base no WSN e na rede Zigbee.	O método pode resolver muitos abusos, como o complexo de layout, os cabos entre cruzados e trazer a conveniência para a supervisão e gestão

			moderna do armazém.
<i>Research on the Intelligent Logistics Warehouse Management System based on Internet of Things</i>	Identificação de rádio frequência; Internet das Coisas; Gerenciamento de armazenagem; Logística Inteligente.	Projetar uma estrutura de sistema, com base nas dificuldades de identificação do produto e na aquisição de informações em tempo real.	O sistema pode adquirir informações de produtos, de estoque, implementar supervisão inteligente, realizar a intelligentização, a ciberização e a aplicação de TI da administração de depósito.
<i>Design of Logistics Warehouse Environment monitoring and Management system based on WSM</i>	Armazéns; Rede de sensores sem fio; Sistema de monitoramento e gestão.	Projetar um sistema inteligente de controle e medição ambiental baseado em redes Zigbee, WSM.	O sistema tem uma medição estável e precisa, fácil de operar e traz a conveniência para o monitoramento e gerenciamento do ambiente de depósito.
<i>A RFID case-based logistics resource management system for managing order-picking operations in warehouses</i>	Identificação por radiofrequência (RFID); Sistemas de gerenciamento de armazém.	Adotar a tecnologia RFID para facilitar a coleta e o compartilhamento de dados em um armazém.	Uma simplificação do procedimento de adoção de RFID; uma melhoria na visibilidade das operações; e um aumento da produtividade do armazém.
<i>A RFID Logistics Resource Management System for the Warehouses</i>	Identificação por radiofrequência (RFID); Logística; Sistemas de gerenciamento de armazém.	Adotar a tecnologia RFID para facilitar a coleta e o compartilhamento de dados em um armazém.	Desenvolvido um sistema de gerenciamento de recursos logísticos baseados em caixas de identificação por radiofrequência.
<i>Warehouse logistics control and management system based on RFID</i>	Gestão de Logística; Rótulo inteligente; RFID.	Explicar o método de implementação do hardware e software do sistema de gerenciamento e controle de logística de armazém com base em RFID.	São realizadas as operações como armazenagem, entrega de armazém e movimentação de mercadorias.

A figura 04 mostra os anos iniciais de publicação de artigos relacionados com o sistema WMS.

Número de citações por ano



**Figura 4: Anos de citações dos artigos envolvendo WMS**  
**Fonte: WEB of Science**

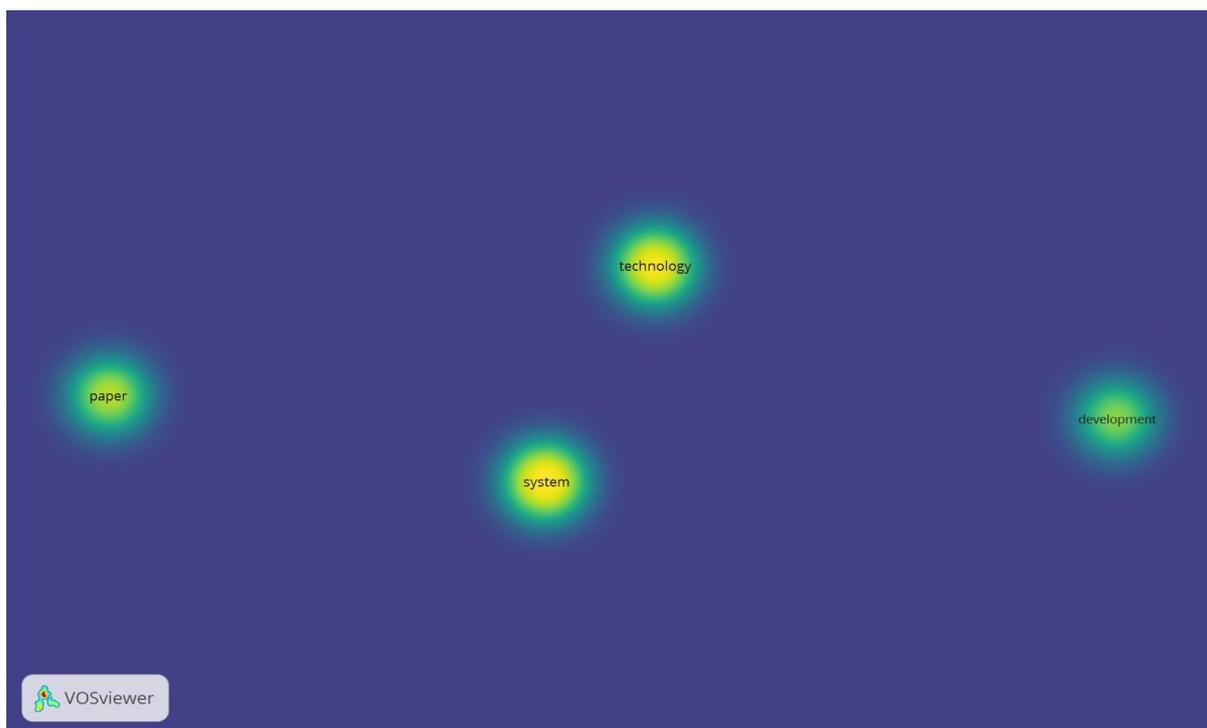
A tabela a seguir apresenta os dados do Tabela 01:

Ano de publicação	Quantidade de citações envolvendo WMS
2009	4
2010	9
2011	14
2012	14
2013	11
2014	14
2015	16
2016	7
2017	18
2018	5

**Tabela 01- Quantidade de Citações envolvendo WMS, por ano de publicação**  
**Fonte: Autoria Própria**

Os primeiros artigos citados foram publicados em 2009, começando timidamente com 4 artigos publicados, em 2010 foram um total de 9 artigos, o crescimento se manteve até 2012 chegando aos 14 artigos publicados tendo um queda considerável em 2013 com 11 publicações citadas, em 2014, o crescimento retomou chegando aos 16 artigos em 2015. 2016 ouve uma queda relevante para 7 artigos citados, e só em 2017 registrou o maior número de citações chegando os 18 artigos citados sobre WMS. Até o termino dessa pesquisa, 2018 registrou 5 artigos citados.

O gráfico 02, retirado do software VOSviewer, analisa quais foram as palavras chaves mais citadas que são elas: **paper, systems, technology, development (papel, sistemas, tecnologia, desenvolvimento)** mostrando que são palavras chaves que apresentam ligação direta com o WMS, tendo seu papel muito bem desenvolvido aliando sistemas e tecnologias, e contribuindo para o desenvolvimento de ambos.



**Figura 5: Análise das palavras chaves mais citadas**  
Fonte: VOSviewer

O estudo bibliométrico dos artigos onde cita o termo WMS e logística apontam o quão é importante o estudo e comprovação do sistema para as empresas, abordou-se na pesquisa 15 artigos onde apresenta com estudos e fatos concretos como o WMS mudou significativamente o processo de gestão, mostrando que a tecnologia voltada para a área de logística é um dos pilares para o crescimento e fortalecimento nas empresas.

## 5. Conclusão

O software WMS é um sistema de gerenciamento muito utilizado nos setores armazenagem, permitindo a execução de tarefas e o controle de entrada, saída e transporte de mercadorias e produtos, de uma forma precisa e eficiente. Toda essa qualidade atribuída ao WMS se deve principalmente ao fato do mesmo organizar adequadamente todas as informações de uso no planejamento, auxiliando nos roteiros e realizando múltiplas tarefas em todos os processos dentro dos pavilhões de estoque. Outro fator relevante no WMS são os seus processos que envolvem núcleos de operação e administração de todo um setor de armazenagem, sendo eles: entradas e saídas, monitoramento, indicação de endereços, separação de pedidos, embalagens, carregamento, expedição, transporte e todo o controle na documentação necessária.

Diante desse contexto, a análise bibliométrica em questão possibilitou a identificação do sistema como auxiliador do processos operacionais dos armazéns, tendo como ressalva as palavras desenvolvimento, papel, sistemas, tecnologia, ligando e aliando os artigos explorados. Os mesmos apresentam a aplicação do sistema WMS e os resultados obtidos com a implantação, o que se observa é um contexto de eficiência, aumento produtivo, agilidade, além do entrelaçamento do sistema com fatores da atualidade como a Internet das Coisas e a Indústria 4.0. E também um fator de

surgimento na Segunda Guerra Mundial, mas que auxilia na alimentação do sistema, de forma ágil, com informações, que é o caso do sistema de RFID (Identificação por Rádio Frequência) que foi objeto de pesquisa de 5 artigos dos 15 verificados através do estudo bibliométrico, esse sistema trata-se de um leitor de sinal que transfere informação ao contato com etiquetas presentes em produtos, animais, embalagens, entre equipamentos diversos. O WMS é bem mais que um software, ele é uma ferramenta que quando a empresa consegue executá-lo perfeitamente pode alavancar os lucros e diminuir os custos de manuseio dentro e fora da empresa, garantindo maior praticidade, mais agilidade nos processos e menos burocracia. Esses pontos em destaque demonstram que o investimento no sistema WMS se faz necessário e é recomendado para as empresas que desejam aperfeiçoar os processos logísticos, alinhando-o com fatores mais avançados, esperando resultados mais positivos, que também auxiliam no cenário competitivo.

### Referências

- Alves, T. E. M., Carneiro, C. M. B., Fonseca, R. de C., Nunes, R. V. & Santos, G. P. Sistema de informação logística (SIL) para o gerenciamento eficiente de centros de distribuição – estudo de caso em um Centro de Distribuição do ramo varejista. Publicado no XV Congresso Brasileiro de Custos. Curitiba - PR, 2008.
- Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em questão*, 12 (1), 11-32.
- Arozo, R. Softwares de supply chain management: Definições, principais funcionalidades e implantação por empresas brasileiras. In: FIGUEIREDO, K. F..
- Ballou, Ronald H. Gerenciamento da cadeia de suprimentos: planejamento, organizações e logística empresarial. São Paulo: Bookman, pp26, 2001.
- Bezerra, A. M., Bezerra, A. L. B. Vantagem Competitiva Logística Empresarial Baseada em Tecnologia de Informação. VI SEMEAD. 2009.
- Bowersox, D. J. Closs, D. J. Logística empresarial: o processo de integração da cadeia de suprimentos. 8. ed. São Paulo: Atlas, 2010.
- Breznik, Lidija. Can Information Technology Be a Source of Competitive Advantage? *Economic and Business Review*: Vol. 14. Nº 3. 2012.
- Fabrimatel. SISTEMA DE ARMAZENAGEM: WMS: O que essa sigla significa? 2014. Disponível em: <<http://www.sistemadearmazemagem.com.br/sistemawms-2/>>. Acesso em: 28 set. 2018.
- Fleury, P. F. & Wanke, P. Logística e gerenciamento da cadeia de suprimentos: Planejamento do fluxo de produtos e dos recursos. São Paulo: Atlas, 2003.
- Guarnieri, PATRÍCIA et al. WMS – Warehouse Management System: adaptação proposta para o gerenciamento da logística reversa. 2006. Disponível em: <<http://www.scielo.br/pdf/%0D/prod/v16n1/a11v16n1.pdf>>. Acesso em: 28 set. 2018.
- Las Casas, Alexandre L. “Plano de Marketing”. 2. ed. São Paulo: Atlas, 2007
- Leite, P. R. Logística reversa: meio ambiente e competitividade. 2009.

- Lima, Adilson C., Vasconcelos, Gabriela. Análise Bibliométrica da Produção Científica acerca da Controladoria. *Management Control Review*, v. 2, n. 1, 2017
- Mgitech. WMS : Rastreabilidade que vai além da remessa e entrega. 2018. Disponível em: <<http://mgitech.com.br/pt/solucao/wms/>>. Acesso em: 11 out. 2018.
- Platt, Allan Augusto; Nunes, Rogério da Silva. Logística e cadeia de suprimento . 1. ed. Florianópolis: UFSC, 2007. 23 p. v. 01. Disponível em: <<https://pt.slideshare.net/JanserGomes/livro-logistica-em-cadeia-de-suprimentos>>. Acesso em: 28 set. 2018.
- Porter, M.E., 1985. *Competitive advantage: Creating and sustaining superior performance*. Nova Iorque: The Free Press.
- Souza, C. D. de. (2013). A organização do conhecimento: estudo bibliométrico na base de dados ISI Web of Knowledge. *Biblios Online*, 51, 20-32.
- Turban, E.; Ranier JR., R. K.; Potter, R.E. *Introdução a sistemas de Informação uma abordagem gerencial*. Tradução Daniel Vieira. Rio de Janeiro: Elsevier, 2007.
- Van Eck, N. J., Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, v. 84, p. 523-538, 2010.
- Vanti, N. A. P. (2002). Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, 31 (2), 152-162.
- Wagner, JOHN A., 1952. *Comportamento Organizacional: Criando Vantagem Competitiva*. Tradução Sílvio Floreal Antunha. São Paulo: Saraiva, 2012.
- Wiggins, R.R. & Ruefli, W.T. (2005). *Schumpeter's ghost: Is hypercompetition making the best of times shorter?* *Strategic Management Journal*, 26 (10), 887-911.

## Análise da Logística Integrada da Empresa Natura Cosméticos S.A.

Amanda T. G. da Silva<sup>1</sup>, Bruna F. Paulino<sup>1</sup>, Geovana A. Pereira<sup>1</sup>, Juliane M. da Silva<sup>1</sup>, Jussara G. da Silva<sup>1</sup>, Gabriela de Vasconcelos<sup>1</sup>

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

(terezinha-amanda@hotmail.com) ,  
(brunafpaulino@gmail.com) ,  
(geovana0alcantara@hotmail.com) ,  
(julianemayara17@gmail.com) ,  
(jussaragabryelle1@gmail.com) ,  
(devasconcelos.gabriela@gmail.com)

**Resumo.** *As empresas estão buscando uma diferenciação amparada na eficiência da logística integrada e com isso se tornando cada vez mais competitiva no mercado. Nesse sentido, o presente artigo tem como principal objetivo explicar um pouco sobre a logística integrada da empresa de cosméticos Natura S.A a fim de analisar com se dá a logística integrada dessa organização. Este levantamento de dados sobre a logística integrada da Natura contribuiu para que possamos aprender mais sobre a cadeia logística da empresa em estudo e para ajudar em pesquisas futuras.*

**Abstract.** *The companies are seeking a differentiation based on the efficiency of the integrated logistics and with that becoming more and more competitive in the market. In this sense, the main objective of this article is to explain a little about the integrated logistics of the Natura S.A cosmetics company in order to analyze the integrated logistics of this organization. This survey of Natura's integrated logistics data has helped us to learn more about the logistics chain of the company under study and to assist in future research.*

### 1. Introdução

O Brasil ocupa a 3ª posição no ranking mundial no consumo de cosméticos segundo a empresa Euromonitor International. A cada ano, o país se aproxima de Estados Unidos e Japão, que são os atuais líderes mundiais. Esse é um mercado crescente que tem se favorecido da cultura da beleza, do envelhecimento da população e do aumento do poder de compra da classe C do Brasil [SEBRAE,2011].

O setor de beleza está entre os dez principais segmentos do varejo, e isso não é apenas mérito do público feminino, pois o público masculino tem modificado seus hábitos e investido um pouco mais em produtos de beleza, o Brasil ocupa o 2º lugar neste nicho específico [SEBRAE, 2018].

Nesse sentido, surge à empresa Natura com a missão de promover o bem-estar-bem – relações harmoniosas do indivíduo consegue mesmo, com os outros e com a

natureza [NATURA, 2018]. Para tanto a empresa conta com uma cadeia logística altamente estruturada.

Neste sentido, o presente estudo de caso objetiva analisar a logística integrada Natura, pois a empresa apresenta-se como referência na área de logística integrada. Que é de suma importância para conhecimentos acadêmicos. Diante disso, busca-se responder a seguinte questão de pesquisa: Como se dá a logística integrada da empresa Natura?

Nos próximos tópicos o estudo desenvolve o referencial teórico explorando o conceito de logística e logística integrada. Na sequência, o artigo apresenta as definições metodológicas e a análise e discussão de resultados, caracterizando a unidade de análise e descrevendo as especificidades evidenciadas.

## **2. Referencial Teórico**

### **2.1. Logística**

Schier, Lombardo e Cardoso (2012) comenta que a logística esteve presente desde os tempos mais remotos até hoje. Percebe-se que a logística existiu na construção das pirâmides do Egito, nas guerras para o transporte de alimentos, máquinas de guerra, armamentos, transporte de pessoas. Então a logística sempre esteve entre a humanidade e vem ajudando de diversas formas nas empresas.

O Council of Logistics Management tem uma visão muito bem representada do termo “logística”, que foi citado por Neto e Júnior (2002, p.40): “logística é o processo da cadeia de suprimentos que planeja, implementa e controla o fluxo e o estoque de bens e serviços e as informações relativas, do ponto de origem ao ponto de consumo, de maneira eficiente e eficaz, buscando a satisfação das necessidades do cliente”.

Segundo Rodrigues, Nagano e Musetti (2005), administrar a logística significa adequar a estrutura organizacional com linhas de comando e responsabilidade específicas às atividades logísticas que são consideradas com graus de importância diferentes conforme a atividade principal da empresa. O processo exige intensa interação inter e intra-empresas, formando o canal logístico que, de ser gerenciado para balancear os custos de seus membros, porque os canais devem beneficiar a todos os elos.

Silva, Silva, Benvenuto e Santos (2010), ressalta que “diante de sistemas com natureza tão flexível, a logística pode ser entendida como sendo a junção de quatro atividades básicas: as de aquisição, movimentação, armazenagem e entrega de produtos. Para que essas atividades funcionem, é de suma importância que tenha planejamento logístico [...]”. Desse modo com a logística permite-se aos empresários avaliar e analisar os pontos fortes e fracos de parte da cadeia de fornecimento da empresa, dando suporte em tomadas de decisões na estratégia da empresa.

Muitas empresas estão investindo em um sistema logístico eficiente, dessa forma proporciona um aumento do lucro e crescimento da empresa. A logística da Natura mostra-se como um processo que é bem gerenciado fornece um diferencial único perante o mercado, servindo dessa forma como um suporte na estratégia competitiva da empresa. Segundo Silva, Silva, Benvenuto e Santos (2010), o papel da logística no negócio aumentou tanto em escopo, quanto em importância estratégica. Integração do fluxo de material, produção e distribuição, vêm revolucionando, não somente a forma

como as empresas gerenciam suas atividades logísticas, mas também como gerenciam todo o seu negócio.

Para Silva, Silva, Benvenuto e Santos (2010, p. 3), “ao incorporar conceitos de marketing, qualidade, finanças e planejamento, a logística passou a ser uma disciplina multifuncional, aumentando a sua contribuição para a eficiência e eficácia na gestão”. Dessa forma, é possível identificar através da logística tanto as necessidades dentro da empresa, quanto os desejos que os clientes têm e demonstram preocupação com a sincronização das tarefas a serem executadas pela empresa como a racionalização dos fluxos de materiais e informações, procurando assim reduzir o máximo possível no processo logístico.

A logística deve ser aplicada para obtenção de vantagem competitiva, através da disponibilização do produto certo, na quantidade certa, no local certo, no momento certo, nas condições adequadas para o cliente certo a preço justo, atingindo assim a eficiência e eficácia do processo. [Schier, Lombardo e Cardoso 2012].

## **2.2 Logística Integrada**

A logística integrada é o meio no qual a empresa vincula os seus clientes e fornecedores com o processo de fluxo de informações e fluxo de materiais. Esse meio é o caminho que flui para a diferenciação de uma empresa perante o mercado, os clientes e os fornecedores gerando assim valor para a empresa, dessa forma aumentando a lucratividade e adquirindo uma boa visão do mercado [Silva, Silva, Benvenuto e Santos 2010].

Hoje em dia a logística é uma das principais fontes de vantagem competitiva dentro de organizações produtivas. Envolvendo a parte estratégica, tática e operacional da empresa, visando uma integração de todo o processo produtivo e garantindo vantagens propícias [Silva, Silva, Benvenuto e Santos 2010].

É necessário definir um trabalho de rede de relacionamento que proporcione valor ao cliente, de forma que todos os componentes da cadeia sejam capazes de unir os objetivos estratégicos [Rodrigues, Nagano e Musetti 2005]. Então, com a logística integrada à empresa consegue transformar o seu benchmarking, e dessa forma mostrando que com o planejamento logístico integrado a empresa vai agregar valor ao produto percebido pelo cliente.

## **3. Metodologia**

O seguinte artigo adotou uma abordagem de cunho qualitativo, cujo foco central foi à análise da logística integrada da Natura. Segundo Denzin e Lincoln (2006), a pesquisa qualitativa envolve uma abordagem interpretativa do mundo, o que significa que seus pesquisadores estudam as coisas em seus cenários naturais, tentando entender os fenômenos em termos dos significados que as pessoas a eles conferem. Seguindo essa linha de raciocínio, Vieira e Zouain (2005) afirmam que a pesquisa qualitativa atribui importância fundamental aos depoimentos dos atores sociais envolvidos, aos discursos e aos significados transmitidos por eles. Nesse sentido, esse tipo de pesquisa preza pela descrição detalhada dos fenômenos e dos elementos que o envolvem.

A pesquisa pode ser considerada descritiva, visto que levou a cabo coleta de dados e observação para descrever as características do sistema logístico integrado da

Natura, verificando assim a sua estratégia competitiva nesse âmbito. Segundo Vergara (2000), a pesquisa descritiva expõe as características de determinada população ou fenômeno, estabelece correlações entre variáveis e define sua natureza. A autora coloca também que a pesquisa não tem o compromisso de explicar os fenômenos que descreve, embora sirva de base para tal explicação.

O presente artigo caracteriza-se como levantamento, de modo que empreendeu-se a coleta de dados por meio do site da natura, onde esses dados foram utilizados para a construção deste artigo.

Neste estudo, também se utilizou de uma pesquisa bibliográfica a fim de obter um embasamento sólido quanto aos assuntos ligados à logística e logística integrada gerando subsídios teóricos para a interpretação dos resultados.

#### **4. Análise e Discussão de Resultados**

Neste levantamento de dados, será focado na logística integrada da empresa de cosméticos Natura S.A. Onde dá-se início no plano estratégico, no qual serão realizadas as decisões, que irá se relacionar com a gestão estratégica da empresa. A logística atua em decisões sobre serviços, produtos, alocação de investimentos e recursos.

Já no plano das atividades táticas irá relacionar ao desenvolvimento das metas estratégicas e planejamento do sistema logístico feito pelo plano de atividades estratégicas. As atividades táticas atuam sobre as decisões de fornecedores, sistemas de CPC, rede de distribuição e subcontratação de serviços.

No plano das atividades operacionais é onde se relaciona a gestão do dia a dia da rede logística. Atua em melhorar o sistema e também em sua manutenção para solucionar problemas. É no plano das atividades operacionais que os vendedores têm o contato com o cliente e sabe bem o que ele necessita.

A logística integrada vai englobar os setores de produção, suprimentos, gestão de estoque, marketing, finanças, recursos humanos e gestão de uma empresa para que todos trabalhem de forma sinérgica e que todos se comuniquem para uma performance de qualidade, que chegue ao cliente dentro do prazo e para que o cliente esteja satisfeito com o produto ou serviço feito pela empresa.

A estratégia competitiva na implementação da logística integrada traz consigo resultados a longo prazo para a empresa, dessa forma acumulando competências com funcionalidades futuras de valor agregado.

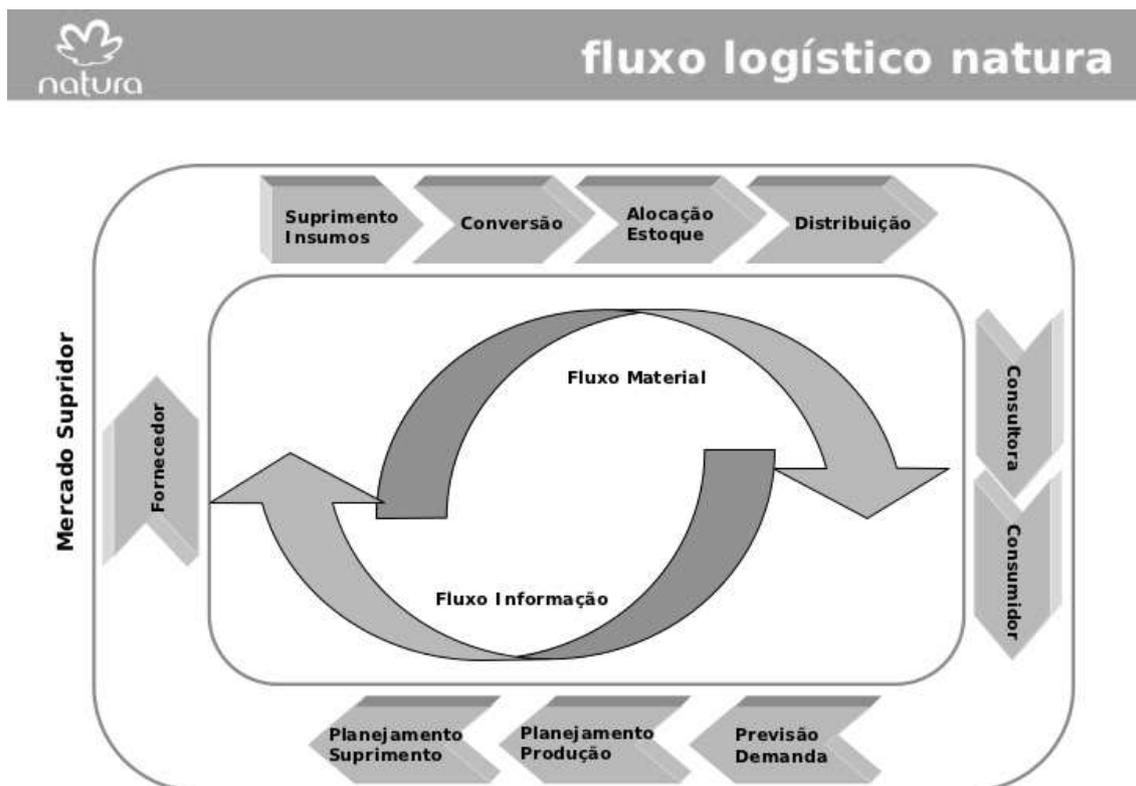
A logística utilizada pela Natura agrega valor para a empresa, pois com o planejamento adequado em sua cadeia de suprimentos gera diferenciação perante os concorrentes, como podemos citar a Avon e o Boticário, que também trabalham no mesmo ramo de cosméticos.

Em grande parte das cadeias logísticas, assim como na Natura não é diferente, o processo de transformação inicia-se junto à fonte de recursos naturais (produtor, neste exemplo), em seguida a matéria-prima é armazenada e na sequência encaminhada à unidade transformadora, conhecida como fábrica ou indústria. Uma vez transformada a matéria-prima em produto acabado, esse é alocado em embalagens adequadas para sua proteção e transporte, para ser enviado ao centro de distribuição. Quando a venda ocorre através de redes varejistas, o produto acabado ainda é encaminhado ao varejista para,

então, ser oferecido ao consumidor final. Como ocorre na empresa Natura, do centro de distribuição, o produto acabado é encaminhado diretamente ao consumidor (neste caso, para as Consultoras de vendas as quais se responsabilizam pela entrega ao cliente final).

Deste modo, conforme Martel & Vieira (2008), considera-se que a cadeia logística é formada pelo conjunto de redes logísticas internas das empresas membro, e os fluxos nessas redes são associados às etapas do processo de transformação. Assim, para conseguir gerenciar os fluxos de materiais e informações e aumentar a competitividade no mercado, há necessidade de se recorrer a uma rede de planos fundamentados nas decisões de curto (planos operacionais), médio (planos táticos) e longo (planos estratégicos) prazos.

As informações necessárias para se iniciar o planejamento da produção são recebidas de clientes e fluem pela empresa na forma de atividades de vendas, previsões e pedidos e, são filtradas em planos específicos de compras e de produção. O processo apresenta duas ações inter-relacionadas: fluxo de materiais e fluxo de informações. Onde esse processo passa das informações sobre a demanda, decorrendo planejamentos de produção e suprimentos, para daí serem requisitados aos fornecedores os suprimentos que serão transformados em produtos que seguirão para ser entregue pelas consultoras as clientes. Segue o fluxograma da cadeia logística integrada da Natura:



**Figura 1. Fluxograma da logística da Natura**

No caso da empresa, cuja estratégia de vendas é a direta, é possível chegar até suas consumidoras de forma personalizada e explorando os principais atributos de seus produtos de modo adequado e convincente. Outro aspecto relevante nesse sistema de distribuição é a não-dependência de grandes redes atacadistas e varejistas para a distribuição dos produtos, o que torna a empresa imune ao crescente aumento do poder de negociação dos intermediários da cadeia de comercialização, fato que afeta,

atualmente, a maioria das empresas de produtos de consumo que deles dependem para escoar sua produção. Além desses fatores, o sistema de venda direta, com a amplitude do construído pela empresa Natura, constitui uma forte barreira de entrada para novos competidores nesse setor.

Então, dessa forma o levantamento de dados possibilitou o conhecimento sobre a formação de uma logística integrada utilizada na Natura. Esse processo logístico contribuiu de forma que toda a empresa esteja conectada, vantagens competitivas perante os concorrentes e satisfação dos clientes.

## 5. Considerações Finais

A Natura mostra com bastante evidencia a questão sustentabilidade em suas campanhas de produtos, mas não podemos nos esquecer de que ela também realiza um processo de logística integrada, onde mostra com eficiência sua cadeia logística de maneira formidável e sem sombra de dúvidas um bom trabalho de planejamento.

O gerenciamento da logística requer o conhecimento de toda a cadeia produtiva. E a cada ano que passa a Natura aumenta a sua produção e conseqüentemente seu volume de vendas, fruto da necessidade do mercado que a cada dia está mais exigente.

Este levantamento de dados sobre a logística integrada da Natura contribuiu para que possamos aprender mais sobre a cadeia logística da empresa em estudo e para ajudar em pesquisas futuras. A logística da natura mostra-nos que cada fluxo de informação é importante para que toda a cadeia logística possa ser feita corretamente e assim atingindo o sucesso.

## Referências

- Augusto, C. A.; Souza, J. P.; Dallagnelo, E. H. L.; Cario, S. A. F. Pesquisa Qualitativa: rigor metodológico no tratamento da teoria dos custos de transação em artigos apresentados nos congressos da Sober (2007-2011). Revista de Economia e Sociologia Rural, volume 51, nº 4. Brasília, 2013.
- Ballou, R. H. Logística Empresarial: Transportes, Administração de materiais e Distribuição física. São Paulo: Atlas, 2009.
- Braga, Moema Mesquita da Silva. Campanhas da natura: a comercialização do bem-estar, uma nova tendência de marketing do século XXI. Disponível em: <[http://fido.palermo.edu/servicios\\_dyc/encuentro2007/02\\_auspicios\\_publicaciones/actas\\_diseno/articulos\\_pdf/A6030.pdf](http://fido.palermo.edu/servicios_dyc/encuentro2007/02_auspicios_publicaciones/actas_diseno/articulos_pdf/A6030.pdf)>. Acesso em: 28 de setembro de 2018.
- Ferraz Neto e Kühne Júnior . Suplemento de Logística. Revista FAE, 2000.
- Gioanelli, Anderson. NATURA. Disponível em: <<http://logisticaversa.org/2015/08/30/natura/>>. Acesso em: 28 de setembro de 2018.
- Melo, Mariana Pereira. Cadeia de Suprimentos: Curso Técnico em Logística: Educação a distância / Mariana Pereira Melo. – Recife: Secretaria Executiva de Educação Profissional de Pernambuco, 2016.
- Musetti, M. A. A engenharia e as capacitações para a logística integrada. Cobenge, 2001.

- Neto, Ferraes.; Júnior, Kühne. Suplemento de logística. Revista FAE, 2000.
- Peña, B. K.; Mori, H.; Reis, G. O.; Lima, I. X.; Azevedo, E.; Meleiro, M.V.; Campello, M. Logística Reversa da Empresa Natura Cosméticos S. A. Simpósio de Excelência em Gestão e Tecnologia, Rio de Janeiro, 2017.
- Portal São Francisco. História da Natura. Disponível em: <<https://www.portalsaofrancisco.com.br/curiosidades/historia-da-natura>>. Acesso em: 28 de setembro de 2018.
- Rodrigues, E. de A.; Nagano, M.S.; Musetti, M. A. Um estudo sobre as especificidades da logística integrada dos Correios. SIMPEP – Bauru, São Paulo, 2005.
- Schier, C.U. da C.; Lombardo, M. A.; Cardoso, S. Logística integrada na cadeia de suprimentos (supply chain). Revista eletrônica dos cursos de Administração e Ciências Contábeis. Revista nº 07, 2012.
- Silva, A. T.; Silva, W. L. V.; Benvenuto, S. R. dos S.; Santos, Z. A. da S. A Logística Integrada Como Fonte de Vantagem Competitiva: O caso de uma empresa do setor de mineração. Encontro Nacional de Engenharia de Produção, São Paulo, 2010.
- SILVA, I. C. D. Venda, Logística e Distribuição da Natura Cosméticos. **Slide Sare**, 28 março 2012. Disponível em: <<https://pt.slideshare.net/expogestao/itamar-correia>>. Acesso em: 19 outubro 2018.
- Silva, Vanina M. D. Sistema de Logística da empresa brasileira de cosméticos NATURA. Disponível em: <<http://producaologistica.blogspot.com/2009/08/sistema-de-logistica-da-empresa.html>>. Acesso em: 21 de setembro de 2018.
- Soares, D. S.; Souza, E. R.; Soares, F. F.; Soares, M. F.; Custódio, E. R. S. Logística Integrada: satisfação dos clientes e redução de custos. Revista Científica do Universitário – Lins – São Paulo, ano 2, nº 4, 2011.
- Strada. Logística Integrada: O que é e como pode beneficiar sua empresa. Disponível em: <<http://www.stradasolucoes.com.br/logistica-integrada-beneficios/>>. Acesso em: 3 de outubro de 2018.
- Teixeira, Larissa Xavier Natario. Mercado de Cosméticos. Disponível em: <<http://www.sebraemercados.com.br/mercado-de-cosmeticos/>>. Acesso em: 28 de setembro de 2018.

## **Análise da logística de uma organização do ramo varejista como instrumento de vantagem competitiva**

Cleiciane Josivania Ventura, Gislayne Barbosa Morais, Maria Jullyana de Alves Lima, Sara Verissimo Silva, Vanessa das Graças Santos Ramos, Gabriela de Vasconcelos

[cleiciiventura@gmail.com](mailto:cleiciiventura@gmail.com), [laynemorais15@gmail.com](mailto:laynemorais15@gmail.com),  
[marialimajullyana@gmail.com](mailto:marialimajullyana@gmail.com), [vanessa.dgsr@gmail.com](mailto:vanessa.dgsr@gmail.com)

[devasconcelos.gabriela@hotmail.com](mailto:devasconcelos.gabriela@hotmail.com)

### **Resumo**

**Contexto:** A dinâmica do mercado ocasiona uma forte concorrência entre as empresas, e a logística pode ser uma das mais importantes áreas da organização que pode garantir a sua vantagem competitiva, sendo capaz de criar mecanismos que agregam valor entre seus produtos e serviços para o seu cliente final. **Objetivo:** Este trabalho propõe através de uma metodologia exploratória com abordagem qualitativa, o objetivo de analisar a importância da logística como vantagens competitivas de uma empresa varejista. **Conclusão:** Por meio do trabalho apresentado foi possível concluir-se que a empresa na qual foi realizada o estudo de caso, apresenta um processo de logística empresarial que é estruturado como ferramenta estratégica para obtenção de vantagem competitiva e isso auxilia em um maior entendimento do tema e sua relevância para as empresas, mais especificamente aos processos da cadeia de suprimentos da empresa.

### **Abstract**

**Context:** The dynamic of the market causes a very strong competition between the companies and the logistic can be one of the most important areas of the organization wich can provide a competitive advantage, being able to create the mecanisms that accreate value in their products and services for the final client. **Objective:** This work proposes through a exploratory methodology with qualitative approaches, the goal on analysing the significance of the logistics as an competitive advantage of one of the retailer company. **Conclusion:** Through the work stated it has been possible to conclude that the company wich were used as part of the study, shows a business logistic that is structured as strategic tool to acquire a competitive advantage, that also provides a better understanding of the theme and their relevancies for the company, more especifically to the chain process of suplies of the company.

### **Introdução**

O mercado se encontra cada vez mais competitivo e com constantes mudanças socioeconômicas, a partir de então, surge a necessidade das organizações em buscar um diferencial para se manter mais atrativa no mercado, atender a essas mudanças e as demandas do seu público-alvo.

A área da logística exerce uma grande influência em todos os processos organizacionais, ela está interligada a todas as áreas da empresa, otimizando os processos que vão desde a sua formação até a entrega ao cliente final. Além disso, a logística impacta diretamente no marketing empresarial sendo capaz de gerar valor aos produtos e serviços ofertados, sem atrasos, com qualidade, satisfação do cliente suprimindo a sua demanda. Como explica Santos (2017) a missão da logística é colocar os produtos ou serviços certos, no lugar e momento certo, e nas condições desejadas pelos seus clientes, com o objetivo de criar uns mecanismos de entregar os produtos aos seus clientes finais com menos tempo, custo e mais qualidade.

A logística empresarial tem como objetivo organizar os processos de materiais e estoque da empresa, com foco operacional na cadeia de abastecimento. De acordo com Pedrosa (2008), ela é responsável por planejar, implementar, gerenciar o fluxo de matérias-primas, produtos e informações ao longo da cadeia de suprimento. Desse modo, a logística empresarial é capaz de garantir vantagem competitiva às organizações agregando valor a cadeia logística.

Diante disso, a logística empresarial pode ser um grande diferencial competitivo para essas empresas proporcionando a elas a possibilidade de poder se destacar através de estratégias competitivas e aprimorando o nível de serviços oferecido aos seus clientes, e se tornando uma das principais peças do planejamento estratégico das organizações.

Diante desse contexto, este estudo tem como finalidade analisar a importância da logística como vantagem competitiva de uma empresa varejista, a fim de responder a seguinte problemática: Como os processos logísticos podem conferir vantagem competitiva para a organização?

## **Referencial Teórico**

### **Logística**

Historicamente a logística teve sua origem nos campos militares, onde foi se desenvolvendo e teve o intuito em ajudar os militares a vencerem as batalhas, isto é, para obter que as estratégias usadas pelos generais tivessem o resultado positivo. Para isso era fundamental a locomoção das tropas através do campo de batalha fosse mantida da melhor forma possível, sendo indispensável, assim, todo um mecanismo logístico. Nos dias atuais, diante da globalização a logística tem a finalidade de conquistar batalhas comerciais. Antigamente nas empresas, a logística era visualizada como uma ação de defesa, que não lhes acrescia lucro, nem tinha consequências sobre as técnicas e criações de negócios, essa seção era vista como um custo comum.

Para Leite (2003), a logística engloba novas áreas de atuações, a logística faz parte da gestão de fluxos entre as funções do negócio. O conceito de logística atualmente integra maior grandeza de fluxos que no passado. Historicamente as empresas incluíam a simples entrada de matérias-primas ou a saída dos produtos acabados em sua definição de logística. Atualmente, no entanto, essa definição desenvolve-se e inclui todas as formas de movimentos de produtos e informações.

De modo geral, a logística está ligada a todos processos que cercam a compra, transporte, depósito e entrega de produtos. Assim sucede em todas as organizações que tentam aumentar sua lucratividade utilizando esse método. A logística se enquadra em todos os procedimentos que pretendem ter eficiência organizacional, por meio de afirmar

que as exigências do consumidor sejam respondidas, bem como sustentar outras áreas da organização.

A logística empresarial se trata da gestão de estoque de uma empresa, organizar os pedidos de produtos e transportar o mesmo até o consumidor final. É uma das funções fundamentais dentro da organização e que pode ocasionar em grande vantagem competitiva caso bem dirigida, ou em prejuízo caso não tenha sua importância levada em conta.

Novaes (2007) certifica que a logística se relaciona a crescer valor de espaço, período, qualidade e informação, e explica que as atividades logísticas geralmente são confundidas apenas com transporte e armazenagem. Desta forma, a logística empresarial que na atualidade é desempenhada nas empresas implicada na otimização dos recursos, subtraindo processos que causem custos elevados e não implementam valor ao consumidor final. O Council of Supply Chain Management Professional dos Estados Unidos da América (EUA) considera a logística como fazendo parte da gestão da cadeia de suprimentos que objetiva, produz e sistematiza de forma hábil o fluxo de estoque de produtos, serviços e informações relacionadas, entre o ponto de origem e o ponto de consumo com intenção de atender às necessidades dos clientes (CSCMP, 2014).

Na literatura e estudos atuais é possível encontrar o conceito de logística empresarial agregado como parte da gestão da cadeia de suprimentos, também referenciada como Supply Chain Management – SCM. Simchi-Levi D., Kaminsky e Simch-Levi E. (2010) defendem que “ambas as definições destacam a integração dos diferentes componentes da cadeia de suprimentos”. Para Ching (2009) a SCM é a junção da organização e direção de fluxo de itens, informações e recursos desde os fornecedores até o consumidor final, com a cadeia logística sendo dirigida de forma cooperadora objetivando a melhoria de todos os envolvidos. Um dos temas fundamentais para a compreensão do conceito de gestão da cadeia de suprimentos é a inclusão entre as peças incluídas.

### **Gestão de Estoque**

Gestão de estoque é um processo fundamental para as empresas que pretendem reduzir seus custos de estoque, no entanto sem perder vendas por estar faltando produto na loja. Isto acontece quando a empresa tem o domínio de seu estoque através de uma eficiente gestão e faz o acompanhamento das vendas dos produtos. Agindo assim é possível controlar o volume de demanda e oferta dos produtos (Pires, 2013).

De acordo com Nogueira (2012) para se tomar a decisão de estocar ou não produtos, deve-se pensar nos aspectos diante da disponibilidade de aquisição ou a variação no que está sendo demandado. Estas ideias devem ser examinadas pela(s) pessoa(s) responsável(is) pela gestão de estoques, este é um fator relevante para a qualidade no nível de serviço prestado ao cliente interno ou externo.

Com isto, uma gestão de qualidade na demanda ajuda a empresa na redução dos custos com manutenção de estoques, reduz os riscos da obsolescência dos materiais e reduz o tempo de resposta ao mercado. Para que as empresas tenham uma gestão de demanda competente é necessário a integração entre os setores de marketing, produção e logística. Sendo assim é possível prever a demanda que será capaz de otimizar a utilização do espaço físico de seus centros de distribuição.

Percebesse que a gestão de estoques é um motivo predominante para o êxito ou o insucesso de uma organização. Possuir colaboradores habilitados e sistemas que facilitem

no controle de estoques é um ponto importante para o sucesso empresarial. A expressão estoque está aplicado com muito destaque nas organizações deste século, contudo é uma expressão que há muito tempo já existe, desde o período mais distante da existência humana.

De acordo Accioly et Al (2008), as pirâmides e outras maravilhas do mundo antigo, assim como a passagem dos alpes por Aníbal com cerca de 40 mil soldados, foram desenvolvidos por meios avançados de projetar demanda e controlar os estoques e a produção. Desse modo, a gestão de estoque do período, foi efetivo apenas porque havia um sistema de organização dos processamentos produtores.

Deste então, surgiram sistemas e métodos com o objetivo de sanar os prováveis impasses descobertos na gestão de estoques, devido à produção que aumentavam. Em seguida, com o surgimento do lote econômico, iniciou-se a ter maior relevância ao sistema de previsão.

Um divisor de águas na evolução da gestão dos estoques foi a definição de demanda dependente e independente, quando ficou esclarecido que a demanda independente deveria ser tratada pelo método de ponto de reposição estatístico (ROP), e a demanda dependente por planejamento de necessidades de material (MRP).

Brevemente, no Japão, surgiu o conceito Just in Time (produção no momento certo), que rapidamente foi englobado como um instrumento de gestão de estoques. Nos dias atuais o conceito do Just In Time tem apresentado grande importância para as empresas formarem o planejamento das necessidades de materiais (MRP), aumentando para o planejamento da necessidade de distribuição dos materiais.

### **Logística Empresarial como Ferramenta Estratégica para Vantagem Competitiva**

A logística auxilia as empresas a otimizarem seus processos e conseguir resultados melhores de uma forma mais clara. A maioria das empresas trabalham fortemente para conseguir resultados melhores a cada dia, essa é uma tarefa que não é fácil, e vai permanecer na frente aquele que usar da melhor vantagem competitiva. Segundo Council of Logistic Management (1996), a logística pode ser definida como o processo de planejar, executar e controlar a eficiência, o fluxo e depósitos de mercadorias, serviços e informações relacionadas a partir da origem ao ponto de consumo, com a finalidade de atender às exigências dos clientes.

Para Ballou (1993), a finalidade da logística é aprimorar o nível de serviço oferecido ao cliente, onde o poder do serviço logístico impacta positivamente na qualidade do fluxo de produtos e serviços gerenciado. A Logística Empresarial hoje está ligada ao fato de uma empresa se relacionar com seu cliente, interagindo da melhor forma com a cadeia produtiva para alcançar o objetivo de estar à frente na competição do mercado.

Para conseguir maior vantagem competitiva, as empresas contam com o auxílio de sistemas integrados de informação, tornando mais rápido e prático a coleta de informações da empresa, onde o processo produtivo é guiado por algumas tecnologias do tipo: Electronic Data Interchange (EDI), o Warehouse Management System (WMS), tecnologia de código de barras e o Vendor Managed Inventor (VMI). Os Sistemas de Informação dão suporte para as empresas aumentarem sua eficiência.

Para adquirir vantagem competitiva numa empresa varejista é necessário, entre outras coisas, que haja o serviço ao cliente de um modo que agrade o máximo seu público,

esta é uma parte primordial da estratégia de marketing. A relação entre o marketing e a logística, estão totalmente interligadas quando se trata de focar no consumidor final, uma análise com os 4Ps do marketing possibilita que o produto certo esteja no lugar certo, na hora certa, para o cliente certo. Salientando que um dos fatores importantes para a conquista de clientes fiéis é o ponto de venda da empresa ou centro de distribuição, é através dele que a empresa vai estipular o prazo de entrega e os custos do serviço logístico oferecido, até a entrega da mercadoria ao cliente (Grewal, 2012). A vantagem competitiva é a competência da empresa em mostrar seus produtos ou serviços de tal maneira que os concorrentes não possam alcançar.

Segundo Fleury (2012), os serviços logísticos são uma oportunidade de mercado para que as empresas que querem seus produtos ou serviços à frente da concorrência, consigam se destacar no mercado, auxiliando consumidor final na sua decisão de compra. A agilidade na entrega dos produtos influencia fortemente na satisfação do cliente, mesmo que preço e qualidade sejam relevantes, o cliente vai à loja em busca de um produto e quer que o produto esteja lá, então a agilidade na entrega, a disponibilidade na loja e a junção de um bom atendimento fazem com que o cliente sai da loja satisfeito.

Desempenhar vantagem competitiva não depende só do varejista, depende também de todos os serviços agregados ao longo do processo da cadeia de suprimentos. A vantagem competitiva tem como pontos importantes: a redução de custos, o aperfeiçoamento e clareza nos processos, a qualidade dos produtos e os serviços oferecidos ao consumidor final.

A implantação da logística empresarial auxilia para que a empresa tenha a possibilidade de se ajustar às mudanças externas. Tendo a capacidade de entregar ao seu consumidor final o produto certo, na quantidade correta, com variedade e qualidade, e tudo isto dentro do prazo ao qual o cliente está disposto a esperar.

Para Marques (2012) a estratégia empresarial refere-se a desenhar estratégias para cada um dos setores visando atingir o principal objetivo da empresa. Estratégia de produção é traçada a partir da visão geral da organização diante dos resultados que pretendem alcançar ao final de um período estabelecido. A estratégia é o modo de planejar e executar as ações da empresa, visando alcançar uma melhor posição diante da concorrência ou ainda, atingindo o objetivo da empresa dentro de um prazo pré-estabelecido. Em razão disso, a empresa é direcionada para aumentar sua lucratividade, empregando os seus recursos de forma mais vantajosa, auxiliando a empresa a atingir diferentes metas planejadas diante do mercado (Ribeiro, 2012).

O diferencial diante das demais empresas deve ser percebido pelos consumidores como uma vantagem a eles, as empresas devem focar em beneficiar o seu público alvo, para aumentar o fluxo de vendas e lucratividade da empresa através da satisfação do consumidor (Kotler; Keller, 2006).

### **Metodologia**

O presente estudo caracteriza-se pela sua abordagem qualitativa já que foi realizado uma análise cujo intuito foi explorar as particularidades de uma empresa do ramo varejista. O principal propósito foi primeiramente realizar um estudo de caso para delimitar a coleta de dados, que segundo a linha de pensamento de Cervo e Bervian (2009) pode ser retratado como sendo a pesquisa sobre um determinado indivíduo, família, grupo ou comunidade que seja representativo do seu universo, para examinar aspectos variados de sua vida.

Esse estudo de caso que conforme Zanella (2009), nada mais é que a pesquisa voltada para explicar uma realidade específica que pode ser descrita na análise dos dados que na presente pesquisa foi realizada de forma intuitiva e indutiva pelo pesquisador, não requerendo o uso de técnicas e métodos estatísticos. E por fim, teve como preocupação maior a interpretação de fenômenos e a atribuição de resultados.

Posterior ao estudo de caso, foi realizada uma abordagem qualitativa já que se tratava de uma pesquisa exploratória executada com a gerente da empresa em questão. Essa foi a abordagem selecionada para a pesquisa justamente pelo fato de que a principal finalidade era coletar o máximo de informações plausíveis de um determinado sujeito que estivesse diretamente ligado(a) as experiências vivenciadas no dia a dia da corporativa, fazendo com que fosse agregado o devido prestígio.

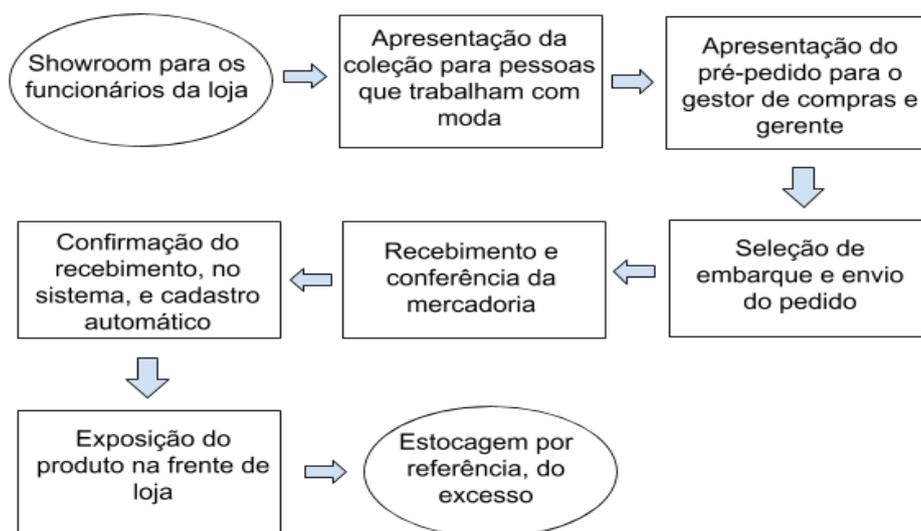
Para tanto, adotou-se como procedimento de coleta de dados a entrevista estruturada, que serviu para analisar o modelo adotado pela gerente da loja analisada, onde foram realizadas seis perguntas para a funcionária, feitas de forma formal e objetiva.

### **Análise e discussão dos resultados**

Conforme observado a importância da logística e organização do estoque, foi aplicado um questionário para entender como funciona o processo, do pedido da mercadoria, até a parte final onde é vendida ao consumidor. A observação aconteceu durante o período de funcionamento da loja com ênfase em como as colaboradoras que operam o estoque agem diante das situações, e a compra que a responsável pelo pedido faz durante o período de abastecimento. O Estoque de segurança funciona como apoio quando o estoque atual está em nível baixo, e a mercadoria do pedido está em trânsito. A entrevista foi realizada observando e seguindo os conceitos sobre gestão de estoque, foi aplicado um questionário com a gerente que respondeu as seguintes perguntas de acordo com a prática aplicada na loja.

- A) Quem faz a organização do estoque?
- B) Como é feita a localização de mercadorias?
- C) São utilizados sistemas na entrada e saída de mercadoria?
- D) A loja possui estoque de segurança?
- E) Os produtos possuem prazos de validade?
- F) Quais as estratégias utilizadas quando o estoque está com excesso de mercadorias?

De acordo com a entrevista e as respostas obtidas as respostas vão ser apresentadas logo abaixo, seguindo os conceitos dos autores mencionados que falam sobre vantagem competitiva na organização do estoque em empresas. O mapeamento abaixo mostra como funciona as fases iniciais do pedido da mercadoria, até o momento em que os produtos chegam no estoque.



**Figura 1. Mapeamento do pedido de produtos que compõem o estoque**

**Fonte: Autoria Própria.**

O produto quando é armazenado e setorizado de forma correta, o processo fica mais organizado e a visualização do produto facilita a eficiência ao giro do estoque, para quem opera diariamente, ajudando assim a uma visão ampla, de como está o estoque atual e o estoque de segurança, podendo ajudar na programação de futuras compras, evitando o excesso de compras, e não vencendo os prazos de validades. Pois assim como o setor alimentício, os produtos do varejo também têm seus prazos relacionados a obsolescência da mercadoria.

Conforme foi observado no questionário aplicado, a estrutura de organização do estoque segue os padrões abordados para a vantagem competitiva nas empresas quando se trata de armazenagem. Os produtos são primeiros expostos na área de venda e em seguida organizados e setorizados no estoque.

A exposição no PV (ponto de venda) é de vital importância, pois é onde são vistos aos olhos dos clientes. Seguindo a linha de organização, a arrumação e setorização desses produtos são feitos pela estoquista de loja, com auxílio da gerente e vendedores. Foi observado que são utilizados códigos e referências para setorizar os produtos, e isso facilita no processo das próprias vendedoras ao precisar localizar peças para reposição na área de vendas.

As entradas e saídas de mercadorias funcionam como gestão do que está disponível o que foi vendido e saiu da loja, conforme foi analisado a loja utiliza sistemas para fazer esse processo. Empresas que são maiores e trabalham com estocagem de produtos, necessitam de um bom sistema para executar esse gerenciamento, pois ele manual fica inviável se tratando de uma empresa de médio/grande porte. De acordo com o questionário respondido a loja utiliza um sistema prático para calcular o que entrou e saiu em tempo real, o sistema utilizado é o LINX, obtendo assim um controle maior sobre as peças e fornecendo informações claras e objetivas aos seus clientes.

Novaes (2007) enfatiza que para obter vantagem competitiva é importante além da qualidade dos produtos e clareza nos processos, é preciso também observar os custos. Antes o fato do estoque está lotado representava muito dinheiro envolvido, mas hoje com a grande concorrência e o consumo de novidades acelerado, as organizações estão se preocupando e investido mais em estocagens organizadas, trabalhando em cima do que é

necessário, evitando problemas com produtos que possam chegar ao prazo de validade em estoque.

No questionário aplicado, foi informado que a empresa trabalha com estoque de segurança, ou seja, não pode faltar o produto em loja, mas não existe estocar produtos em excesso e acabar chegando em seu prazo de validade. Assim como empresas que trabalham no setor alimentício, o varejo também precisa analisar os prazos de validade, de acordo com a gerente da loja recebe seis coleções diferentes anualmente, e isso faz com que cada coleção tenha sua particularidade e preferência. Então as antigas remarcam e entram com preço de liquidação, o que não é tão favorável a empresa, pois são vendidas quase que com preço de custo.

Ballou (1993) nos faz entender a importância da logística, e que vários são os fatores atrelados a ela, foi observado no questionário que além de fazer um resumo de vendas passadas, a empresa realiza inventário semestral, auxiliando bastante para a noção dos produtos disponíveis, esse é um dos objetivos dos inventários nas grandes empresas, nortear uma base de quantos produtos a empresa possui em números reais.

As famosas liquidações servem exatamente para baixar o nível de produtos e estoque, grande parte das empresas fazem isso durante o ano, foi constatado que a empresa analisada utiliza essa tática pelo menos 3 vezes ao ano, o sistema acusa excesso em estoque e são pensadas estratégias de como manter o nível bem alinhado para não sofrer consequências futuras. Como argumenta Fleury (2012) os processos logísticos são oportunidades que o mercado encontra em ganhar cada vez mais vantagem competitiva, e de acordo com os estudos feitos sobre o tema e pontos analisados, a empresa encontra-se em vantagem competitiva por conseguir abranger os principais argumentos para se manter um estoque com organização e giro produtivo.

### **Conclusão**

O presente estudo teve como principal objetivo analisar como funciona a logística de uma determinada empresa, que segue o ramo varejista. E para isso se verificou a necessidade de realizar um estudo de caso para coletar melhor os dados necessários. Após a averiguação de todos os processos, pode-se dizer que o objetivo geral deste estudo foi realizado com deleite, pois os resultados esperados na pesquisa demonstraram um maior entendimento do tema e sua relevância para as empresas, mais especificamente aos processos da cadeia de suprimentos da empresa objeto de estudo, atuante no ramo varejista.

Quanto ao problema de pesquisa, que visa compreender como a logística empresarial pode ajudar para obtenção de vantagem competitiva, concluiu-se que mesmo não havendo a aplicação de todas as ferramentas de logística empresarial estudadas no presente artigo, a empresa, através da utilização de algumas destas ferramentas consegue obter diferencial competitivo conforme resultados obtidos nas análises. Assim sendo, é possível dizer que através da análise dos resultados realizada por meio de estudo de caso e entrevista com a atual gerente da loja, chegamos à conclusão de que quando a logística é utilizada como um ícone de vantagem competitiva, podemos notoriamente constatar que a maior eficiência e eficácia no serviço prestado, podendo assim acarretar diretamente na redução de custos.

Mediante esses ajustes, é perceptível que a logística empresarial estuda as possibilidades reais de prover um nível melhor de rentabilidade nos serviços de distribuição aos clientes e consumidores, através de excelente planejamento (porque é

aqui a etapa inicial, onde surgem todas as ideias, e destacam-se as que serão viáveis e utilizadas), organização (não basta apenas pensar, precisamos organizar para poder colocar em prática) e por fim, controle (é um procedimento indispensável, pois a partir dele pode avaliar como as adaptações estão reagindo).

Dessa forma, pode-se concluir que esta foi uma pesquisa de extrema valia já que demonstrou de maneira direta como uma boa administração traz para os estabelecimentos uma vantagem competitiva em termos de serviços, redução de custos, preços acessíveis, boa qualidade dos produtos/serviços e um elevado aumento na rentabilidade do investimento, fazendo com que as empresas que utilizam esse tipo de processo se tornem ainda mais bem-sucedidas.

### **Referências Bibliográficas**

- Accioly, Felipe; AYRES, P. A. Salmeron; SUCUPIRA, Cezar. Gestão de estoques. Rio de Janeiro: Editora FGV, 2008.
- Ballou, Ronald H. Logística empresarial: transportes, administração de materiais e distribuição física. tradução Hugo Yoshizaki1 ed. – São Paulo: Atlas, 2012.
- Cervo, A. L; Bervian, P. A. Silva, R. Metodologia científica. – 6ª ed. São Paulo: Pearson. 2009.
- Ching, Hong Y. Gestão de estoques na cadeia de logística integrada – Supply Chain – 3ª. Ed. São Paulo: Atlas, 2009.
- Christopher, Martin. Logística e Gerenciamento da Cadeia de Suprimentos: criando redes que agregam valor. 2. Ed. São Paulo: Cengage Learning, 2008.
- Cscmp. Council of Supply Chain Management Professionals. Disponível em: <<http://cscmp.org/>>. Acesso em: 30 de Setembro de 2018.
- Erdmann, Rolf Hermann. Organização de sistemas de produção. Florianópolis: Insular, 1998.
- Leite, Paulo Roberto. Logística reversa. Pearson. São Paulo, 2003.
- Novaes, Antônio G. Logística e gerenciamento da cadeia de distribuição – 13ª. Ed. Rio de Janeiro: Elsevier, 2007
- Pedrosa, André Sousa. A logística reversa como uma ferramenta gerencial: um novo diferencial competitivo para as organizações. Qualitas Revista Eletrônica, 2008.
- Santos, Adriana Barbosa; Do Santos, Jarbas Queiroz. O processo logístico como estratégia empresarial: um estudo de caso em uma indústria alimentícia de animais. Revista Formadores, 2017.
- Zanella, Liane Carly Hermes. Metodologia da Pesquisa científica na prática. – 2ª Ed. Olinda: Livro Rápido, 2012.

## Um estudo do lead time no e-commerce internacional na perspectiva do público nacional

Sandriele R. da Silva<sup>1</sup>, Ingrid C. da Silva<sup>1</sup>, Jonas T. dos Santos<sup>1</sup>, Aldynny T. C. B. de Carvalho<sup>1</sup>, Gabriela de Vasconcelos<sup>1</sup>

<sup>1</sup>Universidade de Pernambuco (UPE)

Caruaru – PE – Brasil

{sandriele\_silva, ingrid.csilva}@outlook.com

aldynnyburgos\_@hotmail.com

{devasconcelos.gabriela, trajanojonas} @gmail.com

**Resumo.** Este artigo trata-se de um estudo do lead time atrelado ao e-commerce internacional. O estudo buscou identificar através de um questionário e uma pesquisa bibliográfica os fatores que geram os atrasos de mercadorias vindas do exterior, na percepção dos consumidores sobre a relação de compra oriunda do meio digital e qual o real fator que acarreta o atraso. Com a análise de resultados da pesquisa, foi enfatizado que o fator decisivo na demora de entrega possui relação direta com o processo de desembaraço aduaneiro, estando associado com a grande quantidade de produtos que chegam ao Brasil, juntamente com a fiscalização rigorosa dos importados realizada pela Receita Federal brasileira.

**Abstract.** This article is about a lead time study linked to international e-commerce. The study sought to identify through a questionnaire and a bibliographical research the factors that generate the delays of goods coming from abroad, in the perception of the consumers about the relation of purchase from the digital medium and what the real factor that causes the delay. With the analysis of the results of the research, it was emphasized that the decisive factor in the delay of delivery has a direct relation with the process of customs clearance, being associated with the large quantity of products that arrive in Brazil, together with the strict inspection of the imports carried out by of the Brazilian Federal Revenue Service.

### 1. Introdução

A logística é uma estratégia empresarial que consiste na integração e coordenação de uma série de atividades ao longo da cadeia de negócios, bem como as demais funções internas e os fatores externos que exercem influência na organização. Assim, a logística está posicionada para se tornar, junto com um sistema de informações bem estruturado, o novo modelo de gestão da empresa, permitindo coordenar e integrar todos os processos ao longo da cadeia produtiva, ampliando em muito as atividades e fronteiras de uma organização. Soma-se a isto, conforme Novaes (2001), as oportunidades geradas por sua utilização passam a criar valor para o cliente, e, dessa forma, proporciona um diferencial no que tange a administrar e gerenciar empresas.

As organizações estão cada vez mais preocupadas com a concorrência e estão procurando um meio de obter uma vantagem competitiva. Uma das soluções para esse problema seria o uso do gerenciamento logístico aliado à redução do lead time. De acordo com Ericksen e Stoflet (2007) lead time é a quantidade de tempo desde o recebimento do pedido até que ao menos um dos itens do pedido esteja disponível no estoque e seja entregue ao cliente.

Portanto, é necessário analisar onde os processos sofrem atraso para que seja desenvolvida uma estratégia para redução do lead time. Pois segundo Cunha e Estender (2017) o cumprimento de prazos é, sem dúvidas, um dos mais importantes fatores de sucesso de uma organização, independente do seu ramo de atuação. Esse é um fator que tem relação direta com a satisfação dos clientes e pode se tornar um diferencial competitivo quando é bem gerenciado.

De acordo com Nascimento *et al* (2012) a maior parte das organizações enfrenta problemas em controlar o tempo que leva para comprar a matéria-prima, fabricar e entregar o produto acabado ao cliente. Pois esse tempo muitas vezes é maior do que o cliente está disposto a esperar. Mapear os canais proporciona uma base fundamental para reorganizar a engenharia logística, conseguindo melhorar o desempenho produtivo e reduzindo o lead time.

Nesse contexto, surgem as empresas do comércio on-line que muitas vezes são afetadas pela falta de logística ou simplesmente pelo mau gerenciamento do lead time, entretanto diante dos problemas apresentados o comércio on-line está evoluindo e muitas empresas estão investindo nesta nova modalidade de varejo, proporcionando economia financeira e comodidade aos consumidores que utilizam a internet para pesquisar e adquirir bens e serviços [Mansano e Gorni 2014]. Herzer (2013) complementa que no e-commerce a logística busca se preocupar com a agilidade de entrega, cumprindo prazos, e se possível, efetuando a entrega com o tempo menor do que o prometido.

O presente trabalho se torna relevante, uma vez que irá contribuir no entendimento do público em geral informando por que os produtos demoram muito tempo para chegar até o cliente, ressaltando também a interferência dessa demora na vantagem competitiva das empresas.

Diante do exposto, este artigo tem por objetivo responder o seguinte questionamento: Quais os fatores que impactam na demora de entrega dos pedidos das empresas de e-commerce? Por esta razão, o objetivo geral do presente artigo é analisar o lead time sob a perspectiva das empresas de e-commerce internacional. Logo como objetivos específicos da pesquisa têm-se: analisar o lead time no e-commerce sob a perspectiva dos clientes e identificar os fatores geradores da demora de entrega dos pedidos. O trabalho consiste em estudo exploratório, para que seja possível com os dados coletados a partir de questionários aplicados a consumidores ativos do e-commerce e de uma pesquisa bibliográfica alcançar o objetivo deste artigo.

## **2. Referencial Teórico**

### **2.1 Logística: Conceitos e Definições**

Um sistema logístico é formado por uma rede composta de etapas que funcionam de forma sincronizada. Aperfeiçoar os processos logísticos tornou-se um requisito essencial para as organizações que desejam sobreviver no atual ambiente competitivo. (BALLOU, 2006).

Segundo a definição do CSCMP (2010), logística é a parte da Gestão da Cadeia de Suprimentos responsável por planejar, desenvolver, e controlar de forma eficiente e eficaz o fluxo e a armazenagem de produtos, bem como os serviços e informações relacionados. Este pensamento é reforçado pela definição de Christopher (2002) e Novaes (2001) que acrescentam que a logística tem como objetivo o aumento da lucratividade e a redução dos custos ao longo do processo, tendo em vista o atendimento das necessidades e preferências do consumidor final.

De acordo com Uelze (1974), qualquer organização que use ou produza bens, tem que se preocupar com os problemas de movimentá-los de um local para outro, envolvendo assim de alguma forma de transporte.

Segundo Ballou (2006, p. 25), “A Logística Empresarial é um campo relativamente novo de estudo da Gestão Integrada, das áreas tradicionais das finanças, marketing e produção”. Percebe-se que o autor interliga esses três setores da Administração como um meio ligado à resolução e otimização dos processos Logísticos.

## 2.2. Logística de Distribuição

A logística de distribuição, como evidencia Novaes (2007), é uma parte importante da logística como um todo, por seu cunho estratégico e pela relevância econômica dada pelos governantes e empresários quando principal objetivo dela é disponibilizar a quantidade de mercadorias certa, no momento certo e no lugar certo. Além disso, é preciso otimizar processos para que as operações sejam rentáveis e lucrativas.

De acordo com Wanke (2011, p. 256), a logística de distribuição, “[...] [é] o desenvolvimento de um sistema que abranja todas as atividades, desde a saída da linha de produção até a entrega”.

Em síntese, Ching (2001) abordou os processos que se relacionam com a logística de distribuição, e que acabam contribuindo para a sua existência em termos de fluxo de processos para gerar organização adequada em termos de resultados harmoniosos. O autor explicou que possibilidades de minimização dos custos são elevadas quando a logística de distribuição encontra o equilíbrio entre a qualidade de serviço, os custos e o capital investido.

Vale salientar que a Logística de transporte trabalha de mãos dadas com a Logística de distribuição pois todos os produtos necessitam ser transportados de um local a outro. (Nogueira 2012 p. 83).

A logística de transporte tem por funções básicas a organização das operações e a redução de custos. O transporte representa um estímulo da estratégia de negócio em termos de competitividade, principalmente quando a empresa está definindo as necessidades-alvo dos clientes, como o nível de serviço ou a redução de custos (Nogueira, 2012).

## 2.3. Lead Time

O *lead time* em tradução para o português significa “prazo de entrega”, e se caracteriza pelo tempo total entre a emissão do pedido, até a entrega efetiva ao consumidor final.

De acordo com Simchi-Levi e Kaminsky(2003), o *lead time* total é composto pelo tempo dedicado ao processamento de pedidos, fornecimento, manufatura dos itens, transporte, e todos os estágios da cadeia de suprimentos. Nas situações onde os pedidos precisam ser produzidos, montados, ou adquiridos de fornecedores externos, esse tempo

será ainda maior. Para cada negócio existem fatores que influenciam diretamente em seu lead time, como por exemplo, fornecimento e compra de matéria-prima, tempo de produção do produto e também o transporte até o cliente.

A empresa que consegue uma combinação perfeita entre o *lead time* logístico, e o ciclo do pedido do cliente não precisa de previsões, nem estoque, sendo assim possuindo uma economia na gestão de estoque. Gilberto Kosaka (2016), diretor-presidente do jornal São Paulo Shimbun, entende o *lead time* como um componente do planejamento e, por isso, diferencia as práticas do *lead time* em áreas: na de Compras o *lead time* é caracterizado pelo intervalo de tempo de compra dos materiais necessários para sua produção, desde a emissão do pedido, até estar pronto para uso. Já na Produção, é caracterizado pelo intervalo de tempo que o material passa pela fabricação, e por fim o *lead time* de entrega, que se caracteriza pelo intervalo de tempo desde o momento de efetivação da compra, até o recebimento desta.

Na perspectiva de fornecedor, lead time é o tempo do processamento da ordem de compra, compreendendo o período desde o seu recebimento até sua entrega. Há alguns autores que inclui nesses períodos o lead time financeiro, onde é acrescido nesse tempo o prazo para a liquidez do pagamento. Na perspectiva do cliente, lead time é o tempo de processamento total do pedido, ou seja, desde o momento em que é identificada uma necessidade e gerada uma ordem de compra até o momento em que esta é recebida e a necessidade suprida. (Ballou, 2006; Tubino 1999).

Atualmente a maior dificuldade das empresas é a redução desse tempo, que só é possível se a organização tiver uma produção extremamente padronizada e eficiente. Diante disso, o lead time é um dos elementos mais relevantes para assegurar a vantagem competitiva de uma empresa e expandir as suas possibilidades de alcance de mercado.

Conforme Peteraf (1993), vantagem competitiva é criar um valor para os compradores que exceda o custo do que está sendo vendido a eles; logo, vantagem competitiva é fornecer um produto de valor agregado maior do que seu custo. Portanto, para que a empresa não perca em vantagem competitiva, esta deve ter algum diferencial que faça com que os clientes vejam vantagem em comprar nela. Esse diferencial pode ser, por exemplo, o lead time, que é compreendido desde o recebimento do pedido até a entrega do mesmo.

De acordo com Barney (1991), uma empresa possui vantagem competitiva quando detém de capacidade de criação de valor maior que seus concorrentes. Então, pode-se obter essa vantagem por meio da redução do lead time, que pode ser reduzido de diversas formas, inclusive otimizando o tempo de produção dos funcionários e a entrega do produto final ao cliente.

### 3. Metodologia

O método de pesquisa utilizado para a realização deste trabalho se caracteriza como qualitativa, pois segundo Silva e Menezes (2000), esse tipo de pesquisa visa descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis. Envolve o uso de técnicas padronizadas de coleta de dados: questionário e observação sistemática e assume, em geral, a forma de levantamento. Nesse sentido o presente artigo se enquadra nessa modalidade, uma vez que visa descrever qual a frequência de compra dos consumidores, quais os principais produtos comprados, média de valores gastos, dentre outras características.

O presente trabalho refere-se a dois procedimentos metodológicos clássicos, o primeiro baseado na aplicação de uma Survey, cujo questionário foi elaborado através de uma pesquisa bibliográfica onde um dos principais tópicos abordados está diretamente ligado com o nosso objetivo de pesquisa, que são as principais causas na demora de entrega dos produtos na concepção dos consumidores nacionais; segundo Babbie (1999) a pesquisa tipo survey, termo em inglês que se destina a pesquisa em grande escala caracteriza-se por ser uma abordagem quantitativa, que visa apresentar as opiniões das pessoas por meio de questionários ou entrevistas.

Dessa forma, procedeu-se a aplicação de questionários com consumidores do e-commerce, afim de verificar quais são as empresas que mais demoram a entregar o pedido ao consumidor final e qual fator gerador da demora em sua concepção. No segundo, analisaremos qual é o fator real que impacta na demora da entrega dos produtos, sob a perspectiva macroeconômica e demográfica.

Primeiramente foi efetuada uma pesquisa de levantamento que de acordo com Fonseca (2002) é usada para confirmar ou negar uma hipótese, baseado no método de pesquisa descritiva, que consiste na exploração de um problema para prover critérios e a compreensão de determinado problema. A seguir, foi realizada coleta de dados secundários através de pesquisa bibliográfica a fim de solucionar o problema de pesquisa citado na parte introdutória. Segundo Chiara *et al* (2008) a pesquisa bibliográfica visa analisar as principais teorias de um tema a fim de analisar, produzir ou explicar um objeto sendo investigado. Dessa forma, nossa pesquisa nos ajudou a analisar e explicar quais os reais fatores que impactam na demora de entrega dos produtos internacionais.

#### 4. Análise e discussão dos Resultados

A análise do cenário nacional no que concerne ao lead time logístico do e-commerce internacional nos permite observar sobre a perspectiva dos consumidores que a principal insatisfação em relação a compras de produtos importados, está na demora em que levam para chegar até o consumidor final, foi então definido como público-alvo os consumidores desses produtos, aplicamos então uma survey com intuito de alcançar o objetivo deste trabalho citado anteriormente, onde foi possível avaliar algumas preferências dos consumidores sobre determinadas compras, como por exemplo, os sites mais confiáveis, os produtos mais comprados, entre outros pontos que serão apresentados posteriormente. Foi então definida uma amostra de 100 respondentes para o questionário, iniciamos a pesquisa com a pergunta filtro onde 50 dessas pessoas afirmaram ser consumidoras de e-commerce internacional:



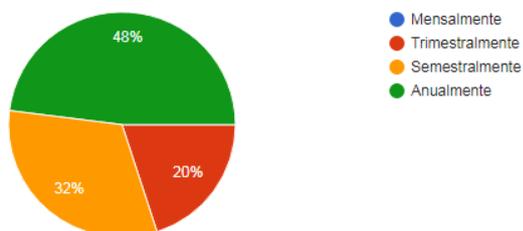
Figura 1: Pergunta filtro sobre ser consumidor de e-commerce internacional.

Fonte: Google Forms

Em relação ao fator frequência, 48% dos respondentes fazem esse tipo de compra anualmente, 32% semestralmente e 20% trimestralmente.

Com que frequência você faz essas compras?

50 respostas



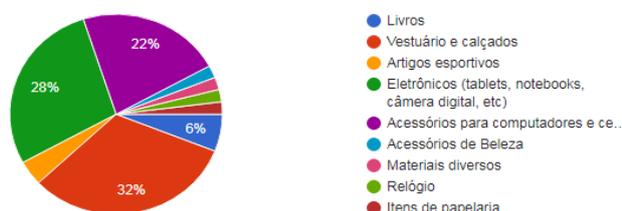
**Figura 2: Representação da frequência de compras em sites internacionais de e-commerce.**

Fonte: Google Forms

De acordo com a categoria de produtos que mais são compradas, as que ficaram mais em evidência foram: 32% vestuário e calçados, 28% Eletrônicos como tablets, notebooks, câmera digital e etc, 22% acessórios para computadores e celulares e 6% livros.

Qual a categoria de produto de você mais compra?

50 respostas



**Figura 3: Categorias de produtos mais comprados pela amostra de consumidores nacionais.**

Fonte: Google Forms

Em relação ao preço médio gasto por cada consumidor em cada compra, destacaram-se as porcentagens 70% que gastam até R\$ 300,00 e 18% que gastam de R\$ 301,00 a R\$ 500,00.

Qual a média de preço que você gasta com esse tipo de compra?

50 respostas



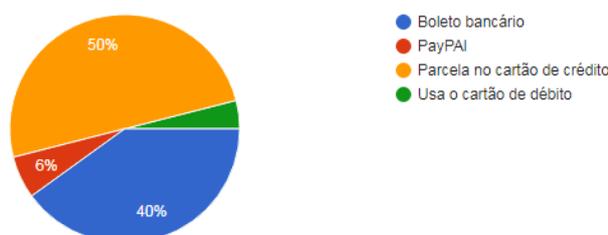
**Figura 4: Média de preço gasto nas compras pelos e-commerces estrangeiros.**

Fonte: Google Forms

As formas de pagamento que mais utilizadas foram parcelamento no cartão de crédito (50%), boleto bancário (40%) e PayPal (empresa de pagamentos online - 6%).

Qual a forma de pagamento?

50 respostas



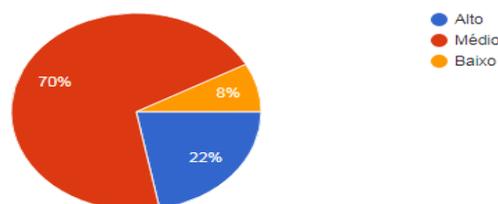
**Figura 5: Forma de pagamento.**

**Fonte: Google Forms**

O nível de confiança que os consumidores têm na procedência dos produtos comprados em sites internacionais se deu da seguinte forma: 70% responderam ter nível de confiança médio, 22% nível alto e 8% nível baixo. Esse resultado remete a citação Barney (1991), mencionado anteriormente, que afirma que o lead time é um fator que interfere na vantagem competitiva e no valor agregado ao produto, consequentemente na satisfação do consumidor final.

Qual o nível de confiança que você tem na procedência do produto?

50 respostas



**Figura 6: Nível de confiança dos compradores.**

**Fonte: Google Forms**

Em relação ao tempo que demora até os produtos chegarem aos respondentes, as respostas se deram da seguinte maneira: 68% recebem seus produtos entre 2 e 6 meses e 28% em menos de 2 meses. De acordo com Nogueira (2012) o transporte representa um estímulo da estratégia de negócio em termos de competitividade, dessa forma, o prazo de entrega tem que satisfazer a necessidade do consumidor.

Quanto tempo demora em média para seus produtos chegarem até você?

50 respostas



Figura 7: Tempo de entrega dos produtos.

Fonte: Google Forms

Entre os sites internacionais mais comprados, de acordo com o gráfico a baixo o AliExpress ficou em primeiro lugar como o site que mais demora a entregar os pedidos aos consumidores, seguido por Wish, em terceiro lugar ficaram empatados Ebay e GearBest.

Em sua opinião quais são os principais sites que mais demoram a entregar os produtos?

50 respostas

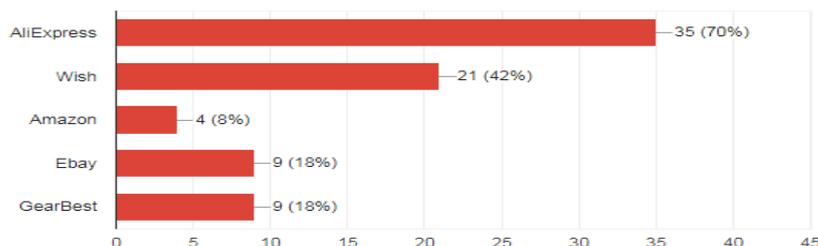


Figura 8: Relação gráfica de sites de e-commerce internacionais que mais demoram para chegar as mercadorias.

Fonte: Google Forms

66% dos respondentes afirmaram que conseguem identificar as causas da demora na entrega dos produtos e 34% afirmaram não saber.

Você consegue identificar as causas da demora na entrega dos produtos?

50 respostas

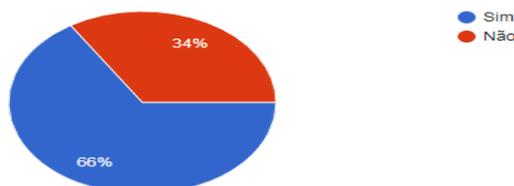


Figura 9: Pergunta filtro para examinar a percepção dos consumidores nacionais sobre as causas da demora.

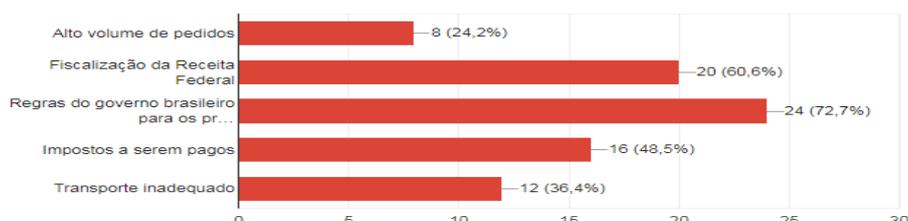
Fonte: Google Forms

Com a análise do questionário aplicado identificamos que segundo a perspectiva dos respondentes as causas que mais impactam na demora da entrega dos produtos,

destacam-se: regras do governo brasileiro para os produtos serem liberados com 66% dos votos, seguida por fiscalização da Receita Federal e impostos a serem pagos. Isto mostra que na concepção dos consumidores estes são os principais fatores que impactam na demora, uma vez que este fator não é decorrente de uma única causa, mas de uma junção de condições que atrasam a entrega dos pedidos dos consumidores.

Identifique então quais as que você conhece:

33 respostas



**Figura 10: Principais fatores que ocasionam e impactam diretamente na demora da entrega dos produtos ao público nacional.**

Fonte: Google Forms

Por conseguinte, para alcançar nosso objetivo, analisamos qual o real fator que impacta na demora de entrega de produtos comprados no e-commerce internacional, diante das pesquisas bibliográficas chegamos à seguinte conclusão, todas as pequenas encomendas postais urgentes e não urgentes (etiquetas que começam com código “L” e “R”, respectivamente) de até 2kg, além das pequenas encomendas sem rastreamento (remessas simples) chegam apenas em um lugar no Brasil: Curitiba.

Por fim, todos são avaliados pela Receita Federal, para saber se os tributos devidos foram cobrados. Assim que o produto é liberado pela Receita, sendo tributado ou não, ele vai para as mãos dos Correios, que irão fazer a distribuição das encomendas pelo Brasil. É aqui que está a maior parte da demora e, infelizmente, não se sabe o motivo exato disso, mas pode ser devido a vários fatores: Lentidão da Receita Federal, Logística dos Correios, Falta de funcionários ou desorganização, Burocracia, Ineficiência dos Correios, Extravio, alguns clientes esquecem detalhes simples, como colocar o nome completo e o CEP do endereço de entrega o que também pode provocar atraso, já que os funcionários precisam checar essas informações. O que realmente acontece, é quase impossível saber. Entretanto, alguns desses motivos podem descrever boa parte do atraso das encomendas internacionais que chegam ao país.

#### 4. Considerações Finais

O resultado de um e-commerce com logística eficiente é a satisfação dos clientes e o aumento das vendas. Por isso, é cada vez mais importante que as empresas estejam preparadas para atender as exigências do consumidor multicanal e entenda a real importância da melhora desses processos, consequentemente aumentando a produtividade das organizações.

As empresas de e-commerce internacional estão sendo cada vez mais procuradas pelos consumidores brasileiros por conta dos valores mais acessíveis, porém a demora da entrega afeta no grau de satisfação dos clientes e consequentemente no nível

competitividade das organizações, que fazem com que muitas pessoas deixem de comprar por conta da demora.

O que poderia mudar essa realidade é a aplicação de políticas de fiscalização e distribuição que otimizassem o processo de entrega dos produtos ao público no território nacional. Os sites internacionais não podem agir diretamente para mudar esse cenário, pois esse atraso é em decorrência ao processo alfandegário que os produtos passam quando chegam no Brasil para poderem chegar na mão do cliente final.

## Referências

- Babbie, Earl. Métodos de pesquisa de survey. 1.ed. Belo Horizonte, MG: Edições UFMG, 1999.
- Ballou, R.H. Gerenciamento de Cadeia de Suprimentos/Logística Empresarial. Porto Alegre: Bookman, 2006.
- Barney, J.B. Firm resources and sustained competitive advantage. *Journal of Management*, v. 1, 1991.
- Chiara, I. D. et al. Normas de documentação aplicadas à área de Saúde. Rio de Janeiro: Editora E-papers, 2008.
- Cunha, Maria Garcia; Estender, Antônio Carlos. O Lead Time Como Aliado da Vantagem Competitiva. *Revista Científica do ITPAC, Araguaína*, v.10, 12 f, fevereiro 2017.
- Ching, H.Y. Gestão de estoques na cadeia de logística integrada: Supply chain. 2ª ed. São Paulo: Atlas, 2001.
- Ericksen, P. D.; Stoflet, N. J.; Manufacturing Critical-path Time (MCT): the QRM metric for lead time. Wisconsin-Madison: Technical Report, Center for QRM, 2007.
- Embarcando na Viagem. Correios Curitiba | Internacional Demora Como Funciona. 2018. Disponível em: <https://www.youtube.com/watch?v=hpRrmG-AIVY>. Acesso em: 05 out. 2018.
- Fonseca, J. J. S. D. Metodologia da Pesquisa Científica. Apostila (Apostila de Disciplina) – Universidade Estadual do Ceará. Fortaleza, p. 127. 2002.
- Herzer, A. Fidelizando clientes no comércio eletrônico. 2013. Disponível em: <http://ecommercenews.com.br/artigos/cases/fidelizando-clientes-no-comercio-eletronico>. Acesso em: 18 set. 2018.
- Kaminsky, G. L.; Reinhart, C. M. The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *International Finance Discussion Paper No. 544*, Board of Governors of the Federal Reserve System, March 2003.
- Mansano, Adriana T. R.; Gorni, Patrícia M. Satisfação do consumidor com o comércio eletrônico: estudo de caso de uma fabricante de tapetes. *Revista de Extensão e Iniciação Científica SOCIESC - REIS*, Santa Catarina, v. 1, n. 1, p.12-22, jun. 2014. Disponível em: <http://www.sociesc.org.br/reis/index.php/reis/article/view/8>. Acesso em: 18 set. 2018.

- Muambator. Por que os seus pacotes internacionais ficam presos em Curitiba? Disponível em: <https://medium.com/muambator/porque-os-pacotes-internacionais-ficam-presos-em-curitiba-1c463ada3120>. Acesso em: 03 out. 2018.
- Nascimento, C. F. S., Brochieri, M., Jesus, N. A. B., Cabrini, T. R. Suprimentos: Gerenciamento Estratégico do Lead-Time. Teclog. 2012.
- Novaes, A. G., (1935). Logística e gerenciamento na cadeia de distribuição: Antônio Galvão Novaes. Rio de Janeiro: Elsevier, 2007.
- Palpite Digital. Quanto tempo demora uma encomenda internacional? Disponível em: <https://www.palpitedigital.com.br/wp/2017/05/21/quanto-tempo-demora-encomenda-internacional/>. Acesso em: 05 out. 2018.
- Peteraf, M. A. The cornerstones of competitive advantage: a resource based view. *Strategic Management Journal*, 14(3), 179-191. Doi: 10.1002/smj.4250140303, 1993.
- Silva, E. L., Menezes, E. M. (2000) Metodologia da pesquisa e elaboração de dissertação. Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2000, 118p.
- Tubino, D. F. Manual de planejamento e controle da produção. São Paulo: Atlas, 2000  
\_\_\_\_\_. Sistemas de Produção: a produtividade no chão de fábrica. Porto Alegre: Bookman, 1999.

## Uma Abordagem para Detecção de Diagnóstico da Doença de Alzheimer Utilizando Extratores de Textura de Haralick

Iago R. R. Silva<sup>1</sup>, Cristiano S. de Oliveira<sup>2</sup>, Lucas H. Cavalcanti<sup>2</sup>,  
Renato S. Bezerra<sup>2</sup>, Roberta A. de A. Fagundes<sup>1</sup>, Wellington P. dos Santos<sup>1</sup>

<sup>1</sup>Departamento de Engenharia da Computação – Escola Politécnica de Pernambuco –  
Universidade de Pernambuco (UPE)  
Rua Benfica, 455 – Madalena – Recife – PE – Brasil

<sup>2</sup>Centro de Informática – Universidade Federal de Pernambuco (UFPE)  
Cidade Universitária – Recife – PE – Brasil

{irrs,wps}@ecomp.poli.br

{roberta.fagundes}@upe.br

{cso, lhcs, rsb5}@cin.ufpe.br

**Abstract. Objective:** To propose an approach for the classification and diagnosis of Alzheimer's disease from magnetic resonance imaging. **Methodology:** Relevant texture features are extracted in magnetic resonance images for use in classifiers. From this, class balancing is performed. The method of dividing the training/test data was cross validation with 10 folds. The Random Forest, SVM and K-NN algorithms were used to classify the data. **Results:** The best result obtained after an average of 30 executions was obtained through the Random Forest algorithm, close to 80 % accuracy, hypothesis tests and ROC curve analysis also emphasize this obtained result. **Conclusion:** It can be concluded that only the use of texture descriptors is insufficient to classify patients with Alzheimer's, it is necessary to use shape extractors to improve the results.

**Resumo. Objetivo:** Propor um uma abordagem para classificação e diagnóstico da doença de Alzheimer a partir de imagens de ressonância magnética. **Metodologia:** São extraídas características de textura relevantes nas imagens de ressonância magnética para utilização nos classificadores. A partir disso, é realizado o balanceamento das classes. O método de repartição dos dados treino/teste foi cross validation com 10 folds. Para classificação dos dados foram utilizados os algoritmos Random Forest, SVM e K-NN. **Resultados e Discussão:** O melhor resultado obtido após uma média de 30 execuções foi obtido através do algoritmo Random Forest, próximo a 80% de acurácia, testes de hipótese e análise da curva ROC também enfatizam este resultado obtido. **Conclusão:** Pode-se concluir que só a utilização de descritores de textura é insuficiente para classificação de pacientes com Alzheimer, se faz necessário a utilização de extratores de forma para aprimoramento dos resultados.

### 1. Introdução

A doença de Alzheimer é uma das doenças neurodegenerativas mais comuns que afetam os idosos devido às alterações que ocorrem no cérebro. A doença afeta diretamente a capacidade cognitiva e memória do indivíduo, causando a perda da mesma

[Kandel et al. 2014]. Atualmente, não existe cura para Alzheimer, porém quando a mesma é diagnosticada precocemente, existem tratamentos que podem amenizar a perda cognitiva e controle na memória [Kandel et al. 2014].

Estudos comprovam que a presença da doença de Alzheimer em um indivíduo está diretamente associada a diversas características, como a atrofia da região do hipocampo, presença das proteínas Amiloides e Tau, e emaranhados neurofibrilares [Kandel et al. 2014]. Todas as características anteriormente citadas podem ser detectadas através de análise de exames de imagem, como ressonância magnética e tomografia computadorizada por emissão de pósitrons (PET-SCAN).

A análise do exame de imagem por um especialista em neurologia é utilizado na medicina para diagnóstico da doença. Para auxílio ao especialista, foram desenvolvidos métodos utilizando técnicas de Inteligência Computacional, a fim de classificar se o paciente é saudável ou possui a doença. Atualmente diversas técnicas de Inteligência Computacional têm sido utilizadas com a finalidade de classificação e diagnóstico (precoce ou não) da doença de Alzheimer, como Redes Neurais Multilayer Perceptron [Munteanu et al. 2015], SVM [Xiao et al. 2017], Random Forest [Oppedal et al. 2015], dentre outros. Algumas técnicas de aprendizagem supervisionada, que são técnicas aplicadas a base de dados rotuladas, ou seja, com a natureza dos indivíduos presentes devidamente conhecida, oferecem rápida criação de modelo e classificação da doença com uma acurácia aceitável.

Para utilização da maioria destas técnicas é necessário definir um modelo de extração das características mais relevantes presentes nas imagens que diferenciam os indivíduos de diferentes naturezas, no caso desta aplicação, essas natureza são definidas através do estado de saúde do paciente, em relação à doença de Alzheimer. Extratores de textura podem oferecer uma classificação adequada dos exames de imagem [Oliveira et al. 2010]. Além disso, se a base de dados oferecer uma quantidade desbalanceada de indivíduos de diferentes naturezas, pode ocasionar um fenômeno nomeado *overfitting*, que produz incorretas classificações das classes minoritárias. Para resolução deste problema, se faz necessário a utilização de técnicas para o balanceamento da base de dados em relação as classes.

Diante desse contexto, este trabalho objetiva a investigação, aplicação e análise de algoritmos de classificação supervisionada aliado à duas outras técnicas: (I) Texturas de Haralick: para extração de características relevantes nas imagens e (II) *Synthetic Minority Over-sampling Technique* (SMOTE): para balanceamento das classes utilizadas, para o diagnóstico precoce da doença de Alzheimer. Os algoritmos de classificação selecionados para este estudo são alguns que possuem um rápido tempo de treinamento.

## **2. Referencial Teórico**

Esta seção aborda o referencial teórico deste trabalho, conceituando os conteúdos necessários para a compreensão do mesmo.

### **2.1. Revisão da Literatura**

Abordagens para detecção e diagnóstico da doença de Alzheimer utilizando uma proposta similar a deste trabalho têm sido frequentemente criadas na literatura [Bansal et al. 2018].

Algoritmos inteligentes que aprendem a detectar padrões em dados de imagens de pacientes são utilizados para tal finalidade, como Random Forest [Sarica et al. 2017], K-NN [Nozadi et al. 2018] e SVM [Marwa and Nadia 2018]. Dependendo da abordagem utilizada na extração de características, podem alcançar uma porcentagem de acertos na classificação dos dados. A acurácia dos trabalhos atualmente presentes na literatura é considerada boa acima da faixa de 80%, pois é o que se encontra em parte dos trabalhos presentes. Diversas abordagens para extração de características são utilizadas na literatura para diagnóstico da doença de Alzheimer, dentre elas destacam-se os descritores de forma [Gorji and Haddadnia 2015], que necessitam da realização de um pré-processamento nas imagens, a segmentação. É alvo deste trabalho propor uma abordagem de extração de textura, pois é possível obter características quantificáveis dos objetos presentes nas imagens sem a necessidade da realização da segmentação de imagens. Os algoritmos anteriormente citados serão utilizados para avaliar este modelo.

## **2.2. Extração de Texturas de Haralick**

As texturas caracterizam esse efeito de variação de tonalidades presentes em uma imagem e podem ser descritas como medidas que são capazes de quantificar características contidas em uma imagem. Haralick et al. (1973) propôs uma abordagem estatística para criação de descritores para extração de textura, baseado na matriz de co-ocorrência [Haralick et al. 1973]. A matriz de co-ocorrência contém a probabilidade de cada intensidade das 256 tonalidades possíveis - neste caso no sistema de cores RGB - ocorrerem na imagem. Valores obtidos com a extração de textura como energia, entropia, correlação, inércia e homogeneidade são obtidos através de cálculos na matriz de co-ocorrência, sendo estes utilizados como descritores de características.

## **2.3. Random Forest**

É um algoritmo de aprendizagem supervisionada para problemas de classificação de dados. Este algoritmo é amplamente utilizado na literatura devido a sua simplicidade para ajuste de seus parâmetros. Como é explícito etimologicamente, o algoritmo cria uma floresta aleatória [Oshiro et al. 2012], que corresponde a uma combinação de árvores de decisão para obtenção de uma maximização da acurácia do modelo de classificação.

Dado uma quantidade de dados de um determinado conjunto, o algoritmo Random Forest separa os dados em subconjuntos aleatórios de acordo com as características presentes nesses subconjuntos de dados, permitindo uma alta capacidade de generalização do algoritmo nos modelos criados a partir da geração dos dados [Oshiro et al. 2012]. O principal parâmetro a ser utilizado em um processo de aprendizado de máquina utilizando este algoritmo é a quantidade de árvores a serem utilizadas no experimento.

## **2.4. Support Vector Machines (SVM)**

SVM, ou máquinas de vetor de suporte é um algoritmo de aprendizado de máquina que consiste na criação de vetores de suporte para classificação dos dados [Xiao et al. 2017]. Os dados de entrada são mapeados em um espaço multidimensional e é gerado um hiperplano, superfície capaz de separar os dados de acordo com classes de entrada. O treinamento consiste na busca do melhor hiperplano que atenda melhor a demanda do modelo preditivo. Um dos objetivos do SVM é encontrar a menor distância do primeiro ponto de cada classe e o hiperplano, gerando um parâmetro nomeado folga.

## 2.5. K-Nearest Neighbors (K-NN)

O K-NN ou método dos K vizinhos mais próximos é um algoritmo de classificação fundamentado nas características das classes. No K-NN, o objeto é atribuído à classe majoritária de acordo com seus K vizinhos mais próximos, e as distâncias entre o objeto e seus vizinhos mais distantes (além dos K vizinhos mais próximos) são desconsiderados [Liu et al. 2013]. Como um método de classificação simples, eficaz e não paramétrico, o K-NN é amplamente utilizado em classificações [Zhou et al. 2010].

Neste algoritmo, primeiro são selecionadas K amostras de treinamento mais próximas para uma amostra de teste e, em seguida, prevê a amostra de teste com a classe principal entre as amostras de treinamento mais próximas. No entanto, o K-NN precisa calcular a distância (ou similaridade) de todas as amostras de treino para cada amostra de teste, sendo esta uma desvantagem do K-NN [Deng et al. 2016].

## 3. Metodologia

Esta seção apresenta a metodologia deste projeto de pesquisa, evidenciando a forma como a pesquisa será conduzida, conforme a Figura 1. As etapas serão descritas abaixo:

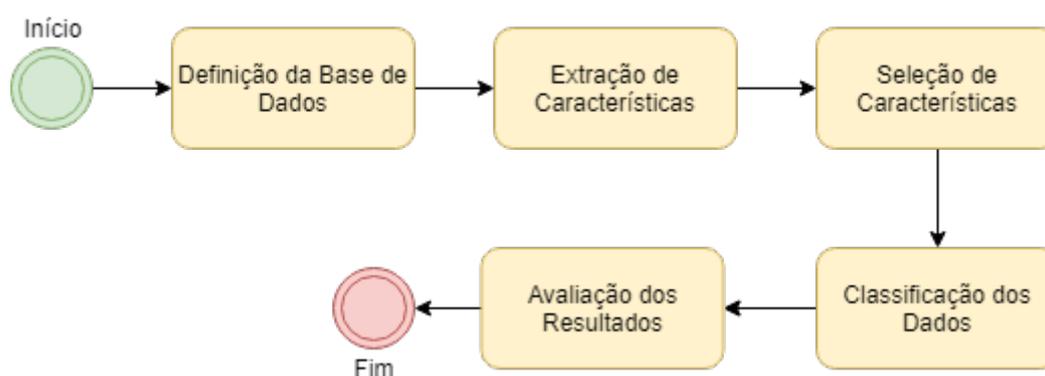


Figura 1. Metodologia com as etapas utilizadas no presente trabalho.

### 3.1. Definição da Base de Dados

A base de dados que será utilizada no projeto é a ADNI (Alzheimer's Disease Neuroimaging Initiative). ADNI contém uma coleção de exames de imagem, como ressonâncias magnéticas e Tomografias por Emissão de Pósitrons, de pacientes: (I) saudáveis (*Normal Controls* - NC), (II) com perda cognitiva leve (*Mild Cognitive Impairment* - MCI), e (III) com Alzheimer (Alzheimer Disease - AD). A coleção de dados a ser utilizada neste trabalho é a ADNI-1 que contém 826 exames de pacientes ao longo de 36 meses, sendo destes 199 exames pertencentes à classe AD, 229 à classe NC e 393 à classe MCI. Todos os dados utilizados desta coleção são correspondentes a exames de ressonância magnética. A base de dados tem sido utilizada na literatura [Belathur Suresh et al. 2018] [Shi et al. 2018]. A escolha desta base de dados é importante para implementação da proposta deste trabalho por ser uma base dedicada exclusivamente ao diagnóstico de Alzheimer, além de proporcionar uma justa comparação com outros trabalhos recentes presentes na literatura.

### 3.2. Extração de Características

As os arquivos gerados pela ressonância magnética são do tipo “.nii”. Cada arquivo geralmente possui 256 fatias, correspondentes à toda região da cabeça. Para diminuição de ruídos e confusão no aprendizado dos algoritmos, para cada exame foram selecionadas 20 fatias correspondentes à região do córtex cerebral superior ao olho. Após a seleção destas fatias, é aplicado o método de extração de texturas, sendo possível obter propriedades quantificáveis de determinadas formas presentes em uma imagem, distinguindo de outras formas. A atrofia do hipocampo e características que definem pacientes com Alzheimer podem ser quantificadas com dados de textura presentes nas imagens. O método de Haralick para extração de textura será utilizado neste trabalho, pois foi comprovado que o método de Haralick apresenta bons resultados para análise de imagens de ressonância magnética [Oliveira et al. 2010].

### 3.3. Balanceamento das Classes

Como explanado na seção 3.1, a base de dados contém uma larga diferenciação da quantidade de dados presentes em cada classe. Isto pode ocasionar dificuldades do aprendizado, gerando um problema nomeado *Overfitting*. Para resolução deste problema, o balanceamento das classes da base de dados foi aplicado, que sugere realizar um aumento no número de instâncias das classes minoritárias (NC e AD), aproximando-se da majoritária (MCI). O algoritmo *Synthetic Minority Over-sampling Technique* (SMOTE) [Chawla et al. 2002] foi utilizado para tal finalidade. Este método consiste na criação de instâncias artificiais através de uma interpolação realizada nos dados que correspondem às classes minoritárias. A técnica foi aplicada nas classes minoritárias (AD e NC), a quantidade de instâncias geradas foi igual a ao complemento necessário para igualar a quantidade de dados da instância majoritária (MCI).

### 3.4. Classificação dos Dados

Para realizar a detecção e diagnóstico na base de dados em estudo, é necessário aplicar algoritmos de classificação. Os dados serão classificados de acordo com a similaridade das características das instâncias, seguindo regras determinadas pelas heurísticas dos algoritmos.

Os algoritmos de classificação utilizados neste projeto são os que foram citados na seção 2: Random Forest, KNN e SVM com Kernel RBF. Os parâmetros utilizados no Random Forest foi a quantidade de árvores igual a 100, enquanto com o K-NN o valor do parâmetro K foi igual a 1, permanecendo sempre a distância euclidiana para os devidos cálculos de distância e por fim o Kernel RBF foi utilizado no SVM.

O método de repartição dos dados utilizado foi *cross-validation* de 10 *folds*, onde cada algoritmo é executado 10 vezes com repartições diferentes entre conjunto de dados de treinamento e conjunto de dados de teste, os resultados são coletados e uma média geral para as 10 execuções é dada como resultado final da execução, aumentando a confiabilidade dos resultados de teste.

### 3.5. Avaliação dos Resultados

Com o objetivo de se realizar uma análise estatística mais robusta, cada cenário de execução do definidos na subseção anterior é executado 30 vezes, os dados de média e

desvio padrão são coletados e juntamente com os resultados, aplicados a testes de hipótese a fim de ratificar qual o classificador que foi capaz de obter melhores resultados.

Diferentemente da regressão linear, que analisa o erro, a análise de resultados obtidos com classificadores é realizada tomando como princípio os acertos. As métricas Acurácia (ACC) e Área sob a curva ROC (AUC) foram utilizadas para avaliação dos resultados. A métrica ACC consiste na razão entre a quantidade de exemplos do conjunto de teste que foram classificados corretamente e a quantidade total de exemplos contidos no conjunto de teste. Esta métrica expõe tão somente a porcentagem total de acertos do algoritmo em relação à quantidade de exemplos presente no conjunto de teste.

Por fim, a métrica AUC é a área correspondente abaixo da plotagem da curva ROC. A curva ROC é caracterizada pela plotagem da quantidade de exemplos positivos classificados corretamente vs a quantidade de exemplos negativos classificados corretamente. Esta métrica representa a probabilidade de que a identificação positiva feita pelo algoritmo para determinada classe escolhida de forma aleatória irá exceder o resultado da mesma classe escolhida aleatoriamente como negativa.

#### 4. Resultados

Esta seção apresenta os resultados obtidos com os experimentos realizados seguindo a metodologia deste trabalho, bem como os testes estatísticos que comprovam as afirmações que serão ditas a seguir.

##### 4.1. Resultados dos Experimentos

Após a aplicação da extração de características, os dados foram gerados em um arquivo CSV e utilizado nos algoritmos de aprendizagem 30 vezes em cada algoritmo, seguindo as mesmas condições expressas na seção 3. A Tabela 1 apresenta os resultados de média e desvio padrão obtidos com os experimentos realizados com as métricas ACC e AUC.

**Tabela 1. Resultados obtidos com as 30 execuções dos algoritmos de classificação na base de dados gerada pela etapa de extração de características.**

	Acurácia		Área sob a curva ROC	
	Média	Desvio Padrão	Média	Desvio Padrão
<b>Random Forest</b>	0,7865	0,0043	0,9204	0,0020
<b>SVM</b>	0,5690	0,0077	0,6756	0,0664
<b>K-NN</b>	0,7319	0,0069	0,8014	0,0075

Ao analisar os resultados explícitos na Tabela 1, pode-se notar que a média das execuções proporcionou resultados mais significativos com a utilização dos algoritmos Random Forest e K-NN. O desvio padrão obtido nas execuções de diferentes algoritmos foi muito próximo, subentendendo-se que os resultados obtidos não apresentaram uma grande dispersão, mantendo-se muito próximos entre si. Os resultados obtidos com estes dois algoritmos ratificam que o modelo proposto com o extrator de características de textura podem apresentar resultados competitivos.

Devido a diferença de valores de média em relação aos demais algoritmos, em ambas as médias, o algoritmo SVM com o kernel RBF foi considerado como inapto para o problema de pesquisa em questão. A escolha do kernel para classificação possui uma

grande influência na tentativa de separação das classes utilizando o hiperplano, e o fato dos resultados utilizando este kernel não terem sido satisfatórios pode se dar ao fato do conjunto de dados gerado através do extrator de características poder ser linear (o kernel RBF gera um hiperplano para dados não lineares). O kernel RBF foi escolhido por ser utilizado em problemas de classificação de Alzheimer [Nozadi et al. 2018], porém é provável que os dados gerados por seus extratores de características sejam não lineares.

Para se ter a certeza que o Random Forest obteve melhor acurácia em relação ao K-NN, é necessário a realização de teste de hipótese, que será apresentado na próxima subseção.

#### 4.2. Testes Estatísticos Realizados

Para escolha do teste estatístico apropriado para determinado problema, é necessário verificar qual distribuição de probabilidade os dados a serem testados seguem.

Neste caso, é necessário verificar a distribuição de todos os dados de resultados obtidos com os experimentos realizados. Para tanto, foi utilizado o teste de Kolmogorov-Smirnov [Lopes 2011], que verifica se dado conjunto de dados segue uma distribuição normal. Com um nível de significância de 5% o teste foi realizado em todas as amostras de dados, e concluiu-se que estes não seguem uma distribuição normal.

Dada esta informação, o teste estatístico de Wilcoxon [Gehan 1965] pode ser utilizado para aplicar teste de hipótese, com a finalidade de verificar qual classificador de fato obteve melhor resultado em relação as métricas adotadas para estudo. A hipótese a ser sempre testada será se um algoritmo obteve melhor desempenho que o outro, logo pode-se formular a hipótese nula como considerando os duas amostras apresentarem valores consideravelmente iguais e a hipótese alternativa (a que é desejável provar) é designada através da amostra apresentar valores consideravelmente maiores que o da outra amostra. A formulação desta hipótese é descrita na Figura 2, que segue.

$$\begin{cases} H_0 : u_1 = u_2 \\ H_1 : u_1 > u_2 \end{cases}$$

**Figura 2. Formulação das hipóteses.**

Diante do fato do algoritmo SVM não ter alcançado resultados competitivos em relação aos demais, apenas os resultados dos algoritmos Random Forest e K-NN foram utilizados no teste de hipótese. Neste caso, devido o fato do Random Forest ter obtido em média valor maior que o K-NN, assume-se que deseja-se provar que de fato o primeiro algoritmo obteve maior rendimento de acurácia e área sob a curva ROC que o segundo algoritmo.

Desta forma, de acordo com a hipótese formulada na Figura 2, assume-se que  $u_1$  seja a média obtida com os experimentos do Random Forest e  $u_2$  seja a média obtida com os experimentos do K-NN. É necessário a realização de dois testes de hipótese, um para a acurácia e outro para a área sob a curva ROC. Os resultados do teste de hipótese são apresentados na Tabela 2, onde a segunda e terceira colunas apresentam os valores de  $W$  e  $p$  - *value* retornados a partir do teste.

Métrica	W	p-value
Acurácia	0	8.831e-05
Área sob a curva ROC	0	8.438e-05

**Tabela 2. Resultados obtidos com os testes de hipótese realizados em cada conjunto de dados (métricas).**

Com os resultados dos testes de hipótese, pode-se afirmar, em ambas as métricas, que a hipótese nula pode ser refutada, e que há evidências suficientes no conjunto de dados apresentados de que o Random Forest apresenta o melhor resultado de classificação para detecção e diagnóstico da doença de Alzheimer, utilizando o descritor de texturas de Haralick.

## 5. Conclusão e Trabalhos Futuros

Foi apresentado um modelo para detecção de diagnóstico da doença de Alzheimer baseado na extração de características através dos descritores de Haralick, balanceamento das classes utilizando SMOTE e validação do modelo utilizando diferentes métodos de classificação, como Random Forest, K-NN e SVM com kernel RBF.

De acordo com os resultados obtidos o modelo proposto apresentou uma boa capacidade de generalização, ou seja, dois dos três algoritmos de classificação testados obtiveram uma acurácia acima de 70%, resultado considerado bom para problemas de reconhecimento de padrões. Destaca-se positivamente o algoritmo Random Forest, que através deste foi possível alcançar uma acurácia média de 78,65% e área sob a curva ROC de 0,9204. Os resultados iniciais obtidos neste trabalho são considerados bons, pois foi utilizado apenas um extrator de características e algoritmos de classificação em sua forma canônica, sem nenhuma adaptação específica para este problema de pesquisa. Destaca-se negativamente os resultados obtidos com o classificador SVM, que obteve uma baixa acurácia, é provável que isto deve-se ao fato dos dados gerados pelo extrator de característica utilizado sejam linearmente separados e o kernel RBF gera hiperplanos para dados não lineares. Esta afirmação toma forma devido a outros trabalhos da literatura que utilizam este kernel alcançarem uma acurácia alta.

Tratando-se de um problema da área de Medicina e é necessário o mínimo de erro no diagnóstico, é notório que ainda é preciso ainda maximizar os resultados de acurácia. A utilização de apenas um extrator de características pode ser a explicação para o valor máximo que foi obtido através do Random Forest. Diante deste fato, pretende-se investigar o impacto da utilização de outro descritor de forma conjunta ao descritor de textura já utilizado no presente trabalho. Também pretende-se investigar o impacto da quantidade de características extraídas para a classificação, utilizando meta-heurísticas como Algoritmos Genéticos ou Partículas de Exames.

## Agradecimentos

Os autores agradecem a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento da pesquisa.

## Referências

- Bansal, D., Chhikara, R., Khanna, K., and Gupta, P. (2018). Comparative analysis of various machine learning algorithms for detecting dementia. *Procedia Computer Science*, 132:1497–1502.
- Belathur Suresh, M., Fischl, B., Salat, D. H., and (ADNI), A. D. N. I. (2018). Factors influencing accuracy of cortical thickness in the diagnosis of alzheimer's disease. *Human brain mapping*, 39(4):1500–1515.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.
- Deng, Z., Zhu, X., Cheng, D., Zong, M., and Zhang, S. (2016). Efficient knn classification algorithm for big data. *Neurocomputing*, 195:143–148.
- Gehan, E. A. (1965). A generalized wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, 52(1-2):203–224.
- Gorji, H. and Haddadnia, J. (2015). A novel method for early diagnosis of alzheimer's disease based on pseudo zernike moment from structural mri. *Neuroscience*, 305:361–371.
- Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.
- Kandel, E., Schwartz, J., Jessell, T., Siegelbaum, S., and Hudspeth, A. (2014). *Princípios de Neurociências-5*. AMGH Editora.
- Liu, Z.-G., Pan, Q., and Dezert, J. (2013). A new belief-based k-nearest neighbor classification method. *Pattern Recognition*, 46(3):834–844.
- Lopes, R. H. (2011). Kolmogorov-smirnov test. In *International encyclopedia of statistical science*, pages 718–720. Springer.
- Marwa, Z. and Nadia, S. (2018). Comparative study of two classification methods for the detection of alzheimer's disease. *Current Medical Imaging Reviews*, 14(1):88–94.
- Munteanu, C. R., Fernandez-Lozano, C., Abad, V. M., Fernández, S. P., Álvarez-Linera, J., Hernández-Tamames, J. A., and Pazos, A. (2015). Classification of mild cognitive impairment and alzheimer's disease with machine-learning techniques using 1h magnetic resonance spectroscopy data. *Expert Systems with Applications*, 42(15-16):6205–6214.
- Nozadi, S. H., Kadoury, S., Initiative, A. D. N., et al. (2018). Classification of alzheimer's and mci patients from semantically parcelled pet images: A comparison between av45 and fdg-pet. *International journal of biomedical imaging*, 2018.
- Oliveira, M. S. d. et al. (2010). Análise de textura em imagens cerebrais: aplicações em acidente vascular cerebral isquêmico, epilepsia mioclônica juvenil, doença de machado-joseph, déficit cognitivo leve e doença de alzheimer.
- Oppedal, K., Eftestøl, T., Engan, K., Beyer, M. K., and Aarsland, D. (2015). Classifying dementia using local binary patterns from different regions in magnetic resonance images. *Journal of Biomedical Imaging*, 2015:5.

- Oshiro, T. M., Perez, P. S., and Baranauskas, J. A. (2012). How many trees in a random forest? In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, pages 154–168. Springer.
- Sarica, A., Cerasa, A., and Quattrone, A. (2017). Random forest algorithm for the classification of neuroimaging data in alzheimer’s disease: A systematic review. *Frontiers in Aging Neuroscience*, 9:329.
- Shi, J., Zheng, X., Li, Y., Zhang, Q., and Ying, S. (2018). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer’s disease. *IEEE journal of biomedical and health informatics*, 22(1):173–183.
- Xiao, Z., Ding, Y., Lan, T., Zhang, C., Luo, C., and Qin, Z. (2017). Brain mr image classification for alzheimer’s disease diagnosis based on multifeature fusion. *Computational and mathematical methods in medicine*, 2017.
- Zhou, L., Wang, L., Ge, X., and Shi, Q. (2010). A clustering-based knn improved algorithm clknn for text classification. In *Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference on*, volume 3, pages 212–215. IEEE.

## Fatores que influenciam o desempenho acadêmico no ENADE

Roberta A. A. Fagundes<sup>1</sup>, Geovanne O. Alves<sup>1</sup>

<sup>1</sup>Instituto de Informática – Universidade de Pernambuco (UPE)  
– Caruaru – PE – Brasil

{raaf, geovanne.oliveira}@upe.br

***Resumo.** O Exame Nacional de Desempenho de Estudante (ENADE) é uma das formas de avaliar os concluintes dos cursos de graduação, o que permite conhecer a qualidade dos cursos. Os resultados do exame aliados com o questionário do estudante concede dados importantes que, através de técnicas de mineração de dados, podem gerar informações valiosas para o desenvolvimento de políticas públicas. Este trabalho busca identificar os aspectos que mais influenciam os participantes do ENADE, utilizando de métodos estatísticos no software R. De acordo com a metodologia CRISP-DM, a presente pesquisa encontra-se na etapa de preparação dos dados.*

### 1. Introdução

De acordo com Griboski [Griboski 2012], em relação às considerações sobre a aplicação do ENADE, o principal aspecto a ser considerado envolve muito mais do que o resultado a ser alcançado, pois para a autora, é a participação do estudante no processo avaliativo, o aspecto que deve se constituir como elemento central para saber a qualidade do curso que está sendo ofertado. Tem-se, portanto, além da avaliação do aluno com o ENADE, o subsídio para a produção de indicadores que visam aferir a qualidade e os processos de avaliação dos cursos das Instituições de Ensino Superior(ies).

Segundo Brito [Brito et al. 2008], o ENADE, que é um exame em larga escala, não avalia a ênfase do curso, mas sim os itens que constam das diretrizes curriculares nacionais, comuns a todos os cursos e a partir dos quais os projetos são construídos. Portanto, com uma análise em nível micro e com o auxílio de técnicas de mineração de dados, é possível utilizar as informações disponibilizadas pelo Inep para retratar uma realidade social mais específica do desempenho dos estudantes, fornecendo para os professores e instituições um *feedback* sobre os métodos utilizados.

Este artigo tem como objetivo identificar os fatores que influenciam o desempenho acadêmico dos alunos no ENADE, com ênfase nos cursos de ciências da computação e sistemas de informação das ies públicas e privadas do estado de Pernambuco.

### 2. Referencial teórico e trabalhos relacionados

#### 2.1. Regressão Linear

Regressão Linear clássica (RL) é um dos mais antigos modelos estatísticos preditivos na área de aprendizagem de máquina. Os métodos supervisionados baseados neste tipo de regressão tentam modelar a relação entre entradas (variáveis explicativas) e saídas (variável resposta) na forma de equações paramétricas, onde os parâmetros são estimados a partir do método dos mínimos quadrados [Helene 2006]. Este método minimiza a soma

de quadrados das distâncias entre os pontos do diagrama e os respectivos pontos na curva da equação, tentando obter uma relação funcional entre  $X$  e  $Y$  para o modelo escolhido com o menor erro possível. O modelo de regressão linear simples pode ser descrito da seguinte forma:

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

Esta configuração representa o modelo de regressão linear simples. A variável  $X_i$  a variável independente da equação enquanto  $Y_i$  a variável dependente das variações de  $X$ ,  $\alpha$  é uma constante, que representa a interceptação da reta com o eixo vertical e  $\beta$  é outra constante, que representa o declive da reta.

O modelo de regressão é chamado de simples quando envolve uma relação causal entre duas variáveis. O modelo de regressão é multivariado quando envolve uma relação causal com mais de duas variáveis. Isto é, quando o comportamento de  $Y$  é explicado por mais de uma variável independente  $X_1, X_2, \dots, X_n$ .

## 2.2. Correlação

A correlação do termo é usada com funções estatísticas, para referir-se ao movimento de duas ou mais variáveis em consequência fornecida pela função em estudo. A correlação basicamente ocorre quando dois elementos tem harmonia em variação, esta harmonia é dependente, ou seja, estabilização de um depende da posição do outro [Culturama 2013]. Esse método é utilizado quando precisa verificar se existe relação entre duas ou mais variáveis, quanto a força e a direção o que pode ser demonstrado em um gráfico de Dispersão, porém, existe um número que mede a relação conhecido como coeficiente de correlação. O coeficiente de correlação pode variar de  $-1,00$  a  $+1,00$ , com um coeficiente de  $+1$ , indicando uma correlação linear positiva perfeita [Viali 2014].

Assim dadas duas amostras, uma da variável  $x$  e outra da variável  $y$ , sendo  $n$  o tamanho total, o coeficiente de correlação amostral poderá ser calculado através da seguinte expressão:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Considerando as amostras  $x$  e  $y$ , onde  $x$  assumindo valores individuais para  $x_1, x_2, x_3, \dots, x_n$  e  $y$  assumindo valores individuais  $y_1, y_2, y_3, \dots, y_n$ . Assim, após aplicar a fórmula é obtido o coeficiente de correlação que fornece o grau de relacionamento linear entre duas variáveis.

## 2.3. Stepwise

No conjunto de variáveis independentes pode haver variáveis com o coeficiente de correlação tão baixo que não influencia significativamente as variáveis dependentes (saída). O método *stepwise* é usado para selecionar quais variáveis mais influenciam o conjunto de saída podendo, assim, diminuir o número de variáveis a compor a equação

de regressão. Uma das modificações de seleção do *stepwise* é o método *backward*. O método *backward* inicia o processo com todas as variáveis auxiliares do modelo proposto e depois, por etapas, cada uma pode ou não ser eliminada [Action]. A decisão de retirada da variável é tomada baseando-se no método AIC, que são calculados para cada etapa, referentes àquele modelo. O Critério de Informação de Akaike (AIC) é definido como:

$$AIC_p = -2 \log(L_p) + 2[(p + 1) + 1]$$

Em que  $L_p$  é a função de máxima verossimilhança do modelo e  $p$  é o número de variáveis explicativas consideradas no modelo.

### 3. Metodologia

A metodologia *Cross-Industry Standard Process for Data Mining* (CRISP-DM) [Reinartz et al. 1998] permite atacar o problema de extração do conhecimento de forma progressiva e organizada, partindo de uma análise de alto nível, que busca a compreensão do negócio, e indo em direção à definição e implantação de modelos que permitam efetivamente atingir os objetivos da mineração. Essa metodologia é constituída de seis etapas: Entendimento do negócio, Entendimento dos dados, Preparação dos dados, Modelagem, Avaliação e Aplicação conforme figura 1.

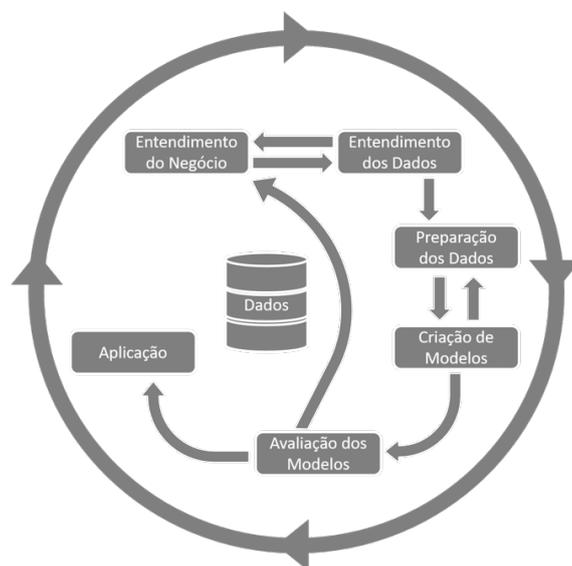


Figura 1. Metodologia CRISP-DM

De acordo com essa descreve-se as seguintes etapas:

- **Entendimento do Negócio:** Inicialmente, é vital entender qual problema deve ser resolvido. Apesar de parecer óbvio, muitos projetos podem chegar de forma ambígua, pré-empacotados e até mesmo ser de uma área de negócios desconhecida pelo especialista. Assim, é preciso análise o contexto em que a demanda está inserida, quais são os cenários previstos (ou não) para que o projeto comece da melhor forma possível.

- **Entendimento dos Dados:** Resolver o problema é o principal objetivo. A compreensão dos dados é a matéria prima para que esta solução seja construída. Aqui, vamos entender as características e limitações das bases de dados, o histórico, sua composição, seu tipo e se os dados realmente são suficientes para entender o problema proposto.
- **Preparação:** dos Dados Existem inúmeras técnicas e tecnologias para se trabalhar com dados. E, com isto, surgem diferentes formatações para as bases de dados. Pensamos aqui em transformar dados não estruturados em estruturados, fazer tratamento de *missing*, converter diferentes tipos de dados de acordo com a necessidade, entender se os dados são categóricos, contínuos, se devem ser normalizados ou não, dentre muitas outras tarefas. Nesta etapa temos também a construção de variáveis, fundamental para o sucesso de qualquer modelo. Até este ponto, é dispendido entre 70% e 80% do tempo do projeto. Isso mesmo, você não leu errado. As quatro primeiras etapas do *CRISP-DM* devem ser tratadas com muita atenção, pois sua boa execução é determinante para o sucesso do projeto (e economizará muito tempo e dinheiro).
- **Modelagem:** Neste ponto serão utilizadas técnicas que sejam mais aderentes ao objetivo do projeto, seja ele uma predição, classificação, agrupamento ou regressão. Nesta etapa temos o saída do nosso projeto, o que será entregue ao nosso cliente. A etapa de modelagem pode "conversar" livremente com a etapa de preparação de dados, seja para readequação dos dados ou mesmo para criação de novas variáveis que ajudem a explicar o fenômeno.
- **Avaliação:** Nesta etapa realizamos um rigoroso *assessment* dos resultados, para que haja confiança no projeto que está a uma etapa de ser entregue. Se necessário, podemos voltar para a primeira etapa (Entendimento do Negócio), caso seja entendido que o objetivo do projeto ainda não tenha sido alcançado.
- **Aplicação:** Nesta etapa, o projeto é encerrado com a entrega feita ao cliente. É a etapa menos técnica do processo de *data mining*, mas não a menos importante. Aqui o resultado é entregue ao cliente em forma de um relatório ou na forma de uma implantação de um sistema para acesso aos dados em tempo real.

## 4. Resultados

De acordo com a Metodologia apresentada, segue a discussão dos resultados em cada etapa do *CRISP-DM*.

### 4.1. Entendimento do negócio

Nessa etapa procurar identificar quais os fatores influenciam o desempenho acadêmico no ENADE, considerando os cursos de Sistemas de informação e Ciências da computação das ies públicas e privadas do estado de Pernambuco.

### 4.2. Entendimento dos dados

O Ministério da Educação define, anualmente, as áreas propostas pela Comissão de Avaliação da Educação Superior (Conaes), órgão colegiado de coordenação e supervisão do Sinaes, e a periodicidade máxima de aplicação do ENADE em cada área será trienal [INEP 2015]. A base que contém os dados atuais sobre os cursos de Sistemas de informação e ciências da computação, que está disponível no site do INEP, é referente ao exame aplicado no ano de 2014, portanto nessa pesquisa foi utilizada a base de dados do ENADE 2014 [INEP 2018].

### 4.3. Preparação dos dados

Nessa seção está descrito todo o processo realizado no pré processamento da base de dados.

#### 4.3.1. Limpeza e filtragem dos dados

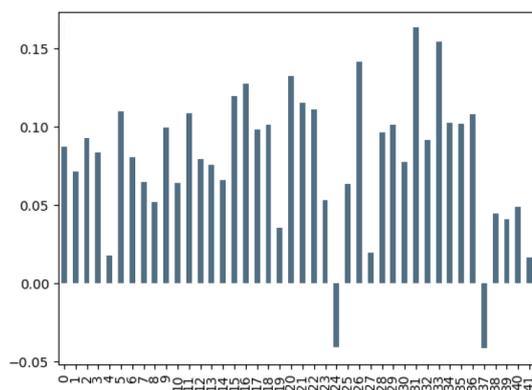
- Nessa etapa foi necessário filtrar os dados da base para atender os objetivos do projeto. A partir da base no estado original, que contém os dados de todas as IES do Brasil, foi realizada uma seleção dos registros referente ao estado de Pernambuco.
- Em seguida foram selecionados, dentre os dados de Pernambuco, os referentes aos cursos de sistemas de informação e ciência da computação.
- Logo depois foram removidas as colunas que representavam variáveis categóricas e nominais, das quais descreviam informações sobre o estudante e as IES, permanecendo somente as variáveis discretas do questionário do estudante.
- Para que possa explorar e inferir uma relação entre os dados foram selecionadas as colunas pertinentes à nota geral e as assertivas numéricas da Organização Didático-Pedagógica que está presente no questionário do estudante, onde o participante indica o grau de concordância com cada afirmação, segundo uma escala que varia de 1 (discordância total) à 6 (concordância total), aos candidatos que não souberam responder ou não se aplicava ao seu curso seus dados foram desconsiderados.
- Portanto depois de todos esses procedimentos citados acima, a base utilizada para aplicar os procedimentos estatísticos foi formada por 861 registros e com 43 colunas referentes às assertivas e à nota geral (ver Tabela 1).
- Nesses procedimentos foram utilizados *scripts*, na linguagem *Python* para realizar os processos de filtragem, e a ferramenta *Excel*, do pacote *office* da *Microsoft*, para visualização e entendimento dos dados.

**Tabela 1. Estado da base de dados em cada procedimento de filtragem  
ENADE 2014**

Estado da Base	Linhas	Colunas
Original	481720	143
Pernambuco	13879	143
Cursos (sistemas de informações e ciências da computação)	1073	143
Colunas	1073	43
Dados Inválidos (N/A)	861	43
<b>Base Final</b>	<b>861</b>	<b>43</b>

#### 4.3.2. Seleção das variáveis

Com a base organizada foi possível aplicar os métodos estatísticos para extração de informações. No primeiro momento foram utilizados os dados disponíveis para verificar a correlação entre os dados. Os resultados estão dispostos abaixo através de um gráfico (Figura 2).



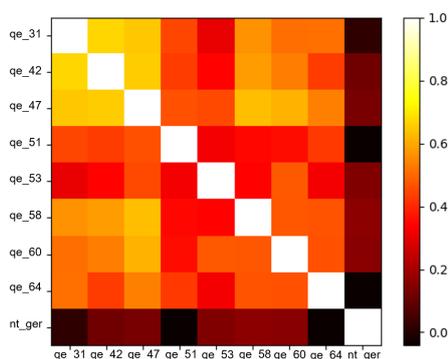
**Figura 2. Correlação das variáveis com a nota geral**

De acordo com o resultado apresentado, o maior coeficiente de correlação foi de 16%, que segundo Paulo Justiniano [Justiniano] indica-se uma correlação bem fraca. Como há uma baixa correlação e uma dependência com muitos fatores, utilizando o software R, foi aplicado o método de seleção de variáveis *Stepwise* com direção *backward*, com a função `step()` utilizando como entrada a direção e a Regressão Linear (RL). Para fazer o RL, foi utilizado a função `lm()`, como entrada a nota geral (variável resposta), e as assertivas disponíveis no questionário do estudante (variáveis explicativas).

Aplicando o método *stepwise*, das 42 variáveis que foram utilizadas, resultou na seleção de 8 variáveis das quais estão discriminadas na tabela abaixo (Tabela 2) com as questões e a matriz com os coeficientes de correlação (Figura 3).

**Tabela 2. Assertivas selecionadas pelo stepwise**

Identificador	Assertivas
qe_31	O curso contribuiu para o desenvolvimento da sua consciência ética para o exercício profissional.
qe_42	O curso exigiu de você organização e dedicação frequente aos estudos.
qe_47	O curso favoreceu a articulação do conhecimento teórico com atividades práticas.
qe_51	As atividades realizadas durante seu trabalho de conclusão de curso contribuíram para qualificar sua formação profissional
qe_53	Foram oferecidas oportunidades para os estudantes realizarem intercâmbios e/ou estágios fora do país.
qe_58	Os professores utilizaram tecnologias da informação e comunicação (TICs) como estratégia de ensino (projutor multimídia, laboratório de informática, ambiente virtual de aprendizagem).
qe_60	O curso disponibilizou monitores ou tutores para auxiliar os estudantes.
qe_64	A biblioteca dispôs das referências bibliográficas que os estudantes necessitaram.



**Figura 3. Matiz de correlação das variáveis escolhidas pelo método *stepwise***

#### 4.4. Trabalhos Futuros

De acordo com os resultados obtidos com a utilização da metodologia *CRISP-DM* na seção 3. As próximas etapas para o desenvolvimento da pesquisa serão:

- Desenvolver, na etapa de modelagem, os modelos de regressão aplicados a base de dados tratada conforme a etapa de preparação dos dados;
- Na etapa de validação, será utilizado o método estatístico *holdout* para realização da simulação de Monte Carlo, ao final será calculado a métrica do erro médio relativo para validação dos modelos de regressão;
- Criar cenários de simulação de dados e utilizar as etapas de modelagem e validação para ratificar os resultados encontrados com a base de dados real do INEP 2014;
- Para confirmar os achados dessa pesquisa realizada com a base de dados do INEP 2014, utilizaremos a metodologia *CRISP-DM* para a base do INEP 2017 disponível em 09/10/2018.

#### 5. Conclusão

Foi possível com esse estudo, identificar quais os fatores que influenciaram no desempenho acadêmico no ENADE, considerando os cursos de Sistemas de informação e Ciências da computação das ies públicas e privadas do estado de Pernambuco. Os fatores que menos influenciaram na construção dos modelos de regressão não foram considerados para previsão da estimativa de erro. Os fatores com maiores relevâncias estão citados abaixo:

- O curso contribuiu para o desenvolvimento da sua consciência ética para o exercício profissional.
- O curso exigiu de você organização e dedicação frequente aos estudos.
- O curso favoreceu a articulação do conhecimento teórico com atividades práticas.
- As atividades realizadas durante seu trabalho de conclusão de curso contribuíram para qualificar sua formação profissional
- Foram oferecidas oportunidades para os estudantes realizarem intercâmbios e/ou estágios fora do país.
- Os professores utilizaram tecnologias da informação e comunicação (TICs) como estratégia de ensino (projeter multimídia, laboratório de informática, ambiente virtual de aprendizagem).

- O curso disponibilizou monitores ou tutores para auxiliar os estudantes.
- A biblioteca dispôs das referências bibliográficas que os estudantes necessitaram.

Esses fatores foram importantes pois observa-se uma integração de competências entre os três pilares da educação (aluno, escola e professor). A escola precisa de investimentos para melhorar a qualidade no que diz respeito a presença de bibliotecas, monitorias, tutores, programas de intercâmbio, como também a presença de tecnologias de informação aplicadas ao ensino. O Professor precisa realizar, durante as aulas, articulação entre a teoria e a prática, bem como, a qualificação da formação profissional do aluno. Para isso, é necessário o investimento em tecnologias como os ambientes virtuais de aprendizagem. Já o aluno através desses investimentos no âmbito da escola e do professor, se sentirá motivado e exercerá sua profissão com ética, organização e dedicação.

### Referências

- Action, P. Seleção backward. [Online; accessed 11-Outubro-2018].
- Brito, M. R. F. d. et al. (2008). O sinaes e o enade: da concepção à implantação. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*.
- Culturama, W. (2013). Definição, conceito, significado, o que é correlação. [Online; accessed 11-Outubro-2018].
- Griboski, C. M. (2012). O enade como indutor da qualidade da educação superior. *Estudos em avaliação educacional*, 23(53):178–195.
- Helene, O. (2006). *Métodos dos Mínimos Quadrados*. Editora Livraria da Física.
- INEP (2015). Perguntas frequentes. [Online; accessed 11-Outubro-2018].
- INEP (2018). Microdados. [Online; accessed 11-Outubro-2018].
- Justiniano, P. 8 correlação. [Online; accessed 11-Outubro-2018].
- Reinartz, T., Wirth, R., Clinton, J., Khabaza, T., Hejlesen, J., Chapman, P., e Kerber, R. (1998). The current crisp-dm process model for data mining. In *Proceedings of the Annual Meeting of the German Machine Learning Group FGML-98*, pages 14–22.
- Viali, L. (2014). Texto v: Correlação e regressão. *Série estatística básica*. Disponível em: <http://www.mat.pucrs.br/~lori>.

## **Análise da Logística Reversa sob a perspectiva sustentável dos 5R's de uma empresa de economia criativa**

**Bárbara Valéria Felix de Bourbon Nogueira e Silva<sup>1</sup>, Fernanda Maria de Melo<sup>1</sup>,  
Gabriela de Vasconcelos<sup>1</sup>, Roberta Mikaella Xavier de Farias<sup>1</sup>, Suzana Thaís  
Batista de Lima<sup>1</sup>**

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

{barbarabourbon.BrB, fernandamelo472, devasconcelos.gabriela,  
robertamxfarias, suzanathaisbl}@gmail.com

### **Resumo**

A consolidação do conceito de sustentabilidade a nível mundial ensejou o surgimento de políticas, estratégias e ações organizacionais a fim de alcançar o desenvolvimento sustentável e contribuir para o alcance dos ODS's estabelecidos pela ONU. Nesse sentido, surge a PNRS que propõe práticas de logística reversa (LR) e a operacionalização dos 5r's. A LR tem como finalidade gerenciar o fluxo de materiais de forma inversa, ou seja, do seu ponto de consumo até o de origem. Diante disso, o objetivo deste artigo é analisar a logística reversa a partir dos 5R's no contexto de uma empresa de economia criativa. A metodologia foi utilizada de modo a coletar informações para um levantamento sobre o processo produtivo da empresa supracitada, a partir do redirecionamento de resíduos sólidos e enfatizando a política dos 5 R's.

**Palavras-chave: Sustentabilidade. Logística Reversa. Economia Criativa. 5R's.**

### ***Abstract***

*The consolidation of the concept of sustainability worldwide has led to the emergence of policies, strategies and organizational actions in order to achieve sustainable development and contribute to the achievement of the ODS established by the UN. In this sense, the PNRS appears that proposes practices of reverse logistics (LR) and the operation of the 5r's. The purpose of LR is to manage the flow of materials in reverse, that is, from its point of consumption to its source. Therefore, the objective of this article is to analyze the reverse logistics from the 5R's in the context of a creative economy company. The methodology was used in order to collect information for a survey on the production process of the aforementioned company, from the redirection of solid waste and emphasizing the policy of the 5 R's.*

***Key-words: Sustainability. Reverse Logistics. Creative Economy. 5R's.***

## 1. Introdução

O avanço empresarial tem provocado impactos na sociedade, principalmente ambientais, levando-se em conta a utilização do meio ambiente para obtenção de matéria prima e descarte de resíduos. De acordo com estudo do IPEA (2017), a geração de resíduos sólidos no Brasil gira em torno de 160 mil toneladas diárias, e deste total 30% a 40% são considerados passíveis de reaproveitamento e reciclagem. Para Donaire (1995) com o aumento acelerado da população e da indústria, o mundo anda sofrendo impactos ambientais, por isso as empresas têm se ajustado com relação a essas questões ambientais para o desenvolvimento dos seus produtos. Dessa forma, as empresas passaram a buscar alternativas para a diminuição desses impactos, considerando fatores como competitividade e preocupação ambiental.

Diante desse cenário, a logística reversa se apresenta como uma área funcional da organização com o intuito que as empresas direcionam os resíduos que seriam descartados pelos consumidores pelo desgaste ou simplesmente pelo descarte para a obtenção de novos produtos, ou seja, o ciclo de vida do produto passou a não terminar no momento da entrega. Além da característica sustentável, esse viés da logística promove uma diminuição nos custos no que diz respeito a otimização da produção, pois reutiliza materiais já existentes. Em estudo, Goldsby e Closs (2000) fizeram uma pesquisa em torno da cadeia de suprimentos de cervejas e refrigerantes, que levou seus integrantes a terceirizarem todo o processo de coleta e retorno para reciclagem de embalagens usadas e com isto obtiveram uma economia anual superior a US\$ 11,4 milhões. Segundo os autores, a correta implementação do custeio por processo na cadeia de suprimentos apontou oportunidade para redução de custos totais.

A concepção de logística reversa vem sendo explorado desde os anos 1970, quando as questões como canais e fluxos reversos se relacionavam mais com aspectos ambientais e com reciclagem de materiais [Kinobe *et al.*, 2012]. Ela começa a ser aplicada nas organizações a partir de então, sendo considerada uma ferramenta estratégica de competitividade no mercado empresarial.

A Lei Federal nº 12.305/2010, que institui a Política Nacional de Resíduos Sólidos (PNRS), tem como principal objetivo a resolução de problemas sociais e econômicos vinculados ao manejo e disposição incorreta de resíduos sólidos. Nesse cenário, os Objetivos de Desenvolvimento do Milênio (ODM) estabelecidos pela ONU tem a finalidade de tornar o mundo um lugar mais justo, solidário e melhor para se viver. Tendo em vista a evolução da humanidade com relação a preocupações ambientais e sociais, os Objetivos do Desenvolvimento Sustentável (ODS) surgem, então, como um aprimoramento e adequação à realidade vivida, de modo que as metas sejam atualizadas e o espírito de mudança dos países membros aprimorados. Entre os ODS, destaca-se o objetivo 12 que assegura padrões de produção e de consumo sustentáveis, ressaltando o seu poder de utilidade em questão dos 5 R's (Repensar, Reduzir, Recusar, Reciclar e Reutilizar).

Diante do exposto, o presente estudo se justifica pela importância da logística reversa como meio para a reutilização de resíduos sólidos, que seriam descartados mas através da mesma, são destinados para uma nova forma de uso. Dessa forma, esse artigo visa contribuir para o estudo da mesma enquanto estratégia sustentável, considerando seu papel como potencializador de inovação e contribuinte na relação direta entre

sociedade e meio ambiente.

Diante desse contexto, busca-se responder à seguinte questão: Como se dá a logística reversa sob a perspectiva sustentável dos 5R's da empresa Saulo M? Essa empresa tem um processo produtivo baseado principalmente na coleta de resíduos que são recolhidos e utilizados na confecção de produtos específicos. O objetivo geral do trabalho é analisar o processo de logística reversa sob a perspectiva sustentável dos 5R's na empresa em questão. Quanto a metodologia, aplica-se uma pesquisa qualitativa de caráter exploratório-descritiva baseada em um levantamento de informações a partir de entrevista estruturada.

## **2. Referencial Teórico**

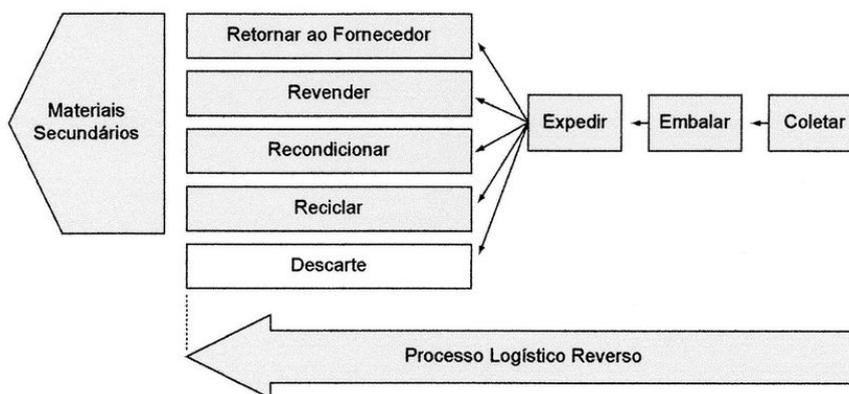
### **2.1 Sustentabilidade no âmbito organizacional**

A logística é ramo extremamente importante para as empresas pois permite uma melhor alocação dos processos, que vai desde a obtenção de insumos até a entrega do produto ao consumidor final, através de uma otimização dos recursos. Além disso, tem um papel determinante pois contribui para que os produtos não fiquem estagnados no estoque já que trabalha em uma ideia “just in time”, onde os mesmos devem ser produzidos de acordo com sua demanda e na hora exata, para que assim não provoque acúmulo na estocagem. De acordo com Ballou (1993), o objetivo da logística é adquirir níveis de excelência e baixo custo, através de um bom planejamento e uma boa coordenação de atividades.

Considerando a logística como agente de movimentação e distribuição de materiais entre o ponto de origem ao de consumo, a logística reserva surge com um propósito totalmente inverso. Pohlen e Farris (1992) dizem que a logística reversa é a atividade que faz com que o bem adquirido pelo consumidor volte para o seu produtor, ou seja, o caminho oposto da logística direta. Em outras palavras, de acordo com Lacerda (2002), pode ser entendida como um processo complementar à logística tradicional, pois enquanto a mesma tem a função de levar produtos dos fornecedores até os clientes, a logística reversa deve concluir o ciclo, levando os produtos que já foram utilizados de volta para sua origem.

A logística reversa vem sendo utilizada com frequência em livros modernos. Pode ser classificada como uma versão contrária da logística que conhecemos. Para Leite (2003, p. 16-17) a logística reversa se define como a “área da logística empresarial que planeja, opera e controla o fluxo e as informações logísticas correspondentes, do retorno dos bens de pós-venda e pós-consumo ao ciclo de negócios ou ao ciclo produtivo, por meio dos canais de distribuição reversos, agregando-lhes valor de diversas naturezas: econômico, ecológico, de imagem corporativa, entre outros.” A verdade é que um planejamento de logística reversa utiliza os mesmos processos que um planejamento convencional. Os dois tratam de nível de serviço, armazenagem, transporte, nível de estoque, fluxo de materiais e sistema de informação. A diferença é onde o processo começa e acaba.

De acordo com Zikmund e Stanton (2005), a ideia mais antiga sobre logística reversa apareceu a partir do início dos anos 70, onde se aplicou os conceitos de distribuição voltados para o método de forma inversa. Atualmente, essa técnica vem ganhando cada vez mais força nas empresas, principalmente pelo seu potencial de diferencial competitivo. Barbosa (2005, p. 2) diz que “Empresas que possuem um processo de logística reversa, bem gerido, tendem a se sobressair no mercado, uma vez que estas podem atender seus clientes de forma melhor e diferenciada de seus concorrentes”.



**Figura 1. Atividades do Fluxo Logístico Reverso**

Fonte: Lacerda (2002)

A figura acima retrata como se dá o processo da logística reversa, iniciando na coleta de produtos já existentes, até a nova destinação dos mesmos. Esse fluxo de materiais que voltam para as empresas pelos consumidores, pode por vários motivos, como: devolução de produtos com defeito, para o processo de reciclagem, destiná-los a vender para outra empresa, ou até mesmo, descarte no lixo. A decisão final será definida pela própria organização.

Baseando-se nas ideias apresentadas, pode-se dizer que a logística reversa adota estratégias e ações sustentáveis no âmbito organizacional tendo em vista a preservação do meio ambiente. A Lei nº 12.305/10, que institui a Política Nacional de Resíduos Sólidos (PNRS), é atual e contém ferramentas que permitem enfrentar os principais problemas ambientais, sociais e econômicos resultantes do manuseio inapropriados dos resíduos sólidos. Atraído aos Objetivos de Desenvolvimento Sustentável (ODS) que é uma recente iniciativa das Organizações das Nações Unidas (ONU), onde trata-se de uma agenda criada para combater questões do desenvolvimento sustentável. Nesse contexto, a logística reversa possui papel fundamental para o desenvolvimento sustentável, ligadas às ações à serem citadas adiante.

## **2.2 Políticas, estratégias e ações sustentáveis: PNRS, LR, ODS, 5R's**

Sustentabilidade se define como uma prática da utilização de materiais visando a não agressão ao meio ambiente. De acordo com Bellen (2005, p.23) “sustentabilidade é o emprego de tecnologia e de organização social, atendendo a necessidade das gerações

presentes sem comprometer a possibilidade das gerações futuras atenderem as suas próprias necessidades”. Já com relação ao desenvolvimento sustentável, a Comissão Mundial sobre o Meio Ambiente e Desenvolvimento (1987), define como “a capacidade de prover o crescimento e suprir as necessidades da geração presente sem afetar a habilidade das gerações futuras”. No âmbito empresarial, a sustentabilidade proporciona que os produtos e/ou serviços atendam as necessidades dos clientes, evitando danos ao meio ambiente, ou seja, extraindo apenas os recursos necessários, de acordo com sua capacidade de reposição.

Em 2015, foram estabelecidos os Objetivos de Desenvolvimento Sustentável (ODS), a partir de uma Cúpula das Nações Unidas com o intuito de auxiliar no processo de sustentabilidade. Entre os 17 objetivos assegurados para serem atingidos até 2030, destaca-se o Consumo e Produção Sustentáveis. Esse objetivo tem relação direta com a questão de conscientização sobre o consumo de modo a preservar a natureza para o alcance de desenvolvimento econômico, buscando a responsabilidade dos consumidores sobre as políticas sustentáveis e o melhor aproveitamento de resíduos sólidos. Conforme estabelecido pela lei que trata da Política Nacional de Resíduos Sólidos, cuja busca promover a prevenção e a redução na geração de resíduos sólidos, tendo como meta a prática de consumo sustentável e maneiras para propiciar o aumento da reciclagem e da reutilização desses resíduos e a destinação ambientalmente adequada dos rejeitos.

Nesse cenário, a política dos 5 R's ganha força no mercado devido a tentativa de minimizar a quantidade de resíduos descartados. Com relação ao Ministério do Meio Ambiente (2017), os 5 R's fazem parte de um procedimento educativo que tem o objetivo de melhorar os hábitos cotidianos dos cidadãos, através de ações como Repensar, Reduzir, Recusar, Reciclar e Reutilizar. Fazendo uma análise sobre cada um desses pontos, chega-se às seguintes conclusões: Repensar, está relacionado com a real necessidade de adquirir certo produto, ou seja, a consciência do consumidor sobre seu tipo de consumo. Segundo Alkmin (2015, p. 35) “É o ato de pensar de novo, pensar duas vezes, refletir. Antes de efetuar qualquer compra, é preciso refletir sobre a real necessidade de tal aquisição, se não está sendo levado a comprar por impulso”. Recusar, tem a ver com dar preferência a um consumo consciente, com relação a produtos que não agridam o meio. Para Alkmin (2015, p. 35) “Quando se recusa produtos que prejudicam a saúde e o meio ambiente, contribui-se para um mundo mais limpo”. Reduzir trata-se da redução de consumo e porventura, a produção de lixo. Assim, é necessário optar por produtos que sejam mais duráveis. Em síntese, seria consumir menos, dando predileção aos produtos têm uma maior durabilidade. Reutilizar consiste em reaproveitar materiais que seriam descartados, mas ganha um novo direcionamento de função. Louredo (2017) afirma que ao reutilizar a vida útil do produto fica maior. Já reciclar, é o conceito que mais tem se falado nos dias atuais. Está relacionado em transformar algo que já não tinha mais tanta usabilidade em algo totalmente novo. Além disso, Alkmin (2015, p. 35) também entra no ponto sobre a questão da reciclagem com relação a redução do consumo de água, energia e insumo.

As práticas sustentáveis passaram a ser utilizadas principalmente como uma fonte estratégica de potencial competitivo, visto que há uma crescente preocupação dos próprios consumidores com o processo produtivo das empresas, cujo mau uso e descarte irresponsável dos materiais pode prejudicar o meio ambiente. Dessa forma, as organizações que têm sua imagem atrelada a práticas socialmente e ambientalmente viáveis estão sendo mais bem vistas no mercado. Além disso, as práticas de

sustentabilidade podem ser bem mais econômicas, principalmente quando os materiais utilizados na produção são reciclados, como pode ser observado nas atuações da logística reversa.

### **2.3 Economia Criativa como estratégia de sustentabilidade**

O termo “Indústria Criativa” é relativamente novo. No entanto, já é bem reconhecido no mercado. Sua definição rodeia as vertentes da arte, cultura e criatividade, sendo essas as matrizes de tal economia. Segundo Oliveira, Araujo e Silva (2013), a designação “Indústria Criativa” se desenvolveu e ampliou o âmbito das indústrias culturais para além das artes e marcou uma mudança na abordagem a potenciais atividades comerciais que até recentemente eram consideradas puramente ou predominantemente em termos não econômicos.

Esse segmento passou a ganhar espaço pelo seu vínculo com a cultura de uma forma criativa, considerando também questões econômicas. Em outras palavras, significa um olhar pro mundo de forma inovadora e que isso resulte em economia para a sociedade. No Brasil, essa temática passou a chamar atenção na Conferência Ministerial da UNCTAD XI em 2004, porém sem grandes avanços na área. Apenas a partir de 2011, o governo brasileiro deu atenção a partir da abertura de uma secretaria, a Secretaria de Economia Criativa (SEC), com ligação ao Ministério da Cultura. De acordo com dados levantados pela ONU (2013), 10% do PIB mundial já era em cima da economia criativa. Desse modo, essa vertente trabalha com o capital intelectual, ou seja, a criatividade passou a ser considerada como matéria prima. Howkins (2001) deu condição para a economia criativa, ao comercializá-la como uma forma das pessoas transformarem ideias em dinheiro. Esse cenário tem a capacidade para que se desenvolvam formas de compreender a importância do meio ambiente e não o utilizar apenas para a obtenção de matérias-primas. De acordo com Deheinzelin (2017) os recursos intangíveis se multiplicam com o uso e só se conseguirá chegar a sustentabilidade se utilizar da economia criativa, pois ela gera riquezas a partir de recursos que não se esgotam, como cultura e conhecimento. O estímulo a competências criativas não se limita à capacidade técnica, como também abrange atitudes e posturas de empreendimento, habilidades sociais e de comunicação, além da compreensão de dinâmicas socioculturais e de mercado, capacidade de articulação e sustentabilidade [SEC, 2012].

Diante desse contexto, a economia criativa tem o potencial de ser utilizada em vista da promoção de desenvolvimento sustentável, pois o processo produtivo decorrente dessa ferramenta tende a ser fruto do equilíbrio entre o econômico e o social, pois busca que o ser humano possa produzir e consumir de forma criativa, gerando alternativas aos modelos praticados na atualidade. Ao considerar que cultura, criatividade e conhecimento são recursos inesgotáveis, estes se tornam recursos estratégicos para a sustentabilidade dos negócios, e portanto, das empresas. A inovação sustentável presume que a produção, a assimilação ou a exploração de produtos, processos produtivos, táticas de gestão ou negócios modernos ou melhorados que tragam benefícios econômicos, sociais e ambientais [BARBIERI *et al.*, 2010].

### 3. Metodologia

O presente artigo foi elaborado a partir de uma abordagem qualitativa, que consiste em um trabalho de campo, com vistas a coletar dados diretos de um determinado ambiente. Sobre a mesma, Godoy comenta que os estudos denominados qualitativos têm como preocupação fundamental o estudo e a análise do mundo empírico em seu ambiente natural. Nessa abordagem valoriza-se o contato direto e prolongado do pesquisador com o ambiente e a situação que está sendo estudada. [GODOY, 1995]

Tal modalidade foi escolhida tendo em vista a oportunidade de coletar material necessário para o desenvolvimento da pesquisa de forma detalhada, com o intuito de obter informações mais fiéis possíveis, relacionadas a temática em questão. Quanto ao objetivo, a pesquisa é exploratória-descritiva, pois se propôs a explorar e descrever a dinâmica de produção da empresa Saulo M que faz uso da logística reversa como matriz para a captação de insumos produtivos. A pesquisa exploratória, segundo Raupp e Beuren (2006), consiste no aprofundamento de conceitos preliminares sobre determinada técnica não contemplada de modo satisfatório anteriormente. Assim, contribui para o esclarecimento de questões superficialmente abordadas sobre o assunto. Com relação ao tipo descritiva, segundo Gil (1999), as pesquisas descritivas têm como intenção principal a descrição das características de determinada população ou fenômeno. No caso específico, buscou-se abordar como se dava o método de logística reversa na prática.

Quanto ao procedimento para levantamento de dados, utilizou-se uma entrevista estruturada relacionada abaixo realizada com o proprietário, baseada em um roteiro com perguntas específicas, para analisar e entender os processos da sua empresa a partir de suas concepções. Essa coleta se deu a partir da elaboração de um questionário com perguntas abertas, cujo intuito foi investigar o processo produtivo de tal empresa enfocando o abastecimento de matéria prima e relacionando com a políticas do 5 Rs.

**Tabela 1. Entrevista Estruturada**

1-	Como você descreve a sua empresa quanto aos processos produtivos?
2-	Quais os produtos que você oferece e qual seu diferencial quanto a eles?
3-	Como funciona o processo de reabastecimento da matéria prima?
4-	Como é feito o processo de criação a partir da matéria prima que você consegue armazenar?
5-	Na sua concepção, sua empresa se encaixa no conceito de economia criativa?
6-	Quais os produtos que você oferece e qual seu diferencial quanto a eles?
7-	Quais as dificuldades enfrentadas ao escolher uma modalidade de produção diferente da tradicional?
8-	Qual o diferencial competitivo do seu modelo de negócio?
9-	Sabendo que a logística reversa é uma técnica onde os produtos fazem o caminho inverso ao tradicional, onde saem do consumidor e voltam ao produtor para um novo direcionamento, como se dá esse processo na sua empresa?
10-	Considerando os 5R's (Repensar, Reduzir, Recusar, Reciclar e Reutilizar), como é feita essa abordagem dentro da sua empresa?
11-	Como é feito o processo de descarte?

#### 4. Resultados e Discussão

A partir da coleta de dados qualitativos por meio de uma entrevista de profundidade com o proprietário da Saulo M, Saulo Moura, localizada na cidade de Caruaru, Pernambuco. Tal empresa tem como base o conceito de economia criativa, na qual o principal capital utilizado, tanto no início de suas atividades como atualmente, é o capital intelectual. A mesma imprime o conceito artístico do próprio idealizador, como também tendências de moda nas diversas linhagens de produtos desenvolvidos, são eles: vestuário, mobiliário, bolsas, embalagens, objetos de decoração e em seus serviços prestados de aluguel de mobília para eventos. Pode-se dizer que a partir de seus produtos o criador transforma sua arte em algo rentável, esse contexto está intimamente ligado com a definição de indústria criativa de Oliveira, Araujo e Silva (2013).

O empresário afirma que o início do seu processo produtivo é baseado na busca por matéria prima e a partir dela, o proprietário passa a desenvolver o processo criativo. Para a produção, o mesmo desenvolveu vários meios alternativos para compor a base de seus produtos, em todas as escolhas existe a preocupação com o custo, que consequentemente recai no valor do produto. O processo de captação de matéria prima funciona como uma logística reversa, ao passo que Ballou (1993) define. Outra característica eminente a respeito da matéria prima é o fato de praticamente tudo ser material reutilizado. Tal fator incide sobre o que foi abordado anteriormente a respeito de sustentabilidade.

A reutilização se apresenta a medida em que o empresário montou uma estratégia logística em conjunto com profissionais do setor de catadores no qual estes são pagos pelo recolhimento de insumos que são reutilizados de modo que estes continuem a exercer sua função em objeto produzido no qual anteriormente anulada. Fator que remete a afirmação de Louredo quanto a vida útil do produto.

A partir da matéria prima, desenvolve-se o processo produtivo para que se tenha o produto acabado, no caso da empresa estudada, a mesma é reabastecida através do próprio empresário, visto que o mesmo usa madeira de reflorestamento que normalmente seria direcionada para a construção civil. Após o uso da mesma, os restos são vendidos para padarias, essas utilizam para queima no processo de fabricação do pão.

Outra maneira de re-abastecimento utilizada pela empresa consiste em coletas terceirizadas através de catadores de material reciclável, onde estes são remunerados de acordo com a quantidade de aviamentos ou utensílios que recolhem, sendo eles vindos de bolsas velhas, móveis, eletrodomésticos etc, encontrados no ambiente de trabalho desses profissionais ou partindo de outras fontes, a partir do momento em que há um objeto que atende uma demanda de material utilizado na fabricação que não tenha vindo necessariamente do lixo.

O processo de reciclagem, como afirmado anteriormente por Alkmin (2015), se manifesta na mudança de função de um objeto a partir do momento em que ele transformado em algo novo, ela se mostra na rotina de produção da empresa ao passo de que muitos de seus insumos tinham uma roupagem anterior com funções completamente nulas e passam a exercer um novo papel.

Assim, o desenvolvimento criativo é definido por Saulo como um planejamento em conjunto com o cliente, o que ocorre nesses casos é o fato de que a venda segue um

processo de consultoria, no qual o objeto a ser produzido é desenhado de acordo com as necessidades do cliente em questão, visando aderir o máximo de funcionalidade e identidade. Esse processo de criação acontece de maneira inversa ao mais visto, na qual ele obtém o insumo a partir da visualização do produto finalizado adaptando-se às características e aprimorando-as de forma inovadora ao produto bruto.

Partindo da ideia dos 5R's, enfatizando a reutilização e reciclagem, o diferencial competitivo da empresa consiste inicialmente no baixo valor de custo de produção, no qual é possível alimentar estoque de matéria prima para três ou quatro clientes a partir de apenas metade do pagamento de um cliente quando o produto é demandado por encomenda. É importante salientar que o produto não se distancia da realidade de fabricação para venda de pronta entrega, na qual há a preferência por insumos de valor de obtenção baixo onde a principal força trabalhada consiste na forma de produção focada na qualidade de execução. O empresário destaca a qualidade como sua principal característica no mercado.

## **5. Conclusão**

Diante do exposto, observou-se que o negócio analisado têm ênfase no conceito de economia criativa tomando como base nos pressupostos dos 5r's. Enfatizando a reutilização e reciclagem de insumos a partir da coleta em ambientes de descarte ou do resgate através de outros meios como terceirizar essa coleta através de catadores ou até mesmo doações das pessoas, a fim de alcançar metas estabelecidas dentre os Objetivos do Desenvolvimento Sustentável (ODS), mais precisamente no objetivo 12 que é relacionado com a produção e o consumo sustentáveis e considerando a Política Nacional de Resíduos Sólidos (PNRS) que busca promover a prevenção e a redução na geração de resíduos sólidos.

Como destacado por Deheinzelin (2017), a economia criativa é um grande meio para alcançar a sustentabilidade, a partir da utilização de recursos inesgotáveis como criatividade e conhecimento, ou seja, transformá-los em potencial de mercado. Além disso, esse tipo de economia aproxima-se da logística reversa justamente pelo viés sustentável, ao utilizar resíduos que seriam descartados, tendo assim, a possibilidade de redução de custos na produção. Com isso, faz englobar a política dos 5 R's, cujo estudo pertence a ideias de proteção ao meio ambiente, com medidas como Repensar, Reduzir, Recusar, Reciclar e Reutilizar.

Com o levantamento da pesquisa conclui-se que é possível que uma empresa de economia criativa desenvolva a logística reversa a fim de captar matéria prima de forma inovadora. Ao passo que inversamente a forma tradicional de logística, o proprietário da empresa em questão consegue administrar uma produção de produtos que são de cunho artístico, como também tem um forte conceito de moda baseado no slow fashion, assim como se mantém quase que totalmente de forma sustentável. Esse último fator é concluído devido ao uso iminente e corriqueiro dos 5R's em todo seu processo produtivo. Assim, a empresa ganha diferencial competitivo pela imagem positiva de empresa consciente.

## Referências

- Catallão, B. and Fogolin, M. (2011) “Logística Reversa e Marketing Verde”
- Cavalcanti, A. and Lima, G. and Silva, A. Ferreira, C. (2017) “Aplicação do Diagnóstico de Inovação e Gestão a Entidades do Setor de Economia Criativa”
- Costa, A. and Santos, E. (2001) “Economia Criativa: Novas Oportunidades Baseadas no Capital Intelectual” Shibao, F. and Moori, R. and Santos, M. (2010) “A logística reversa e a sustentabilidade empresarial”
- Furlan, J. and Zanotto, M. and Nascimento, R. and Bertollo, D. and d'Avila, A. and Tondolo, V. (2015) “Mapeamento da Produção Científica Internacional sobre Aprendizagem Organizacional”
- Gerhardt, T. and Souza, A. (2009) “Métodos de Pesquisa”
- Jesus, D. (2017) “Economia Criativa, Desenvolvimento e Sustentabilidade: O Caso do Rio de Janeiro”
- Leite, P. (2002) “Logística Reversa Nova Área da Logística Empresarial”
- Maiello, A. (2018) “Implementação da Política Nacional de Resíduos Sólidos”
- Ministério das Relações Exteriores. “Objetivos de Desenvolvimento Sustentável” <http://www.itamaraty.gov.br/pt-BR/politica-externa/desenvolvimento-sustentavel-e-meio-ambiente/134-objetivos-de-desenvolvimento-sustentavel-ods>, Outubro.
- Oliveira, G. and Dilly, P. and Altaf, J. and Troccoli, I. (2017) “Logística Reversa como Ferramenta para Sustentabilidade: O Caso Unisão Transportes LTDA”
- Organização das Nações Unidas no Brasil “17 Objetivos para Salvar o Mundo” <https://nacoesunidas.org/pos2015/>, Outubro.
- Ronconi, E. and Poffo, G. “Sustentabilidade: O Caminho para as Presentes e Futuras Gerações”
- Silva, I. and Paraíso, L. and Junior, I. (2017) “Análise Crítica Política Nacional de Resíduos Sólidos: Principais pontos e aplicabilidade”
- Silva, S. and Ferreira, E. and Roesler, C. and Borella, D. and Gelatti, E. and Boelter, F. and Mendes, P. (2017) “Os 5 R's da Sustentabilidade”
- Vasconcelos, G. (2018) “Análise do Desempenho Sustentável dos Sistemas de Esgotamento Sanitário Municipais de Pernambuco”

## **Análise da Logística Reversa da empresa Baterias Moura**

**Cintia Ferreira<sup>1</sup>, Claudiane Silva<sup>1</sup>, Gabriela de Vasconcelos<sup>1</sup>, Giovana Thais<sup>1</sup>,  
Mayara Silva<sup>1</sup>, Rebeca Barros<sup>1</sup>**

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

{cintiaferreira1996, devasconcelos.gabriela, giovanathaisantos,  
}@gmail.com

{claudiane\_admsilva, mayara\_larissa46, rebecabarros123456}@hotmail.com

### **Resumo**

*O presente artigo tem por objetivo apresentar uma análise da Logística Reversa da empresa Baterias Moura, bem como demonstrar a preocupação da mesma com o meio ambiente. Foi analisado o processo de retorno e reciclagem das baterias usadas, sendo averiguado que a empresa recolhe 100% das baterias produzidas e colocadas no mercado, fazendo com que a empresa reduza custos devido o reaproveitamento de materiais. Foi verificado também um forte empenho com relação a preservação do meio ambiente, sendo que a Baterias Moura cumpre minuciosamente a lei que institui a Política Nacional de Resíduos Sólidos (PNRS), e muito mais do que cumprir uma lei, a empresa de fato tem a responsabilidade de preservar o meio em que ela está inserida.*

**Palavras-chave:** Logística Reversa, Sustentabilidade, PNRS.

### **Abstract**

*The purpose of this article is to present an analysis of Reverse Logistics of the company Baterias Moura, as well as to demonstrate its concern with the environment. It was analyzed the process of return and recycling of used batteries, being verified that the company collects 100% of the batteries produced and placed on the market, causing the company to reduce costs due to the reuse of materials. A strong commitment to environmental preservation was also verified, with Baterias Moura complying with the law that establishes the National Solid Waste Policy (PNRS), and much more than complying with a law. responsibility to preserve the environment in which it is embedded.*

**Keywords:** Reverse Logistics, Sustainability, PNRS.

## INTRODUÇÃO

Decorrente do aumento dos problemas ambientais, a sociedade passa a preocupar-se mais com o meio ambiente. Isso fez com que muitas empresas de mobilizem a favor de ter uma empresa mais sustentável. Para Leite (2002) “A obsolescência e a descartabilidade de crescentes dos produtos observados nesta última década têm-se refletido em alterações em alterações estratégicas empresariais, dentro da própria organiza e principalmente em todos os elos de sua rede operacional”. Isso faz com que os produtos tenham um prazo de vida mais curto, causando o descarte desses produtos, prejudicando o meio ambiente.

A Política Nacional de Resíduos Sólidos (PNRS), lei nº 12.305/10, é bastante atual e contém informações de grande importância, que permite o avanço necessário ao País no enfrentamento dos principais problemas ambientais, sociais e econômicos, decorrentes do manejo inadequado dos resíduos sólidos. A PNRS, prevê a prevenção e a redução na geração de resíduos, tendo como proposta aumentar a reciclagem a reutilização dos resíduos sólidos, tudo aquilo que têm valor econômico e pode ser reciclado ou aproveitado, e a destinam adequada daquilo que não irá precisar ser reciclado ou reutilizado. Criar metas também é um destaque da PNRS, que irão contribuir para a diminuição dos lixões e também coloca o Brasil em patamar de igualdade aos principais países desenvolvidos.

A logística reversa é um dos instrumentos para a aplicação da responsabilidade compartilhada pelo ciclo de vida dos produtos. Segundo o site, Ministério do Meio Ambiente, a PNRS define a logística reversa como um instrumento de desenvolvimento econômico e social caracterizado por um conjunto de ações, procedimentos e meio destinados a viabilizar a coleta e a restituição dos resíduos sólidos ao setor empresarial, para reaproveitamento, em seu ciclo ou em outros ciclos produtivos, ou outra destinação final ambientalmente adequada. Ou seja, a lógica reversa tem como característica um conjunto de ações destinados a viabilizar a coleta e a restituição dos resíduos sólidos ao setor empresarial, para reaproveitamento, em seu ciclo ou em outros ciclos produtivos, ou outra destinação final ambiental adequada.

A preocupação ambiental tornou-se um diferencial competitivo pois as empresas são bem vistas no mercado, e passam a ser valorizados pelos clientes gerando economia e confiança nos investidores. As empresas passam a não se preocupar apenas com os ganhos mas também com respeito aos fatores ambientais e sociais envolvidos em todo processo em que esta inserida. As empresas interessadas em empregar a sustentabilidade empresarial no seu contexto realizam práticas relacionadas à otimização do uso dos recursos ambientais em suas atividades e a diminuição do seu impacto no meio ambiente. Quase metade das cidades brasileiras não tem plano para resíduos sólidos, diz o

IBGE. Quase metade das 5.570 cidades brasileiras não tem atualmente um plano integrado para o manejo do lixo. É importante para as prefeituras terem um plano de resíduos sólidos, assim as mesmas recebem recursos, financiamentos e incentivos da união para a área.

Foi criado o Instituto Brasileiro de Energia Reciclável – IBER, conforme item 7.3 do Termo de Compromisso firmado com o Poder Público Estadual, para conduzir o Sistema de logística reversa de baterias chumbo-ácido, na condição de entidade gestora. O IBGE é uma associação civil sem fins econômicos, cuja estrutura foi criada com base em seu Estatuto e Regimento interno, para apoiar a implementação da logística reversa, conforme as exigências legais.

Este trabalho fará um estudo dos processos envolvidos na logística reversa das Baterias Moura, uma empresa pernambucana que atua no território sul-americano. A Baterias Moura acredita na importância de exercer um papel positivo na sociedade na qual está inserida. Por isso a Moura conduz uma série de ações socioambientais, sempre voltadas para o desenvolvimento de sua cidade sede, Belo Jardim, e do planeta. A Moura promove a reciclagem de 100% das baterias produzidas e vendidas no mercado nacional. O programa Moura Logística Reversa, é praticado há mais de 35 anos, fechando o ciclo das reciclagens das baterias, garantindo que as sucatas terão uma destinação ecologicamente correta.

## **REFERENCIAL TEÓRICO**

### **Logística Reversa**

O bom desenvolvimento da logística é de grande importância para as empresas, pois além de favorecê-las com vantagem competitiva, reduz os custos e aumenta a possibilidade de se manterem no mercado. De acordo com Rogers e Tibben-Lembke (1999) logística reversa é “O processo de planejamento, implementação e controle do fluxo eficiente e de baixo custo de matérias primas, estoque em processo, produto acabado e informações relacionadas, desde o ponto de consumo até o ponto de origem, com o propósito de recuperação de valor ou descarte apropriado para coleta e tratamento de lixo”. A logística reversa envolve o recolhimento do produto usado, seu armazenamento correto, transporte e reciclagem.

Para Lora (2000) as organizações estão tomando uma posição ambiental ativa, transformando uma atitude em oportunidade de negócio, ou seja, os encargos legais não são mais os únicos motivadores para as empresas investirem

em sustentabilidade, mas o fato de que essas atitudes acarretam redução de custos para a empresa e diminuem os impactos negativos no meio ambiente.

Existem algumas diferenças entre os fluxos da logística normal e da logística reversa. Podendo ser comprovada essas diferenças com uma afirmativa de Krikke (1998, p. 154) “A primeira diferença é que a logística tradicional à frente é um sistema onde os produtos são puxados (“pull system”), enquanto que na Logística Reversa existe uma combinação entre puxar e empurrar os produtos pela cadeia de suprimentos.[...] Como resultado de uma legislação mais restritiva e a maior responsabilidade do produtor, na Logística Reversa, a quantidade de lixo produzido (e a distinção entre o que é reciclável do que é lixo indesejado) não pode ser influenciada pelo produtor e deverá ser igualada à demanda de produtos, já que a quantidade de descarte já é limitada em muitos países. “Em segundo lugar, os fluxos tradicionais de logística são basicamente divergentes, enquanto que os fluxos reversos podem ser fortemente convergentes e divergentes ao mesmo tempo. “Terceiro, os fluxos de retorno seguem um diagrama de processamento pré-definido, no qual produtos descartados são transformados em produtos secundários, componentes e materiais. No fluxo normal, esta transformação acontece em uma unidade de produção, que serve como fornecedora da rede. “Por último, na Logística Reversa, os processos de transformação tendem a ser incorporados na rede de distribuição, cobrindo todo o processo de ‘produção’, da oferta (descarte) à demanda (reutilização)”. Essa afirmação prova que apesar do processo logístico em si, ser complexo, é de mastodôntica importância para as organizações.

## **Sustentabilidade**

Sustentabilidade é um termo muito usado quando se refere à preservação da natureza. O site Brasil Escola, define sustentabilidade como o ato de prover as necessidades da geração atual, sem comprometer a geração futura e suas necessidades. Essa ação visa proteger a natureza, sabendo que ela é um recurso limitado, ou seja, se tirar e não colocar pode acabar. Para Veenan e Polytilo (2003), sustentabilidade é difícil conceituar, pois eles consideram a sustentabilidade grande demais para ser medida.

O alto índice de problemas ambientais, fez com que muitas empresas se mobilizassem a favor da causa sustentável. Para Leite (2002) “A obsolescência e a descartabilidade crescentes dos produtos observados nesta última década têm-se refletido em alterações estratégicas empresariais, dentro da própria organização e principalmente em todos os elos de sua rede operacional”. A obsolescência faz com que o produto dure menos e precise ser descartado mais rapidamente, é nesse sentido que entra a descartabilidade, onde esse produto que será substituído por um novo, será jogado no lixo, aumentando

consideravelmente esse lixo. O consumo exagerado incide o ser humano a levar uma desordenada onde ele vive para consumir e consome para viver. Mediante essa situação, como Leite (2002) já afirmou, muitas empresas alteraram as estratégias de marketing, visando diminuir esse lixo sem prejudicar nos rendimentos da mesma.

O movimento sustentável garante a geração atual e a geração futura o direito de desenvolvimento, pois já foi e é debatido que se onde se tira e não repõe, acaba, e a natureza é um recurso limitado, por mais abundante que seja, se não for cuidada pode acabar, e de onde a sociedade tirará matéria prima para sobreviver? Por tal motivo que os órgãos competentes se mobilizaram a se unir junto a empresários e a população com programas de reciclagem, diminuição do lixo e preservação da natureza.

De acordo com Silva (2003), o desenvolvimento sustentável sob o ponto de vista empresarial pode ser definido como: “A busca do equilíbrio entre o que é socialmente desejável, economicamente viável e ecologicamente sustentável”, ou seja, os benefícios são mútuos quando se busca os mesmos objetivos.

### **Política Nacional de Resíduos Sólidos**

De acordo com a Lei nº 12.305/10, ela institui a Política Nacional de Resíduos Sólidos, dispondo sobre seus princípios, objetivos e instrumentos, bem como sobre as diretrizes relativas à gestão integrada e ao gerenciamento de resíduos sólidos, incluídos os perigosos, às responsabilidades dos geradores e do poder público e aos instrumentos econômicos aplicáveis. Onde a prevenção e a precaução é não geração, redução, reutilização, reciclagem e tratamento dos resíduos sólidos, bem como disposição final ambientalmente adequada dos rejeitos, e estímulo à adoção de padrões sustentáveis de produção e consumo de bens e serviços.

Existem 15 objetivos na PNRS, uma delas é proteção da saúde pública e da qualidade ambiental; onde é fundamental definir e implementar políticas públicas adequadas com vistas a garantir a destinação adequada dos resíduos sólido. Vale ressaltar que o poder público não é o único responsável pela cadeia dos resíduos sólidos. Os consumidores, os produtores e os fabricantes, entre outros atores sociais, também precisam se engajar e ter uma participação efetiva no processo. Com isso todos envolvidos devem ter uma responsabilidade a mais com seu

produto em relação à sustentabilidade e criar meios para que seus consumidores também faça parte dessa ação.

Onde também a lei determina que as medidas de logística reversa devem se estender a produtos comercializados em embalagens plásticas, metálicas ou de vidro, e aos demais produtos e embalagens, considerando, prioritariamente, o grau e a extensão do impacto à saúde pública e ao meio ambiente dos resíduos gerados. Ou seja, as empresas devem se preocupar em saber qual será a destinação que o usuário final deu ao seu produto após ser consumido e oferecer opções para reaproveitá-lo em suas cadeias produtivas ou destiná-lo corretamente.

A PNRS tem um grande potencial de reconfigurar as práticas de gestão, o resíduo é um potencial insumo para as cadeias produtivas. Onde passa pela, redução ,reaproveitamento, reciclagem, tratamento e a destinação dos rejeitos.

A PNRS oportuniza alterações substanciais nos padrões hegemônicos de encaminhamento político e econômico de gestão de resíduos por parte dos municípios. Tais padrões hegemônicos, além de inviabilizarem os processos de reciclagem, colocam o objeto de trabalho dos catadores sob disputa com grandes empresas de destinação de resíduos. Não obstante, os espaços por excelência para os diálogos sobre a construção de um modelo de gestão integrada de resíduos consistem em esferas públicas que proporcionem o controle social sobre os processos de manejo de resíduos e sobre seus protagonistas. Além disso, a mudança de paradigmas de gestão de resíduos é uma missão a ser perseguida para a efetivação da PNRS.(Catadores de materiais recicláveis: um encontro nacional, Publicado em 31 de março de 2016 por Erica Sepulveda, pág 241).

A Baterias Moura passa por um processo de recolhimento que a empresa realiza através de seus distribuidores e revendedores, ondem a distribuidora recolhe o material e é levado para a fábrica onde ocorre o processo de reciclagem de bateria que é dividido em trituração da sucata de bateria com separação do plástico; reciclagem propriamente dita deste plástico e recuperação de chumbo ligado. O chumbo segue o processo de: separação, fundição, refino, lingoteamento até a fabricação de novas baterias. O plástico é recuperado e reutilizado na produção de caixas e tampas de novas baterias.

## **METODOLOGIA**

A primeira etapa dessa pesquisa consistiu em realizar a introdução para darmos ênfase ao assunto em questão, obtida através de pesquisa de análise documental na internet e artigos. A partir dos dados obtidos a pesquisa científica foi realizada de forma sistemática, foi utilizado para isso meios estratégicos e técnicos procurando entender e interpretar as informações obtidas para o objeto

de estudo método de pesquisa adotado em circunstâncias mostra-se ser o mais adequado ao problema de pesquisa.

Pesquisa de caráter qualitativo, realizada de forma descritiva e exploratória, foi realizada através de revisão bibliográfica afim de analisar o entendimento da logística reversa e seus processos. Segundo Sueli Liberatti (2011) podemos definir pesquisa qualitativa como uma forma de se fazer pesquisa, no qual o foco, o olhar da pesquisa, encontra-se nas relações que tem significado para o pesquisador.

Quanto aos procedimentos metodológicos a pesquisa caracteriza-se como pesquisa de levantamento, realizado através de um questionário com o intuito de obter informações necessárias para análise pretendida, assim, foi encaminhado um e-mail, o qual obtivemos retorno da empresa Baterias Moura, prospectando informações de como é feito a logística reversa da empresa, e a partir dele foi verificado os processos logísticos e de que forma a empresa trabalha para garantir seu papel de acordo com a legislação, preservando o meio ambiente.

O motivo pelo qual escolhemos trabalhar com a empresa Baterias Moura, se deu através da forma pela qual a empresa é reconhecida, e também por suas políticas que prezam pela sustentabilidade, a mesma é conhecida por prezar e conscientizar a preservação do meio ambiente, evitando assim os gastos desnecessários e a poluição na natureza.

## **ANÁLISE DOS RESULTADOS**

### **Questões sobre a logística da empresa Baterias Moura**

#### **1.O que mais motiva a Moura, no que se refere ao desenvolvimento da logística reversa?**

A Baterias Moura acredita na importância de exercer um papel positivo na sociedade na qual está inserida. Por tal motivo conduz uma série de ações socioambientais, sempre voltadas para o desenvolvimento de sua cidade sede, Belo Jardim, e do planeta.

#### **2.Quais os benefícios do programa de coleta de baterias pós-consumo?**

Com o programa Moura de Logística Reversa, praticado há mais de 35 anos, fechamos o ciclo de reciclagem das baterias, garantimos que as sucatas terão

uma destinação ecologicamente correta, evitamos que toneladas de chumbo e ácido sejam lançados no meio ambiente contaminando o solo, água dos rios, lençol freático, ar e contribuimos para a redução do impacto ambiental em decorrência da extração das reservas minerais.

**3. Como faz para que as baterias automotivas voltem para a empresa para serem recicladas?**

As baterias automotivas inservíveis devem ser encaminhadas aos fabricantes que, por sua vez, são responsáveis pela reciclagem. Esta garante que seus componentes perigosos (metais e ácido) fiquem afastados de aterros e de incineradores de lixo urbano e que o material recuperado possa ser utilizado na produção de novos bens de consumo. Todos os constituintes de uma bateria chumbo ácido apresentam potencial para reciclagem.

**4. Como se dá o armazenamento das baterias depois de usadas?**

A armazenagem de baterias inservíveis de chumbo ácido deverá ser feita local coberto, com piso apropriado (concreto), com canaletas de contenção.

**5. Como se dá o transporte dessas baterias?**

O transporte de baterias inservíveis de chumbo ácido deve estar de acordo com a Resolução ANTT nº 420/2004, que trata do transporte rodoviário de produtos perigosos. Conduzir as baterias com a tampa para cima para evitar vazamento da solução.

**6. Qual a visão da empresa com relação ao descarte inadequado das baterias?**

Uma bateria que tenha sido imprópriamente disposta, ou seja, não reciclada, representa uma importante perda de recursos econômicos, ambientais e energéticos e imposição de um risco desnecessário ao meio ambiente e seus ocupantes.

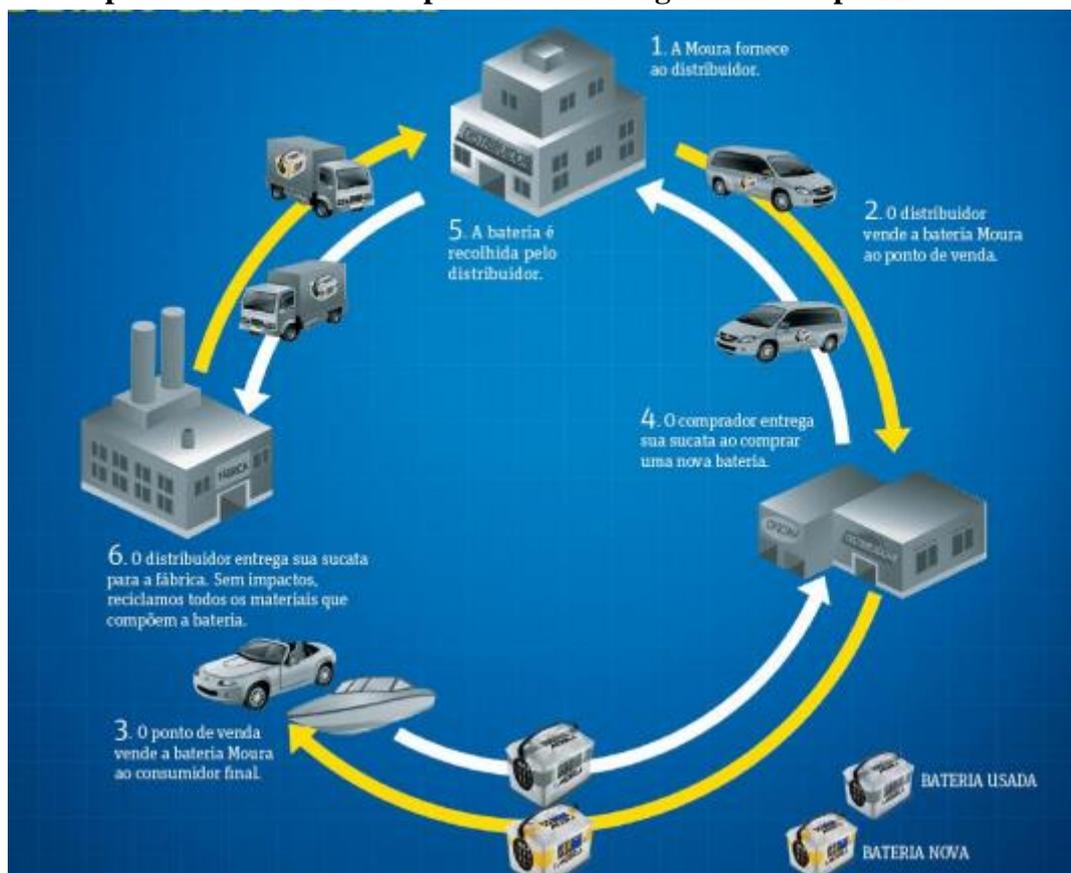
**7. Como se dá o processo de reciclagem?**

A Moura dispõe de tecnologia apropriada, com certificação de comprovação, para efetuar de forma correta a reciclagem da sucata. O processo de reciclagem de bateria é dividido em trituração da sucata de bateria com separação do plástico; reciclagem propriamente dita deste plástico e recuperação de chumbo ligado. O chumbo segue o processo de: separação, fundição, refino, lingoteamento até a fabricação de novas baterias. O plástico é recuperado e reutilizado na produção de caixas e tampas de novas baterias.

**8. Como são distribuídos os pontos de coleta?**

Distribuidores e pontos de venda estão devidamente habilitados para receber as baterias inservíveis de qualquer marca. Com foco no cliente, a Moura mantém uma Rede de Distribuição formada por mais de 80 unidades situadas nos maiores centros de consumo do Brasil.

**9. Como explicar de uma maneira prática o fluxo logístico da empresa?**



**10. Quais as metas desse programa?**

Os signatários do Termo de compromisso se obrigam a coletar a quantidade de Baterias Inservíveis conforme as metas definidas com o setor público. Os percentuais abaixo serão apurados mensalmente utilizarão o seguinte critério:

$$\% = \frac{\text{Baterias coletadas pelos signatários durante o ano (kg)}}{\text{Baterias comercializadas pelos signatários durante o ano (kg)}}$$

Este percentual é definido pela proporção entre baterias inservíveis coletadas progressivamente ao longo dos anos pelo número de baterias chumbo ácido que forem comercializadas no mercado. Seguem abaixo as metas de recolhimento ano a ano:

ANO	I	II	III	IV
Percentual de baterias inservíveis de chumbo ácido recolhidas, no País, por volume gerado (em quilogramas).	75%	80%	85%	90%
Percentual de abrangência do Sistema com relação aos Municípios do País (Estabelecimentos de venda se constituem pontos de coleta).	100%	100%	100%	100%

O percentual de abrangência do Sistema com relação aos Municípios do País, se refere à quantidade de municípios que possuem estabelecimentos de venda os quais se constituem pontos de coleta.

## CONCLUSÃO

Diante da pesquisa realizada da empresa, através das respostas tiradas do documento, vemos que a empresa feita atrás do documento, a empresa tem o comprometimento das normas exigidas no âmbito legal. A Moura sempre visa obter constantes melhorias no sistema ambiental da empresa, desenvolvendo projetos relacionado ao meio ambiente e o que reitera sua preocupação ambiental. Líder em baterias na América do Sul e certificada desde 2004 com a norma ISO 14001, que reflete o direcionamento da empresa para processos sustentáveis, a Moura possui uma diretoria de sustentabilidade, que tem como objetivo identificar oportunidades de melhoria e aperfeiçoar o sistema ambiental da empresa, gerindo projetos referentes ao contexto do meio ambiente. Atualmente, a Moura possui um importante programa de logística reversa, que possibilita reciclar mais de 100% das baterias produzidas.

Além da preocupação ambiental, existe a questão da redução de custo, também apontada como uma das principais vantagens de se implementar um processo de logística reversa em uma empresa. Concluindo, foi possível observar que a Baterias Moura é uma organização responsável, que atua de acordo com as exigências ambientais do Brasil, e que realiza além do que é de sua obrigação, reciclando e reutilizando mais que 100% das baterias que produz. Respeitando a Política Ambiental de Resíduos Sólidos e a Resolução do CONAMA 401 de 2008.

## REFERÊNCIAS

G1 GLOBO. Disponível em: <https://www.google.com.br/amp/s/g1.globo.com/google/amp/economia/noticia/quase-metade-das-cidades-brasileiras-nao-tem-plano-para-residuos-solidos-diz-ibge.ghtml> Acesso em 09 de outubro de 2018.

CIDADES SUSTENTÁVEIS. Disponível em: <http://www.mma.gov.br/cidades-sustentaveis/residuos-perigosos/logistica-reversa> Acesso em 09 de outubro de 2018.

LOGÍSTICA REVERSA SENAI. Disponível em: <http://www.ms.senai.br/programas/logistica-reversa/> Acesso em 09 de outubro de 2018.

KRIKKE, H. 1998, Recovery Strategies and Reverse Logistics Network Design – Holanda: BETA – Institute for Business Engineering and Technology Application.

ROGERS, D S. e TIBBEN-LEMBKE, R S. 1999, Going Backwards: Reverse Logistics Trends and Practices. University of Nevada, Reno - Center for Logistics Management, in <http://equinox.unr.edu/homepage/logis/reverse.pdf> Acesso em 10 de outubro de 2018.

LORA, E. Prevenção e controle da poluição no setor energético industrial e transporte. Brasília: ANEEL, 2000.

O QUE É SUSTENTABILIDADE? BRASIL ESCOLA. Disponível em: <https://brasilecola.uol.com.br/o-que-e/quimica/o-que-e-sustentabilidade.htm> Acesso em 28 de setembro de 2018.

VEEMAN, T. S.; POLYTILO, J. The role of institutions in policy in enhancing sustainable development and conserving natural capital. Environment Development and Sustainability, Netherlands, v. 5, n. 3-4, p. 317-332, September 2003.

LEITE, Paulo Roberto. REVISTA TECNOLÓGICA – MAIO / 2002. SÃO PAULO, EDIT. PUBLICARE Disponível em: <http://meusite.mackenzie.br/leitepr/LOG%20CDSTICA%20REVERSA%20-%20NOVA%20%20C1REA%20DA%20LOG%20CDSTICA%20EMPRESARIAL.pdf> Acesso em 28 de setembro de 2018.

SILVA. V. G. Avaliação da sustentabilidade de edifícios de escritórios brasileiros: diretrizes e base metodológica. 2003. 210 f.. Tese (Doutorado em Engenharia Civil) Departamento de Engenharia de Construção Civil, Escola Politécnica da Universidade DE SÃO PAULO, SÃO PAULO, 2003.

POLÍTICA NACIONAL DE RESÍDUOS SÓLIDOS. Disponível em: <http://www.mma.gov.br/pol%C3%ADtica-de-res%C3%ADuos-s%C3%B3lidos> Acesso em 09 de outubro de 2018.

O QUE É POLÍTICA NACIONAL DE RESÍDUOS SÓLIDOS (PNRS)? Disponível em: <https://www.ecycle.com.br/component/content/article/67-dia-a-dia/3705-o-que-e-politica-nacional-de-residuos-solidos-pnrs-urbanos-descartedanos-saude-meio-ambiente-qualidade-vida-reciclagem-consumo-instrumento-responsabilidade-produto-metas-lixoes.html> Acesso em 09 de outubro de 2018.

BATERIAS MOURA. Disponível em: <https://www.moura.com.br/> Acesso em 10 de outubro de 2018.

CATADORES DE MATERIAIS RECICLÁVEIS : UM ENCONTRO NACIONAL / BRUNA CRISTINA JAQUETTO PEREIRA, FERNANDA LIRA GOES (ORGANIZADORAS) – Rio de Janeiro : Ipea, 2016. 562. p. : il., mapas color. Disponível em: [http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/livros/160331\\_livro\\_catadores.pdf](http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/livros/160331_livro_catadores.pdf) Acesso em 10 de outubro de 2018.

## Desenvolvimento de uma arquitetura para sistemas *e-health* baseada em IoT, fog e cloud

Kayo Monteiro<sup>1</sup>, Élisson Rocha<sup>1</sup>, Émerson Silva<sup>1</sup>, Guto Leoni Santos<sup>2</sup>,  
Wylliams Santos<sup>1</sup>, Patricia Takako Endo<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Engenharia de Computação  
Universidade de Pernambuco (UPE), Recife, Brasil

<sup>2</sup>Universidade Federal de Pernambuco (UFPE), Recife, Brasil

{khcm, esr2, ers2}@ecomp.poli.br, {wbs, patricia.endo}@upe.br

{guto.leoni}@gprt.ufpe.br

**Resumo.** *Estima-se que a expectativa de vida da população mundial, até 2030, seja de 90 anos, em média. As comunidades tem dedicado esforços no desenvolvimento de aplicações que promovam qualidade de vida para os idosos. Assim, serviços de monitoramento de sinais vitais e sistemas de detecção de quedas estão em evidência. Uma parte dos esforços focam nestes serviços utilizando IoT (Internet of Thing). No entanto, devido ao alto fluxo de dados, dispositivos IoT não são capazes de processar, armazenar e garantir os SLAs (Service Level Agreements) destes serviços. Neste contexto, a IoT possui dois grandes aliados: fog e cloud. Este artigo apresenta uma proposta de arquitetura para sistemas e-health utilizando IoT, fog e cloud, e também as características para construção de um protótipo.*

### 1. Introdução

O aumento da população idosa em todo o mundo é uma realidade. De acordo com a Organização Mundial de Saúde (OMS)<sup>1</sup>, até 2020, pela primeira vez na história, o número de pessoas idosas será maior que o de crianças com até cinco anos; e até 80% dos idosos viverão em países de baixa e média renda. Isto desafia os sistemas atuais de saúde. A OMS aborda que embora as pessoas estejam vivendo mais, não significa que estão mais saudáveis que antes. Uma população mais velha é sinônimo de uma população que necessita de cuidado e acompanhamento constante.

Durante assembleia geral a Organização das Nações Unidas (ONU)<sup>2</sup>, o tema de cuidados com os idosos foi colocado em pauta como um dos desafios globais de saúde. Este fato desafia os atuais sistemas de saúde, porque os idosos têm maior predisposição para apresentar doenças crônicas, sofrer algum tipo de deficiência ou limitações físicas, necessitando de cuidados e monitoramento constantes. Por exemplo, ainda segundo a OMS, as estimativas mais recentes indicam o aumento do número de idosos com demência para 135 milhões até 2050, o acompanhamento em tempo integral feito por um profissional de saúde, por muitas vezes, está fora da realidade financeira do paciente ou dos seus familiares levando em consideração que até 80% viverão em países de baixa e média renda.

<sup>1</sup>[https://news.un.org/en/story/2014/11/483012#.VFyq6\\_nF-z4](https://news.un.org/en/story/2014/11/483012#.VFyq6_nF-z4)

<sup>2</sup>[http://www.un.org/en/ga/search/view\\_doc.asp?symbol=A/RES/67/81](http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/67/81)

A utilização de dispositivos IoT é uma das principais alternativas para fazer o acompanhamento de idosos através de sistemas para cuidados de saúde mais acessíveis. Através de dispositivos IoT, os dados vitais podem ser monitorados e enviados para um sistema de processamento e armazenamento disponível na *cloud*, possibilitando que médicos, enfermeiros e familiares consultem os dados do paciente a qualquer momento. Em caso de alguma anormalidade, o sistema gera uma notificação em tempo real; essa notificação só é possível através do processamento garantido pela *fog* [Ghosh et al. ] [Corotinschi and Găitan ]. O aumento da demanda de cuidados nestes pacientes, alinhado com o controle das despesas, fomenta a necessidade de uso de tecnologias de baixo custo. Uma alternativa são *smartphones* e dispositivos IoT, pela popularidade e custo em relação a dispositivos médicos especializados.

Em 2017, estima-se que cerca de 1,57 bilhões de *smartphones* e 310,4 milhões de dispositivos móveis foram vendidos em todo o mundo, representando um aumento de 16,7% em relação a 2016 [Gartner 2018]. Com estes dispositivos IoT é possível coletar diferentes sinais vitais, tais como: frequência cardíaca, temperatura corporal, pressão arterial e se o usuário sofreu algum acidente, como queda. Contudo, estes dispositivos IoT têm capacidade computacional e de armazenamento limitadas. Para solucionar este problema, enviam-se os dados coletados para a *fog* ou *cloud*, os quais são processados e armazenados. Neste trabalho, apresenta-se uma arquitetura de sistema *e-health* baseado em *Internet of Things* (IoT) integrado com *fog* e *cloud*, focando em uma solução de baixo custo para pacientes, cuidadores e familiares.

## 2. Arquitetura

A arquitetura proposta para sistemas de *e-health* baseado em dispositivos IoT utilizando *fog* e *cloud* está representada na Figura 1.

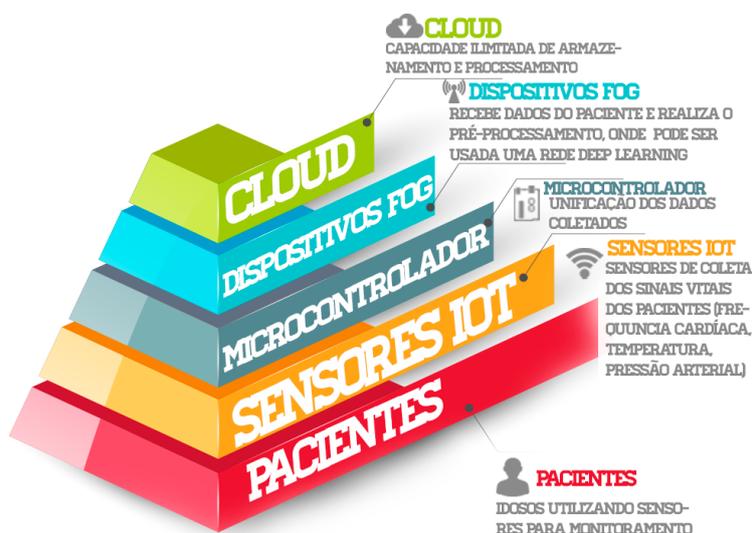


Figura 1. Proposta: Arquitetura para Sistemas de *E-health*

Os sensores IoT coletam os sinais vitais do paciente continuamente e enviam dados para os microcontroladores, assim, pode-se unificar os dados do paciente. Em seguida, os microcontroladores enviam estes dados para pré-processamento na *fog*, garantindo a rápida resposta do sistema no caso de identificação de alguma anormalidade. Por

fim, a *fog* envia os dados do paciente para a *cloud*, no intuito de catalogar e armazenar para análise posterior.

Segundo Miotto [Miotto et al. 2017], a *fog* deve garantir uma resposta rápida do sistema. Portanto, na arquitetura proposta, a camada da *fog* precisa de uma tecnologia implementada para pré-processamento de dados descrita na subseção 2.1. Inicialmente, este trabalho propõe utilizar *deep learning* (DL), devido ao seu potencial em muitas áreas, incluindo a biomédica [Miotto et al. 2017], sendo usado, por exemplo, para detectar quedas humanas, arritmia cardíaca ou ataques de epilepsia.

## 2.1. Protótipo

Baseado na arquitetura apresentada na Figura 1, para construir um protótipo da arquitetura proposta é preciso seguir os seguintes requisitos: (a) Os pacientes devem usar sensores para monitorar sinais vitais. Estes sensores são *smartphones* ou dispositivos *wearable*;<sup>3</sup> (b) a integração de dados é feita por um microcontrolador. Propõe-se o uso de Arduino<sup>4</sup>, com interface para padrões de comunicação com sensores e *smartphones*; (c) na *fog*, os dados padronizados na camada anterior são classificados (por exemplo, por um modelo DL). Para este pré-processamento, utiliza-se um dispositivo com capacidade computacional média para garantir que esta camada não se torne o gargalo do sistema. O Raspberry<sup>5</sup> Pi 3 supre essa necessidade e possui preço acessível; e (d) a *cloud* armazena os dados e realiza análises em larga escala. Para o protótipo, uma *cloud* pública da Amazon<sup>6</sup> é adotada, devido a facilidade de gerenciamento e alta disponibilidade.

## 3. Próximos Passos

Pretende-se implementar o protótipo de baixo custo para realizar comparações com outras arquiteturas da literatura. Também planeja-se desenvolver modelos estocásticos para estimar disponibilidade destes sistemas, assim, encontrar os principais gargalos.

## Agradecimentos

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro desta pesquisa.

## Referências

- Corotinschi, G. and Găitan, V. G. The use of iot technologies for providing high-quality medical services. In *System Theory, Control and Computing (ICSTCC)*.
- Gartner, I. (2018). What will be the most coveted devices for this holiday season? <https://goo.gl/v5dUPv>, note = Accessed: 2018-09-27.
- Ghosh, A. M., Halder, D., and Hossain, S. A. Remote health monitoring system through iot. In *Conference on Informatics, Electronics and Vision (ICIEV)*.
- Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*.

<sup>3</sup><https://www.dfrobot.com/product-1540.html/>

<sup>4</sup><https://www.arduino.cc/>

<sup>5</sup><https://www.raspberrypi.org/>

<sup>6</sup><https://aws.amazon.com/>

## Proposta de uma Taxonomia para Criação de Valor em Desenvolvimento Ágil de Software

Geraldo Torres<sup>1,2</sup>, Wylliams Santos<sup>1,2</sup>, Roberta Fagundes<sup>2</sup>

<sup>1</sup>REACT Research Labs, Caruaru – PE – Brasil  
Universidade de Pernambuco (UPE), Caruaru – PE – Brasil

{gtgn}@ecom.poli.br

{wbs, roberta.fagundes}@upe.br

**Resumo.** Apesar da criação de valor ser enfatizada no manifesto ágil, pouco se sabe como de fato o valor é criado em Desenvolvimento Ágil de Software (DAS). Este estudo tem como objetivo exibir os primeiros passos rumo a criação de uma taxonomia para apontar as práticas ágeis que podem entregar valor em DAS.

### 1. Introdução

As pesquisas envolvendo o DAS estão passando por um processo de transição. O foco dos estudos está sendo direcionado para o valor do produto que está sendo desenvolvido, ao invés de salientar de forma majoritária os aspectos do processo de desenvolvimento [Dingsøyr e Lassenius 2016]. [Bloch 2012] afirmou que no ano de 2012 em média 45% dos grandes projetos de Tecnologia da Informação (TI) foram executados além do orçamento, enquanto fornecem 56% menos valor do que o previsto.

Os autores [Gregory et al. 2016] identificaram vários desafios na perspectiva de profissionais que trabalham com DAS, um dos desafios relatados foi o "valor de negócio" (*Business Value*). Dados do último relatório da [VERSIONONE 2018] apresentam o *business value* como terceiro indicador de sucesso mais mencionado pelos entrevistados, obtendo 53% das respostas no que diz respeito ao sucesso da iniciativa ágil. Para medir o sucesso dos projetos ágeis o valor entregue pelo projeto foi a segunda métrica de sucesso mais indicada pelos participantes da pesquisa, obtendo 43% das respostas. No ano anterior esse número correspondeu a apenas 23% das respostas, em apenas um ano esse número quase duplicou o que mostra a relevância do tema.

### 2. Valor em Desenvolvimento Ágil de Software

O valor ainda não é definida de forma clara no DAS vários estudos definem a criação de valor de formas distintas. Em seu estudo [Racheva et al. 2009] realizou uma Revisão Sistemática da Literatura (RSL) para identificar como o valor é criado em DAS. Entretanto os autores não identificaram com clareza como o valor é definido.

Os autores [Huijgens et al. 2016] definiram o valor em quatro perspectivas: cliente, processo interno, inovação e perspectiva financeira.

Em seu trabalho [Alahyari et al. 2017] identificou 16 aspectos de valor em organizações de desenvolvimento de software. [Conboy 2009] definiu que o valor pode ser definido nos aspectos de economia, qualidade e simplicidade. Pode-se observar a diversidade de definições apresentadas em diferentes estudos.

Dessa forma, este estudo tem como objetivo desenvolver uma taxonomia para classificar a criação de valor em DAS. Em seu estudo [Conboy 2009] criou uma taxonomia para verificar que práticas ágeis possuem agilidade. De forma complementar, [Usman et al. 2017] afirmam que muitas áreas de conhecimento da engenharia de software ainda precisam ser classificadas, um meio para fazer esta classificação pode ser feito utilizando as taxonomias. Portanto, a taxonomia proposta neste estudo tem como objetivo verificar que práticas ágeis podem criar valor em DAS.

### 3. Abordagem de pesquisa

Para o desenvolvimento da taxonomia, este estudo irá utilizar o método proposto por [Usman et al. 2017] exibido na Figura 1. A fase de planejamento é composta de seis etapas. Na primeira etapa deve ser definida a área de conhecimento, nesta etapa já está sendo desenvolvida uma RSL sobre a criação de valor em DAS. O escopo da taxonomia será a criação de valor em DAS. O foco da taxonomia será identificar que práticas podem criar valor em DAS. Por se tratar de uma pesquisa em andamento, as etapas quatro a sete da fase de planejamento ainda não foram definidas, bem como as fases dois a quatro.

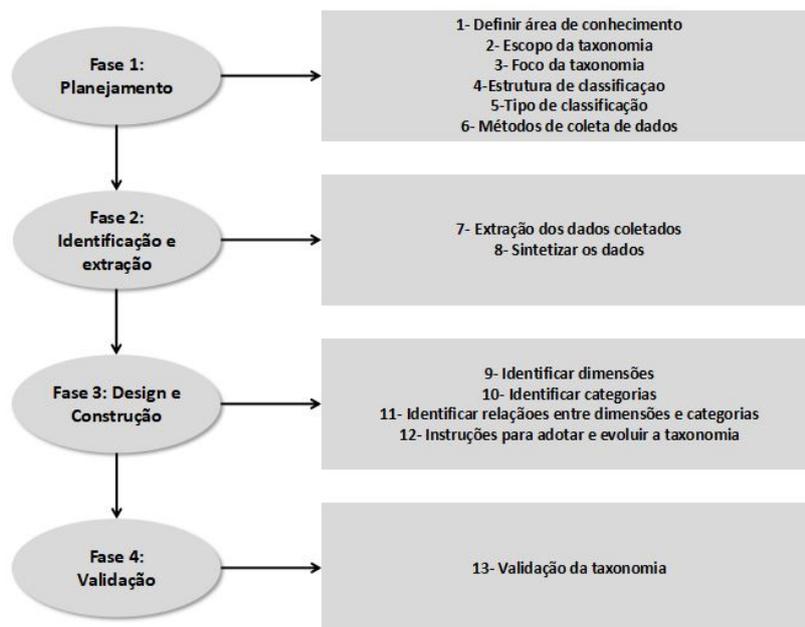


Figura 1. Método proposto por [Usman et al. 2017] para desenvolvimento de taxonomias em Engenharia de Software

### 4. Resultados Parciais e Discussões

Até o presente momento a fase de planejamento da taxonomia está sendo executada. (Figura 1). Nesta fase uma RSL está sendo conduzida, as bases de dados do *IEEE Digital Library (IEEEExplore)* e *ACM Digital Library* foram analisadas. Um número de 224 estudos foram examinados. Apenas três estudos identificados na RSL conceituam a criação de valor em DAS, outros 14 estudos foram selecionados porém identificando práticas e fatores que levam a criação de valor em DAS. O estudo de [Sobiech et al. 2015] conceitua o valor para histórias de usuário, são identificadas sete dimensões de valor. O trabalho

de [Dong et al. 2011] afirma que o valor comercial fornecido dentro de uma iteração ou sprint é composto pela soma do valor de todas as histórias de usuários completas. A pesquisa dos autores [Jonsson et al. 2013] sintetiza o conceito de valor através de uma RSL sobre conceitos de *Lean Software Development*. A partir dos resultados preliminares desta RSL, reafirmamos os resultados obtidos por [Racheva et al. 2009] de que a definição de valor em DAS ainda é desafiadora. Porém, esperamos que ao final da fase de planejamento da taxonomia os conhecimentos obtidos da literatura, juntamente com coleta de dados em organizações de desenvolvimento de software ágil na etapa seis possam fornecer uma visão abrangente do conceito de valor em DAS.

## Referências

- Alahyari, H., Svensson, R. B., e Gorschek, T. (2017). A study of value in agile software development organizations. *Journal of Systems and Software*, 125:271 – 288.
- Bloch, M., B. S. L. J. (2012). Delivering large-scale it projects on time, onbudget, and on value.
- Conboy, K. (2009). Agility from first principles: Reconstructing the concept of agility in information systems development. *Info. Sys. Research*, 20(3):329–354.
- Dingsøyr, T. e Lassenius, C. (2016). Emerging themes in agile software development: Introduction to the special section on continuous value delivery. *Information and Software Technology*, 77:56 – 60.
- Dong, X., Yang, Q., Wang, Q., Zhai, J., e Ruhe, G. (2011). Value-risk trade-off analysis for iteration planning in extreme programming. In *2011 18th Asia-Pacific Software Engineering Conference*, pages 397–404.
- Gregory, P., Barroca, L., Sharp, H., Deshpande, A., e Taylor, K. (2016). The challenges that challenge: Engaging with agile practitioners' concerns. 75.
- Huijgens, H., van Deursen, A., e van Solingen, R. (2016). An exploratory study on the effects of perceived value and stakeholder satisfaction on software projects. In *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*, EASE '16, pages 31:1–31:10, New York, NY, USA. ACM.
- Jonsson, H., Larsson, S., e Punnekkat, S. (2013). Synthesizing a comprehensive framework for lean software development. In *2013 39th Euromicro Conference on Software Engineering and Advanced Applications*, pages 1–8.
- Racheva, Z., Daneva, M., e Sikkel, K. (2009). Value creation by agile projects: Methodology or mystery? In Bomarius, F., Oivo, M., Jaring, P., e Abrahamsson, P., editors, *Product-Focused Software Process Improvement*, pages 141–155, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Sobiech, F., Eilermann, B., e Rausch, A. (2015). A heuristic approach to solve the elementary sprint optimization problem for non-cross-functional teams in scrum. *SIGAPP Appl. Comput. Rev.*, 14(4):19–26.
- Usman, M., Britto, R., Börstler, J., e Mendes, E. (2017). Taxonomies in software engineering: A systematic mapping study and a revised taxonomy development method. *Information and Software Technology*, 85:43 – 59.
- VERSIONONE (2018). 12th annual state of agile report.

## Aplicação de *Learning Analytics* nas Instituições de Ensino Superior: Um Protocolo de Revisão Sistemática de Literatura

Thomás Tabosa de Oliveira<sup>1</sup>, Elyda Laisa Soares Xavier Freitas<sup>1</sup>,  
Vinicius Cardoso Garcia<sup>2</sup>, Fernando da Fonseca de Souza<sup>2</sup>

<sup>1</sup>Instituto de Informática – Universidade de Pernambuco (UPE)  
– Caruaru – PE – Brasil,

<sup>2</sup>Centro de Informática (CIn) – Universidade Federal de Pernambuco (UFPE)  
Recife – PE – Brasil

{thomas.oliveira, elyda.freitas}@upe.br, {vcg, fdfd}@cin.ufpe.br

**Abstract.** *Learning Analytics (LA) arose from the need to improve teaching, to understand the best approach to learning. Currently, LA is one of the most widely used means of analyzing student information. Despite this, preliminary studies expose some lack of practical information on the subject. The goal of the project is to find practical results of how LA was applied. This article presents the preliminary results of the ongoing research, describing its Protocol of Systematic Review of Literature.*

**KEY WORDS:** *Learning Analytics, Systematic Literature Review, Indicators.*

**Resumo.** *Learning Analytics (LA) surgiu da necessidade de melhorar o ensino, de compreender qual a melhor abordagem para o aprendizado. Atualmente, LA é uma dos meios mais usados para analisar informações sobre estudantes. Apesar disso, estudos preliminares expõem certa carência de informações práticas sobre o assunto. O objetivo do projeto é encontrar resultados práticos de como LA foi aplicada. Este artigo apresenta os resultados preliminares da pesquisa em andamento, descrevendo o Protocolo de Revisão Sistemática de Literatura.*

**PALAVRAS-CHAVE:** *Learning Analytics, Revisão Sistemática de Literatura, Indicadores.*

### 1. Introdução

De acordo com a *Society for Learning Analytics Research*, *Learning Analytics (LA)*, é “a medição, coleta, análise e descrição de dados sobre alunos e seus contextos, para fins de compreender e otimizar a aprendizagem e os ambientes em que ocorre”[SoLAR 2018]. Em síntese, LA é o processo de coleta e análise dos dados gerados pelos alunos a fim de melhorar os processos de ensino e aprendizagem.

No entanto, [Ferguson 2012] explica que “existe uma grande lacuna entre os papéis potenciais para a *Learning Analytics* que foram identificados na literatura de pesquisa como um todo e os temas dominantes em *Learning Analytics* como colocados em prática”. Muitas vezes, há apenas expectativa de realização de determinada análise, sem estudos concretos que a implementem. Assim, o objetivo para realização dessa pesquisa é identificar resultados práticos da aplicação de LA, buscando entender as consequências dessas iniciativas. Este artigo, por sua vez, apresenta o protocolo da Revisão Sistemática de Literatura a ser executada para atingir tal objetivo.

## 2. Protocolo de Revisão Sistemática de Literatura

O protocolo de Revisão Sistemática de Literatura (RSL) que será utilizado nessa pesquisa segue o modelo proposto por Kitchenham [Kitchenham e Charters 2007]

### 2.1. Pergunta de Pesquisa

Para a realização dessa pesquisa, a seguinte pergunta foi definida: **Como foram aplicadas as técnicas de Learning Analytics nas Instituições de Ensino Superior?** Portanto, esta pesquisa visa investigar como *Learning Analytics* é executada de forma prática nas instituições de ensino e quais técnicas e ferramentas têm sido utilizadas, bem como os resultados obtidos. A seguinte estrutura foi identificada para esta pergunta:

População (P): Instituições de Ensino;

Intervenção (I): técnicas e ferramentas para *Learning Analytics*;

Resultado (O): ?

Para auxiliar a análise dos dados, as seguintes sub-perguntas de pesquisa foram definidas:

1. Quais as ferramentas/técnicas utilizadas?
2. Em que aspectos a LA melhorou o ensino e aprendizagem na instituição?
3. Quais as dificuldades na implantação de LA na instituição?
4. A instituição segue algum padrão para armazenamento de dados?

### 2.2. Estratégia de Pesquisa dos Estudos Primários

Com o objetivo de realizar uma busca abrangente, que retorne os mais relevantes trabalhos publicados, as seguintes decisões foram tomadas:

#### 2.2.1. Fontes de Informação e Estratégia de Busca

As fontes de pesquisas escolhidas foram: IEEEExplore, ACM Digital Library, Scopus e Scielo. Será utilizada uma busca automática nesses engines de busca, em que apenas artigos científicos revisados por pares deverão ser utilizados. A busca será realizada no artigo completo e metadados. Por fim, será utilizado o método snowball para encontrar artigos que foram citados por ou que citem os artigos selecionados.

#### 2.2.2. Equipes de Busca

As equipes de busca ficaram divididas em duas pessoas, a orientadora e o aluno. Além disso, um terceiro membro ficará responsável por resolver as questões de empate dos debates.

A seguinte metodologia será utilizada para executar a RSL:

- Cada pesquisador ficará responsável por realizar a busca em todos os engines de busca. Os pesquisadores deverão realizar a busca no engine e aplicar o primeiro filtro (leitura de título e *abstract*). A inclusão ou não de um artigo deve ser decidido em um processo de união de conjuntos;
- Em seguida, o resultado da união deverá ser avaliado pelos dois pesquisadores, que deverão ler o conjunto de artigos, aplicando os critérios de inclusão e exclusão. Neste momento, conflitos serão resolvidos pela decisão do terceiro membro. Só então será realizada a síntese dos dados.

### 2.2.3. String de Busca

A String de Busca definida a partir da pergunta de pesquisa foi:

(*"techniques"*) AND (*"tools"*) AND (*"learning analytics"*) AND (*"higher education"*)

Adicionando os sinônimos, tem-se:

((*"technique"*OR *"method"*OR *"performance"*OR *"approach"*OR *"practice"*OR *"manner"*OR *"procedure"*) AND (*"application"*OR *"tool"*OR *"implementation"*) AND (*"higher education"*OR *"college education"*OR *"tertiary education"*) AND (*"learning analytics"*))

### 2.2.4. Critérios de Inclusão e Exclusão

A inclusão ou não de um trabalho é determinada pela relevância deste em relação à questão de pesquisa. Os seguintes critérios de inclusão e exclusão foram definidos na Tabela 1.

**Tabela 1. Critérios Inclusão e Exclusão**

Critérios de Inclusão	Critérios de Exclusão
Estudos que tratem sobre Learning Analytics; Estudos que apresentam aplicações práticas de LA	Artigos que não estejam em língua inglesa; Artigos de Grey Literature; Estudos indisponíveis em formato digital; Short Papers com menos de 6 páginas; Estudos em bibliotecas pagas; Estudos repetidos; Será incluída apenas a versão mais atual do artigo

## 3. Conclusões Parciais

O projeto ora descrito encontra-se em fase de execução. O desenvolvimento do protocolo permitiu a estruturação da pesquisa e avaliação de sua viabilidade. A *string* de busca foi testada e está sendo aplicado o primeiro filtro na busca nos engenhos selecionados. Após finalizada a busca dos artigos, estes serão analisados e os resultados sintetizados para responder à pergunta de pesquisa.

### Referências

- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6):304–317.
- Kitchenham, B. e Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.
- SoLAR (2018). Society for learning analytics research. [Online; accessed 5-Julho-2018].

## **Desenvolvimento de um Módulo para Geração Mensagens Construtivas a partir da Análise de Mineração de Dados Educacionais**

**Pedro Macêdo<sup>1,2</sup>, Wylliams Santos<sup>1,2</sup>, Alexandre Maciel<sup>1</sup>**

<sup>1</sup>REACT Research Labs, Caruaru – PE – Brazil  
Universidade de Pernambuco (UPE), Caruaru – PE – Brazil

{phrm, amam}@ecomp.poli.br, wbs@upe.br

***Resumo.** Visando o uso da tecnologia no contexto educacional, este trabalho tem como objetivo aperfeiçoar as ferramentas de ensino a distância e motivar o aluno através de mensagens pedagógicas de acompanhamento. Para isso, serão utilizados dados gerados por estas ferramentas para automatizar o processo.*

### **1. Introdução**

O uso da tecnologia vem afetando a sala de aula, que sofre modificações em como o conteúdo é ministrado. Segundo Ferreira [Ferreira 2014] a aula não esta mais contida apenas dentro do ambiente escolar, os alunos utilizam-se de meios tecnológicos para estudar em qualquer ambiente.

Com objetivo de atender as demandas crescentes de aulas que utilizam de meios tecnológicos, muitas ferramentas para Ensino a Distância (EAD) surgiram. Este trabalho tem como objetivo analisar os dados gerados por estas ferramentas com a finalidade de gerar mensagens pedagógicas auxiliando os alunos em seus estudos.

### **2. Análise de Mineração de Dados Educacionais**

Para auxiliar um estudante, é necessário que o professor possa monitorar o seu aluno, verificando suas principais dificuldades e seus pontos fortes. Esse tipo de acompanhamento é uma prática normal em sala de aula, visto que o professor está presente junto ao aluno. Porém, Alves [Alves e Jaques 2014] explica que em plataformas de EAD os professores recebem muitos questionamentos e por esse motivo fica difícil de acompanhar o desempenho de cada aluno, o que pode gerar desinteresse ou até mesmo desistência por parte dos alunos.

Mas quando o estudante passa a utilizar ferramentas de EAD, fica mais difícil fazer esse acompanhamento, como explica Medeiros [de Medeiros e Gomes 2012].

Medeiros ainda explica que [de Medeiros e Gomes 2012] para auxiliar o professor nessa tarefa, os Ambientes Colaborativos Virtuais de Aprendizagem costumam gerar relatórios de monitoramento de cada aluno. Porém, esses dados não são organizados e muitas vezes são incompreensíveis para o professor.

Utilizando de técnicas de Mineração de Dados é possível organizar essas informações e gerar melhores relatórios, otimizando assim o funcionamento dessas ferramentas.

## 2.1. Mensagens Construtivas

A utilização de Ambientes Virtuais de Aprendizagem gera benefícios para os alunos, mas ainda há diversos desafios relacionados a essa prática. Medeiros [de Medeiros e Gomes 2012] explica que uma das maiores dificuldades destas ferramentas é a falta de suporte devido a baixa interação entre o aluno e o professor.

Segundo Bugs [Bugs et al. 2011], os *softwares* educativos podem ensinar através de experiências, oferecendo suporte ao aluno com o uso de mensagens pedagógicas, ou seja, mensagens que facilitam o aprendizado do estudante estimulando-o a continuar tentando, indicando-o onde errou ou qual caminho seguir. Sendo assim, podemos utilizar de o *software* para ensinar, mesmo que não haja uma interação direta entre o aluno e o professor.

Com isso, é possível utilizar mensagens pedagógicas baseadas em atividades onde o aluno apresentara dificuldades. É citado por de Paula [de Paula e Valente 2015] que os erros podem ser explorados como recursos didáticos, contanto que esses ofereçam ao usuário uma experiência de que aquilo é transitório e poderão ser melhorados.

## 3. Arquitetura Proposta

Este trabalho propõe uma integração de plataformas propostas nos trabalhos de [Maciel et al. 2014] (Figura 1). Nessa figura podemos identificar três *frameworks* que estarão trabalhando de forma relacionada.

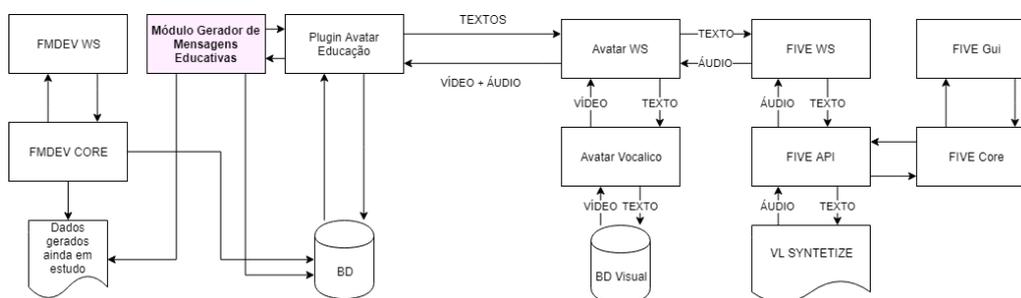


Figura 1. Arquitetura Proposta

O primeiro *framework*, localizado na parte mais a esquerda da Figura 1 é o FIVE. Com ele, Maciel [Maciel e Carvalho 2010] propõe um *software* que tem como função transformar mensagens de texto em áudio. Estas mensagens de áudio geradas poderão ser utilizadas pelo segundo *framework* o Avatar.

Seguindo o diagrama apresentado na Figura 1 temos, ao meio, o Projeto Avatar. O Avatar é um *framework* que utiliza dados recebidos de uma ferramenta de educação a distância, o Moodle. Com esse *framework* há um personagem que poderá informar ao estudante como está seu desempenho, se há novas atividades ou se existe alguma atividade incompleta.

O Avatar é dividido em duas partes, a primeira é referente a geração de vídeos. Nessa etapa o programa receberá informações textuais e gerará um vídeo utilizando imagens com desenhos de bocas em posições que representam os fonemas existentes na língua portuguesa, como é explicado por Maciel [Maciel et al. 2014]. Essa técnica é utilizada para gerar maior aproximação com a realidade ao projeto final.

A segunda parte do projeto avatar é a parte que será acessada pelo estudante, o Avatar Plugin, nela aparecerá o vídeo gerado na primeira parte juntamente com o áudio recebido do FIVE. As mensagens são faladas pelo personagem que representa o Avatar.

Por fim, temos o FMDev, este *framework* visa unificar os dados educacionais independente de suas origens, fazendo com que toda essa arquitetura não dependa exclusivamente do *moodle*.

O trabalho proposto por esse projeto visa utilizar os dados do *moodle* assim como os gerados pelo FMDev para gerar mensagens pedagógicas de apoio ao estudante. Diferenciando-se do projeto avatar por gerar mensagens que apoiem o estudante e faça com que este supere suas dificuldades.

#### 4. Conclusões e Trabalhos Futuros

É possível perceber que existem falhas relacionadas ao apoio aos alunos quando se trata de plataformas de ensino a distância. Porém, existem boas práticas para construir um bom material de apoio pedagógico, que não necessitam de um professor totalmente presente, como as diretrizes para elaboração de mensagens construtivas propostas por [?].

Com a existência de um problema comum entre os Ambientes Virtuais de Aprendizagem e a existência de diretrizes utilizadas para gerar bons materiais, esse trabalho contribuirá com a geração automática de boas mensagens de suporte, utilizando-se dos dados educacionais fornecidos pelas plataformas de aprendizagem.

#### Referências

- Alves, F. P. e Jaques, P. (2014). Um ambiente virtual com feedback personalizado para apoio a disciplinas de programação. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 3, page 51.
- Bugs, D., Veiga, M. B., e Silveira, M. S. (2011). Diretrizes para elaboração de mensagens de erro construtivas. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 1.
- de Medeiros, F. P. A. e Gomes, A. S. (2012). Monitoramento da experiência do usuário em ambientes colaborativos virtuais de aprendizagem: Um mapeamento sistemático. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 23.
- de Paula, B. H. e Valente, J. A. (2015). Errando para aprender: a importância dos desafios e dos fracassos para os jogos digitais na educação. *RENOTE*, 13(2).
- Ferreira, J. M. M. (2014). Flipped classrooms: From concept to reality using google apps. In *Remote Engineering and Virtual Instrumentation (REV), 2014 11th International Conference on*, pages 204–208. IEEE.
- Maciel, A. e Carvalho, E. (2010). Five-framework for an integrated voice environment. In *Proceedings of International Conference on Systems, Signal and Image Processing, Rio de Janeiro*. sn.
- Maciel, A. M., Rodrigues, R. L., e Carvalho, E. C. (2014). Desenvolvimento de um assistente virtual integrado ao moodle para suporte a aprendizagem online. In *Brazi-*

*lian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE), volume 25, page 382.*

# Uma ferramenta para auxílio à formação de equipes socioafetivas em ambientes híbridos de ensino-aprendizagem: Resultados Preliminares

Denini Silva<sup>1</sup>, Wylliams Santos<sup>1</sup>

<sup>1</sup>REACT Research Labs  
Universidade de Pernambuco (UPE), Caruaru – PE – Brazil

{denini.silva, wbs}@upe.br

**Resumo. Introdução** - Destaca a importância da criação de equipe, com técnicas, visando maximizar o seu desempenho. **Contexto** - Ambiente escolar, seja presencial ou virtual, no qual uma atividade para ser concluída é necessário que se façam equipes entre os alunos. **Objetivo** - Dar suporte à criação de uma ferramenta para auxílio na formação de equipes considerando os fatores socioafetivos para sua formação. **Resultados e Conclusões** - Através da implementação da ferramenta para criação de equipes, tendo como base seus perfis temperamentais, é possível criar-se equipes que tenham os requisitos socioafetivos necessários para que a mesma seja bem sucedida em sua atividade. Serão feitos experimentos em um ambiente híbrido: Moodle(EAD) e sala de aula presencial, com o intuito de pôr em prática o que neste trabalho for discutido e abordado.

## 1. Introdução

Equipes de alunos sempre fizeram parte da educação em sala de aula, seja uma única equipe, toda a turma, ou em equipes menores para a realização de uma atividade. Tais equipes, geralmente são formadas tendo como base a amizade dos participantes. Segundo Lima [Lima 2016] esse critério de amizade parece ser bom, contudo, dependendo dos fatores socioafetivos dos participantes, o desempenho da equipe pode acabar sendo reduzido. Essa relação com desempenho de equipe, pode se dar em função do perfil de cada aprendiz e/ou dos laços de afinidade existentes entre eles, sendo assim, podemos considerar que a tarefa de dividir a classe em equipes no ambiente de aprendizagem colaborativo torna-se de extrema importância.

Na computação, existem diversas ferramentas que podem ser implementadas para criação de grupos, ferramentas as quais são conhecidas como clusters, este trabalho vem propor a criação de uma ferramenta clustering utilizando como base os fatores socioafetivos dos alunos para a criação de equipe. Podendo assim, o professor, fazendo uso dessa ferramenta criar as equipes da turma para que o seu desempenho seja maximizado. Esta ferramenta não se restringirá apenas às salas de aula comuns, presenciais, se estenderá também a ambientes EAD como o Moodle<sup>1</sup>.

## 2. Afinidade social e metodologia de avaliação de temperamentos

Para Filho [Lopes Filho et al. 2011], a afetividade sempre terá papel fundamental na natureza, portanto, é um fator influenciador, nas interações colaborativas de ensino e

---

<sup>1</sup><https://moodle.org>

aprendizagem. Inclusive, estudantes que trabalham juntos em função da afinidade social não precisam negociar as regras de colaboração, pois conseguem estabelecer modos de interação que são explicitamente entendidos pela equipe. De acordo com Jones [Jones e Issroff 2005], estudantes que trabalham juntos sem afinidade social, geralmente, devem negociar as regras da interação, o que se torna desmotivador para alguns participantes. Segundo Justos [Justo 1966] existem oito tipos de temperamentos de caráter, sendo eles: Amorfo, Melancólico, Instável, Social, Líder, Fleumático, Ativo e Apático e cada pessoa pode ser definida por um dos oito.

Visando determinar o temperamento do aluno de uma maneira quantitativa, o questionário Roger Verdier foi escolhido, pois para Lima [Lima et al. 2008] o questionário, tem uma didática apropriada para o contexto da pesquisa, tendo em vista que outros meios mais aprofundados de mensuração, análise e avaliação de características psicológicas exigiriam um profissional vinculado à área, sendo assim necessário uma consulta com tal profissional. O questionário é composto por 15 perguntas<sup>1</sup> de respostas binárias (Sim ou Não), tais perguntas sofreram pequenas modificações para que ficassem mais compreensíveis aos alunos, tornando-as mais claras e objetivas, de maneira que não altere o sentido original da pergunta. O questionário está disponível na web através da ferramenta Google Forms<sup>2</sup>.

### 3. Ferramenta com algoritmo cluster

A Ferramenta para criação das equipes foi desenvolvida em Linguagem JAVA e dado que todos os alunos da turma responderam ao questionário Roger Verdier, a ferramenta funcionará em duas etapas:

Primeira etapa: De acordo com as respostas de cada aluno, determinar o temperamento do aluno seguindo os casos propostos por Justo [Justo 1966] classificando assim o aluno como sendo: Amorfo, Melancólico, Instável, Social, Líder, Fleumático, Ativo ou Apático.

A segunda etapa é a utilização de um algoritmo<sup>3</sup> clustering genérico, para a criação de equipes, que utiliza os fatores sociais dos alunos e tabela de temperamentos afins e antitéticos para a sua criação. Para o início do algoritmo, o professor deve informar o número de equipes que ele deseja formar<sup>4</sup>. A seguir é apresentado o fluxo do algoritmo de criação de equipes:

1. **Definição dos centroides ou cabeças:** Segundo [Lima 2016] os centroides ou cabeças de cada equipe devem ser os alunos de temperamento Líder, caso não exista quantidade de líderes suficiente, os de temperamento Ativo irão completar, caso esses ainda sejam insuficientes, os Sociais também tomarão o seu lugar como centroide se ainda não for suficiente será escolhido um aluno aleatório.
2. **Agregação dos grupos:** A equipe será completada por pessoas que têm temperamentos afins do centroide e evitando pessoas com temperamento antitético, este passo é feito por rodadas, em cada rodada as equipes recebem uma nova pessoa, essas rodadas duram até acabar o número de alunos afins dos grupos, garantindo

<sup>1</sup>As questões estão disponíveis em: <http://twixar.me/nqd3>.

<sup>2</sup>Disponível em: [goo.gl/forms/goWtev57xJUfJoft2](https://goo.gl/forms/goWtev57xJUfJoft2)

<sup>3</sup>Acesso ao código-fonte disponível em: [github.com/ReactLabs/ClusterAfinidade](https://github.com/ReactLabs/ClusterAfinidade)

<sup>4</sup>Tabela de temperamentos afins e antitéticos: <https://imgur.com/a/8gma9SI>

assim que as equipes fiquem com um número proporcional de integrantes. Há de se considerar que os líderes, quando em mesmo grupo, não se relacionam bem, pois são pessoas que antes gostam de comandar, tendo dificuldade em aceitar comandos [Justo 1966].

3. **Formação dos grupos:** Por fim, se não existir mais nenhum aluno com temperamento afim com os dos centroides, o algoritmo irá agrupar os alunos restantes de forma aleatória, com exceção dos temperamentos antitéticos do centroide.

#### 4. Conclusões e Trabalhos Futuros

Este trabalho ainda está em andamento, contudo, para alcançar o objetivo de fornecer esta ferramenta na web, já foram concluídas etapas extremamente necessárias. Para alcançar o desejo de criar equipes baseadas não na afinidade dos integrantes, mas sim no seu tipo socioafetivo, maximizando o desempenho dessa equipe, foi-se criado o questionário Roger-Verdier através da plataforma Google formulários. Após todos os alunos da turma, seja ela online ou presencial, responderem o questionário, será colocado em funcionamento o algoritmo cluster, para isso é necessário apenas que o professor entre com os dados da quantidade de equipes que ele deseja criar.

O experimento será feito em salas de aula na UPE, tal teste será crucial para observar a eficiência do algoritmo e todos os aspectos referentes à formação das equipes, também será rodado um experimento no ambiente do Moodle, para fins de comparação com o teste presencial, os aspectos observados serão os mesmos que no presencial. Para o futuro, desejamos oferecer essa ferramenta em uma plataforma na web, para isso, enumeramos alguns processos que vamos seguir, portanto esses serão os nossos trabalhos futuros:

- I Evoluir o algoritmo cluster para um já consolidado na mineração de dados.
- II Experimento na sala de aula UPE.
- III Experimento na sala de aula virtual, Moodle.
- IV Colocar toda a ferramenta na web facilitando, assim, o seu acesso.

#### Referências

- Jones, A. e Issroff, K. (2005). Learning technologies: Affective and social issues in computer-supported collaborative learning. *Computers & Education*, 44(4):395–408.
- Justo, F. (1966). Teste de caráter ao alcance de todos. *Editora Escola Profissional La Salle, Canoas Rio Grande do Sul*.
- Lima, M. R., Quarto, C. C., Labidi, S., e Schivitz, I. M. (2008). Considerando fatores sócio-afetivos para a formação de grupos em ambientes colaborativos de aprendizagem apoiados por computador utilizando algoritmos genéticos. *Scientia*, 19(2):128–143.
- Lima, T. D. V. (2016). Desenvolvimento de uma ferramenta com algoritmo de clustering para auxílio à formação de grupos de trabalho em ambientes de ensino-aprendizagem levando em consideração o fator socioafetivo. *13th International Conference on Information Systems Technology Management - CONTECSI*.
- Lopes Filho, J. A., Quarto, C., e Martins França, R. (2011). Um algoritmo de clustering no auxílio à formação de grupos sócio-afetivos afins para o ensino em ambientes cscl. *Sexta Conferencia Latinoamericana de Objetos de Aprendizaje y Tecnologías para la Educación - LACLO 2011*.

## **Análise da Logística Reversa de uma Instituição de Ensino Superior do Agreste Pernambucano**

**Nathália, C. de A. Moraes<sup>1</sup>, Tatiana, A. Silva<sup>1</sup>, Amanda, Duarte<sup>1</sup>, Gabriela de Vasconcelos<sup>1</sup>**

<sup>1</sup>Universidade de Pernambuco (UPE)  
Caruaru– PE – Brasil

{nathyamoraes@gmail.com, tatianaalexandresilva@live.com,  
amandaduarte.cn@gmail.com e devasconcelos.gabriela@gmail.com}

**Resumo.** *O presente artigo tem como objetivo analisar a logística reversa de uma Instituição de Ensino Superior no Agreste Pernambucano. A Logística Reversa está atrelada ao fluxo do consumo de materiais consumidos, que precisam ser gerenciados, um modo reinserir os produtos pós-consumidos afim de preservar o meio ambiente e a vida. O estudo de caso foi feito na Universidade de Pernambuco, com o intuito de verificar se a mesma possui desenvolvimento sustentável. Este trabalho analisa os casos descritos por autores da literatura com o objetivo de destacar o conceito da logística reversa e a preocupação com o meio ambiente.*

### **1. Introdução**

A humanidade produz milhões de toneladas de resíduos sólidos diariamente de diversas origens, decorrentes de um sistema de produção, circulação e consumo de bens e serviços, que aumentam cada vez mais a quantidade de lixo gerada. O acelerado desenvolvimento tecnológico experimentado pela humanidade tem permitido a introdução constante e com velocidade crescente de novas tecnologias, que contribuem para a redução de preços e dos ciclos de vida de grande parcela dos bens de consumo. Duráveis e semiduráveis, além da busca de diferenciação mercadológica que gera obsolescência e aumenta a velocidade de descarte dos produtos [Leite 1998].

Com o ciclo de vida útil dos produtos reduzidos os números de resíduos sólidos aumentam extrapolando a capacidade dos ecossistemas de sustentar os moldes de produção e consumo da sociedade contemporânea, o que resulta em consequências negativas para o meio ambiente [Cidin e Silva 2007]. Desse modo, surge a logística reversa como uma alternativa ao modelo tradicional de disposição final dos resíduos sólidos, os aterros sanitários e os lixões, que têm se mostrado inviáveis dentro do parâmetro do desenvolvimento sustentável. Os lixões continuam sendo o destino da maioria dos resíduos urbanos produzidos no Brasil, trazendo prejuízos à qualidade de vida da população. O Brasil produz 241.614 toneladas de lixo por dia, destes, 76% são depositados a céu aberto em lixões, 13% são depositados em aterros controlados, 10% em usinas e 0,1% são incinerados (IBGE, 2000). Portanto, o destino final do lixo têm sido um desafio a ser superado a fim de se alcançar a sustentabilidade e contribuir para o

alcance dos Objetivos de Desenvolvimento Sustentável (ODS) do Milênio estipulados pela ONU para serem alcançados até 2030 [Vasconcelos 2018].

Diante do aumento significativo dos resíduos sólidos a crescente conscientização da população quanto à preservação do meio ambiente e contra o desperdício, impulsionou-se a preocupação das organizações a Logística Reversa, uma tecnologia limpa, sua aplicação permite que a organização priorize sua atuação sobre os resíduos gerados nos processos, definindo o seu gerenciamento. A logística reversa traz de volta ao processo produtivo, materiais que podem ser reparados, reaproveitados, reciclados ou adequadamente descartados segundo [LEITE, 2000, p.1], ensejando a prática do conceito dos 5R's (repensar, reduzir, recusar, reutilizar e reciclar).

Nessa perspectiva, o presente estudo buscou analisar a logística reversa de uma Instituição de Ensino Superior (IE's) do Agreste pernambucano a fim de identificar se a universidade analisada põe em prática os 5R's, contribuindo para sustentabilidade local, partindo-se do seguinte problema de pesquisa: Como se dá o processo de logística reversa dentro da universidade analisada?

## 2. Metodologia

O presente estudo tem caráter exploratório-descritivo, visto que pretende analisar como se dá o processo de logística reversa dentro da Universidade objeto de estudo e descrever os achados. De acordo com Gil (2010) a pesquisa exploratória enseja maior familiaridade com o objeto de estudo, enquanto que a descritiva se propõe a analisar, observar e registrar os fatos ou fenômenos, sem manipulá-los. Desse modo, a presente pesquisa explora o cenário de uma IE's a fim de familiarizar-se como o processo de logística reversa dessa entidade para poder descrevê-lo.

Quanto a abordagem, a pesquisa caracteriza-se como qualitativa, na medida em que se preocupa em analisar e interpretar aspectos mais profundos dos fenômenos observados, descrevendo-os, tecendo uma análise detalhada das tendências de hábitos, atitudes e comportamentos [Marconi e Lakatos 2011]. Nesse sentido, o estudo propõe-se a analisar o processo de logística reversa de uma IE's a partir dos 5'Rs buscando descrever as ações adotadas pela instituição.

Ademais o estudo enquadra-se quanto ao procedimento metodológico como uma pesquisa de campo do tipo exploratória-descritiva, posto que consiste na observação de fatos e fenômenos com o objetivo de descrevê-los [Marconi e Lakatos 2011]. Ademais, pode ser enquadrado como um estudo de caso, uma vez que lança um olhar sobre a logística reversa de uma IE's do Agreste Pernambucano, tecendo considerações específicas sobre o caso analisado. O estudo de caso caracteriza-se como um método de pesquisa que investiga-se os fatos após a sua ocorrência [Gil 2010].

O instrumento de coleta de dados da presente pesquisa será um questionário que será aplicado ao gestor da universidade estudada. Tal instrumento adotará a seguinte estrutura:

**Tabela 1. Questionário da Pesquisa**

<b>Objetivo de Desenvolvimento</b>	<b>1. A universidade desenvolve ações de consumo responsáveis? Se sim, quais são essas ações?</b>
------------------------------------	---

<p><b>Sustentável 12: Consumo e Produção Responsáveis</b></p>	<p><b>2. A universidade faz uso eficiente dos recursos naturais? Se sim, como se dá o uso eficiente desses recursos?</b></p>
	<p><b>3. A universidade busca alcançar o manejo ambientalmente adequado dos produtos químicos e de todos os resíduos, ao longo de todo o ciclo de vida destes, de acordo com os marcos internacionalmente acordados, e reduzir significativamente a liberação destes para o ar, água e solo, para minimizar seus impactos negativos sobre a saúde humana e o meio ambiente? Se sim, como é feito o manejo adequado?</b></p>
	<p><b>4. A universidade busca reduzir substancialmente a geração de resíduos por meio da prevenção, redução, reciclagem e reuso? Se sim, quais ações baseadas nos 5R's (reduzir, reciclar, reutilizar, reaproveitar e repensar) são adotadas?</b></p>
	<p><b>5. A universidade promove práticas de compras públicas sustentáveis, de acordo com as políticas e prioridades nacionais?</b></p>
<p><b>Política Nacional de Resíduos Sólidos - PNRS</b></p>	<p><b>6. A PNRS prevê a prevenção e a redução na geração de resíduos, tendo como proposta a prática de hábitos de consumo sustentável e um conjunto de instrumentos para propiciar o aumento da reciclagem e da reutilização dos resíduos sólidos (aquilo que tem valor econômico e pode ser reciclado ou reaproveitado) e a destinação ambientalmente adequada dos rejeitos (aquilo que não pode ser reciclado ou reutilizado). Diante disso, pode-se dizer que a universidade segue a PNRS na gestão dos resíduos? Se sim, quais ações são desenvolvidas com base na PNRS?</b></p>
	<p><b>7. A universidade institui a responsabilidade compartilhada dos geradores de resíduos? Se sim, quais os stakeholders (fabricantes, importadores, distribuidores, comerciantes) envolvidos e quais as responsabilidades de cada um?</b></p>
	<p><b>8. A universidade possui Coleta Seletiva?</b></p>
	<p><b>9. Pode-se dizer que a universidade possui uma postura sustentável? Justifique a resposta.</b></p>

### 3. Referências

[http://coral.ufsm.br/seminarioeconomia/images/anais\\_2017/OS\\_5\\_RS\\_DA\\_SUSTENTABILIDADE\\_OS\\_5\\_RS\\_DA\\_SUSTENTABILIDADE\\_OS\\_5\\_RS\\_DA\\_SUSTENTABILIDADE\\_OS\\_5\\_RS\\_DA\\_SUSTENTABILIDADE\\_OS.pdf](http://coral.ufsm.br/seminarioeconomia/images/anais_2017/OS_5_RS_DA_SUSTENTABILIDADE_OS_5_RS_DA_SUSTENTABILIDADE_OS_5_RS_DA_SUSTENTABILIDADE_OS_5_RS_DA_SUSTENTABILIDADE_OS.pdf)

SOTO, M.M.T. Dissertação de Mestrado. Centro de Ciência e Tecnologia da Universidade Estadual do Norte Fluminense, 2006. Disponível em: <[http://www.uenf.br/Uenf/Downloads/POS-ENGPRODUCAO\\_2397\\_1160058817.pdf](http://www.uenf.br/Uenf/Downloads/POS-ENGPRODUCAO_2397_1160058817.pdf)> Acesso em: 2018 set. 2018.

IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional de Saneamento Básico 2000. Leite, P.; (2000). A Logística Reversa e a Responsabilidade Empresarial. Revista Tecnológica , Dez/2000. Disponível em: < [www.tecnologica.com.br](http://www.tecnologica.com.br)> Acesso em 2018 set. 2018.

Leite, P.; (1998). Canais de Distribuição Reversos. Revista Tecnológica, Abril/ 1998. [www.tecnologica.com.br](http://www.tecnologica.com.br)

de Vasconcelos, G., & Pimentel, M. S. (2018). Análise Do Disclosure Dos Custos E Investimentos Ambientais Das Empresas Potencialmente Poluidoras Que Compõem O Índice De Sustentabilidade Empresarial Da Bovespa. *Revista de Gestão Ambiental e Sustentabilidade*, 7(2), 210-229.

file:///C:/Users/Rosa%20Vasconcelos/Desktop/artigo\_Revista\_GeAs.pdf

[http://www.ppgc.ufrpe.br/sites/ww3.ppgc.ufrpe.br/files/documentos/gabriela\\_de\\_vasconcelos\\_dissertacao\\_oficial\\_com\\_assinaturas.pdf](http://www.ppgc.ufrpe.br/sites/ww3.ppgc.ufrpe.br/files/documentos/gabriela_de_vasconcelos_dissertacao_oficial_com_assinaturas.pdf)

Cidin, R. D. C. P. J., & da Silva, R. S. (2007). Pegada ecológica: instrumento de avaliação dos impactos antrópicos no meio natural. *Estudos Geográficos: Revista Eletrônica de Geografia*, 2(1), 43-52.

[http://www.ifba.edu.br/professores/armando/eng531/Unid%20I/Artigo\\_Pegada\\_ecologica.pdf](http://www.ifba.edu.br/professores/armando/eng531/Unid%20I/Artigo_Pegada_ecologica.pdf)

## **Ferramentas de aprendizagem em Problem Based Learning: Resultados Parciais de uma Revisão Sistemática da Literatura**

**Aline S. Gomes<sup>1</sup>, Wylliams Barbosa Santos<sup>1</sup>**

<sup>1</sup>Agile Software Development and Education (REACT)  
Universidade de Pernambuco (UPE), Caruaru – PE – Brazil

{aline.gomes@upe.br, wbs@upe.br}

**Resumo. Contexto:** A necessidade de ferramentas e recursos de aprendizagem usados em conjunto com o PBL, que proporcionem um maior grau de satisfação e conhecimento por parte dos estudantes durante a experiência PBL. **Objetivo:** verificar quais ferramentas e recursos de aprendizagem usados em conjunto com o PBL, a fim de garantir a eficácia de sua aplicação. A possível proposta de um modelo pedagógico para um Ambiente Virtual de Aprendizagem (AVA) poderá surgir ao fim do estudo, que só será possível devido a produção de uma revisão sistemática da literatura (RSL) no contexto de Ferramentas de aprendizagem em Problem based Learning, realizada como parte deste estudo. **Método:** Por meio de uma RSL, buscou-se refletir acerca da importância de ferramentas e recursos de aprendizagem trabalhados em conjunto ao PBL. **Resultados:** A partir das pesquisas iniciais, foi observada a tendência em importância dessas ferramentas e recursos como facilitadores, que auxiliam a aplicação do PBL, apresentando resultados positivos na literatura para a apreensão do PBL. **Conclusão:** Espera-se possibilitar a reflexão sobre as ferramentas e recursos mais utilizados, assim como, novas estratégias de aprendizagem com o uso do PBL e novos direcionamentos para investigações futuras.

### **1. Introdução**

Vivemos em um mundo em constantes mudanças e com grandes avanços tecnológicos, ao qual a educação não fica alheia, passando por várias mudanças de paradigmas, onde as aulas expositivas já não produzem o mesmo efeito no aprendizado. Nos últimos anos, o processo tradicional de ensino em sala de aula tem gerado uma extensa discussão e pesquisa no campo educacional quanto a sua total eficácia. E isso se deve ao surgimento de outros processos de ensino, que proporcionam algo novo e que demonstram que o grau de satisfação e realização dos estudantes podem ser muito maiores com seu uso, sendo necessário repensar as metodologias de ensino. Diante da observação de inúmeras concepções metodológicas fundamentadas em Aprendizagem Baseada em Problemas (PBL), houve o interesse em revisitar o referencial teórico relativo à temática, possibilitando contemplar as ferramentas e os recursos utilizados durante a aplicação do PBL. Ferramentas que facilitam e possibilitam o sucesso da execução do PBL.

A inclusão da tecnologia nesse processo tem trazido grandes benefícios. Logo a introdução das TIC tem criado novas plataformas virtuais, assim como, diversas ferramentas e recursos capazes de auxiliarem o desenvolvimento de novas metodologias de aprendizagem, como o PBL.

## 2. Problem-Based Learning - PBL e Ferramentas de Aprendizagem

Na definição dada por [Delisle e Oliveira 2000](p. 5), PBL é “*uma técnica de ensino que educa apresentando aos alunos uma situação que leva a um problema que tem de ser resolvido*”. [Lambros 2002] afirma que PBL é um método de ensino que se baseia na utilização de problemas como ponto inicial para adquirir novos conhecimentos. Portanto, o PBL surge como uma dessas estratégias inovadoras em que os estudantes trabalham com o objetivo de solucionar um problema real ou equivalente a partir de um contexto. Trata-se, portanto, de um método de aprendizagem centrado no estudante, que elimina o papel deste de receptor do conhecimento, para assumir o lugar de protagonista de seu próprio aprendizado.

A tecnologia, no entanto, não deve apenas distribuir e permitir acesso ao conhecimento, mas deve melhorar a aprendizagem no sentido de torná-la significativa para os estudantes [Lauretta e Partridge 2007].

## 3. Revisão Sistemática da Literatura

“Uma revisão sistemática de literatura é uma forma de estudo secundário que utiliza uma metodologia bem definida para identificar, analisar e interpretar todas as evidências disponíveis a respeito de uma questão de pesquisa particular de maneira imparcial e Repetível” [Kitchenham and Charters, 2007].

A equipe para esta RSL é composta por 2 participantes, um discente em Sistemas de informação(UPE) e o outro Doutor em ciência da computação(UFPE). Ambos os pesquisadores trabalham na Busca, Seleção, Extração e Síntese dos dados.

**Questões de Pesquisa:** Perguntas de pesquisa foram desenvolvidas para que fosse possível dar início a revisão sistemática. O trabalho parte da seguinte questão central: “*Quais as ferramentas e recursos de aprendizagem usados na aplicação do PBL?*”. A partir desta, três perguntas de pesquisa específicas foram também desenvolvidas: (PP1) Qual a importância da aplicação dessas ferramentas e recursos em conjunto com o PBL? (PP2) Quais os AVAs mais utilizados? (PP3) Quais os desafios apresentados na implementação destas ferramentas?.

**Estratégia de Pesquisa:** Para a realização das buscas dos documentos, foram utilizadas as seguintes bases automáticas: ScienceDirect<sup>1</sup>, IEEE Xplore Digital Library<sup>2</sup>, ACM Digital Library<sup>3</sup> e Wiley Library<sup>4</sup>. De modo a obter todos os documentos relevantes que apresentassem respostas significativas às questões de pesquisa, duas *strings* de busca foram construídas, sendo uma em português e outra em inglês: (“*problem based learning*” OR *pbl*) AND (“*virtual environment*” OR “*virtual learning environments*” OR “*LMS*” OR “*VLE*” OR “*LCMS*” OR “*tools*” OR “*learning tools*” OR “*learning environments*”) AND (*computer* OR “*engineering computer*” OR “*science computer*” OR “*information technology*”). E “*Aprendizagem baseada em problemas*” OR “*pbl*” AND “*computação*” OR “*engenharia da computação*” OR “*ciência da computação*”.

**Critérios de Inclusão:** (CII) Estudos que apresentem relação com a questão

<sup>1</sup><http://www.sciencedirect.com>

<sup>2</sup><https://ieeexplore.ieee.org/Xplore/home.jsp>

<sup>3</sup><http://dl.acm.org>

<sup>4</sup><https://onlinelibrary.wiley.com/>

(e subquestões de pesquisa); (CI2) Estudos que retratam a experiência de PBL dentro do contexto de ensino; (CI3) Pesquisas que identifiquem recursos/ferramentas/AVAs utilizados para desenvolvimento do PBL. (CI4) Abstract menciona explicitamente recursos/ferramentas/AVAs utilizados para desenvolvimento do PBL.

**Crterios de Exclusão:** (CE1) Estudos fora do contexto de ensino da computação; (CE2) PBL não ser parte das contribuies do estudo e não ser mencionado no abstract; (CE3) Documentos incompletos, rascunhos, documentos de compilação dos anais de conferências (proceedings), documentos apenas acessíveis por meio da compra e apresentações em slides, documentos que não são artigos, não acessíveis pela internet; (CE4) Pesquisas escritas em outras lnguas com exceção do Inglês e do português; (CE5) Documentos Duplicados; (CE6) Estudos que não identifiquem recursos/ferramentas/AVAs utilizados para desenvolvimento do PBL no abstract; (CE7) Pesquisas que se refiram a PBL e a identificação das ferramentas apenas como projetos futuros;

#### 4. Resultados Parciais e Análise

Esta seção descreve a principal contribuio deste trabalho, a aplicao das tcnicas de Reviso Sistemática durante a fase de conduo que é apresentada conforme protocolo descrito na seo anterior. Entre os meses de Agosto e Setembro de 2018, foram realizadas buscas experimentais com base nos critrios de seleo (Incluso e Excluso) anteriormente evidenciados. Até o momento, 387 artigos já passaram pela primeira etapa de avaliao (análise pelo título e resumo). A análise por título e resumo continua sendo executada a fim de contabilizar os estudos que sero pré-selecionados para a segunda etapa (análise pela introduo). Os próximos pré-selecionados seguirão para a última etapa (Análise completa). A reviso segue em processo de finalizao da primeira etapa. Os resultados obtidos em cada base são demonstrados na tabela 1.

**Tabela 1. Dados obtidos das Buscas e Resultados Pós-Análise**

Avaliao	ScienceDirect	IEEE Xplore	Wiley	ACM DL	SBIE	RBIE
Documentos Retornados	19007	316	1122	59	11	4
1. Título e <i>Abstract</i>	-	Avaliado 273	avaliado 40	25	7	2

#### 5. Concluso e Trabalhos Futuros

Com relao aos desafios para o desenvolvimento do trabalho, o levantamento dos dados foi inicialmente difcil, pois há necessidade de strings diversificadas para cada base. Além disso, apesar da string específica, sempre há retorno de documentos que não são relevantes para o estudo. Apesar dos limites e das possibilidades enfrentadas até o momento, há grande expectativa sobre os resultados da reviso, que terá importncia para o desenvolvimento e aplicao de alternativas metodológicas, que fundamentem um ensino-aprendizagem com novas possibilidades de situaes desafiadoras e ativas aos estudantes no ensino.

#### Referências

Delisle, R. e Oliveira, V. (2000). *Como realizar a aprendizagem baseada em problemas*.

Lambros, A. (2002). *Problem-based learning in K-8 classrooms: A teacher's guide to implementation*. Corwin press.

Kitchenham, B. and Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report (2007).

Lauretta, S., Partridge, H. (2007) "E-Learning Learning Objects: Learning Information Searching in a E-Learning Environment", In: Harman, K. Koohang, A. (Eds.) *Learning Objects: Applications, Implications, Future Directions*. Information Science Press, Santa Rosa, Califórnia.

## **Monitoria de Banco de Dados: Um projeto para capacitação integral do discente de Sistemas de Informação e seus Resultados Parciais**

**Sóstenes Gomes de Souza<sup>1</sup>, Elyda Laisa Soares Xavier Freitas<sup>1,2</sup>**

<sup>1</sup>Sistemas de Informação - Universidade de Pernambuco (UPE)  
Caruaru- PE – Brasil

<sup>2</sup>Centro de Informática (CIn) – Universidade Federal de Pernambuco (UFPE)  
Recife – PE - Brasil

sostenesg7@gmail.com, elyda.freitas@upe.br

***Abstract.** Monitoring provides extra support to learners so that they maintain a post-lesson interaction to increase their school performance. With extra support and addressing classic and modern issues, it is possible to collaborate for a full qualification of students. This paper exposes the experience of a monitoring project for the course of Databases II at the University of Pernambuco.*

***Resumo.** A monitoria oferece suporte extra aos discentes, de modo que eles mantenham uma interação pós-aula, a fim de elevar seu aproveitamento escolar. Com o suporte extra e abordando temas clássicos e atuais da disciplina, é possível colaborar para a capacitação integral dos discentes. Este artigo expõe a experiência de implantação do projeto de monitoria para a disciplina de Banco de Dados II na Universidade de Pernambuco.*

### **1. Introdução**

De acordo com Frison e Moraes (2010), a monitoria é uma técnica que visa apoiar o ensino, de forma que o monitor, uma vez que está mais avançado nos programas disciplinares, colabora no processo de obtenção de conhecimento dos monitorados e com o qual os estudantes podem sanar dúvidas, aprimorando o aprendizado tanto por parte do aluno monitorado quanto por parte do monitor.

Além disso, Costa e Baltar (2015), explicam que essa técnica de ensino-aprendizagem proporciona ao aluno monitor o aprimoramento de sua prática docente, uma vez que este recebe apoio constante nesse processo. Diante disso, o trabalho em questão tem por objetivo expor a experiência de monitoria da disciplina de Banco de Dados II em um curso de Sistemas de Informação, fornecendo informações sobre o progresso de sua execução, assim como os resultados obtidos ao seu término.

### **2. Andamento das Aulas**

Na dinâmica das monitorias, foram empregadas atividades cujo nível de dificuldade aumenta gradativamente, possibilitando a construção de conhecimento dos estudantes e a identificação de falhas no aprendizado. Outra estratégia foi a utilização de problemas reais, em conjunto com tecnologias amplamente difundidas, como a ferramenta EER Case, Oracle e MongoDB, o que permite que esse estudante ingresse no mercado de trabalho com o conhecimento necessário para execução de suas atividades. Além das

atividades presenciais, o monitor realizou atividades complementares por meio das redes sociais, a fim de atingir também os alunos que não podiam comparecer ao encontro presencial por razões diversas.

### 3. Resultados e Discussão

Ao final do primeiro semestre letivo foi realizada uma avaliação, por meio de questionário disponibilizado na rede social *Facebook*, da execução da monitoria. É preciso ressaltar que não é possível identificar uma relação de causa e efeito da monitoria sobre as notas dos alunos. No entanto, o questionário permite conhecer melhor o papel da monitoria sob a visão dos monitorados, identificando seus pontos fortes e eventuais falhas a serem ajustadas para os semestres seguintes.

#### 3.1. Quanto à Opinião dos Alunos

Oito alunos (dos 13 alunos que concluíram a disciplina) responderam ao questionário a eles direcionado. Os resultados obtidos estão apresentados abaixo, com base nos seguintes questionamentos, entre outros:

- Qual o nível de dificuldade da disciplina?
- Em que momento a monitoria mais ajudou?
- Quais fatores que mais influenciaram a decisão de procurar a monitoria?

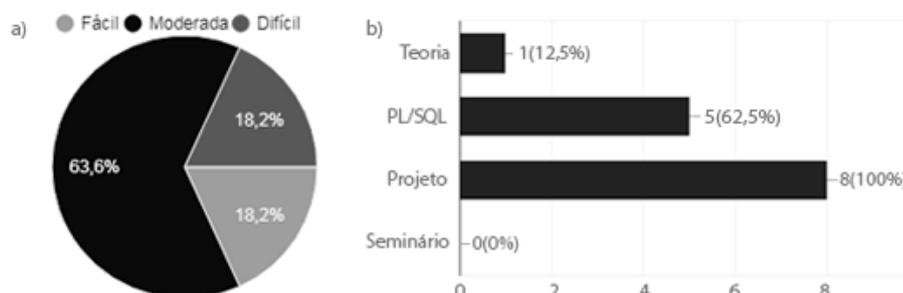


Figura 1. a) Nível de dificuldade da disciplina; b) Momento que a monitoria mais ajudou.

Os resultados da Figura 1 expõem a opinião dos alunos que, em sua maioria, consideram Banco de Dados 2 uma disciplina de moderada (63,6%) a difícil (18,2%). No tocante ao momento em que a monitoria se fez mais necessária, sendo estes durante a parte teórica, prática com PL/SQL, projeto e seminário, os estudantes afirmaram, respectivamente, 12,5%, 62,5%, 100% e 0%.

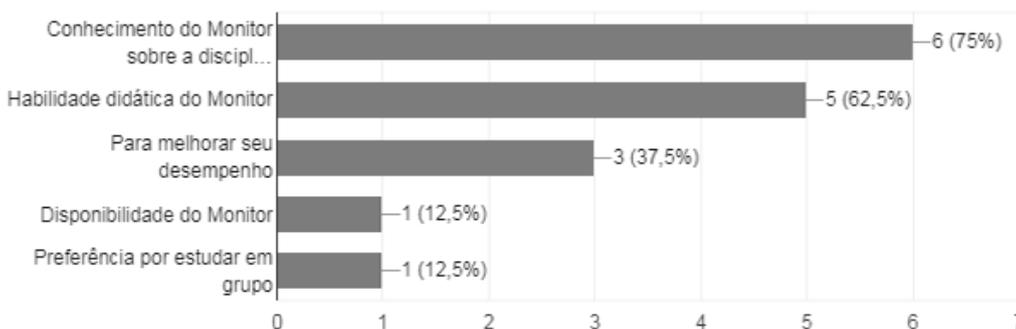


Figura 2. Quais fatores que mais influenciaram na decisão de procurar a monitoria.

De acordo com as informações apresentadas na Figura 2, os fatores mais relevantes para a ida à monitoria, por parte dos monitorados, foram: o conhecimento do monitor sobre a disciplina (75%), a habilidade e didática do monitor (62,5%) e para melhorar seu desempenho (37,5%). O estudante poderia escolher mais de uma opção dentre as listadas.

### **3.2. Quanto ao Aproveitamento Escolar dos Alunos**

No que se refere ao desempenho dos estudantes nas atividades de avaliação referentes ao 1º EE (Exercício Escolar) e 2º EE, respectivamente, 52,9% e 76,5% dos discentes foram aprovados, de um total de 17 alunos cursando a disciplina. Dos dois alunos que prestaram o exame final, 1 (um) obteve aprovação.

## **4. Conclusão**

Por meio dos resultados obtidos pela aplicação do questionário, percebe-se que os alunos possuem maior interesse pelo auxílio do monitor quando se inicia a prática do conteúdo apresentado em sala. Quanto aos fatores determinantes para a participação dos discentes nas monitorias, os que mais se destacaram são o conhecimento e a didática do monitor em conjunto com o desejo do monitorado em melhorar seu desempenho. No ano anterior, o monitor acompanhou a turma com a disciplina de Banco de Dados I, de modo que os estudantes já o conheciam.

Os resultados de desempenho escolar se apresentaram satisfatórios, uma vez que pôde-se observar a manutenção no baixo índice de reprovação com relação ao semestre anterior, quando apenas um aluno não obteve aprovação por média (excetuando os que reprovaram por falta); e ainda a diminuição significativa no número de alunos que precisaram realizar o exame final, de 11 para 2. Além disso, o questionário confirmou o sucesso da abordagem de uso de redes sociais: Os monitorados responderam que as mesmas foram importantes (37,5%) ou os ajudaram bastante (62,5%) a participar da monitoria.

Por fim, pode-se concluir que a prática da monitoria, mais uma vez, cumpriu seu objetivo de dar suporte ao aprendizado, tendo sua importância sido reconhecida pelos estudantes que dela participaram, visto que a totalidade dos estudantes respondeu que a monitoria foi muito importante para o seu aprendizado. Este projeto de monitoria continua em andamento no semestre atual, com uma turma diferente, na disciplina de Banco de Dados I.

## **Referências**

- COSTA, J. S.; BALTAR, S. L. S. M. A. A importância e concepção da monitoria de estágio supervisionado para alunos do curso de licenciatura em Biologia. *Revista Iniciação e Formação Docente*. v. 1, n. 2, 2015.
- FRISON, L. M. B.; MORAES, M. A. C. As práticas de monitoria como possibilitadoras dos processos de autorregulação das aprendizagens discentes. *Póiesis Pedagógica*, Goiás, v. 8, n. 2, p. 126-146, ago./dez. 2010. Disponível em: <<http://www.revistas.ufg.br/index.php/poiesis/article/view/14064>>. Acesso em: 11 out. 2018.

## Priorização de Requisitos em Startups de Software Pernambucanas: Resultados Preliminares

Jonatan Lins<sup>1,2</sup>, Kayo Renato<sup>1,2</sup> e Wylliams Santos<sup>1,2,3</sup>, Maria Lencastre<sup>3</sup>

<sup>1</sup>Universidade de Pernambuco (UPE), Caruaru – PE – Brasil

<sup>2</sup>REACT Research Labs, Caruaru – PE – Brasil

<sup>3</sup>Universidade de Pernambuco (UPE), Poli, Recife – PE – Brasil

{wbs, jonatan.lins}@upe.br, kayorenatocontato@gmail.com,  
mlpm@ecom.poli.br

**Resumo.** *O objetivo principal desta pesquisa é identificar os métodos de priorização de requisitos utilizados pelas startups pernambucanas, assim como descobrir uma possível correlação entre esses métodos e o sucesso das startups.*

### 1. Introdução

Jovens com grande potencial para criação de tecnologia estão cada vez mais buscando formas alternativas de se inserir no mercado de trabalho, tomando iniciativas empreendedoras para conquistar seu próprio espaço por meio de startups, muitas vezes antes mesmo da conclusão de um curso superior.

Segundo [Giardino *et al.* 2014] as startups possuem como característica a alta incerteza, além do fato de serem menos robustas que empresas consolidadas. Além disso deve-se observar que por muitas vezes, a startup é o primeiro ambiente profissional em que seus integrantes trabalham, portanto os mesmos nem sempre têm conhecimentos sobre Engenharia de Requisitos.

Aliando o fato de que algumas pesquisas recentes mostraram que em média, 60% das startups não sobrevivem mais que cinco anos, fica claro que cada processo de Engenharia de Software de uma startup deve ser realizado de forma cuidadosa. Um desses processos é o de priorização de requisitos, que pode ser fundamental para determinar a aceitação de um projeto, e consequentemente, seu sucesso.

### 2. Startups

Startups são empresas novas, até mesmo embrionárias ou ainda em fase de constituição, que contam com projetos promissores [SEBRAE 2017], inseridas em cenários diferenciados, onde as ideias de produtos ou serviços são trabalhadas em condições de extrema incerteza [Patz 2013], com tecnologias e mercados dinâmicos, múltiplas influências dos stakeholders, recursos limitados, pouca história de operação, juventude e imaturidade [Sutton 2000].

Outra característica é que em geral, os fundadores de startups são mais jovens que os administradores de empresas consolidadas, conseqüentemente sua inexperiência torna-se um fator de risco para a o sucesso de qualquer projeto. Além de que quase sempre os recursos são extremamente limitados [Sutton 2000].

### 3. Método

O instrumento utilizado para pesquisa é um questionário online composto por doze perguntas que englobam a forma de priorização de requisitos, o sucesso e dados de um projeto específico e da startup, assim como do colaborador. O mesmo foi submetido a um piloto para fins de calibração, e com base nas sugestões feitas, algumas questões foram aprimoradas.

O questionário contém perguntas a respeito de detalhes do processo de priorização de requisitos, incluindo a metodologia de desenvolvimento utilizada, para que haja um entendimento sobre o modo de trabalho dessas startups. Existem também perguntas que indicam se houve ou não sucesso ao realizar o projeto, para indicar se a forma de trabalho da startup foi eficiente. São feitas também perguntas demográficas em relação ao respondente, de forma a entender o nível de conhecimento no assunto.

O presente estudo está em andamento e o questionário será submetido para diversas startups de software de Pernambuco, para que um dos seus membros o responda. Com base nas respostas será possível ter um entendimento amplo sobre a forma de trabalho nas startups de software de Pernambuco. O questionário encontra-se disponível em <https://bit.ly/react-pr-survey>.

### 4. Conclusão e Próximos Passos

Esse trabalho visa realizar ampliar a amostragem e refinamento do instrumento apresentado no trabalho de Silva *et al.* (2018). Silva e outros pesquisadores exploraram as práticas de priorização de requisitos utilizadas em startups pernambucanas, assim como os desafios associados a essas práticas. O presente estudo realiza uma análise mais ampla, com foco nas startups pernambucanas, Na fase de análise dos dados, técnicas de análise temática [Cruzes e Dybå, 2011] serão utilizadas para analisar possíveis relações entre os dados coletados, na tentativa de identificar técnicas têm uma maior probabilidade de fornecer resultados positivos para um projeto.

Com base nos dados coletados, pode haver a oportunidade de arquitetar um método para auxiliar o processo de priorização de requisitos, aliando conceitos da teoria da simplicidade, com o intuito de tornar o processo mais simples, rápido, acessível e com mais chances de sucesso.

## Referências

- C. Giardino, M. Unterkalmsteiner, N. Paternoster, T. Gorschek, e P. Abrahamsson, “What do we know about software development in startups?” *IEEE software*, vol. 31, no. 5, pp. 28–32, 2014.
- Cruzes, D. S. e Dybå, T. (2011). Recommended Steps for Thematic Synthesis in Software Engineering. In *International Symposium on Empirical Software Engineering and Measurement*.
- SEBRAE. (2017) “O que é uma empresa startup?”, <https://www.sebraemg.com.br/atendimento/bibliotecadigital/documento/texto/o-que-e-uma-empresa-startup>, acesso em: 11 de outubro de 2018.
- G. Silva, W. Santos, e M. Lencastre, “Priorização de Requisitos em Startups de Software - Um Estudo Qualitativo em Startups do Armazém da Criatividade (Porto Digital)”.
- J. Maurício, M. Lencastre, “G-4REPrioritization: Um Guia para Apoio à Escolha de Técnicas de Priorização de Requisitos”.
- M. Patz, “Lean startup: adding an experimental learning perspective to the entrepreneurial process,” Master’s thesis, University of Twente, 2013.
- S. M. Sutton, “The role of process in software start-up” *IEEE software*, vol. 17, no. 4, pp. 33–39, 2000.
- Silva, G. Santos, W., Lencastre. M. Priorização de Requisitos em Startups no Armazém da Criatividade (Porto Digital): Resultados Preliminares. Anais do WER18 - Workshop em Engenharia de Requisitos, Rio de Janeiro, Brasil, Setembro 05-06, 2018.

## Detecção de crises epilépticas utilizando *deep learning* com *fog e cloud computing*

Élisson Rocha<sup>1</sup>, Kayo Monteiro<sup>1</sup>, Émerson Silva<sup>1</sup>, Guto Leoni Santos<sup>2</sup>,  
Wylliams Santos<sup>1</sup>, Patricia Takako Endo<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Engenharia de Computação  
Universidade de Pernambuco (UPE), Recife, Brasil

<sup>2</sup>Universidade Federal de Pernambuco (UFPE), Recife, Brasil

{esr2, khcm, ers2}@ecomp.poli.br, {wbs, patricia.endo}@upe.br

{guto.leoni}@gpert.ufpe.br

**Resumo.** *Hoje, existem mais de 900 milhões de pessoas que sofrem com algumas doenças crônicas em todo o mundo. Por essa razão, sistemas de e-health estão sendo desenvolvidos para prover uma melhor qualidade de vida para os pacientes. Deep learning tem sido uma técnica importante incorporada nesses sistemas para prever e classificar dados sem a necessidade de um especialista atuando 24 horas. No entanto, a combinação de sistemas de e-health e técnicas de deep learning traz vários desafios que precisam ser superados. Com base nesse contexto, este trabalho em andamento propõe um sistema e-health baseado em fog e cloud computing, usando o deep learning para prever convulsões epilépticas.*

### 1. Introdução

A epilepsia é uma das doenças neurológicas mais comuns no mundo, com aproximadamente 50 milhões de pessoas diagnosticadas. Essa doença é um distúrbio cerebral crônico que afeta pessoas de todas as idades [Megiddo et al. ] e também influencia a qualidade de vida, pois as mesmas precisam conviver com a possibilidade de uma convulsão ocorrer a qualquer momento. Diante desse cenário, um sistema para detecção automática de crises pode ajudar a qualidade de vida das pessoas. Considerando o desenvolvimento de um sistema que prevê de forma segura os períodos com maior probabilidade de ocorrência de crises convulsivas, novas oportunidades para desenvolvimento de mecanismos para notificar os pacientes podem surgir e, com isso, evitar atividades arriscadas, como nadar ou dirigir [Hosseini et al. 2016].

A *brain-computer interface* (BCI) permite a aquisição dos dados cerebrais do paciente usando eletrodos no couro cabeludo, capturando atividades elétricas espontâneas do cérebro baseadas em eletroencefalogramas (EEG) [Hosseini et al. ]. Para o desenvolvimento de um sistema para detecção de crises epilépticas, existem vários desafios. O primeiro são os dados; o EEG é um sinal não estacionário, de modo que padrões e convulsões normais podem mudar de paciente para paciente [Hosseini et al. 2016]. O volume de dados gerados por um EEG também é outro desafio; por um lado, o espaço-tempo de alta resolução, ajuda na previsão. Por outro lado, gera o problema de *big data*, que é necessário um local seguro para armazenar os dados do paciente continuamente [Hosseini et al. 2016]. Como consequência, a largura de banda da rede para enviar dados

também é um desafio [Hosseini et al. 2017], mas ambos, rede e armazenamento de dados, podem ser resolvidos por uma solução baseada em *cloud*. Por fim, a aplicação para detecção de crises epiléticas em tempo real requer um diagnóstico rápido para evitar que o paciente realize atividades arriscadas [Hosseini et al. ]. Com base nesses desafios, este trabalho em andamento propõe um sistema *e-health* que utiliza um modelo de *deep learning* (DL) para prever convulsões aproveitando as infraestruturas de *fog* e *cloud computing*.

## 2. Detecção de crises epiléticas com DL

Vários artigos propuseram sistemas para detectar convulsões epiléticas [Hosseini et al. 2016][Hosseini et al. 2017][Hosseini et al. ], no entanto, nenhum utilizou *fog* e *cloud computing* para analisar disponibilidade. Neste trabalho, é proposto um sistema de *e-health* para prever convulsões epiléticas com análise em tempo real, usando as vantagens da *fog* e *cloud computing*, como mostrado na Figura 1.

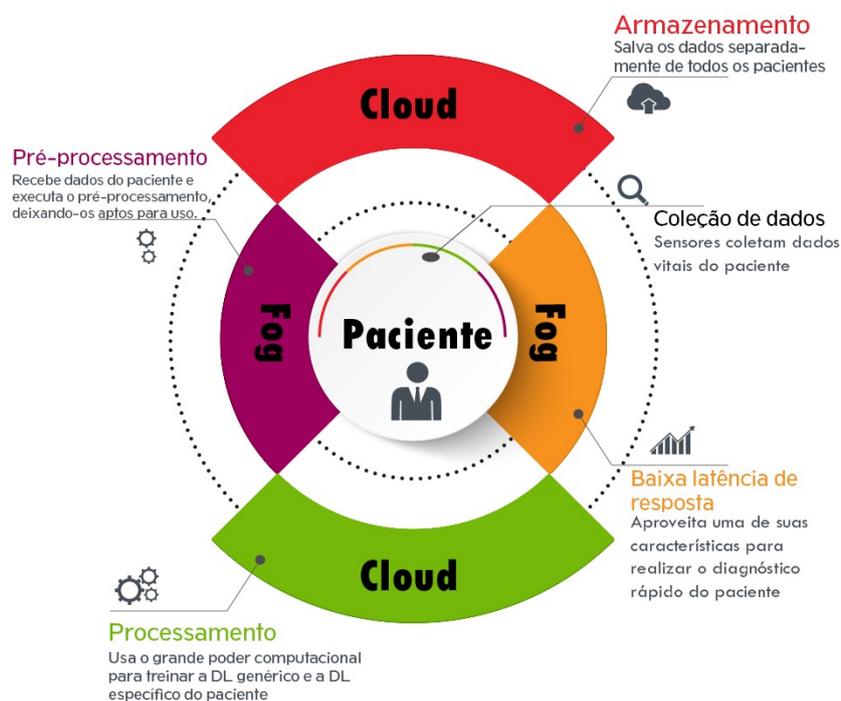


Figura 1. Proposta: Sistema de *e-health* para detecção de epilepsia

Usa-se a *cloud* devido à sua grande capacidade de armazenamento e de processamento, para armazenar dados genéricos existentes na literatura e também novos dados coletados de pacientes individuais. Além desses, também usa-se *cloud* para treinar o modelo de DL, uma vez que este processo requer alta computação. O *fog* é responsável pelo pré-processamento dos dados do paciente e também pelo diagnóstico, utilizando o modelo DL previamente treinado na *cloud*. Dessa maneira, o *fog* e a *cloud* são usados para fornecer análise em tempo real, baixa latência, escalabilidade e alta disponibilidade. Basicamente, a proposta consiste em três camadas: camada de coleta de dados, camada de *fog computing* e camada de *cloud computing*, conforme detalhado abaixo.

- **Camada de coleta de dados:** Nessa camada, os sensores que estão em contato direto com o paciente, coletando seus sinais vitais, como atividades elétricas es-

pontâneas do cérebro, frequência cardíaca e movimento. Depois de coletados, os dados são enviados para a camada de *fog computing*.

- **Camada de *fog computing*:** Esta camada é responsável por diagnosticar e alertar o paciente ou sua família. Recebe os dados do paciente, realiza o pré-processamento dos dados, executa o modelo DL para classificação e, se necessário, gera um alerta. As camadas de *fog* também são responsáveis por enviar os dados pré-processados e seus resultados de classificação para a camada de *cloud computing*. A camada de *fog computing* é implantada perto do cliente, para que possamos evitar problemas com a largura de banda da rede e obter resposta com baixa latência.
- **Camada de *cloud computing*:** Esta camada realiza o treinamento da DL, armazenamento dos dados do paciente e monitoramento de longo prazo. O processo de treinamento da DL requer alta computação, de modo que é possível aproveitar as vantagens da *cloud*, como alta capacidade de processamento e alta disponibilidade. Inicialmente, o modelo DL é treinado com dados genéricos existentes. Com o tempo, pode-se gerar uma DL específica para cada paciente, alimentando o modelo DL com dados do paciente monitorado.

O sistema *e-health* proposto é capaz de generalizar o modelo de DL para obter previsões satisfatórias em novos pacientes. Considerando que o EEG não tem um padrão definido para todos os pacientes, é importante que o modelo de DL aprenda e se torne especialista no paciente observado, diminuindo os falsos positivos. Este objetivo é realizado através dos dados armazenados na *cloud* e seu monitoramento contínuo.

### 3. Próximos passos

Como próximos passos, foi planejado analisar quais serão os desafios para implementação dessa proposta. Com isso, distribuir tarefas para que os desafios sejam superados, e ao fim, comparar a proposta desenvolvida com trabalhos da literatura com ênfase na eficiência.

### Agradecimentos

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro desta pesquisa.

### Referências

- Hosseini, M.-P., Pompili, D., Elisevich, K., and Soltanian-Zadeh, H. (2017). Optimized deep learning for eeg big data and seizure prediction bci via internet of things. *IEEE Transactions on Big Data*, 3(4):392–404.
- Hosseini, M.-P., Soltanian-Zadeh, H., Elisevich, K., and Pompili, D. (2016). Cloud-based deep learning of big eeg data for epileptic seizure prediction. In *IEEE GlobalSIP*, pages 1151–1155. IEEE.
- Hosseini, M.-P., Tran, T. X., Pompili, D., Elisevich, K., and Soltanian-Zadeh, H. Deep learning with edge computing for localization of epileptogenicity using multimodal rs-fmri and eeg big data. In *IEEE ICAC*.
- Megiddo, I., Colson, A., Chisholm, D., Dua, T., Nandi, A., and Laxminarayan, R. Health and economic benefits of public financing of epilepsy treatment in india: An agent-based simulation model. *Epilepsia*, 57(3).