

## Verocímetro: uma proposta para combater as notícias falsas

*Verocímetro: a proposal to combat fake news*

Elisama Santos Silva<sup>1</sup>

Centro Universitário Maurício de Nassau, Salvador - BA, Brasil

Matheus Santos de Jesus<sup>2</sup>

Centro Universitário Maurício de Nassau, Salvador - BA, Brasil

Vitor Everton Lima dos Santos<sup>3</sup>

Centro Universitário Maurício de Nassau, Salvador - BA, Brasil

Fabio Fonseca Barbosa Gomes<sup>4</sup>

Faculdade Visconde de Cairu, Salvador - BA, Brasil

Celso Barreto da Silva<sup>5</sup>

Universidade Senai Cimatec - Salvador, Brasil

Rodrigo P. V. Bôa<sup>6</sup>

Centro Universitário Maurício de Nassau, Salvador - BA, Brasil

**Resumo:** A disseminação de uma grande quantidade de notícias diariamente tem gerado uma crescente preocupação quanto à veracidade das informações veiculadas, especialmente diante da proliferação de notícias falsas. Existem exemplos de notícias falsas, conhecidas como fake news, que são publicadas em conjunto com notícias reais, com informações distorcidas ou manipuladas. Um dos pontos focais deste tipo de notícia são as redes sociais. Elas se destacam por ser um meio bastante popular de troca de mensagens, como resultado de um acesso a internet mais frequente para grande parte da população. Com a popularização dos dispositivos móveis e a expansão da conectividade, essa tendência tende a se intensificar ainda mais. Diante desse cenário, o principal objetivo deste artigo é propor uma ferramenta de acesso público que permita verificar a autenticidade de determinadas notícias. Inicialmente, a verificação será feita manualmente pela equipe de suporte da plataforma Verocímetro,

---

<sup>1</sup> Bacharel em Sistemas de Informação – Centro Universitário Maurício de Nassau. E-mail: [eli.sama15@hotmail.com](mailto:eli.sama15@hotmail.com)

<sup>2</sup> Bacharel em Sistemas de Informação – Centro Universitário Maurício de Nassau. E-mail: [matheus\\_santosdejesus@outlook.com](mailto:matheus_santosdejesus@outlook.com)

<sup>3</sup> Bacharel em Sistemas de Informação – Centro Universitário Maurício de Nassau. E-mail: [vitorlimasw@gmail.com](mailto:vitorlimasw@gmail.com)

<sup>4</sup> Professor Assistente – Faculdade Visconde de Cairu. Mestre em Sistemas e Computação. E-mail: [fabio@cairu.br](mailto:fabio@cairu.br)

<sup>5</sup> Professor – Centro Universitário Senai Cimatec. Mestre em Sistemas e Computação. E-mail: [profelsobarreto@hotmail.com](mailto:profelsobarreto@hotmail.com)

<sup>6</sup> Professor – Centro Universitário Maurício de Nassau. E-mail [rodrigoboa@uninassau.br](mailto:rodrigoboa@uninassau.br)

por meio de uma análise detalhada das informações publicadas. Posteriormente, será apresentada a proposta de desenvolvimento de um algoritmo de aprendizado de máquina para automatizar esse processo de verificação.

**Palavras-chaves:** Notícias falsas. Aprendizagem de máquina. Processamento de linguagem natural. Inteligência artificial. Rede neural artificial.

**Abstract:** The dissemination of a large amount of news every day has generated growing concern about the veracity of the information being published, especially in light of the proliferation of fake news. There are examples of fake news, which are published together with real news, with distorted or manipulated information. One of the focal points of this type of news is social media. They stand out for being a very popular means of exchanging messages, as a result of more frequent internet access for a large part of the population. With the popularization of mobile devices and the expansion of connectivity, this trend tends to intensify even more. Given this scenario, the main objective of this article is to propose a public access tool that allows the authenticity of certain news items to be verified. Initially, the verification will be done manually by the Verocômetro platform support team, through a detailed analysis of the published information. Subsequently, the proposal for developing a machine learning algorithm to automate this verification process will be presented.

**Keywords:** Fake news. Machine learning. Natural language processing, Artificial intelligence. Artificial neural network.

## Introdução

Para Campos (2020), as notícias falsas ou *fake news* são publicadas por veículos de comunicação como se fossem informações reais. Este assunto não é novidade nem tão pouco é algo desconhecido, rumores e histórias adulteradas tomaram forma desde que o ser humano vive em grupo. Coutinho (2018) relata que essas notícias muitas vezes contêm títulos atraentes e até mesmo apelativos, fabricados para chamar a atenção e garantir maior alcance de leitura. Sabe-se que extinguir é algo impossível, pois, qualquer pessoa pode transmitir conteúdo sem nenhum tipo de filtragem, verificação de fatos ou julgamento editorial através das mídias sociais (Allcott; Gentzkow, 2017).

Porém é necessário conseguir identificá-la a fim de minimizar o impacto que pode ser causado (Burkhardt, 2017). Para Allcott e Gentzkow (2017), o conceito *fake news* indica histórias falsas que, ao manterem a aparência de notícias jornalísticas, são disseminadas pela Internet (ou por outras mídias), sendo normalmente criadas

para influenciar posições políticas, ou como piadas. Como mencionado, existiu fake news, mesmo antes da internet surgir ou qualquer tipo de mídia informativa. Com o surgimento desses recursos, as notícias falsas aumentaram, principalmente vindo de pessoas que querem se beneficiar com elas.

Para Peduti (2018), todo tipo de boato e informação falsa, tem duas dimensões diferentes, a primeira delas é avaliar quando quem compartilha é apenas uma vítima, que foi enganada pela notícia falsa, e quando quem compartilha/cria possui a intenção de enganar e divulgar informações falsas que podem denegrir a reputação de terceiros. O grande impasse é avaliar o dolo, a intenção de causar dano, e distingui-la do simples desconhecimento de quem compartilha uma notícia falsa, o que é algo difícil de comprovar sem a existência de outras provas, por exemplo, histórico de postagem da pessoa.

É muito comum pessoas serem pegas lendo apenas os títulos das notícias sem ao menos abri-la e ler a matéria por completo. Muitas vezes o título é tendencioso/mentiroso e na matéria está escrita a notícia verdadeira, mesmo que esteja na última linha. Dessa forma, como é possível classificar se esse tipo de notícia é falsa ou verdadeira? (Peduti, 2018). Para Leetaru (2016), existem sites denominados satíricos, estes têm como objetivo entreter e não enganar, porém o problema que há nisso é: os usuários que acessam estas notícias apesar de terem ciência que se trata de uma inverdade compartilham estas notícias de forma imediata sem medir as suas consequências.

Uma grande quantidade de sites satíricos e jornais dificultam a identificação positiva de que um site específico é satírico sem pesquisas adicionais. Em sua essência, esses tipos de sites se intitulam humorísticos e não tem intenção de enganar, como por exemplo o *site* Sensacionalista (2025). Além disso, existem os sites maliciosos que realizam “*clickbait*”, cujo propósito é publicar histórias virais que gerem uma grande quantidade de cliques para vender a publicidade., como mencionado acima sobre os títulos tendenciosos. Na área cinzenta, Leetaru (2016) menciona que entre os sites satíricos e de *clickbait*, existem aqueles referentes a focos e boatos, que tem como principal pauta as celebridades e produtos.

Esses boatos, geralmente, surgem de uma única fonte baseada em uma conversa com um informante confidencial, o que torna difícil a identificação, pois muitas empresas espalham falsos rumores internos para tentar dissuadir vazamentos.

Devido a quantidade de informações compartilhadas por segundo na internet, é quase impossível identificá-las individualmente. Por isso, é importante cada pessoa ao compartilhar desenvolver uma mentalidade crítica sobre o assunto, verificar a fonte, pesquisar se existem mais notícias sobre ele, examinar as evidências e não se contentar apenas com o título (Leetaru, 2016).

Um dos meios de propagação das fake news, na atualidade, são as redes sociais. Por se tratar de um local onde todas as pessoas podem livremente compartilhar publicações, notícias, ou qualquer outro tipo de postagens (que estejam dentro das diretrizes da rede), as redes sociais e mensageiros em geral acabam se tornando o meio mais comum de se espalhar as *fake news*.

Um bom exemplo é o aplicativo WhatsApp, que em uma de suas atualizações limitou o número de encaminhamento de mensagens, que antes era de 20 pessoas e agora passou a ser para apenas 5. Um número expressivo. Isso surgiu efeito e reduziu em 70 por cento a propagação de Fake News (Delmazo; Valente, 2018). O WhatsApp tem visto uma grande redução nas mensagens virais desde que impôs novos limites ao encaminhamento de mensagens, de acordo com o Singh (2020), a disseminação de mensagens “altamente encaminhadas” no serviço caiu 70% globalmente, disse o porta-voz.

Existe uma grande dificuldade nas *fake news*, visto que elas transmitem difamação ou calúnia, circulando entre os meios de comunicação com frequência, aparentemente diversas pessoas não se atentam ao impacto que pode ser gerado ao ser um vetor de transmissão (Pedroso, 2019). A desinformação gerada por elas causa turbulências comprometendo a integridade do jornalismo podendo até incitar o caos e a desordem. Identificá-las de forma manual não é algo simples, exige real interesse e dedicação em encontrar a verdade (Wang, 2017). Linhares et al. (2017) citam que a Constituição Federal do Brasil declara que são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação.

Segundo Coutinho (2018), uma informação falsa pode ser classificada como:

- a) Sátira ou Paródia: não é intencionalmente nocivo, mas pode levar à confusão do leitor;
- b) Conexão Falsa: título não corresponde fielmente ao conteúdo, gerando uma espécie de “clickbait” para aumentar o acesso;

- c) Contexto Falso: uma determinada informação quando fora de contexto pode se tornar inapropriada ou inválida com o passar do tempo;
- d) Conteúdo Manipulado: seja por adulteração de texto e/ou imagens, ou por tendenciar determinada opinião/visão política/ponto de vista;
- e) Conteúdo Enganoso: a informação é utilizada de forma a difamar a pessoa ou o assunto a que se refere;
- f) Conteúdo Impostor: informação é mal utilizada, moldando uma situação e criando uma inverdade com informações falsas de marcas ou pessoas
- g) Conteúdo Fabricado: todo o seu conteúdo é falso, criado para enganar e prejudicar.

Este artigo visa descrever as ferramentas que serão utilizadas para desenvolvimento de um algoritmo baseado em aprendizado de máquina cujo objetivo é treiná-lo a fim de identificar a veracidade de uma notícia enviada através da aplicação web chamada de Verocímetro. Atualmente o fluxo para análise da veracidade da informação acontece da seguinte forma: o usuário pode enviar uma notícia para análise da sua fidedignidade ou buscar na opção Perguntas Frequentes se a notícia se encontra cadastrada na base de dados.

Caso deseje enviar a notícia para análise, é necessário:

- 1) Acessar a plataforma através do endereço *web*;
- 2) Preencher os campos obrigatórios que são: o *link* da notícia, o título da notícia ou fragmentos da notícia a ser analisada, *e-mail* e CEP. Após informar estes dados basta clicar no botão Enviar.

Ao realizar esta ação, o usuário receberá uma mensagem de alerta informando que dentro de até 48 horas úteis receberá no seu endereço eletrônico o *feedback* da análise realizada. Atualmente esta análise é realizada de forma manual pela equipe de suporte, porém, o objetivo é adequar um algoritmo de aprendizagem a plataforma web a fim de tornar essa análise dinâmica, assertiva e possibilitando um feedback com um menor tempo. Este projeto não visa nenhum tipo de integração com a plataforma *mobile* ou *chatbots*.

## **Inteligência artificial**

O termo IA (Inteligência Artificial) é relativamente novo na história da tecnologia da informação, segundo Russel e Norvig (2013) “O primeiro trabalho reconhecido

como IA foi realizado por Warren McCulloch e Walter Pitts (1943). Eles se basearam em três fontes: o conhecimento da fisiologia básica e da função dos neurônios no cérebro; uma análise formal da lógica proposicional criada por Russell e Whitehead; e a teoria da computação de Turing.” (Russel; Norvig, 2013, p. 41). Ainda de acordo com os autores, às definições de Inteligência Artificial estão dispostas ao longo de duas dimensões: as que se referem a processos de pensamento e raciocínio e as que se referem ao comportamento.

Quanto à primeira dimensão (pensamento e raciocínio) a Inteligência Artificial é a associação de atividades do pensamento humano que são automatizadas, com o caso de tomada de decisões, por exemplo (Bellman, 2010). Quanto à segunda dimensão (comportamento), a IA é definida como máquinas capazes de simular o pensamento humano (Kurzweil, 1992). Ainda na procura de um conceito um pouco mais generalista, Sperandio (2018, p. 98) nos traz uma definição que mais se aproxima à perspectiva deste projeto.

[...] entende-se por IA, como um conjunto de ferramentas desenvolvidas para que sistemas computacionais possam executar tarefas que requeiram capacidade racional do ser humano, especialmente quando faz-se necessária a interpretação de dados.” (Sperandio, 2018, p. 98).

Ao longo dos anos, a IA vem sendo explorada nos mais variados ambientes, um campo de aplicação bem conhecido é a rede social, onde algoritmos inteligentes buscam o melhor conteúdo para cada usuário específico. Outro campo mais recente foi o residencial, onde a IA contribui para a automatização das mais variadas atividades. Além desses, diversos outros setores são impactados com o avanço desta tecnologia, setor industrial, jurídico, financeiro, entre outros.

### **Processamento de Linguagem Natural (PLN)**

O ser humano se destaca sob as demais espécies por evoluir de acordo ao meio que está inserido, desta forma desenvolve a capacidade de comunicar-se através da linguagem verbal, escrita e corporal (Russell; Norving, 2013). Ainda de acordo com Russell e Norving (2013), a compreensão da linguagem exige a compreensão do assunto e do contexto, não apenas a compreensão da estrutura das

frases. Para que haja esta compreensão é necessário identificar algumas questões como por exemplo: reconhecer o idioma, verificar a ambiguidade, detectar a existência de algum tipo de figura de linguagem, figuras de pensamento ou o uso de recursos estilísticos transmitido pelo emissor do texto (Silva, 2020).

A assimilação da linguagem não é algo simples, porém é necessário para que os computadores desenvolvam a capacidade de aprender a reconhecer o contexto, fazer análise sintática, semântica, lexical e morfológica, criar resumos, extrair informação, interpretar os sentidos, analisar sentimentos, compor textos e até aprender conceitos com os textos processados (Rodrigues, 2017). A área de pesquisa da Inteligência Artificial (IA) responsável por tal demanda é chamada de Processamento de Linguagem Natural (PLN), ou Natural Language Processing (NLP) (Sanches, 2017).

Silva (2019) define o Processamento de Linguagem Natural (PLN) como sendo uma subárea de ciência da computação, inteligência artificial e linguística que estuda os problemas da geração e compreensão da linguagem humana. O PLN é uma forma dos computadores analisarem, compreenderem e derivarem o significado da linguagem humana de uma maneira inteligente e útil. Rodrigues (2017) descreve que o PLN estuda a capacidade e as limitações de uma máquina em entender a linguagem dos seres humanos. Barbosa et al. (2017) menciona que a linguagem humana é raramente escrita ou falada corretamente, entender a linguagem humana é entender não apenas as palavras, mas os conceitos e como eles estão ligados entre si para criarem significados.

Apesar da facilidade que temos, como humanos, de aprender a linguagem natural, a ambiguidade de uma linguagem é o que torna o PLN um problema difícil para computadores (Barbosa et al., 2017). Para realizar o processamento da linguagem natural pode-se fazer uso de alguma ferramenta como por exemplo: Natural Language Toolkit (NLTK), Google Speech To TextS entre outras. Também existem técnicas de processamento como por exemplo: Word Embedding (word2vec) (Sanches, 2017) (Silva, 2019). O uso de PLN está cada vez mais presente em nosso cotidiano pode-se mencionar as plataformas de busca online como por exemplo: ChatGPT, Gemini e assistentes virtuais como a Alexa (Take, 2019).

## Machine learning (ML)

Aprendizado de máquina, *machine learning* ou aprendizado automático é um ramo da Inteligência Artificial que tem por objetivo utilizar técnicas e algoritmos que possibilitam ao computador desenvolver a habilidade de aprender a partir de um volume de dados (Silva, 2019). Estes algoritmos melhoram seu desempenho por intermédio das informações utilizadas para treino e das rotinas executadas (Carvalho, 2020). Dessa forma aprende com erros e conseqüentemente produz resultados fundamentados e desenvolve a capacidade de tomar decisões, estabelecer métricas ou fazer previsões caso seja necessário (Gomes, 2019).

Os métodos de aprendizagem são divididos em: (i) Aprendizado supervisionado (AS), (ii) Aprendizado não supervisionado (ANS), (iii) Aprendizado semi supervisionado (ASS) e (iv) Aprendizado por reforço (AR). Rolim et al. (2017), descreve que o aprendizado supervisionado tem como objetivo compreender a relação entre as entradas e saídas fornecidas, para classificar, ou rotular, uma determinada instância em um conjunto de categorias pré-definidas. Gomes (2019) menciona que este tipo de aprendizado utiliza a identificação de padrões para estabelecer previsões.

Segundo Rolim et al. (2017) o aprendizado não supervisionado visa agrupar elementos com características similares. Filho (2017) menciona que este é elaborado com dados não rotulados, ou seja, não se sabe quantas ou quais classes existem. É uma técnica usada para encontrar subgrupos de objetos. O objetivo é classificar as entidades em grupos mutuamente exclusivos baseados na similaridade entre as instâncias. Os grupos não são pré-definidos e não há um “professor”, por isso o elemento não é supervisionado.

Segundo Caldas (2017) o aprendizado semi supervisionado explora a capacidade de combinar dados rotulados com dados não rotulados, e o aprendizado por reforço (iv) é definido por Teixeira (2016) como um campo do aprendizado de máquinas com foco na criação de agentes capazes de tomar decisões acertadas em um ambiente sem que se tenha qualquer conhecimento prévio sobre o tal ambiente. Segundo Russel e Norving (2013) um agente é capaz de perceber seu ambiente por meio de sensores e de agir sobre esse ambiente por meio de atuadores. Gomes

(2019) relata que nesse método, o computador é estimulado a aprender com base em tentativas e erros.

### **Rede neural artificial (RNA)**

O cérebro humano, é capaz de processar informações a uma velocidade extraordinária, isto é estudado há muitos séculos. Com o surgimento da computação, é natural aproveitar esse tipo de processamento de pensamento e tentar aplicar em uma máquina. Apesar de ser muito complexo, muitos estudos e tecnologias mostram que isso é capaz. O primeiro passo em direção às redes neurais artificiais aconteceu em 1943, quando Warren McCulloch, um neurofisiologista e um jovem matemático chamado Walter Pitts, escreveu um artigo sobre como os neurônios podem funcionar. Eles modelaram uma rede neural simples com circuitos elétricos (Anderson, 1992).

À medida que os computadores e a tecnologia avançaram na década de 1950, tornou-se possível começar a modelar os rudimentos dessas teorias sobre o pensamento humano. A primeira tentativa de simular uma rede neural fracassou. Isso aconteceu nos laboratórios da IBM (International Business Machines Corporation), liderado por Nathaniel Rochester. Embora a primeira tentativa tenha fracassado, as posteriores foram bem-sucedidas (Anderson, 1992).

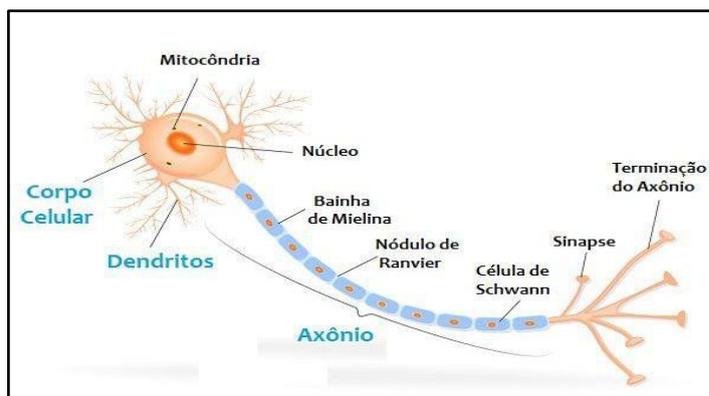
De acordo com Dayhoff e DeLeo (2001) as redes neurais artificiais são metodologias computacionais que realizam análises multifatoriais. Inspirados por redes de neurônios biológicos, os modelos de redes neurais artificiais contêm camadas de nós de computação simples que operam como dispositivos de soma não linear. Esses nós são ricamente interconectados por linhas de conexão ponderadas e os pesos são ajustados quando os dados são apresentados à rede durante um processo de “treinamento”. O treinamento bem-sucedido pode resultar em redes neurais artificiais que realizam tarefas como prever um valor de saída, classificar um objeto, aproximar uma função, reconhecer um padrão em dados multifatoriais e completar um padrão conhecido.

As Redes Neurais Artificiais são ferramentas utilizadas no processo de aprendizado de máquina. São sistemas baseados no neurônio real que se destinam a replicar a maneira como o ser humano aprende. Conforme a figura 2, os neurônios se ligam pelas dendrites, que são as entradas, enquanto os axônios dos outros

neurônios são a saída do neurônio anterior. Existem diversas entradas e saídas num mesmo neurônio e conforme as entradas são ativadas, uma ou mais saídas também são ativadas (Toda Matéria, 2020).

Parte principal do artigo, que contém a exposição ordenada e pormenorizada do assunto tratado. Divide-se em seções e subseções, conforme a NBR 6024, que variam em função da abordagem do tema e do método.

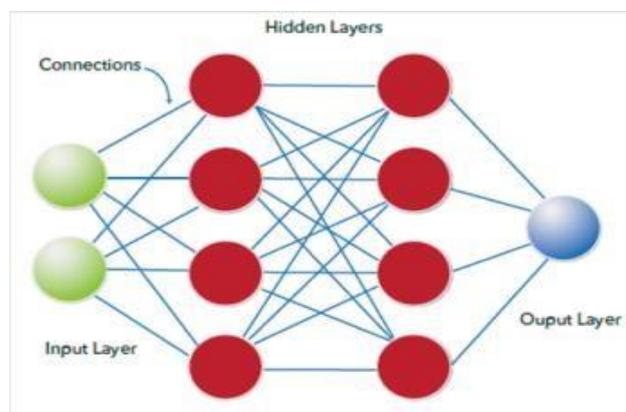
**Figura 1 – Estrutura dos Neurônios**



Fonte: Toda Matéria (2020)

Pode-se notar que as redes neurais artificiais também consistem em camadas de entrada (*Input*) e saída (*Output*), bem como uma camada oculta (*Hidden Layer*) que consiste em unidades que transformam a entrada em algo que a camada de saída pode usar, conforme ilustrado na figura 3. Eles são ferramentas excelentes para encontrar padrões que são muito complexos ou numerosos para um programador humano extrair e ensinar a máquina a reconhecer (Luke Dormehl, 2019).

**Figura 2 – Modelo e uma rede neural**



Fonte: SAS (2020)

As redes neurais artificiais também são conhecidas como Perceptron, ela possui 4 partes: 1) valores de entrada ou uma camada de entrada; 2) pesos e bias; 3) soma líquida e 4) função de ativação. Apesar de existir desde a década de 1940, recentemente tornou-se uma parte importante da Inteligência Artificial. Isso se deve à chegada de uma técnica chamada “retropropagação”, que permite que as redes ajustem suas camadas ocultas de neurônios em situações em que o resultado não corresponde ao que o criador espera (Luke Dormehl, 2019).

O algoritmo de retropropagação é um algoritmo utilizado no treinamento de redes neurais multicamadas com uma ou mais camadas escondidas. Andrade (2001) relata que o algoritmo de retropropagação consiste em dois passos de computação: o processamento direto e o processamento reverso. No processamento direto, uma entrada é aplicada à rede neural e seu efeito é propagado pela rede, camada a camada. Durante o processamento direto, os pesos da rede permanecem fixos. Já no processamento reverso, um sinal de erro calculado na saída da rede é propagado no sentido reverso, camada a camada, e ao final deste processo os pesos são ajustados de acordo com uma regra de correção de erro. O treinamento de uma rede através de retropropagação é realizado de forma supervisionada, ou seja, é apresentada à rede uma determinada entrada e é disponibilizada a resposta desejada para aquela entrada (Andrade, 2001).

As redes neurais artificiais podem realizar várias tarefas, desde fazer os carros dirigirem automaticamente, gerar rostos CGI (*Computer Graphic Imagery* ou Computação Gráfica) realistas, tradução automática, reconhecimento facial e outros trabalhos que podem ser desenvolvidos com as redes neurais artificiais, como: aplicação das Redes Neurais Artificiais na detecção, classificação e localização de defeitos em linhas de transmissão; Aplicação de Redes Neurais Artificiais na Construção de Modelos de Fragilidade Ambiental; Utilização de Redes Neurais Artificiais para Classificação da Spam; Aplicação da Inteligência Artificial em Sistemas de Gerenciamento de Conteúdo; Utilizando Redes Neurais Artificiais e Regressão Logística na Predição de Hepatite A, entre outras aplicações (Lima; Junior, 2020).

O objetivo original da abordagem de rede neural era criar um sistema computacional capaz de resolver problemas como um cérebro humano. No entanto, com o passar do tempo, os pesquisadores mudaram o foco e passaram a usar redes neurais para resolver tarefas específicas, desviando-se de uma abordagem

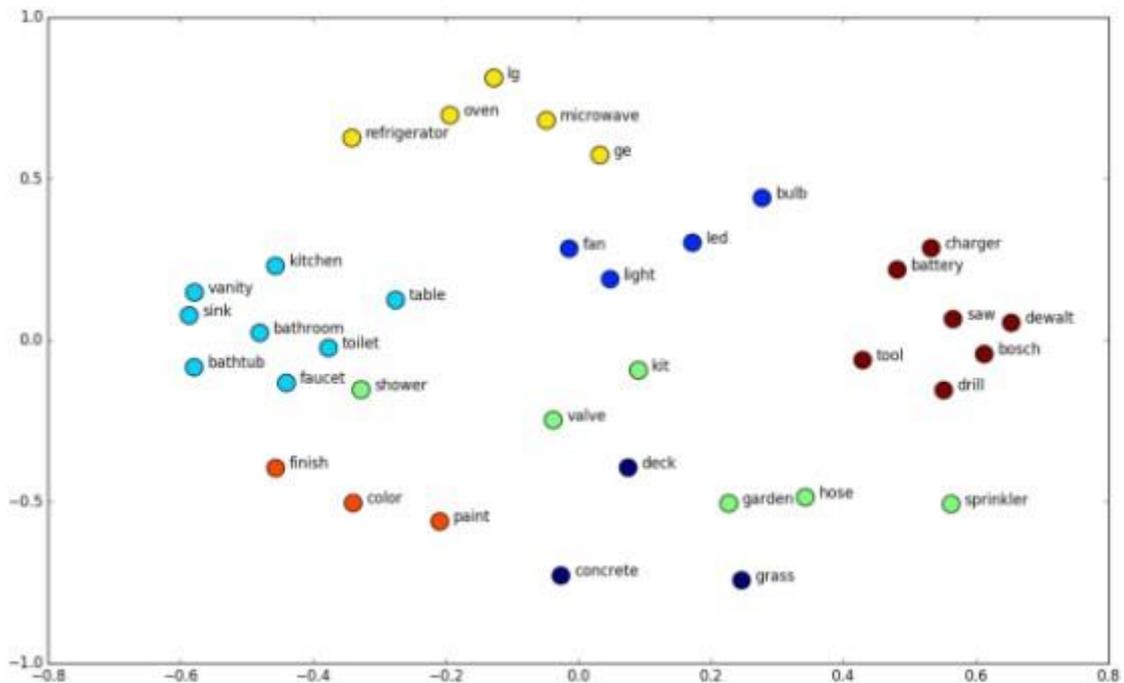
estritamente biológica. Desde então, as redes neurais têm oferecido suporte às mais diversas tarefas, incluindo visão computacional, reconhecimento de fala, tradução de máquina, filtragem de redes sociais, jogos de tabuleiro ou videogame e diagnósticos médicos (Sas, 2020). Conforme o volume de dados estruturados e não-estruturados aumentou a níveis de big data, as pessoas desenvolveram sistemas de *deep learning* que são, essencialmente, redes neurais com muitas camadas. *Deep Learning* permite a captura e a mineração de mais e maiores volumes de dados, incluindo dados não-estruturados (Sas, 2020).

### **Verocímetro: uma proposta para combater notícias falsas**

A primeira proposta é utilizar o método Word Embedding por ser um tipo de representação de palavras que permite que palavras com significado semelhante sejam compreendidas por algoritmos de aprendizado de máquina. Essas representações são mapeadas em forma de vetores densos de alta dimensionalidade, onde cada entrada é uma medida entre a associação entre a palavra e o contexto em que ela se encontra (Goldberg; Levy, 2014). Palavras com alto grau de similaridade são mapeadas em vetores próximos, por exemplo Garden (jardim) e Hose (Mangueira), tendem a estar mais interligadas semanticamente (Agrawal, 2020).

Pode-se notar, na figura 3, que a palavra shower (chuveiro), a depender do contexto ela pode ter uma grande relação com as palavras Garden e Hose, mas tendem a estar mais próximas de palavras como bathtub (banheira) e toilet (banheiro) (Agrawal, 2020). Além da similaridade entre as palavras, estes vetores podem ser utilizados para diversas outras atividades da PLN (processamento de linguagem natural), como reconhecimento da linguagem natural, recuperação de informação, classificação de documentos, análise de sentimentos, parsing, dentre outras (Carvalho, 2018).

**Figura 3 – Word Embeddings**



Fonte: Rohit Agrawal (2010)

A segunda proposta é utilizar a técnica Word2Vec por ser uma ferramenta open-source desenvolvida por Tomás Mikolov em 2013 no Google (Karani, 2018) sendo utilizada para calcular representações de palavras como vetores. O Word2Vec através de um conjunto de frases transforma cada palavra em um vetor numérico denso de tamanho fixo, ou seja, de 300, 50 ou de 60 posições que a representa semanticamente (Contratres, 2018). Este método absorve o contexto de cada palavra e tenta prever a palavra correspondente ao contexto (Karani, 2018).

Ele é baseado em aprendizagem profunda preditiva para calcular e gerar representações vetoriais densas de palavras distribuídas e contínuas de alta qualidade, que capturam semelhanças contextuais e semânticas. Essencialmente, esses são modelos não supervisionados que podem receber corpo textual massivo, criar um vocabulário de palavras possíveis e gerar encaixes de palavras densas para cada palavra no espaço vetorial que representa esse vocabulário. Normalmente, pode-se especificar o tamanho dos vetores de incorporação de palavras e o número total de vetores é essencialmente o tamanho do vocabulário. Esta técnica possui dois tipos de abordagem: *Skip-ram* ou *Common Bag Of Words* (CBOW) (Sarkar, 2018).

O modelo Skip-gram faz um loop nas palavras de cada frase e tenta usar a palavra atual para prever seus vizinhos (Ali, 2019). De acordo com Carvalho (2018)

neste modelo dada uma palavra pressupõe-se as palavras mais prováveis ao contexto. Contrates (2018) descreve que o objetivo é prever o contexto do qual esta palavra veio, para isto também pode-se utilizar uma rede neural como por exemplo o algoritmo Naive Bayes (Silva, 2019). O modelo CBOW utiliza o contexto para prever uma palavra-alvo (Ali, 2019). De acordo com Mikolov (2013) este modelo visa maximizar a classificação de uma palavra com base em outra palavra na mesma frase. Esta arquitetura é também conhecida como modelo do saco contínuo de palavras que tenta prever a palavra-alvo atual (a palavra central) com base nas palavras do contexto de origem (palavras ao redor) (Sarkar, 2018).

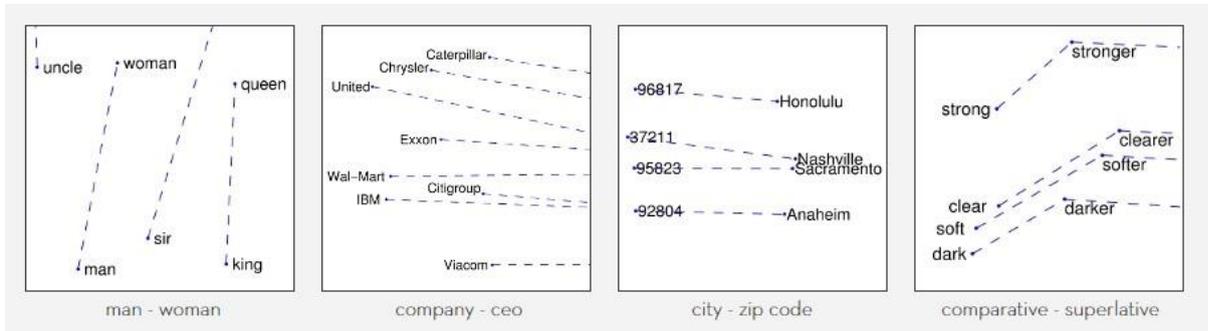
A terceira proposta é utilizar a técnica *Global Vector for Word Representation* (GloVe) pois necessitamos comparar conjuntos de palavras e analisar as mesmas a fim de treinar o algoritmo e fazê-lo aprender a definir, a partir desses conjuntos, a veracidade da informação. GloVe é um modelo global de regressão *log-bilinear* descrito como uma combinação dos pontos fortes dos métodos *global matrix factorization* e *local context window* (Carvalho, 2018). Os modelos de espaço vetorial semântico da linguagem representam cada palavra com um vetor de valor real. Estes vetores podem ser usados como recursos em uma variedade de aplicações, como recuperação de informações (Manning et al., 2008), classificação de documentos (Sebastiani, 2002), resposta a perguntas (Tellex et al., 2003), reconhecimento de entidade nomeada (Turian et al., 2010) e análise (Socher et al. 2013).

A maioria dos métodos de vetor de palavras dependem da distância ou ângulo entre pares de vetores de palavras como o método principal para avaliar a qualidade do conjunto de representações de palavras. Recentemente, Mikolov (2013) introduziu um novo esquema de avaliação baseado em analogias de palavras que investiga a estrutura mais fina do espaço vetorial de palavras, mas sim suas várias dimensões de diferença. Por exemplo, a analogia “rei é para rainha como o homem é para a mulher” deve ser codificada no espaço vetorial pela equação vetorial rei – rainha = homem – mulher. Este esquema de avaliação favorece modelos que produzem dimensões de significado, capturando assim a ideia de multi-agrupamento de representações distribuídas (Bengio, 2009).

Pode-se observar na figura 4, temos um modelo de subestruturas lineares que faz a associação de palavras e o possível significado especificado pela posição dessas palavras. Notamos que no quadrado 1, man está para woman; quadrado 2,

company para CEO; quadrado 3, código postal para cidade e quadrado 4, comparativo para superlativo.

**Figura 4 – GloVe Embedding Space**

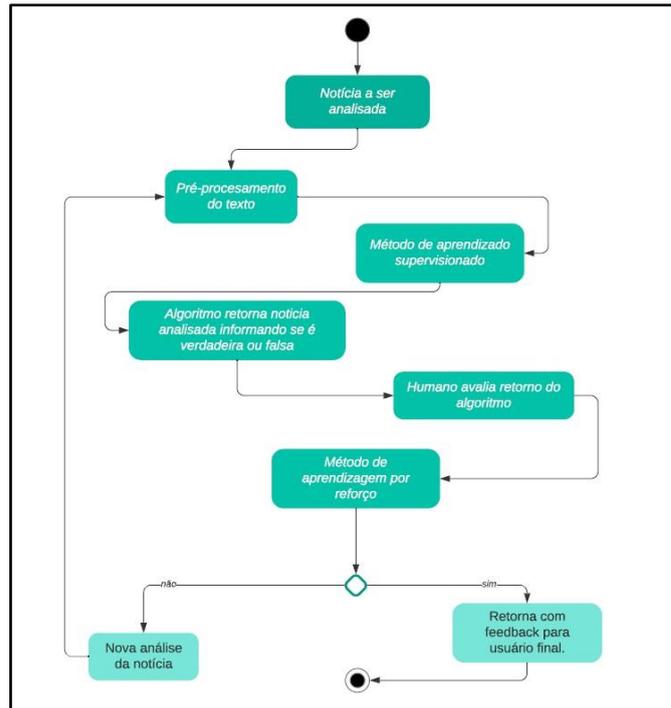


Fonte: Pennington *et al.* (2014)

Em relação a linguagem de programação utilizada será Python por ser uma linguagem de alto nível, interpretada, de script, multiparadigma; possui um modelo de desenvolvimento comunitário aberto e gerenciado pela Python Software Foundation; É a linguagem mais usada em aplicações de *Machine Learning* que é *open-source*; possui um rico ecossistema de bibliotecas para aprendizagem de máquina e análise de dados; rápida adoção pela comunidade científica entre outras características.

No fluxo para análise da veracidade da informação o algoritmo irá buscar no banco de dados a notícia enviada através da plataforma web, o texto será pré-processado através dos métodos: Word Embedding, Word2Vec e Global Vector for Word Representation (GloVe). Será criado uma tabela na base de dados contendo notícias verdadeiras que servirá para comparar com a informação enviada pelo usuário final realizando desta forma o aprendizado supervisionado.

Após análise da notícia o algoritmo retorna "fato" ou "fake", com base nesse retorno um humano revisa a análise realizada, após esse procedimento o humano interage com o algoritmo informando se houve acerto ou erro realizando desta forma o aprendizado por reforço. Se houver acerto a análise realizada é enviada para o usuário final, se houver erro uma nova análise é realizada, conforme figura 5.

**Figura 5** – Fluxo para análise da veracidade da informação

Fonte: Autoria própria

## Considerações finais

Neste ponto, percebe-se que as *fakes news* são informações falsas repassadas como se fossem reais. Este tema tem se tornado cada vez mais comum em todos os tipos de mídias, seja alertando para não credibilizar e não compartilhar notícias falsas ou instruindo para reconhecer se a informação é verdadeira por meio de propagandas. A todo instante, notícias falsas são espalhadas por diversos tipos de mídias, principalmente pela internet, onde todos têm acesso livre.

Notícias falsas existem antes da criação de mídias de informações digitais. Atualmente existe facilidade das pessoas publicarem qualquer coisa através da internet e compartilhar para outras milhares de pessoas. Principalmente em mídias sociais, esse ciclo de disseminação contamina como se fosse um vírus, em milésimos de segundos, milhares de pessoas têm o conhecimento de determinada notícia. A plataforma Verocímetro visa combater a *Fake News*, oferecendo por benefício proporcionar ao usuário acesso a uma análise da veracidade de determinada informação de forma clara, objetiva e gratuita.

Tendo por objetivo alcançar leitores e pessoas que estão em busca de informações verdadeiras, o projeto Verocímetro apresenta uma solução onde o usuário, por meio de um link, título ou fragmentos da notícia, descobre se a notícia divulgada é verdadeira ou falsa. Para tornar o projeto mais dinâmico, futuramente usaremos Inteligência Artificial (IA) em nossa base de dados para detectar automaticamente e informar dentro de alguns segundos a resposta para o leitor.

Para trabalhos futuros, é necessária a implementação utilizando as técnicas anteriormente apresentadas no capítulo descrição do projeto e alternativas que possam contribuir para uma maior efetividade na busca pela veracidade da informação. Além disso, o projeto tende a ser implementado em uma plataforma de fácil acesso ao usuário final, estimulando a busca pela confiabilidade da notícia ao invés de acreditar em qualquer tipo de informação.

## Referências

AGRAWAL, R. **Using fine-tuned Gensim Word2Vec Embeddings with Torchtext and Pytorch**. 2020. Disponível em: [https://medium.com/@rohit\\_agrawal/using-fine-tuned-gensim-word2vec-embeddings-with-torchtext-and-pytorch-17eea2883cd](https://medium.com/@rohit_agrawal/using-fine-tuned-gensim-word2vec-embeddings-with-torchtext-and-pytorch-17eea2883cd). Acesso em: 20 jun. 2023.

ALI, Z. **A simple Word2vec tutorial**. 2019. Disponível em: <https://medium.com/@zafaralibagh6/a-simple-word2vec-tutorial-61e64e38a6a1>. Acesso em: 20 jun. 2023.

ALLCOTT, H.; GENTZKOW, M. "Social Media and Fake News in the 2016 Election". **Journal of Economic Perspectives**, v.31, n. 2, p. 211–236, 2017.

ANDERSON, D; MCNEILL, G. Artificial Neural Networks Technology. 1992. Disponível em: [http://andrei.clubcisco.ro/cursuri/f/f-sym/5master/aac-nnga/AI\\_neural\\_nets.pdf](http://andrei.clubcisco.ro/cursuri/f/f-sym/5master/aac-nnga/AI_neural_nets.pdf). Acesso em: 20 jun. 2023.

ANDRADE, J. F. A. de et al. **Geração de Trajetórias para Robôs Móveis Autônomos Usando Redes Neurais Artificiais**. 2001. Disponível em: [http://abricom.org.br/wp-content/uploads/2016/03/5cbrn\\_122.pdf](http://abricom.org.br/wp-content/uploads/2016/03/5cbrn_122.pdf). Acesso em: 20 jun. 2023.

BARBOSA, J. L. et. al Introdução ao Processamento de Linguagem Natural usando Python. In: **III Escola Regional de Informática do Piauí**, v. 1, n. 1, p. 336-360, 2017.

BELLMAN, R. E. **"Dynamic Programming"**. Princeton University Press, 2010, Princeton.

BENGIO, Y. **Learning Deep Architectures for AI**. Boston: Now Publishers Inc, 2009. 131 p. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=cq5ewg7FniMC&oi=fnd&pg=PA1&dq=Learning+Deep+Architectures+for+AI+bengio+bengio+referencia&ots=KpebRTohKx&sig=6a55Cn4yGKuKhEX0l2pDH62U7qA#v=onepage&q&f=false>. Acesso em: 20 jun. 2023.

BURKHARDT, J. M. "History of Fake News". **Library Technology Reports**, v. 53, n. 8, 2017.

CALDAS, W. L. **Proposta de dois Métodos Semi-Supervisionados Baseados na Máquina de Aprendizagem Mínima Utilizando Co-Training**. 2017. Disponível em: [http://www.mdcc.ufc.br/teses-e-dissertacoes/dissertacoes-de-mestrado/doc\\_download/333-252-wesley-lioba-caldas](http://www.mdcc.ufc.br/teses-e-dissertacoes/dissertacoes-de-mestrado/doc_download/333-252-wesley-lioba-caldas). Acesso em: 20 jun. 2023.

CAMPOS, L. V. "O que são fake news?". 2020. Disponível em: <<https://brasilecola.uol.com.br/curiosidades/o-que-sao-fake-news.htm>>. Acesso em dezembro de 2020.

CARVALHO, F; ALMEIDA, L. **Aprendizagem de Máquina**. 2020. Disponível em: <https://www.cin.ufpe.br/~fatc/AM/AM-Introducao-2020-1.pdf>. Acesso em: 20 jun. 2023.

CARVALHO, M. H. **Estudo Comparativo dos Métodos de Word Embedding na Análise de Sentimentos**. 2018. Disponível em: [https://www.cin.ufpe.br/~tg/2018-2/TG\\_CC/tg\\_mhc.pdf](https://www.cin.ufpe.br/~tg/2018-2/TG_CC/tg_mhc.pdf). Acesso em: 20 jun. 2023.

CONTRATRES, F. **Similaridade entre títulos de produtos com Word2Vec**. 2018. Disponível em: <https://medium.com/luizalabs/similaridade-entre-t%C3%ADtulos-de-produtos-com-word2vec-5e26199862f0>. Acesso em: 20 jun. 2023.

COUTINHO, R. "Fake News: como identificar e evitar a disseminação". 2018. Disponível em: <<https://www.security.ufrj.br/geral/fake-news-como-identificar-e-evitar-a-disseminacao/>>. Acesso em 20 jun. 2023.

DAYHOFF, J. E.; DELEO, J. M. **Artificial neural networks**. 2001. Disponível em: [https://acsjournals.onlinelibrary.wiley.com/doi/10.1002/1097-0142\(20010415\)91:8+%3C1615::AID-CNCR1175%3E3.0.CO;2-L](https://acsjournals.onlinelibrary.wiley.com/doi/10.1002/1097-0142(20010415)91:8+%3C1615::AID-CNCR1175%3E3.0.CO;2-L). Acesso em: 20 jun. 2023.

DELMAZO, C.; VALENTE, J. C.L. "Fake news nas redes sociais online: propagação e Reações à desinformação em busca de cliques". **Media & Jornalismo**, Portugal, v. 18, n.32, 2018.

FILHO, O. F. "O que é falso sobre fake News". **Revista USP**, n.116, p. 39-44, 2018. <https://doi.org/10.11606/issn.2316-9036.v0i116p39-44>

GOLDBERG, Y; LEVY, O. **Word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method**. 2014. Disponível em: <https://arxiv.org/pdf/1402.3722.pdf>. Acesso em: 20 jun. 2023.

- GOMES, P. C. T. **Introdução ao aprendizado de máquina**. 2019. Disponível em: <https://www.datageeks.com.br/aprendizado-de-maquina/>. Acesso em: 20 jun. 2023.
- KARANI, D. **Introduction to Word Embedding and Word2Vec**. 2018. Disponível em: <https://towardsdatascience.com/introduction-to-word-embedding-and-word2vec-652d0c2060fa>. Acesso em: 20 jun. 2023.
- KURZWEIL, R. **“The Age of Intelligent Machines”**. MIT Press; Cambridge, 1992.
- LEETARU, K. **“Why Stopping ‘Fake’ News Is So Hard”**. Forbes, 2016.
- LINHARES, L. C. et. al. **“A inviolabilidade à privacidade (intimidade, vida privada, honra e imagem): CF/88 x atual realidade”**. 2017. Disponível em: <https://jus.com.br/artigos/60125/a-inviolabilidade-a-privacidade-intimidade-vida-privada-honra-e-imagem-cf-88-x-atual-realidade>. Acesso em: 20 jun. 2023.
- MANNING, C. D. et al. **Introduction to Information Retrieval**, Cambridge University Press. 2008.
- MIKOLOV, T. et al. **Efficient Estimation of Word Representations in Vector Space**. 2013. Disponível em: <https://arxiv.org/pdf/1301.3781.pdf>. Acesso em: 20 jun. 2023.
- PEDROSO, A. **Fake News Analisando Textos e Descobrimo Verdades**, USP, São Paulo, 2019.
- PEDUTI Advogados. **“Porque é tão difícil combater o Fake News”**. 2019. Disponível em: <https://peduti.com.br/blog/porque-e-tao-dificil-combater-o-fake-news/>. Acesso em: 20 jun. 2023.
- PENNINGTON, J. et al. **GloVe: Global Vectors for Word Representation**. 2014. Disponível em: <https://nlp.stanford.edu/projects/glove/>. Acesso em: 20 jun. 2023.
- RODRIGUES, J. **“O que é o Processamento de Linguagem Natural?”**. 2017. Disponível em: <https://medium.com/botsbrasil/o-que-%C3%A9-o-processamento-de-linguagem-natural-49ece9371cff>. Acesso em dezembro de 2020.
- ROLIM, V. et al. **Utilização de Técnicas de Aprendizado de Máquina para Acompanhamento de Fóruns Educacionais**. 2017. Disponível em: [https://www.researchgate.net/publication/320743305\\_Utilizacao\\_de\\_Tecnicas\\_de\\_Aprendizado\\_de\\_Maquina\\_para\\_Acompanhamento\\_de\\_Foruns\\_Educacionais](https://www.researchgate.net/publication/320743305_Utilizacao_de_Tecnicas_de_Aprendizado_de_Maquina_para_Acompanhamento_de_Foruns_Educacionais). Acesso em: 20 jun. 2023.
- RUSSELL, S. J.; NORVIG, P. **Inteligência artificial**. 3. ed. Rio de Janeiro: Elsevier Editora Ltda, 2013.
- SANCHES, M. F. **“Processamento e Entendimento de Linguagem Natural no Gerenciamento de Emergências para Obtenção de Consciência Situacional”**. 2017. Centro Universitário Eurípedes de Marília, Marília.

SARKAR, D. **A hands-on intuitive approach to Deep Learning Methods for Text Data** — Word2Vec, GloVe and FastText. 2018. Disponível em: <https://towardsdatascience.com/understanding-feature-engineering-part-4-deep-learning-methods-for-text-data-96c44370bbfa>. Acesso em: 20 jun. 2023.

SEBASTIANI, F. **Machine Learning in Automated Text Categorization**. 2002. Disponível em: <http://nmis.isti.cnr.it/sebastiani/Publications/ACMCS02.pdf>. Acesso em: 20 jun. 2023.

SILVA V. N. (2020) Processamento de Linguagem Natural (PLN). Disponível em <https://operdata.com.br/blog/pln-processamento-de-linguagem-natural/>. Acesso em: 20 jun. 2023.

SENSACIONALISTA. Sensacionalista. Disponível em: [<https://www.sensacionalista.com.br/>](https://www.sensacionalista.com.br/). Acesso em: 28 jun. 2025

SILVA, Jacson Rodrigues Correia da. **Processamento de Linguagem Natural (PLN)**. 2020. Disponível em: <https://inbot.com.br/apoio/artigos/Processamento-de-Linguagem-Natural-PLN-Jacson-Rodrigues-UFES.pdf>. Acesso em: 20 jun. 2023.

SILVA, L. C. **“Aprendizado de Máquina com Treinamento Continuado Aplicado à Previsão de Demanda de Curto Prazo: O caso do Restaurante Universitário da Universidade Federal de Uberlândia**. 2019. Universidade Federal de Uberlândia, Uberlândia.

SINGH, M. **“WhatsApp’s new limit cuts virality of ‘highly forwarded’ messages by 70%”**. Disponível em: [< https://www.niemanlab.org/reading/whatsapps-new-resending-limit-cuts-the-number-of-highly-forwarded-messages-by-70/>](https://www.niemanlab.org/reading/whatsapps-new-resending-limit-cuts-the-number-of-highly-forwarded-messages-by-70/). Acesso em 20 jun. 2023.

SOCHER, R. et al. **Parsing with Compositional Vector Grammars**. 2013. Disponível em: <https://www.aclweb.org/anthology/P13-1045.pdf>. Acesso em: 20 jun. 2023.

SPERANDIO, H. R. C. **“Desafios da Inteligência Artificial para a Profissão Jurídica”**. 2018. Disponível em: <http://bibliotecadigital.fgv.br/dspace/handle/10438/23977>>. Acesso em: 20 jun. 2023.

TAKE. **Tudo sobre NLP: o que é processamento de linguagem natural e seus desafios na Inteligência Artificial**. 2019. Disponível em: <https://take.net/blog/devs/nlp-processamento-linguagem-natural>. Acesso em: 20 jun. 2023.

TEIXEIRA, L. A. **Métodos de Regressão para Aprendizado por Reforço**. 2016. Disponível em: <https://www.ufjf.br/getcomp/files/2013/03/M%C3%89TODOS-DE-REGRESS%C3%83O-PARA-APRENDIZADO-POR-REFOR%C3%87O.pdf>. Acesso em: 20 jun. 2023.

TELLEX, S. et al. **Quantitative Evaluation of Passage Retrieval Algorithms for Question Answering**. 2003. Disponível em:

<https://groups.csail.mit.edu/infolab/publications/Tellex-et-al-SIGIR03.pdf>. Acesso em: 20 jun. 2023.

TODA MATÉRIA. **Neurônios**. 2020. Disponível em: <https://www.todamateria.com.br/neuronios/>. Acesso em: 20 out. 2020.

TURIAN, J. et al. **Word representations**: A simple and general method for semi-supervised learning. 2010. Disponível em: <https://www.aclweb.org/anthology/P10-1040.pdf>. Acesso em: 20 jun. 2023.

WANG, W. Y. “Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection”. **ArXiv**. Cornell University, Cornell, 2017.