



UNIVERSIDADE FEDERAL DO PARÁ  
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

VITOR PINHEIRO ALVES

PREDIÇÃO DE COMPORTAMENTO DE USUÁRIOS ORIUNDOS DO  
MARKETING DIGITAL POR MEIO DE REDES NEURAS ARTIFICIAIS E  
APRENDIZADO SUPERVISIONADO

Tucuruí

2019

VITOR PINHEIRO ALVES

PREDIÇÃO DE COMPORTAMENTO DE USUÁRIOS ORIUNDOS DO  
MARKETING DIGITAL POR MEIO DE REDES NEURAIS ARTIFICIAIS E  
APRENDIZADO SUPERVISIONADO

Texto dissertativo apresentado ao Programa de Pós-Graduação em Computação Aplicada do Núcleo de Desenvolvimento Amazônico em Engenharia, da Universidade Federal do Pará, como requisito para a obtenção de título de Mestre em computação aplicada.

Dr. Otávio Noura Teixeira

Tucuruí

2019

VITOR PINHEIRO ALVES

PREDIÇÃO DE COMPORTAMENTO DE USUÁRIOS ORIUNDOS DO  
MARKETING DIGITAL POR MEIO DE REDES NEURAIAS ARTIFICIAIS E  
APRENDIZADO SUPERVISIONADO

Texto dissertativo apresentado ao Programa de  
Pós-Graduação em Computação Aplicada do  
Núcleo de Desenvolvimento Amazônico em  
Engenharia, da Universidade Federal do

Pará, como requisito para a obtenção de título de  
mestre

Dr. Otávio Noura Teixeira

Aprovada em \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_

BANCA EXAMINADORA

---

Doutor / Otávio Noura Teixeira / UFPA – Orientador

---

Doutor / Cleison Daniel Silva / UFPA

---

Doutor / Edson Koiti Kudo Yasojima / UFRA

Tucuruí

2019

## RESUMO

O sucesso em atrair clientes a partir de técnicas de marketing gera um problema bilionário e um dos maiores problemas em vendas que é justamente escolher entre os muitos interessados, quais possuem maior probabilidade de fechamento. Este trabalho utiliza redes neurais artificiais para analisar o *dataset* gerado a partir de técnicas de marketing digital e classificar quais clientes tem maior chance de fechamento em vendas e quais devem ser descartados. A rede neural acerta aproximadamente 70% dos casos entre 3.541 registros processados.

Palavras Chaves: inteligência artificial, aprendizado de máquina, redes neurais, marketing digital, *inbound* marketing.

## ABSTRACT

Success in attracting customers using marketing techniques creates a billionare problem wich is one of the most difficult that is selecting among the many prospects, which are more likely to become a customer. This work uses artificial neural networks to analyze the dataset generated from digital marketig techniques and classify which prospects have a greater chance to become a customer and which ones should be discarded. The Neural Network scores approximately 70% of cases among 3,541 records processed

Keywords: artificial intelignce, machine learning, neural networks, digital marketing, inbound marketing.

## Lista de Imagens

<b>Figura 1- Funil de vendas. ....</b>	<b>14</b>
<b>Figura 2 – Modelo de neurônio artificial. Fonte: Haykin (2007). ....</b>	<b>18</b>
<b>Figura 3 Modelo de rede neural simplificada. ....</b>	<b>18</b>
<b>Figura 4 - Modelo de rede neural de múltiplas camadas.....</b>	<b>19</b>
<b>Figura 5 Fluxo de trabalho da pesquisa .....</b>	<b>21</b>
<b>Figura 6 – representação de público por idade, fonte: youtube.....</b>	<b>22</b>
<b>Figura 7 – Representação de público por gênero, fonte: youtube .....</b>	<b>22</b>
<b>Figura 8 – Gráfico de comparação de Palavras-Chave. fonte: Google Trends .....</b>	<b>23</b>
<b>Figura 9 – Painel de gestão de vídeos. Fonte: Youtube.....</b>	<b>24</b>
<b>Figura 10 - Gráfico de retenção de público com uso história, Fonte: youtube.....</b>	<b>24</b>
<b>Figura 11 - Gráfico de retenção de público sem uso história, Fonte: Autor.....</b>	<b>25</b>
<b>Figura 12- Vídeo com gatilho de ameaça no título .....</b>	<b>25</b>
<b>Figura 13 - Exemplo do uso do ameaçador para conseguir mais atenção. Fonte: Youtube. .....</b>	<b>26</b>
<b>Figura 14 - Meme do autor com mais de 20 mil de alcance.....</b>	<b>27</b>
<b>Figura 15 – Meme voltado ao público estudado neste trabalho.....</b>	<b>27</b>
<b>Figura 16 – Infográfico de teste A/B para títulos diferentes .....</b>	<b>29</b>
<b>Figura 18 – Configuração base da rede neural.....</b>	<b>35</b>
<b>Figura 19 – Gráfico de crescimento e acessos diários no youtube .....</b>	<b>36</b>
<b>Figura 20 - Gráfico de acessos ao youtube por um período de 3 meses .....</b>	<b>38</b>

## Lista de Tabelas

<b>Tabela 1 – E-mails e gatilhos de persuasão.....</b>	<b>31</b>
<b>Tabela 2 – Tabela de eventos da primeira conversão .....</b>	<b>33</b>
<b>Tabela 3 – Tabela de provedores de e-mail.....</b>	<b>33</b>
<b>Tabela 4 – Tabela de referências para dias da semana.....</b>	<b>34</b>
<b>Tabela 5 – Comparação entre acertos de redes neurais. ....</b>	<b>37</b>

Lista de Equiações

**Equação 1 - Somador ..... 17**



## Sumário

<b>1. Introdução</b> .....	11
1.1 Objetivos .....	13
1.1.1 Objetivo Geral.....	13
1.1.2 Objetivos Específicos.....	13
<b>2 Fundamentação teórica</b> .....	<b>13</b>
2.1 Marketing e Marketing Digital.....	13
2.1.1 Funil de vendas.....	14
2.1.1.1 Atrair .....	15
2.1.1.2 Gerar Leads .....	15
2.1.1.3 Gerar vendas.....	15
2.1.1.4 Fidelizar.....	15
2.1.2 Persuasão para maior atração no topo de funil.....	16
2.2 Neurônio Artificial .....	16
2.2.2 Tipos de rede .....	18
2.2.3 Número de camadas .....	18
2.2.4 Algoritmo de Retropropagação ( <i>BackPropagation</i> ) .....	19
<b>3 Metodologia</b> .....	<b>20</b>
3.1 Definição do público .....	21
3.2 A procura de interesses e palavras-chave.....	22
3.3.2 Histórias de mistério.....	24
3.3.3- O Ameaçador .....	25
3.3.4 O meme .....	26
3.4 Acesso aos próximos passos do funil de vendas .....	28
3.5 Testes A/B .....	28
3.6 Ferramentas .....	31
3.6.1 RD Station Marketing .....	31
3.6.2 Tensor Flow .....	31
3.6.3 PostgreSQL .....	31
3.6.3 Google Trends.....	32
3.7 Hipótese de comportamentos .....	32
3.7.1 Primeiro evento de conversão .....	32
3.7.2 Provedor de email.....	33
3.7.3 Origem da primeira conversão .....	33

3.7.4 Dia da semana da conversão .....	34
3.7.5 Hora da conversão.....	34
3.7.6 Número de conversões. ....	34
3.8 Configuração da Rede Neural .....	35
<b>4 Resultados</b> .....	<b>36</b>
4.1 Abertura de e-mails e gatilhos da persuasão .....	36
4.2 Acertos da Rede Neural.....	36
<b>5 Considerações finais e discussões</b> .....	<b>37</b>
5.1 Preparação para o trabalho .....	37
5.2 Métricas e observação dos comportamentos .....	38
5.3 Pré-processamento dos dados e o uso de ferramentas.....	38
5.4 Considerações gerais.....	39
5.5 Trabalhos futuros .....	40
<b>6 Referências</b> .....	<b>41</b>

## 1. Introdução

É possível resolver problemas sociais a partir de empresas? De acordo com Porter (2013), Empresas podem resolver problemas sociais por meio de empregos, tecnologia e de várias outras maneiras além do pagamento de impostos. Para Augusto (2018), a principal forma de uma empresa gerar receita é vender. Tais pensamentos podem levar a crer que a eficiência em comunicação e vendas são fatores de extrema importância não só para empresas, mas para qualquer entidade que precise de recursos como órgãos do governo ou a própria universidade.

Comparado com meios de comunicação tradicionais, a internet ajuda com muito mais eficácia os comunicadores a atingir públicos alvo específicos. Seja para realizar uma venda ou para conseguir voluntários para um projeto. Ferramentas e técnicas de marketing digital são utilizadas para captação de adeptos cada vez em maior escala Adolpho (2011).

Uma das principais técnicas utilizadas para captar e gerenciar contatos que podem se tornar futuros clientes é a técnica do funil de vendas que nada mais é que um processo que guia o cliente até um objetivo. Ramos (2008), LeeFlang (2014).

A verdade é que pode haver muitos passos para converter um cliente ao utilizar Marketing Digital assim como em uma venda tradicional. Contudo guiar o cliente desde a descoberta do produto até a compra pode ser feito de maneira automática por meio da internet e outros meios digitais.

Quando se trata de atração de clientes, existe uma técnica muito peculiar que dificilmente poderia ser utilizada sem a ajuda da tecnologia. Para Kotler (2017,p 174):

“Marketing de conteúdo é uma abordagem que envolve criar, selecionar, distribuir e ampliar conteúdo que seja interessante, relevante e útil para um público claramente definido com o objetivo de gerar conversas sobre esse conteúdo.”

O processo de criação de conteúdo para a internet é bastante custoso para as empresas. Precisa-se de pesquisa, profissionais, tempo e muito estudo de público para que os materiais sejam de interesse das pessoas as quais são direcionados. Além das pesquisas tradicionais sobre se o assunto do conteúdo é relevante, outros fatores incidem sobre o conteúdo para que ele se torne um sucesso de conversão (Pulizzi, 2017).

Aplicando as técnicas certas de marketing, é possível ter sucesso na captação de possíveis clientes que podem ou não estar maduros para o processo de vendas. No início da atual pesquisa, os motores de marketing digital convertiam por dia cerca de 20 novos possíveis clientes, já último dia de pesquisa o motor de marketing estava convertendo

em torno de 80 registros diários, tendo um aumento de mais de 200% em conversões em menos de um ano.

Empresas com estratégias de marketing digital definidas podem atender esses possíveis clientes que chegam via internet por telefone, e-mail, mensageiros, etc. Esse novo processo de atendimento é chamado de *inside sales*, Holloway & Seley (2008). Por tanto a equipe de *inside sales* é a equipe responsável por separar os melhores contatos prospectados e investir tempo no atendimento desses contatos para torná-los novos clientes.

Entende-se que o quanto maior o sucesso em atrair possíveis clientes da camada de atração de marketing, mais caro fica para atender a quantidade grande de pessoas assim como mais difícil é para separar os clientes mais suscetíveis a vendas dos menos suscetíveis.

O acúmulo de leads piora aos finais de semana, já que as empresas não estão em horário comercial para atendimento e mesmo assim aumenta-se a quantidade de possíveis clientes. Fazer uma análise do funil de vendas é crucial para o sucesso de uma empresa, especialmente se tratando de predição dos melhores e piores clientes, Yan et al (2015). Em pouco tempo analisar cada um dos interessados se torna humanamente impossível.

Para classificação dos melhores entre os possíveis clientes com o decorrer dos anos, tornou-se mais fácil o uso de técnicas de análise de dados, justamente pelo grande número de dados disponíveis. Para Taghizadeh (2017), o advento das grandes bases de dados (*Big Data*) ajuda a aumentar a quantidade de estudos que utilizam técnicas de predição como estatísticas e redes neurais para gerar novos *insights* em diferentes mercados.

Uma das maneiras de detectar se um usuário é um cliente interessado é prever se ele é capaz de clicar em um conteúdo novo, De acordo com McMahan, et al (2013), a predição do CTR (*click through rate*) ou simplesmente taxa de clique é um problema multi bilionário central da indústria de anúncios online. Certamente não só mostrar os melhores anúncios, mas também investir energia para que os anúncios certos apareçam para a pessoa certa dentre bilhões de usuários no intuito de que o click esteja com a maior probabilidade possível. Redes neurais melhoram a análise preditiva de uma grande variedade de aplicações para negócios, Ha & Cho et al (2005).

Existem muitas maneiras dentre as maneiras de predição de comportamento em clientes como o modelo de rede neural artificial com aprendizado supervisionado. As redes neurais são implementações derivadas dos modelos neurofisiológicos que consistem em um conjunto de dados de entrada e saída ligados em forma de rede, sendo que muitas pesquisas apontam que são melhores modelos de previsão que os manuais feitos por humanos, Lee e Shih (2011).

Nesse contexto, este trabalho tem como proposta aplicar redes neurais artificiais no contexto de marketing digital para classificar os melhores contatos dentre possíveis clientes em um *dataset* criado durante a pesquisa.

## **1.1 Objetivos**

### **1.1.1 Objetivo Geral**

- Desenvolver um método capaz de classificar usuários de marketing digital entre possíveis clientes e possíveis fracassos em venda.

### **1.1.2 Objetivos Específicos**

- Medir resultados de conversão entre conteúdos de diferentes configurações.
- Desenvolver uma rede neural capaz de aprender padrões entre comportamentos de clientes mais e menos relevantes.
- Provar a hipótese que o comportamento inicial em meio digital gera um padrão de previsão para sucesso em vendas e fracassos.

## **2 Fundamentação teórica**

### **2.1 Marketing e Marketing Digital**

Para Kotler(2012) marketing pode ter vários conceitos diferentes que se aproximam. Há quem diz que marketing é um processo para vender produtos. Alguns autores, inclusive o próprio Kotler ligam Marketing a lucro, contudo o conceito de marketing pode ser bem mais abrangente, já que pode englobar todo o processo de uma entidade para conseguir sucesso seja ela com ou sem fins lucrativos.

De acordo com a Sociedade Americana de Marketing, em Kotler (2012, p. 4, apud AMA, 2011):

“O Marketing é uma função organizacional e um conjunto de processos que envolvem a criação, a comunicação e a entrega de valor para os clientes, bem como a administração do relacionamento com eles de modo que beneficie a organização e seu público interessado”.

Devido as mudanças tecnológicas, é possível notar que as pessoas buscam mais informações sobre os produtos e serviços na internet e as empresas passam a criar conteúdos para atrair esse público e conduzi-lo a vendas, Järvinen (2016).

Para Adolpho (2011) uma loja de acessórios para hóquei no gelo situada no sertão do Ceará tem uma possibilidade de sucesso muito baixa, pois o público daquela loja raramente existiria lá. Mas se essa loja ainda que situada no mesmo lugar trabalhasse seu marketing para alcançar pessoas do Brasil inteiro com mais de 200 milhões de habitantes assim como o mundo inteiro, as suas chances aumentariam exponencialmente. Dessa forma a internet pode ser considerada um meio de comunicação de massa e de nicho ao mesmo tempo.

Kotler (2017), afirma que o Marketing Digital, nada mais é que uma parte do Marketing tradicional e para que ele funcione é preciso utilizar todos os conceitos dos métodos tradicionais de Marketing. É claro que ao analisar Kotler (2017) e Adolpho (2011), é possível destacar que Marketing Digital é Marketing tradicional com um poder de alcance muito maior e com um custo muito menor.

### 2.1.1 Funil de vendas

Uma das mais eficientes ferramentas de marketing de conteúdo é o funil de vendas, que nada mais é que uma técnica emprestada da área de vendas tradicional e que ajuda o administrador a se situar em que estágio o cliente encontra-se para fechar uma venda e ele pode variar um pouco dependendo do autor, D'Haen & Van den Poel (2013). A representação do funil de vendas do marketing digital utilizado neste trabalho, pode ser visualizado esquematicamente de maneira mais fácil na figura 1.

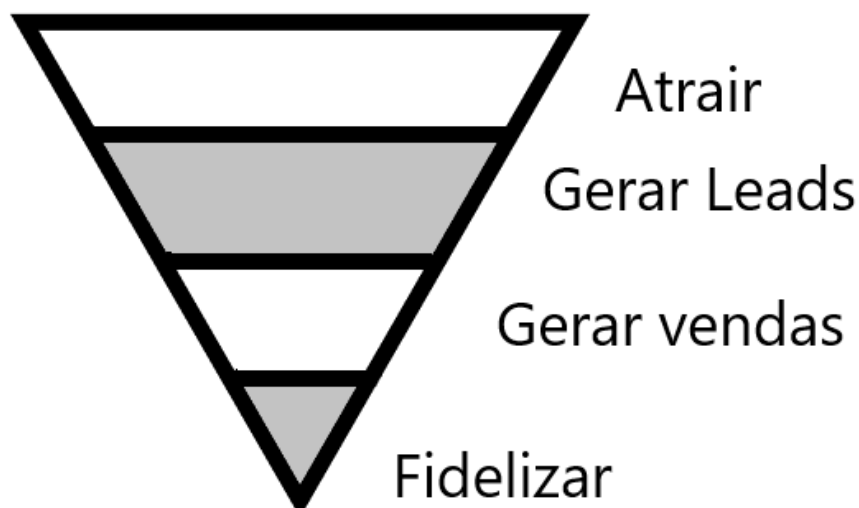


Figura 1- Funil de vendas.

### **2.1.1.1 Atrair**

Prospectar clientes é uma das tarefas mais difíceis e tediosas, Moncrief & Marshall (2015). A atração na técnica de funil de vendas é fazer com que um usuário se interesse por algum conteúdo. Para atrair um usuário, o autor do material precisa estar muito atento para o que gera valor e qual motivo o seu público teria para acessar o conteúdo.

Os usuários podem ter acesso a um conteúdo de várias formas conforme a mídia que usa, assim como é possível acessá-lo em vários formatos como artigos, e-books, áudios, vídeos e até simples imagens como infográficos. Entender quais as melhores mídias e quais as melhores iscas também são parte do estudo desse trabalho.

### **2.1.1.2 Gerar Leads**

*Leads* no jargão do marketing são usuários que se dispõem a deixar seus dados em troca de algo que lhe gere valor. O comportamento varia de acordo com o público alvo que se comporta de maneiras diferentes para se tornar um lead. O importante é que o usuário que foi atraído na primeira etapa entenda que o conteúdo ofertado para ele seja tão útil que valha a pena entregar seus dados em troca desse conteúdo. Na Literatura Lead pode ser chamado de Prospect, ou possível cliente. De acordo com Ryan (2017, p 65): “Leads qualificados da internet convertem em 14,6% em novos clientes. E Leads captados fora dela, ou seja, no mundo real, convertem em apenas 1,7%”.

### **2.1.1.3 Gerar vendas**

À medida que o agora lead consome vários conteúdos é possível guiá-lo a um fundo de funil. Em uma empresa privada, por exemplo, é possível guiar o Lead a se tornar um cliente; já em uma instituição de caridade é possível levar o usuário a se tornar um doador ou um voluntário. Esse processo pode ser finalizado automaticamente ou o cliente pode fechar a venda com um vendedor utilizando qualquer meio, Holloway & Seley (2008). Empresas precisam estar sistematicamente avaliando a performance do funil de vendas e a qualidade dos leads em cada estágio do funil Yan Et Al (2015).

### **2.1.1.4 Fidelizar**

Conteúdos de fidelização são conteúdos lançados para clientes que já consumiram vários outros conteúdos e já fazem parte da rede, esses clientes passam a fazer parte de uma lista de pessoas que vão receber periodicamente conteúdos bem selecionados, pois já navegaram tanto pelas páginas que já é possível entender o que

mais interessa ao cliente. Clientes que já compraram também podem ser potenciais compradores de outros produtos no futuro Järvinen& Taiminen (2015).

### **2.1.2 Persuasão para maior atração no topo de funil**

De acordo com Cialdini (2012), há 6 princípios de persuasão que podem ser utilizados para convencer pessoas a realizarem uma ação como por exemplo abrir um e-mail ou tornar-se um lead, as categorias são: reciprocidade, coerência, aprovação social, afeição ou empatia, autoridade e escassez.

Kahneman (2011) afirma que existem dois sistemas que definem como o ser humano toma decisões. O sistema 1 e o sistema 2. O sistema 1 é herdado de nossos ancestrais e toma decisões baseadas em correr ou enfrentar, esse sistema é automático, qualquer coisa que lhe gere prazer ou dor chama rapidamente sua atenção. Já o sistema 2 é da evolução humana, mais lento, raciocina analiticamente antes de tomar uma decisão. Contudo para economizar energia por envolver menos raciocínio, o organismo usa o sistema 1 a maior parte das vezes.

Em sua obra mais recente, Cialdini (2017) já afirma que existem outros 6 princípios que podem trabalhar antes mesmo da persuasão que o próprio autor chama de **préssuasão**. Neste trabalho o conceito de *préssuasão* e do sistema 1 citado por Kahneman (2011) são usados para atrair atenção do indivíduo antes mesmo que ele seja convencido a entrar em um funil de vendas. Os princípios da préssuasão são: O sexual, o diferente, o ameaçador, a auto-referência, o inacabado e histórias de mistério.

Este trabalho prevê dar ainda mais impulso ao processo de vendas ao classificar automaticamente os melhores prospects entre os leads que ingressam na base de dados durante um dia. Para isso foi utilizada uma rede neural artificial de múltiplas camadas com algoritmo de *backpropagation*.

## **2.2- Rede Neural Artificial**

Redes neurais são modelos que imitam a estrutura funcional de um cérebro dividido e conectado por 3 camadas: Entradas, camadas escondidas e camadas de saída Ha & Cho et al (2005), Lee & Shih et al (2011), Hinton et al (2012).

Assim como o próprio cérebro, Redes neurais artificiais podem ser usadas para detectar padrões de forma automática em uma grande variedade de variáveis que seria muito complexo de identificar as relações entre esses dados de forma manual. “Redes Neurais Artificiais nada mais são que ferramentas de funções de aproximação que aprendem o relacionamento entre variáveis dependentes e independentes”, Smith (2000).

### **2.2.1 Neurônio Artificial**



O neurônio artificial se inspira totalmente no neurônio biológico, inclusive suas estruturas. O neurônio artificial é constituído em basicamente em 4 (quatro) partes: as Sinapses, o somador, o Bias e a função de ativação.

As sinapses no neurônio artificial são a inspiração do encontro entre o axônio e o dendrito do neurônio biológico, sua função é multiplicar o valor enviado pela última camada da rede ao valor do peso encontrado na sinapse. Além dos pesos, no neurônio artificial encontramos 3 características: O Somador, o Bias e a função de ativação (Kasabov, 1996).

O somador é um função que irá reunir todos os pesos e as entradas em uma só variável, Ela na verdade é soma de todas as sinapses.

$$\text{Net}_i(t) = \sum_{j=1}^n w_j \cdot x_j(t)$$

#### **Equação 1 - Somador**

O Bias é uma estrutura semelhante a sinapse, tem peso próprio e tem a função de equilibrar o sinal de saída que será propagado.

A função de ativação é aplicada à saída do neurônio ao final da rede ou antes de ser aplicada em outra camada da rede Hinton et al (2012). A função de ativação é responsável por determinar o valor de saída do neurônio, que restringe esse valor entre -1 e 1 ou entre 0 e 1 ou até mesmo entre -1 e 1. A partir da necessidade da rede, ou o problema que ela resolve, pode ser escolhido a melhor função de ativação, os neurônios de saída vão determinar o resultado do processamento da rede

Um exemplo é uma rede que determina se algo é verdadeiro ou falso, seria interessante utilizar uma função que resultasse em um valor entre 0 e 1, quando que ao mais próximo de 0, a rede tem mais certeza de ser falso, e 1 o resultado da rede é mais certo de ser verdadeiro.

Outro exemplo é quando o resultado da rede precisa determinar algo aproximado como uma previsão de temperatura dado certo ambiente. Então uma função de ativação entre -1 e 1 ajudaria a rede a aproximar-se com facilidade entre temperaturas negativas e positivas.

A função escolhida neste trabalho foi a função softmax, que nada mais é que uma normalização dos resultados em uma ordem probabilística para que todos somem 1, Costa (1996).

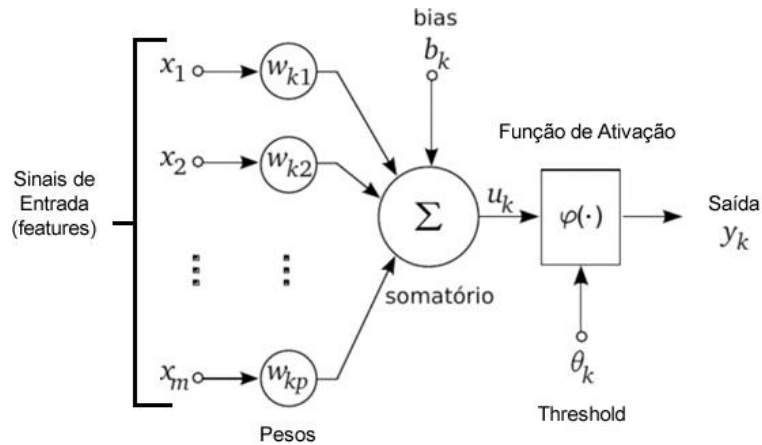


Figura 2 – Modelo de neurônio artificial. Fonte: Haykin (2007).

### 2.2.2 Tipos de rede

Redes neurais nem sempre foram consideradas boas soluções de inteligência artificial, contudo, ao longo do tempo evoluíram para soluções que resolvem problemas mais complexos do mundo real Marini et Al (2008). As redes neurais artificiais podem ser classificadas de acordo com 3 características: Número de camadas, topologia e conectividade.

### 2.2.3 Número de camadas

Para Haykin (2007) as redes podem ter no mínimo 2 camadas, contudo as consideradas redes de uma camada apenas são chamadas assim porque a camada de entrada existe mas não possui processamento. Por tanto, as redes de uma única camada possuem somente a camada de saída.

Para outras soluções existem redes de múltiplas camadas, o quanto maior o número de camadas, maior no processamento da rede e maior o consumo de recurso computacional para processamento da solução

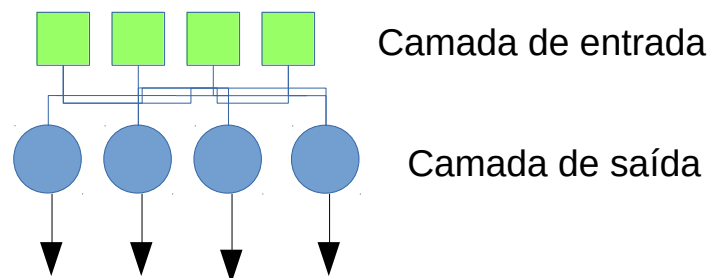


Figura 3 Modelo de rede neural simplificada.

Neste trabalho é usado o modelo de múltiplas camadas onde a rede é totalmente conectada, assim como a propagação do sinal é *feed-foward*, unidirecional do início ao final da rede.

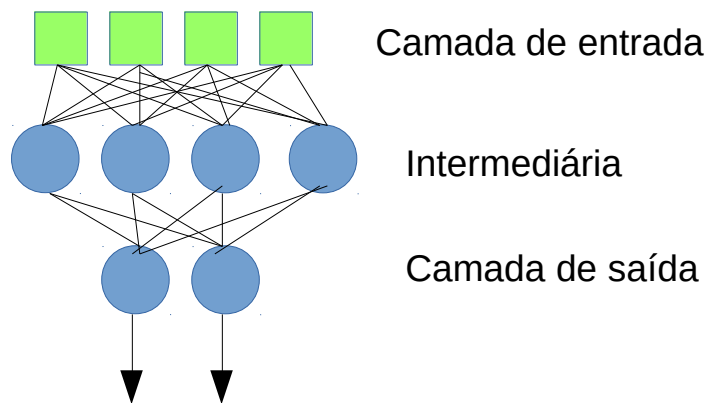


Figura 4 - Modelo de rede neural de múltiplas camadas

#### 2.2.4 Algoritmo de Retropropagação (*BackPropagation*)

*Backpropagation* possui esse nome justamente porque ao final da propagação direta, o sinal é novamente propagado mas em sentido contrário ao de início. Trata-se de um algoritmo de treinamento supervisionado de rede neural artificial de múltiplas camadas que treina a partir da propagação de dados de exemplos até o última camada da rede, ao final calcula-se a diferença entre o que a rede propagou como resposta e a saída esperada, através desse resultado o algoritmo atualiza os pesos das sinapses da última camada que passa pelas camadas ocultas até a primeira camada de neurônios.

Para Graupe(2005) o algoritmo de *back propagation* funciona da seguinte forma:

*“O algoritmo de backpropagation inicia, necessariamente com a computação da camada de saída, que é a única camada onde a saída desejável é acessível, já que as saídas da camada intermediária não o são.”*

E Munakata (2008) afirma que:

*“Nós ajustamos os pesos de forma que o vetor de saída y fique o mais próximo possível do vetor alvo t. Nós começamos isto a partir da camada de saída e voltando para a camada escondida, modificando os pesos, W', e então voltando da camada escondida para a camada de entrada, e modificando os pesos, W.*

O algoritmo de *Backpropagation* com todos os seus passos que vão da propagação do sinal até, medida do erro e atualização dos pesos nas camadas de saída e intermediárias é descrito por Rojas (2015).

### **3 Metodologia**

Este trabalho possui 3 etapas: Etapa de pré-processamento, etapa de treinamento e etapa de classificação dos indivíduos. A etapa de pré-processamento começa com definição de um público alvo específico; O planejamento de conteúdos de marketing digital para ser consumido por esse público e captura de dados dos que fornecem seus dados voluntariamente em troca de mais material aprofundado

A etapa final é a de coleta de todos os dados cadastrados e utilização para treinamento da rede neural no intuito de classificar os clientes mais aptos a atendimento. Ao final do treinamento, é avaliado se a rede acerta mais de 65% e caso o resultado seja insatisfatório, troca-se os campos selecionados para treinamento da rede e treina-se novamente. No processo de aprimoramento da configuração da rede neural, também são trocados quantidades de neurônios nas camadas intermediárias e a quantidade de épocas. A figura 5 representa o fluxo de trabalho

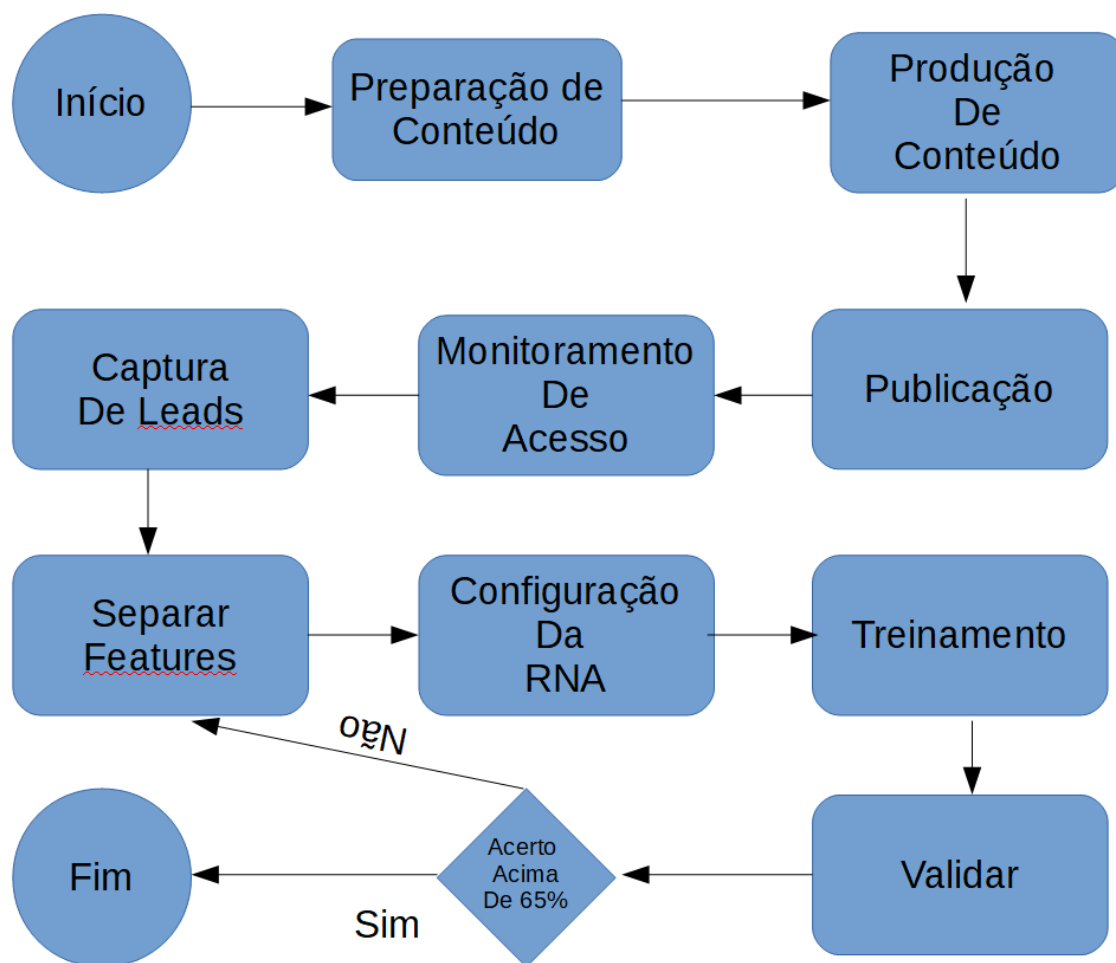


Figura 5 Fluxo de trabalho da pesquisa

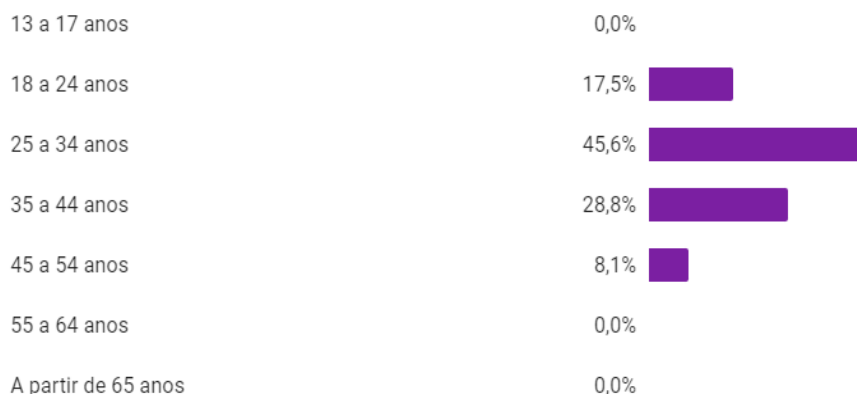
### 3.1 Definição do público

O número de pequenas empresas cresce muito rápido no Brasil, de acordo com o site Empresômetro (2019), há mais de 20 milhões de empresas ativas no país, o que leva a crer que há aproximadamente uma empresa para cada 10 habitantes.

Para melhor definir o público foi analisado o perfil informado pela plataforma youtube referente aos visitantes frequentes do canal disponível em <http://youtube.com/alvesvitor>.

## Idade

Tempo de exibição · Últimos 28 dias



**Figura 6 – representação de público por idade, fonte: youtube**

## Gênero

Tempo de exibição · Últimos 28 dias

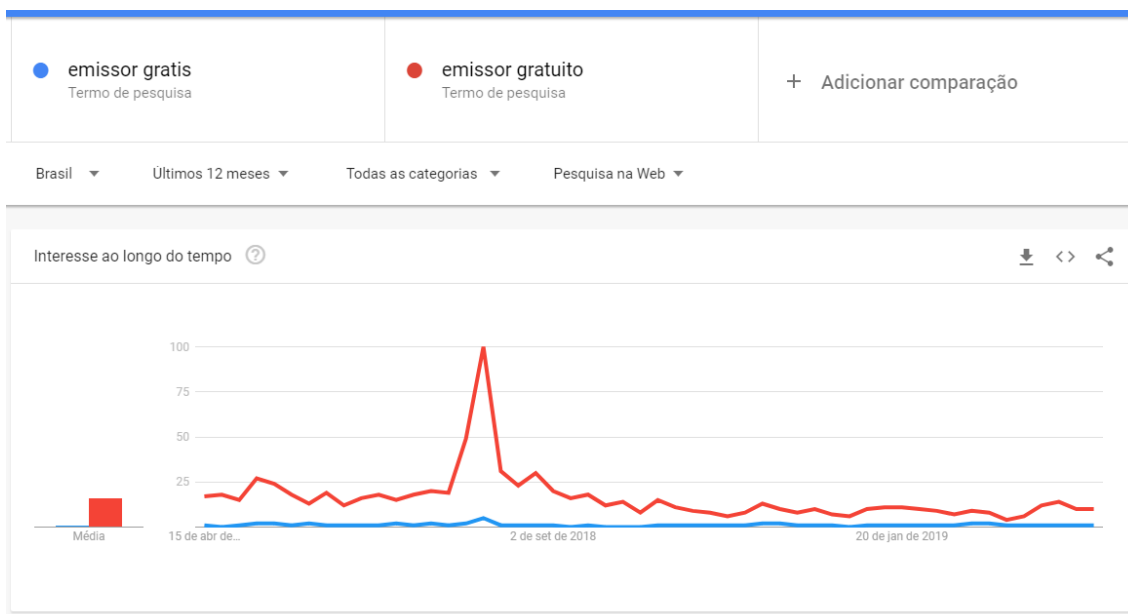


**Figura 7 – Representação de público por gênero, fonte: youtube**

### 3.2 A procura de interesses e palavras-chave

A legislação brasileira obriga que todas as empresas que não estiverem na situação de Microempreendedor individual sejam obrigadas a emitir nota fiscal eletrônica SEFAZ (2019). As notas fiscais eletrônicas são documentos digitais que comprovam um ato de compra, venda ou transferências de produtos e serviços. Sendo uma obrigatoriedade, é possível identificar na internet se pessoas estão interessadas nos termos referentes a emissão de notas.

Para comprovar a hipótese de que emissão de nota fiscal é um termo bastante procurado, ferramentas de marketing digital podem ser utilizadas. A mais usada neste trabalho é o *Google Trends* que permite comparar termos de buscas e identificar potencial público interessado. A Figura 8 é o gráfico que representa a diferença entre intenções de busca para o conjunto de palavras “emissor Gratuito” e “emissor Grátis”.



**Figura 8 – Gráfico de comparação de Palavras-Chave. fonte: Google Trends**

Antes de produzir qualquer conteúdo, é importante entender se o público está realmente interessado no que se tem a oferecer. Esse passo de comparação entre o interesse do público e o que o autor dos conteúdos pode produzir foi feito para todos os materiais produzidos.

### 3.3 Produção de conteúdo

Esse é o primeiro passo para o funil de vendas a Atração: Toda comunicação produzida é pensada em atrair a maior quantidade de público possível e fazê-lo consumir por mais tempo possível, para isso são usados os 12 gatilhos de Cialdini (2012) tanto no título do material quanto no corpo. Nos capítulos seguintes será mostrado que não só a afinidade do lead com o conteúdo faz com que ele abra um e-mail e seguir o processo de venda, mas também como esse e-mail é enviado e com quais gatilhos o público tem mais interesse, muitas vezes o mesmo título escrito de formas diferentes pode atrair mais ou menos interessados.

#### 3.3.1 Produção de vídeos

De acordo com Forbes (2017). O Google é o buscador com maior volume de buscas do mundo, e o Youtube é o segundo maior buscador do mundo e neste estudo é a principal fonte de tráfego. Até a finalização deste trabalho foram produzidos 55 vídeos que já alcançaram mais de meio milhão de visualizações.

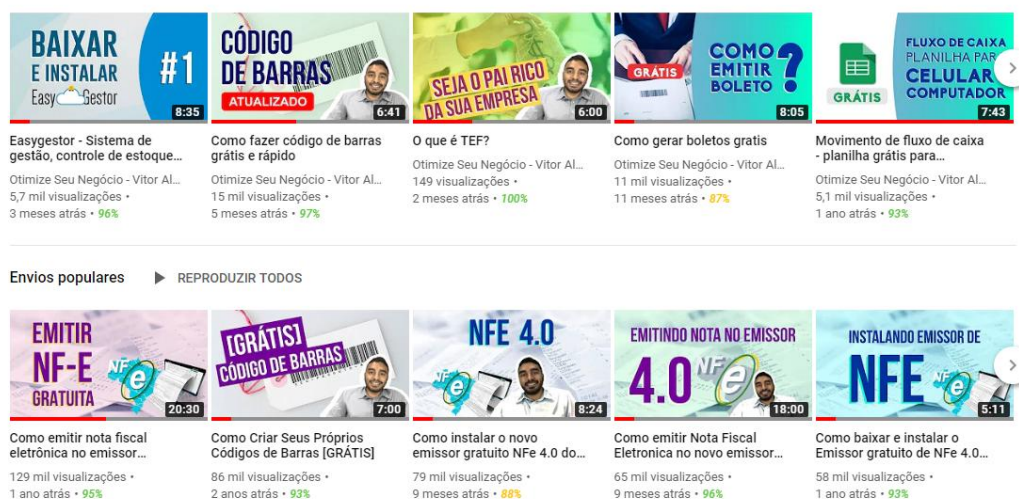


Figura 9 – Painel de gestão de vídeos. Fonte: Youtube

Para chamar atenção e manter a audiência atenta foram usados 3 técnicas básicas: As histórias de mistério, gatilho ameaçador e o meme. Esse último é interessante pois é muito usado na internet ao decorrer da pesquisa notou-se que ele é uma soma de vários gatilhos citados por Cialdini (2017).

### 3.3.2 Histórias de mistério

Para Gallo (2017), histórias não são só métodos de manter atenção das pessoas, mas também métodos de convencimento. Não é a toa que as indústrias mais ricas atualmente são o cinema e o videogame, que são nada mais que mídias que se utilizam de histórias para manter atenção das pessoas e influenciar marcas, modas e costumes. A figura 10 representa um gráfico de retenção de público em um vídeo com história de mistério e a figura 11 representa um gráfico de retenção de público sem o uso de histórias. A Figura 10 mostra que o público assiste até a despedida e na figura 11 o público diminui com o tempo.



Figura 10 - Gráfico de retenção de público com uso história, Fonte: youtube.

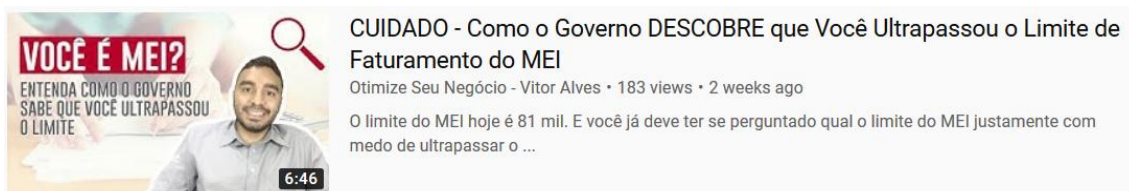




**Figura 11 - Gráfico de retenção de público sem uso história,  
Fonte: Autor.**

### 3.3.3- O Ameaçador

Kahneman (2011) afirma que o cérebro reage automaticamente certos tipos de estímulos, as reações mais comuns do ser humano é reagir no intuito de correr ou atacar. No marketing digital, palavras como CUIDADO e URGENTE são usadas para atrair mais visualizações. Neste trabalho o gatilho do ameaçador é usado tanto na chamada de atenção dos vídeos como nos e-mails encaminhados, e a presença deste aumenta a taxa de abertura em até 30% comparado com disparos sem gatilhos no título.



**Figura 12- Vídeo com gatilho de ameaça no título**



A APOSENTADORIA ACABOU E A PREVIDÊNCIA PRIVADA NÃO É A MELHOR SOLUÇÃO!

**Figura 13 - Exemplo do uso do ameaçador para conseguir mais atenção. Fonte: Youtube.**

### 3.3.4 O meme

Há tempos, é notado um fenômeno de comunicação: o meme. O meme nada mais é que o uso do Diferente + Inacabado + Afeição + O gatilho do prazer pelo humor. Ler a primeira frase de um meme leva automaticamente a ver a imagem e ler a segunda. Em todas as páginas administradas pelo autor desse trabalho, memes são estatisticamente os materiais que geram mais engajamento.

O meme da figura 14, criado pelo autor deste trabalho é uma prova que há tantos compartilhamentos e interações com a imagem que chegam ao alcance de mais de 20 mil pessoas. A empatia está justamente em falar com o público que entende a piada não só por morar na cidade de Belém, mas também por conhecer a série citada na imagem, assim como seus personagens reais e fictícios.

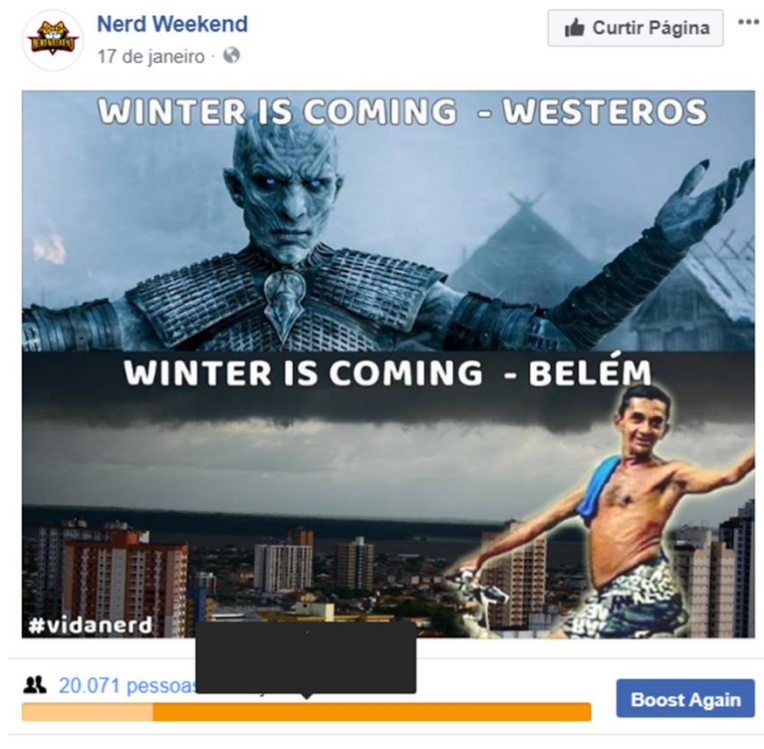


Figura 14 - Meme do autor com mais de 20 mil de alcance.

Já na figura 15, foi postado um meme para o público específico do canal deste trabalho, esse não alcança um público tão expressivo, mas só a sua presença já consegue o triplo do engajamento padrão.



Figura 15 – Meme voltado ao público estudado neste trabalho

### **3.4 Acesso aos próximos passos do funil de vendas**

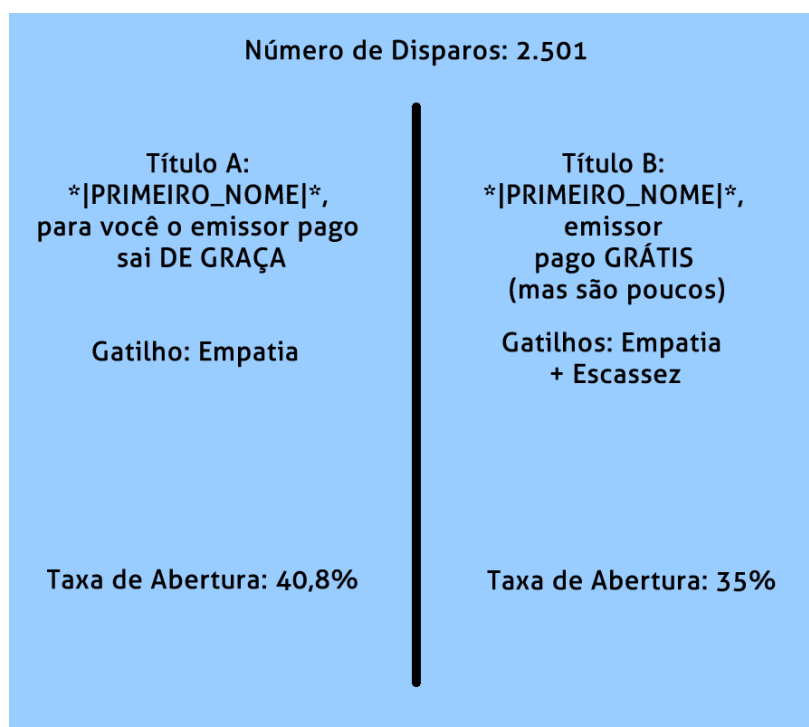
Para mensurar quantos leads consumiram o primeiro conteúdo e quantos conseguiram avançar na comunicação até o fundo do funil foram instaladas nas plataformas algumas ferramentas de rastreamento citadas no capítulo 3.6. Os dois principais indicadores que o lead pode tornar-se um possível cliente é a sua futura interação com novos conteúdos e abertura de emails, se o lead continua a interagir com os conteúdos, significa que além de gerar mais tráfego, esse lead pode seduzir-se por conteúdos de fundo de funil tornar-se cliente.

### **3.5 Testes A/B**

Para aumentar a taxa de abertura dos e-mails são utilizados gatilhos. Sem os gatilhos talvez não seria possível a classificação automática dada por este trabalho, já que a abertura de e-mails sem gatilhos mostra-se inferior aos e-mails enviados sem a técnica, com uma taxa muito baixa de abertura, haveria poucos exemplos de clientes que passaram ao próximo passo do funil de vendas, desbalanceado a proporção de leads com e sem potencial.

Foram feitos testes A/B com os disparos. Os testes consistem em selecionar um público específico e homogêneo, dividi-lo em duas partes iguais e mandar um título de e-mail diferente para cada grupo.

Cada grupo recebe um título com 1 ou 2 gatilhos de persuasão diferentes como mostra na Figura 16



**Figura 16 – Infográfico de teste A/B para títulos diferentes**

O fenômeno mais interessante sobre a abertura de e-mails é o uso dos gatilhos da auto-referência e empatia/afeição. 100 % das vezes que os gatilhos foram utilizados em um teste A/B, sempre houve um aumento de até 30%, sendo para este trabalho o gatilho mais importante para manter o cliente ativo no funil de vendas, maneira mais comum de utilizar os 2 gatilhos é incluir o nome do destinatário do e-mail no título da mensagem.

Toda pessoa fica em estado de alerta instantâneo quando seu nome é chamado. Chamar as pessoas pelo nome ou apenas falar diretamente com elas pelo uso do pronome “Você”, faz com que chame mais atenção, o que gera o gatilho de empatia já citado e a auto-referência. Cialdini (2017) afirma que a marca Coca-Cola ao colocar os nomes mais comuns das pessoas na lata de refrigerante teve um aumento de vendas instantâneo. De acordo com Knackfuss (2018), em alguns países notou-se o aumento significativo nas vendas e essa foi a maior campanha de personalização de produtos físicos em massa da história até então. Testes feitos com gatilhos podem ser notados na tabela de e-mails da tabela 1.

Título	Gatilho de entrada	Entregues Individual	Abertura individual	Taxa de abertura
* PRIMEIRO_NOME *, emita notas fiscais por R\$ 50	Empatia	1927	553	28,697%
* PRIMEIRO_NOME *, emita nota fiscal por R\$ 1	Empatia	1895	543	28,654%
* PRIMEIRO_NOME + Quero te avisar de uma	Afinidade	1251	447	35,73%

multa	+ Ameaçador			
* PRIMEIRO_NOME  + Emissor completo por R\$50,00	Afinidade	1250	353	28,24%
[AJE PARÁ] - Olha * PRIMEIRO_NOME * - Essa quarta, Escola de negócio do CESUPA	Afinidade + Afinidade	78	34	43,59%
[AJE PARÁ] - * PRIMEIRO_NOME * - Essa quarta, Escola de negócio do CESUPA	Afinidade	78	44	56,41%
* PRIMEIRO_NOME *, encontrei um certificado A1 barato!	Afinidade	665	173	26,02%
* PRIMEIRO_NOME *, certificado digital baixou de preço!	Afinidade	664	166	25,00%
* PRIMEIRO_NOME *, ainda quer criar código de barras?	Afinidade + Coerência	721	134	18,59%
* PRIMEIRO_NOME *, faça código de barras como meus clientes	Afinidade + Provação social	720	140	19,44%
[AJE PARÁ] - É hoje, Workshop sobre exportação na prática	Afinidade	77	27	35,06%
[AJE PARÁ] - * PRIMEIRO_NOME *, é hoje, Workshop sobre exportação na prática	Afinidade + Escassez	77	34	44,16%
* PRIMEIRO_NOME *, não compre seguidores no Instagram!	Empatia + Inacabado	4828	446	9,238%
* PRIMEIRO_NOME *, ganhe seguidores REAIS no Instagram!	Empatia	4832	367	7,595%
* PRIMEIRO_NOME *, para você o emissor pago sai DE GRAÇA	Empatia + Empatia	1.158	406	35,060%
* PRIMEIRO_NOME *, emissor pago GRÁTIS (mas são poucos)	Empatia + Provação social	1.161	474	40,827%
[AJE PARÁ] - Amanhã é sobre Investimento anjo	Sem gatinho	64	25	39,063%
[AJE PARÁ] - * PRIMEIRO_NOME *, amanhã é sobre Investimento anjo.	Empatia	63	35	55,556%
Emissor gratuito 4.0 INDISPONÍVEL: Saiba como resolver	Resolução	123	81	65,854%
Emissor 4.0 INDISPONÍVEL: Evite multa com esse substituto	Ameaça	120	54	45,000%
Como evitar multas da SEFAZ	Sem gatinho	783	177	22,605%
Fui multado pela SEFAZ	Empatia	779	215	27,599%
* PRIMEIRO_NOME *, você se atualiza como os melhores empresários?	Empatia + Autoridade	1411	374	26,506%

* PRIMEIRO_NOME *, é sua última chance de entrar nessa lista	Empatia + Escassez	1405	353	25,125%
Pequenas Empresas Grandes Negócios falou de nós!	Sem Gatilho	427	102	23,888%
* PRIMEIRO_NOME *, você nos colocou na Pequenas EmpresasGrandes Negócios	Empatia	434	149	34,332%

**Tabela 1 – E-mails e gatilhos de persuasão**

### 3.6 Ferramentas

Sem ferramentas de automação tanto da parte de marketing como da parte de classificação de indivíduos, este trabalho seria impossível pois programar todos os processos de captura, manutenção de leads até as técnicas de predição e classificação tomariam muito tempo e a garantia da melhor performance não seria dada. A plataforma usada para automatizar o processo de marketing foi RDStation Marketing e as ferramentas usadas para classificação são: Tensor Flow e PostgreSQL, por último é importante lembrar do Google Trends pois é uma das ferramentas mais relevantes para entendimento do público em Marketing.

#### 3.6.1 RD Station Marketing

Disponível em: <https://app.rdstation.com.br>, é um sistema de gestão de leads em marketing digital, com ele é possível criar páginas de captura, monitorar os leads na base de dados e mandar e-mails automáticos.

#### 3.6.2 Tensor Flow

Disponível <https://www.tensorflow.org>, é uma ferramenta de automação de *machine learning*, com ela não é preciso conhecimento em escrever nenhum algoritmo de inteligência artificial manualmente, basta conhecer os conceitos das técnicas e configurar os parâmetros desejados que ele faz todo o processo. Além de automação de redes neurais de múltiplas camadas, também é possível automatizar redes bayesianas, *random forest* e várias outras técnicas.

#### 3.6.3 PostgreSQL

Disponível em <https://www.postgresql.org/>, é uma suíte de aplicativos para criação e manutenção de bancos de dados relacionais. Neste trabalho foi usado para facilitar o processo de tabulação de dados. Dados de usuários, de abertura de e-mail, de hora de acesso normalmente ficam em lugares e fornecedores separados. Principalmente na tarefa de separar os indivíduos que passam para as próximas fases do funil de vendas

dos que não passam, o conjunto de ferramentas do PostgreSQL foi de extrema importância.

### 3.6.3 Google Trends

Disponível em <https://trends.google.com/>, é uma das principais ferramentas de marketing digital que ajuda os produtores de conteúdo a identificarem quais as melhores palavras-chave em destaque no momento ou em um período específico, não está diretamente ligada a este trabalho mas foi citada em vários capítulos e utilizada em todo o processo de produção de conteúdo

## 3.7 Hipótese de comportamentos

Depois da coleta, o lead passa por um processo automático de interação e recebe automaticamente por e-mail indicações de novos conteúdos que o cercam de informações para que possa finalmente realizar uma compra. A hipótese deste trabalho é que o comportamento inicial do lead dá previsão se ele acessará novos conteúdos ou não. O principal meio de comunicação de um lead é o e-mail, portanto esse trabalho mensura o sucesso do próximo passo do funil de vendas pelo grau de certeza que um lead abrirá um e-mail, dado seu comportamento anterior.

As *features* utilizadas para prever se o lead acessará o próximo passo do funil de vendas são 6: **evento da primeira conversão, provedor de e-mail, origem da primeira conversão, dia da semana da primeira conversão, hora da conversão e número de conversões**. Baseado nesses dados foi possível prever se o lead acessará o próximo passo do funil ou não.

### 3.7.1 Primeiro evento de conversão

Os dados foram exportados para uma base de dados Postgres onde foi possível enumerar todos os eventos que ocasionaram uma conversão de lead, foram contados então 73 eventos que na tabela foram enumerados de 1 a 73



id	descricao
1	infografico-de-custos-de-maquinas-de-cartao
2	o-que-voce-nao-sabe-sobre-as-vantagens-de-usar-cartao-de-credito
3	por-que-devo-investir-em-automacao-comercial-na-minha-loja
4	pare-de-perder-dinheiro
5	home
6	pa
7	a-sefa-pa-cancelou-a-prorrogacao-da-obrigatoriedade-da-nfc-e
8	oportunidade-de-renda-para-voce-de-informatica
9	easygestor-emissor-por-50
10	emitir_nota_fiscal_eletronica_programa_gratuito_sebrae
11	lançamento-aje-para
12	como-emitir-nfe-emissor-gratuito-sebrae-4-0

Tabela 2 – Tabela de eventos da primeira conversão

### 3.7.2 Provedor de email

Imaginou-se que tratando-se de empreendedores a maior parte dos endereços e-mails seriam de domínios próprios, contudo mais de 83% da base dos 11.802 leads usa gmail, hotmail, live.com, Outlook ou yahoo. Apenas 1.498 não usam esses principais.

Essa *feature* foi escolhida para ingressar no conjunto de entrada da rede pois e-mails inválidos são automaticamente descartados da lista de bons leads, dessa forma emails como @teste.com ou @naoexiste.com são usados pela rede para identificar leads ruins.

ID	Provedor
1	Gmail
2	Hotmail
3	Outlook
4	Live
5	Yahoo
6	Outros

Tabela 3 – Tabela de provedores de e-mail

### 3.7.3 Origem da primeira conversão

Foi possível detectar de onde veio o usuário antes de converter como lead na página de captura, contudo são muitas origens de tráfego, o arquivo bruto extraído do RD Station exporta apenas um texto com a informação da origem do lead com por exemplo: “youtube”, “google-search” e “email-de-titulo-1”. Para a entrada da rede neural foi utilizado o tamanho desse texto em caracteres entendendo que nenhuma origem possui o mesmo tamanho em caracteres.

As principais fontes de tráfego são youtube, google-search e facebook. Sendo que youtube possui 7 caracteres, google-search possui 13 e facebook 8 caracteres; números esses usados como *feature* para entrada da rede.

### 3.7.4 Dia da semana da conversão

Os dias da semana foram escolhidos para o treinamento especialmente porque durante os finais de semana, o fluxo de pessoas com acesso aos conteúdos cai 70% na média, a chance de um lead abrir seus e-mails, principalmente relacionados a trabalho é menor. Então como algumas automações são programadas para dias após a conversão, a taxa de sucesso pode aumentar ou não de acordo com o conteúdo acessado e do dia de conversão. Para isso foi criada uma tabela de dias da semana a partir da data da conversão.

ID	Dia da Semana
0	Domingo
1	Segunda Feira
2	Terça Feira
3	Quarta Feira
4	Quinta Feira
5	Sexta Feira
6	Sábado

**Tabela 4 – Tabela de referências para dias da semana**

Por exemplo: O dia 15/04/2019 é uma segunda feira, logo a entrada da rede utilizaria o número 1 para caracterizar essa conversão.

### 3.7.5 Hora da conversão

A hora da conversão acompanha a data da conversão, também muito necessária caso o lead converta muito tarde da noite, o email que chegaria após 2h estaria abandonado. A hora da conversão é um número inteiro que varia de 0 a 23.

Voltando ao exemplo do dia 15/04/2019, agora às 15h seria possível o *feature* para hora de conversão seria igual a 15 e dia da conversão igual a 1.

### 3.7.6 Número de conversões.

Vários leads convertem mais de uma vez, muitas vezes por perder o material e buscar novamente ou porque conseguiu conectar-se com outro artigo em pouco tempo,

esse dado também foi coletado. O número de conversões é apenas um número inteiro que varia de 1 a x vezes que o lead converteu até o momento da coleta.

### 3.8 Configuração da Rede Neural

Dessa forma a rede neural foi configurada com 6 *features* na cama de entrada, 1 ou 2 camadas de neurônios que variaram de 50 a 100 neurônios em cada um. Algoritmos de aproximação por erro médio quadrático, entre 20 e 40 épocas de treinamento. A taxa de aprendizado foi sempre dinâmica que varia de acordo como a rede aprende, assim como a função de ativação na camada de saída foi usada a *softmax* que distribui os resultados de forma probabilística somando todas as saídas em 1. A figura 18 ilustra a rede neural utilizada, os dados de entradas estão normalizados para números entre 0 e 1.

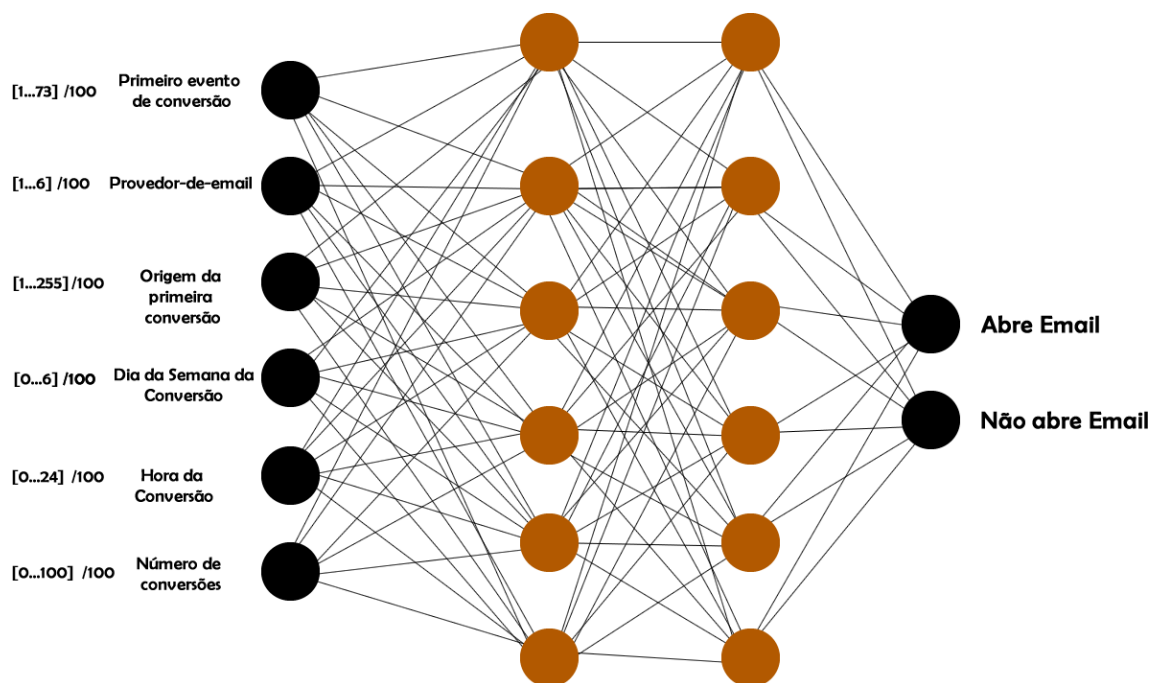


Figura 17 – Configuração base da rede neural

## 4 Resultados



**Figura 18 – Gráfico de crescimento e acessos diários no youtube**

Em 1 ano e 4 meses a base de leads cresceu mais de 1.000% que alcançou aproximadamente meio milhão de usuários dos quais 11.802 deixaram seus dados em troca de conteúdos sobre tecnologia e empreendedorismo.

### 4.1 Abertura de e-mails e gatilhos da persuasão

- Foram disparados 43 e-mails testados com gatilhos de persuasão para melhor abertura.
- A média de abertura de e-mails foi de 30%.
- Somente em dois casos os gatilhos de persuasão não obtiveram mais sucesso na taxa de abertura: Quando o título do e-mail ofertava exatamente o que o cliente esperava como uma redução de preço ou uma vantagem grande.
- Em 100% das vezes que o gatilho de empatia/afeição foi usado contra um e-mail sem gatilhos, a taxa de abertura era maior para emails com empatia.

### 4.2 Acertos da Rede Neural

A proporção de leads estava em 60% para leads que não abrem o e-mail do funil de vendas e 40% para os que abrem no conjunto de treinamento. Dessa forma do total dos 11.802 leads, foram selecionados 30% o conjunto de validação, ou seja, 8.261 leads foram para o conjunto de treinamento e 3.541 leads foram para o conjunto de testes.

A rede Neural que melhor obteve resultado foi a de configuração com 1 camada intermediária de 200 neurônios, configurada a 20 épocas, que acertou 70,743% do conjunto de testes, onde 69,28% são para a classe dos que iriam para os próximos passos do funil e 71,765% dos que não iriam.

Outros testes foram feitos trocando número de neurônios na camada intermediária, cálculo de aproximação, número de épocas e até invertendo o conjunto de treinamento com o conjunto de validação, contudo, a rede nunca acertou menos de 68% e mais que 71%, sendo que os resultados acima de 71% diminuía muito os acertos para a classe dos que abrem o email que também nunca foi menor que 61%. Dessa forma mesmo alterando as configurações a RNA, seus resultados não variam e, mais que 10% para mais ou para menos.

Configuração	Acertos gerais	Acertos abriu e-mail	Acertos não abriu
-1 camada oculta -100 neurônios na camada intermediária -20 épocas	69,726%	68,878%	70,322%
-1 camada oculta -200 neurônios na camada intermediária -20 épocas	70,743%	69,289%	71,765%
-2 camadas ocultas -100 neurônios em cada camada -40 épocas	70,206%	62,585%	75,565%

**Tabela 5 – Comparação entre acertos de redes neurais.**

## **5 Considerações finais e discussões.**

### **5.1 Preparação para o trabalho**

Devido o grande volume de dados acumulados em 12 meses foi possível o desenvolvimento deste trabalho. Assim como Smith (2000) ou Loureiro (2018), foi possível criar um modelo preditivo para comportamentos humanos em meios digitais. Para isso foi necessário estudar tanto técnicas de predição como comportamento do consumidor dentro e fora do ambiente virtual.

Antes de começar o trabalho houve um estudo minucioso do comportamento dos usuários ao consumir conteúdos do marketing digital para entender quais as principais variáveis que possam afetar o comportamento do consumidor. Ao logo do trabalho de mais de um ano, foi possível testar várias hipóteses de alcance e convencimento de público.

Cialdini (2012), e Kunneman (2011) afirmam que o ser humano possui muitos comportamentos instintivos a ponto de agirem de forma automática mediante a provocações específicas. Entender alguns desses processos foi crucial para ter resultados relevantes mesmo antes da aplicação da técnica de predição. Contudo, apesar da alta tendência em repetir comportamentos, os públicos podem ser diferentes ainda

que pertencentes à mesma rede. Neste trabalho apesar de conhecer as técnicas foi necessário realizar testes A/B a cada novo conteúdo lançado para identificar qual a melhor forma de comunicar com cada público.

Um ponto importante a se destacar é que a rede neural desenvolvida neste trabalho ajuda a prever a disponibilidade de clientes a abrir novos e-mails, viabilizando assim a ação do funil de vendas, mas nada garante na totalidade a abertura do e-mail, tão pouco a continuidade do processo de funil de vendas já que foi tratado neste mesmo trabalho que a presença ou não de certas palavras também influenciam o desempenho de campanhas de marketing.

## 5.2 Métricas e observação dos comportamentos

Observou-se que algumas palavras fazem mais efeitos que outras. Em um caso específico a palavra “Exclusivo” utilizada no título de um e-mail, trouxe uma taxa de abertura 30% maior que a taxa do título sem a palavra. Outra observação é que o público costumava acessar os conteúdos em dias e horas específicas. O gráfico da figura 20 expressa uma medição de 3 meses. Nota-se um pico de acessos em dias de semana, especialmente terças e quintas, assim como uma diminuição acentuada aos finais de semana. O Gráfico possui esse formato de “ondas” justamente por ter grandes acessos aos dias de semana e menos da metade desses acessos aos finais de semanas e feriados.

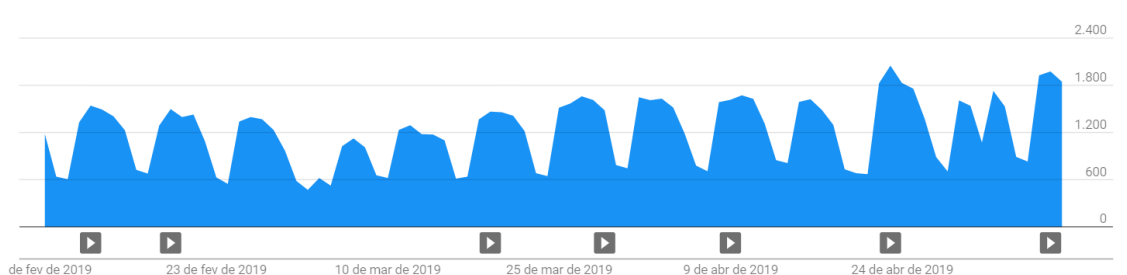


Figura 19 - Gráfico de acessos ao youtube por um período de 3 meses

Por isso entender o comportamento do consumidor é crucial, não adiantaria os esforços de desenvolver uma rede neural capaz de prever a taxa de abertura de um e-mail se esse e-mail foi enviado durante o final de semana para esse público em específico. Um teste foi feito e a taxa de abertura de e-mails durante o final de semana é menor que 10% frente a uma média de 20% a 30% durante a semana.

## 5.3 Pré-processamento dos dados e o uso de ferramentas

Apesar de não ser tema deste trabalho é possível levantar a discussão do uso de sistemas de código aberto. Para coleta dos dados e automação de Marketing foi utilizada a ferramenta RD Station em conjunto com Google Analytics. Houve um

grande prejuízo em relação a ter usado ferramentas em nuvem que limitam o acesso a informação já que o banco de dados fica em um servidor próprio do desenvolvedor da ferramenta, algumas extrações de dados custam muito mais tempo e outras são basicamente impossíveis a medida que os dados aumentam.

Ter acesso a todos os e-mails que um usuário recebeu e não clicou de forma tabelada seria de grande ajuda para conseguir previsões ainda melhores. Contudo para isso teria que abrir cada e-mail enviado e exportar manualmente cada uma das situações (leads que abriram email, que clicaram, que não abriram).

Em sistemas de código de aberto é possível ter acesso total ao banco de dados assim como, se precisar, alterar a própria ferramenta. Em um trabalho futuro seria interessante utilizar para coleta e automação de marketing uma ferramenta de código aberto para melhor qualidade dos dados e menos tempo de pré-processamento.

Como não havia acesso direto banco de dados, foi necessário criar um banco de dados estruturado a partir do arquivo de texto exportado do RD Station, para isso foi utilizado o PostgreSQL e o PGAdmin para manipulação dos dados.

Por último é necessário destacar o TENSOR FLOW, é uma ferramenta de código aberto para inteligência artificial que acompanha um pacote variado de técnicas que vão de redes bayesianas, *multi layer perseptron* e as redes neurais convolucionais que são o que há de mais recente sobre redes neurais mas não foram selecionadas para esse trabalho porque na literatura há muitos exemplos do seu uso para reconhecimento de imagem e não predição nem aproximação de função.

O TENSOR FLOW é de extrema importância para pesquisas na área, muitos dos trabalhos citados nesse texto utilizam a ferramenta. Com ela é possível trocar de algoritmo de predição em alguns segundos, assim como trocar funções de ativação e usar técnicas mais avançadas como taxa de aprendizado dinâmica somente configurando alguns parâmetros, o TENSOR FLOW é tão completo que vem inclusive com métodos de separação de exemplos em teste e treinamento e mensuração automática dos resultados.

#### **5.4 Considerações gerais**

Com a ajuda de gatilhos de persuasão, a média para taxa de abertura de e-mails chegou até aproximadamente 30% usando somente gatilhos de persuasão, com a ajuda de uma rede neural que acerta aproximadamente 70%, é possível aumentar essa taxa de abertura para os mesmos 70%, o que faria com que a porcentagem de e-mails considerados SPAM fosse para praticamente 0%.

Esse trabalho mostra uma clara correlação entre o comportamento prévio e a taxa de abertura de um e-mail que leva o usuário ao fundo de funil, dessa forma é possível em um futuro ainda aumentar essa taxa de abertura que garante o sucesso de visualização da marca do autor do email para acima de 70%.

O ganho em atender o lead melhor qualificado seria de imensa importância, pois seria mais fácil extrair os clientes que mais se adequam as propostas, já que aumenta-se a certeza que o lead tem interesse em consumir um próximo conteúdo, diminui-se então esforço humano e de máquina para futuras conversões em venda.

A questão das vendas propriamente ditas ainda não foram postas em estudo, quando que há menos de 200 vendas para os mais de 11 mil leads, contudo somente a visualização de conteúdo já gera resultado financeiro, hoje cotado em média R\$ 0,01 centavo por cada visualização. Quando os exemplos de leads que compraram aumentar, será possível identificar os melhores compradores entre os muitos usuários.

## **5.5 Trabalhos futuros**

Provado que os clientes que mais interagem com a marca possuem um padrão de comportamento desde o primeiro contato agora é possível prever outros comportamentos como o próprio comportamento de compra, basta ter os dados corretos, na quantidade certa.

Para o aumento de acerto é possível estudar a possibilidade da utilização de uma ferramenta de código aberto que libere o acesso ao banco de dados de forma bruta, dessa forma faz-se assim um melhor aproveitamento dos dados em um menor espaço de tempo.

Finalmente além do uso de técnicas de predição citadas nesse trabalho, outros trabalhos que desenvolveram produtos inovadores especialmente nas universidades poderiam utilizar também as técnicas de marketing digital citadas para alcançar mais público e quem sabe financiar-se com dinheiro próprio das vendas ou das visualizações pagas.



## 6 Referências.

ADOLPHO, Conrado. **Os 9 Os do Marketing Digital**. Primeira edição. São Paulo: Novatec Editora, 2011.

AUGUSTO, F. **Geração De Valor 3**. Primeira edição. São Paulo: Buzz Editora LTDA, 2018

CIALDINI, R. **Armas da persuasão**. Rio de Janeiro. GMT Editores LTDA, 2012

CIALDINI, R. **Présuasão**. Primeira edição Rio de Janeiro: Sextante, 2017

COSTA, M. Probabilistic Interpretation of Feedforward Network Outputs, With Relationships To Statistical Prediction of Ordinal Quantities. **International Journal of Neural Systems**, 1996.

D'HAEN, J; Van den Poel, D. Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. **Industrial Marketing**, (2013).

EMPRESÔMETRO. Inteligência de mercado. **Empresômetro.com**, 2019 – Disponível em: < <http://empresometro.com.br> >. Acesso em 10/05/2019

FORBES. Are You Maximizing The Use Of Video In Your Content Marketing Strategy?. **forbes.com**. Disponível em: <https://www.forbes.com/sites/forbesagencycouncil/2017/05/15/are-you-maximizing-the-use-of-video-in-your-content-marketing-strategy/#379414b23584>> acesso em 30 jun. 2019.

GALLO, C. **Story Telling**. Primeira edição. São Paulo: Alta Books, 2018.

GRAUPE, D.; **Principles of Artificial Neural Networks**. 2ª Edição. World Scientific Publishing Co. Pte. Ltd, 2007.

HAYKIN, S. **Redes Neurais – Princípios e práticas**. 3ª Edição. Porto Alegre: Bookman, 2007.

HINTON, G. E.; Srivastava, N.; Krizhevsky, A.; et al. Improving neural networks by preventing co-adaptation of feature detectors. **Cornell University**, 2012.

HA, K.; CHO, S.; MACLACHLAN, D. Response models based on bagging neural networks. *Journal of Interactive Marketing*, 2005.

JÄRVINEN, J.; Taiminen, H. Industrial Marketing Management Harnessing marketing automation for B2B content marketing . **Industrial Marketing Management**, 2015.

KAHNEMAN, D. **Pensamento rápido e devagar**. Rio de Janeiro: Editora Objetiva LTDA, 2012.

KASABOV, N. **Foundation of Neural Networks, fuzzy systems and knowledge engineering**. Massachusetts Institute of Technology, 1996.

KNACKFUSS, A. VP de transformação digital, Coca Cola. Portal B9, 2018. Disponível em: <https://www.b9.com.br/95153/codigo-aberto-adriana-knackfuss-vp-de-transformacao-digital-coca-cola/> - Acesso em 29 jun. 2019

KOTLER, P.; Keller, K. **Administração de Marketing: A Bíblia do Marketing**. 12<sup>a</sup> edição. São Paulo: Pearson Prentice Hall, 2006.

KOTLER, P.; Kartajaya, H.; Setiawan, I. **Marketing 4.0: Do tradicional ao digital**. Primeira edição. Rio de Janeiro: GMT Editores, 2017.

LEE, W.; Shih, B.; Chen, C. A Hybrid Artificial Intelligence Sales-Forecasting System in the Convenience Store Industry. **Wiley Periodicals** , 2012.

LEEFLANG, P. S. H.; VERHOEF, P. C.; DAHLSTRÖM, P.; FREUNDT, T. Challenges and solutions for marketing in a digital era. **European Management Journal**, 2014.

LOUREIRO, A. L. D.; Miguéis, V. L.; Da Silva, L. F. M. Exploring the use of deep neural networks for sales forecasting in fashion retail. **Decision Support Systems**, 2018.

MCMAHAN , H. B.; Sculley , G. H.; Young D. M; et al. Ad Click Prediction: a View from the Trenches. **Google Research**, 2013.

- MARINI, F.; Bucci, R.; Magrì, A. L.; Magrì, A. D. **Artificial neural networks in chemometrics: History, examples and perspectives**. *Microchemical Journal*, 2008.
- MUNAKATA, Toshinori. **Fundamentals of the New Artificial Intelligence: Neural, Evolutionary, Fuzzy and More**. 2ª Edição. Londres: Springer, 2008.
- MONCRIEF, W. C.; Marshall, G. W. **The evolution of the seven steps of selling**. *Industrial Marketing Management*, 2005.
- POTTER, M. **Why Business Can Be Good at Solving Social Problems**. Ted Talks, 2013. Disponível em: [https://www.ted.com/talks/michael\\_porter\\_why\\_business\\_can\\_be\\_good\\_at\\_solving\\_social\\_problems](https://www.ted.com/talks/michael_porter_why_business_can_be_good_at_solving_social_problems). acesso em 09/05/2019>. Acesso em: 09 jun. 2019.
- Pulizzi, J. **Conteúdo S.A**, primeira edição. São Paulo: DVS Editora, 2017.
- Ramos, A.; Cota, S. **Search Engine Marketing**. McGraw-Hill, New York, 2008.
- Rojas R. **Neural Networks: A Systematic Introduction**. Berlin: Springer-Verlag, 1996.
- RYAN, Damian. **Understanding Digital marketing: Making strategies for engaging the digital generation**, 4ª edição. Nova Iorque: CPI Group, 2017.
- SEFAZ. Conceito, uso e obrigatoriedade da NF-e (26 questões). **Portal do Ministério da Economia**, Disponível em: <http://www.nfe.fazenda.gov.br/portal/perguntasFrequentes.aspx?tipoConteudo=E4+tmY+ODf4=>>. Acesso em 29 jun.2019
- SELEY, A.; & Holloway, B. **Sales 2.0: Improve business results using innovative sales practices and technology**. *Wiley Periodics*, 2008
- SMITH, K. A.; Gupta, J. N. D. Neural networks in business: Techniques and applications for the operations researcher. **Computers and Operations Research**, 2000.
- TAGHIZADEH, E. Utilizing artificial neural networks to predict demand for weather-sensitive products at retail stores. **Cornell University**, 2017

YAN, J.; Gong, M.; Chu, S. M. Sales pipeline win propensity prediction : a regression approach. **Cornell University**, 2015.