CENTRO UNIVERSITÁRIO ALVES FARIA (UNIALFA)
GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO

MATHEUS DE ARAÚJO OLIVEIRA

CLASSIFICAÇÃO DE NOTÍCIAS UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

GOIÂNIA
DEZEMBRO de 2022

MATHEUS DE ARAÚJO OLIVEIRA

CLASSIFICAÇÃO DE NOTÍCIAS UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do grau de Bacharel em Engenharia da Computação do Centro Universitário Alves Faria (UNIALFA), sob a orientação do Prof.ª Ma. Fabiana Rocha de Andrade e Silva.

Orientadora: Profa. Ma. Fabiana Rocha de Andrade Silva

GOIÂNIA

DEZEMBRO de 2022

MATHEUS DE ARAÚJO OLIVEIRA

CLASSIFICAÇÃO DE NOTÍCIAS UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

Este Trabalho de Conclusão de Curso foi analisado e julgado adequado para obtenção do título de Bacharel em Engenharia da Computação e aprovado em sua forma final pelo Orientador(a) e Banca Examinadora designada pelo Curso de Engenharia da Computação do Centro Universitário Alves Faria (UNIALFA).

Prof. Ma. Fabiana Rocha de Andrade e Silva – UNIALFA

(Orientadora)

Nayane Moura Neris (Especialista) – UNIALFA

GOIÂNIA

DEZEMBRO DE 2022

RESUMO

OLIVEIRA, Matheus de Araújo. **Classificação de notícias utilizando técnicas de aprendizado de máquina.** Trabalho de conclusão de curso 24 f. Graduação em Engenharia da Computação. Goiânia, 2022.

Este trabalho tem como objetivo identificar notícias falsas por meio de algoritmos de aprendizado de máquina em um banco de dados contendo uma variedade de notícias, incluindo notícias reais e falsas. Além da identificação, também tenta fornecer métricas através da matriz de confusão, determinar seu desempenho e utilizar mapas de calor para identificar possíveis padrões entre notícias verdadeiras e falsas. Um método de pesquisa exploratória foi usado para realizar este trabalho, com o objetivo de desenvolver um classificador de mensagens. Em termos de métodos e procedimentos de pesquisa, este estudo é bibliográfico e classificado como experimental, utilizou-se um banco de dados existente contendo os dados necessários para o desenvolvimento. Foi alcançado uma precisão de 94% para contribuir no combate a fake news, acredita-se que ainda há bastante espaço para melhorias, visto que se trata de uma área da tecnologia ainda em desenvolvimento.

Palavras-Chaves: Notícias falsas. Classificador. Inteligência artificial.

ABSTRACT

OLIVEIRA, Matheus de Araújo. **News classification using machine learning techniques.** Completion of course work 24 f. University graduate in Computer Engineering. Goiânia, 2022.

This work aims to identify fake news through machine learning algorithms in a database containing a variety of news, including real and fake news. In addition to identification, it also tries to provide metrics through the confusion matrix, determine its performance and use heat maps to identify possible patterns between true and false news. An exploratory research method was used to carry out this work, with the aim of developing a message classifier. In terms of research methods and procedures, this study is bibliographic and classified as experimental, using an existing database containing the necessary data for development. An accuracy of 94% was achieved to contribute to the fight against fake news, it is believed that there is still plenty of room for improvement, as this is an area of technology still under development.

KEY-WORDS: Fake news. Classifier. Machine learning.

SUMÁRIO

1 INTRODUÇÃO **6**

2 FUNDAMENTAÇÃO TEÓRICA **8**
2.1 Fake news **8**
2.2 Inteligência artificial ***9***
2.2.1 Aprendizado de máquina 9
2.2.2 Aprendizagem supervisionada 9
2.2.3 Aprendizagem não supervisionada 9
2.3 Redes neurais artificiais **10**
2.3.1 Rede neural perceptron multicamadas 11
2.4 Redes neurais convolucionais **12**
2.5 Função de ativação **12**
2.6 Métricas de avaliação **13**
2.6.1 Matriz de confusão 14
2.6.2 Acurácia 14
2.6.3 Precisão 15
2.6.4 Recall 15

3 METODOLOGIA **16**

4 CLASSIFICADOR DE FAKE NEWS **17**
4.1 Conjunto de Dados **17**
4.2 Pré-processamento **18**
4.2.1 Conversão para letras minúsculas 18
4.2.2 Stopwords 19
4.3 Treinamento do modelo **20**
4.4 Desempenho do classificador **21**

5 CONCLUSÃO **22**

REFERÊNCIAS **24**

1 INTRODUÇÃO

As *fakes news* têm se tornado cada vez mais notórias nos últimos anos, porém notícias falsas sempre existiram. Há registros de notícias falsas que datam do século VI, como relata o historiador Robert Darnton, professor emérito da Universidade de Harvard (VICTOR, 2017).

Durante as eleições americanas de 2016 e o referendo do *Brexit*, as notícias falsas exerceram grande poder de influência sobre as massas (SILVERMAN, 2016).

O site de notícias *BuzzFeed*[1] analisou as 20 histórias falsas de maior sucesso da campanha eleitoral de 2016 nos Estados Unidos, gerando 8.711.000,00 compartilhamentos, reações e comentários no Facebook. Das 20 principais notícias falsas, apenas 3 endossaram Hillary Clinton, mas 17 endossaram Donald Trump (SILVERMAN, 2016). Vale ressaltar que Donald Trump venceu a eleição e se tornou o 45º presidente dos Estados Unidos.

Durante a campanha do *Brexit*, algumas informações falsas foram usadas pela campanha *Vote to Leave*, que venceu o referendo do Reino Unido sobre a saída da União Europeia (UE). A eleição presidencial brasileira de 2018 também viu a disseminação de notícias falsas (PEREIRA et al., 2018).

O fenômeno da disseminação de notícias falsas é hoje considerado um dos principais fatores de influência dos acontecimentos políticos. Pesquisas mostram que notícias falsas têm mais chances de serem compartilhadas do que conteúdo real. Experimentos realizados na plataforma do Twitter mostram que informações falsas têm 70% mais chances de serem compartilhadas e se espalharem de forma mais rápida e profunda na web (SHARIATMADARI, 2019).

Uma notícia falsa tem o potencial de influenciar pessoas negativamente e, devido ao grande volume de notícias e à dificuldade de categorizá-las, levanta-se questões como:

"É possível classificar mensagens usando técnicas de aprendizado de máquina?".

Para responder a essa pergunta, tem-se como objetivos:

[1]Disponível em: www.buzzfeed.com Acesso em: 19 mar 2022.

Geral

Desenvolver um classificador de notícias em português brasileiro capaz de distinguir notícias falsas usando técnicas de aprendizado de máquina.

Específico

- Relembrar as teorias relacionadas a aprendizagem de máquina, como redes neurais, aprendizagem profunda e classificadores.
- Definir quais algoritmos, linguagens e bibliotecas serão empregadas.
- Definir a base de dados que será utilizada para os testes.
- Desenvolver e treinar um protótipo usando os conhecimentos obtidos, o modelo deverá ser capaz de classificar as notícias utilizando apenas o seu corpo textual.
- Analisar a acurácia do modelo após ser treinada e comparar com trabalhos considerados no estado da arte.
- Validar se o protótipo obteve resultados satisfatórios.

Com isso em mente, é democraticamente importante ter uma maneira de classificar efetivamente as mensagens como verdadeiras ou falsas. Ferramentas para isso já são uma realidade em inglês, mas o português ainda carece de conteúdo acadêmico sobre o assunto. Este estudo, portanto, aborda esta questão.

Trata-se de uma pesquisa aplicada voltada para a construção de um classificador de notícias visando a geração de conhecimento que possa ser aplicado a problemas específicos.

Os métodos de pesquisa utilizados são exploratórios, pois a pesquisa exploratória pode aplicar métodos qualitativos e quantitativos.

A abordagem utilizada foi o método do triângulo, que combina duas ou mais abordagens metodológicas quantitativas e qualitativas. Em termos de métodos e procedimentos de pesquisa, esta pesquisa pode ser dividida em pesquisa bibliográfica e experimental.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 *Fake news*

O termo *fake news* pode ser entendido pela sua tradução livre do inglês para notícias falsas, porém outros autores dão uma maior profundidade ao termo, dessa forma *fake news* não se trata apenas de notícias com informações erradas ou falsas, mas que também tem a intenção de manipular o leitor, por serem intencionalmente falsas (ALLCOTT; GENTZKOW, 2017). Diferenciando-se de notícias de carácter jornalístico, mas com fatos incorretos devido a erros na apuração dos fatos, assim como mostra Figura 1.

Figura 1 — Atributos importantes para definição de *fake news*



Fonte: Elaborada pelo autor (2022).

2.2 Inteligência artificial

Em termos simples, a inteligência artificial (IA) refere-se a um sistema ou máquina que imita a inteligência humana para realizar tarefas e que tem a habilidade de se aprimorar iterativamente com base nas informações coletadas (ORACLE, c2022).

2.2.1 Aprendizado de máquina

O aprendizado de máquina (AM) é um tipo específico de inteligência artificial que treina máquinas para aprender com dados, simplificadamente trata-se de programa de computador que toma decisões com base na experiência acumulada com problemas previamente resolvidos com sucesso (BARANAUSKAS; MONARD, 2003).

As técnicas de AM podem ser amplamente classificadas em aprendizagem supervisionada e aprendizagem não supervisionada.

2.2.2 Aprendizagem supervisionada

A aprendizagem supervisionada lida com problemas conhecidos e, durante o processo de treinamento, a máquina recebe um conjunto de dados rotulados compostos por entradas possíveis correlacionadas com saídas esperadas (TIBCO, c2022).

Por exemplo, um processo de aprendizagem poderia ser o de classificar gatos e cachorros a partir de suas imagens, para isso, na etapa de treinamento, a máquina recebe 100 imagens, contendo 50 imagens de gato e 50 imagens de cachorro rotuladas; baseado nas imagens recebidas a máquina aprende, desse modo a máquina será capaz de diferenciar e classificar gatos e cachorros.

2.2.3 Aprendizagem não supervisionada

Em contraste, o aprendizado não supervisionado é um tipo de AM usado para identificar novos padrões e detectar anomalias, o conjunto de dados que alimenta a

máquina consiste em informações não rotuladas, que a máquina auxiliada por algoritmos, intenta dar semântica aos dados por conta própria, procurando padrões.

Um exemplo de pergunta que o aprendizado de máquina não supervisionado pode responder é “Existem novos padrões de compra?”

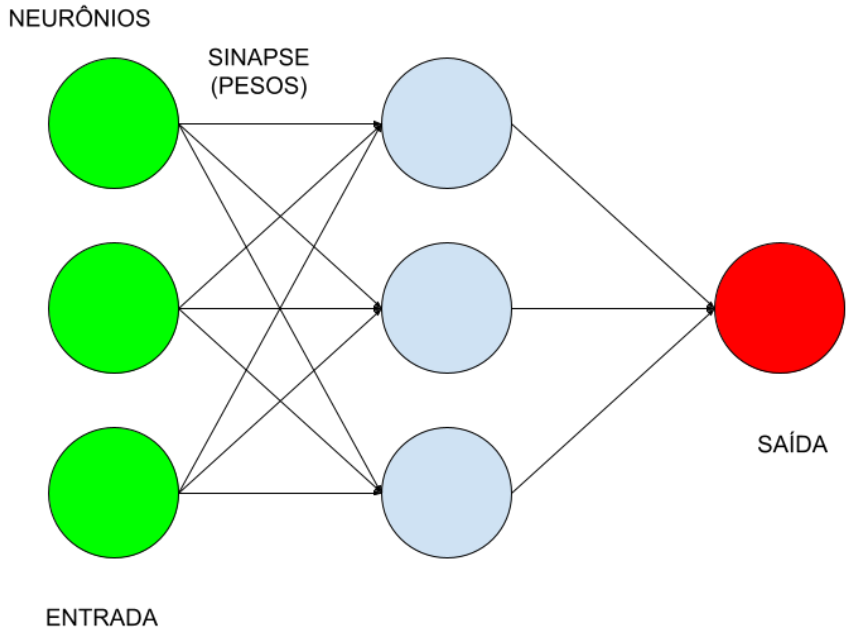
2.3 Redes neurais artificiais

As Redes neurais artificiais (RNA) são um conjunto de algoritmos de aprendizado de máquina e inteligência artificial, que é biologicamente inspirado no cérebro humano, mimetizando a maneira como os neurônios biológicos se comportam (KUMAR, 2020).

Basicamente, as redes neurais são semelhantes ao cérebro de duas maneiras: o conhecimento é adquirido por meio de etapas de aprendizado e é armazenado, o conhecimento usando pesos sinápticos.

Sinapse é o nome dado para a conexão entre os neurônios, nas conexões é atribuído valores, chamados de pesos sinápticos, conforme Figura 2. Isso mostra claramente que uma rede neural artificial tem em sua composição uma série de neurônios artificiais.

Figura 2 — Estrutura de uma rede neural



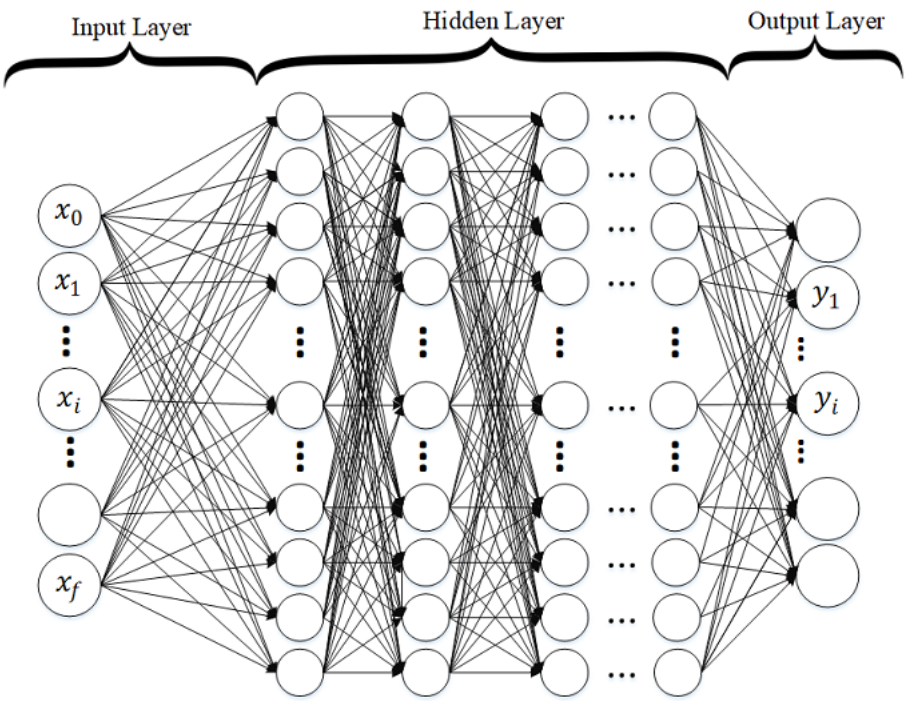
Fonte: Elaborada pelo autor (2022).

Por exemplo, em relação ao sentido da audição é necessário que ruídos externos sejam recebidos através do canal auditivo e enviados aos neurônios para serem processados e como resultado é retornado informações relativas aos ruídos escutados (PRAKASH, 2017).

As Redes neurais artificiais ou RNA são uma teia de neurônios artificiais que enviam sinais uns aos outros, como apresentado na Figura 3. Essas redes funcionam exatamente como o cérebro, aprendendo através da experiência, sem precisar ser especificamente programadas.

Considere como um bebê reconhece um cachorro pela primeira vez, você mostra um cachorro, diz a palavra cachorro, faz isso várias vezes, e agora o cérebro do bebê aprendeu a identificar o padrão entre o visual de um cachorro e o termo usado para isso. Uma máquina aprende exatamente da mesma maneira.

Figura 3 — Arquitetura de uma rede neural artificial



Fonte: Mittal (2019)

2.3.1 Rede neural perceptron multicamadas

O perceptron multicamadas (MLP) é uma variante do modelo perceptron original proposto por Rosenblatt na década de 1950.

Entre as camadas de entrada e saída há uma ou mais camadas ocultas, os neurônios são organizados em camadas, as conexões são sempre de baixo para cima e os neurônios na mesma camada não estão conectados uns aos outros.

O número de neurônios na camada de entrada corresponde ao número de medições no problema padrão, e o número de neurônios na camada de saída corresponde ao número de classes (RAMCHOUN et al., 2016).

2.4 Redes neurais convolucionais

As Redes Neurais Convolucionais (RNC) são semelhantes às Redes neurais artificiais (RNA) tradicionais, pois consistem em neurônios que se auto-otimizam por meio do aprendizado. Cada neurônio ainda recebe entradas e executa operações e retorna uma saída. A única diferença notável entre RNC e RNA tradicionais é que as RNC são usadas principalmente no campo de reconhecimento de padrões de imagem. Isso permite que você codifique recursos específicos da imagem em sua arquitetura e crie redes mais adequadas para tarefas orientadas a imagens (SHEA; NASH, 2015).

Ela tem sido utilizada em modelos voltados para a manipulação de dados de imagem, mas resultados experimentais recentes mostram que as CNNs podem capturar informações presentes em sentenças de forma muito eficiente. Isso é muito importante para manipular e classificar elementos de texto (SHEA; NASH, 2015).

Existem várias maneiras de tratar sentenças como se fossem "imagens" e realizar convoluções. Primeiro, é realizado um processo de tokenização de frase, após o qual as linhas são transformadas em uma matriz de frase, que é uma representação vetorial das palavras em cada token. Para definir as dimensões da matriz final, seja "D" a dimensão do vetor de palavras e "S" o comprimento de cada sentença em número de tokens. Você acaba com uma matriz de dimensão S x D (SHEA; NASH, 2015).

2.5 Função de ativação

As funções de ativação são especialmente usadas em redes neurais artificiais para transformar sinais de entrada em sinais de saída. Os sinais de saída são alimentados como entradas para a próxima camada na pilha.

Uma rede neural artificial calcula a soma dos produtos das entradas e seus pesos correspondentes e, finalmente, aplica uma função de ativação para obter a saída de uma determinada camada, que fornece como entrada para a próxima camada. A precisão da previsão de uma rede neural depende do número de camadas usadas e, mais importante, do tipo de função de ativação usada.

Embora não haja documentação especificando o número mínimo ou máximo de camadas a serem usadas para melhorar os resultados e a precisão de sua rede neural, uma boa regra geral é usar pelo menos duas camadas. O tipo de função de ativação utilizada também não é descrito na literatura (SHARMA et al, 2020).

Um exemplo de função de ativação é a função tangente hiperbólica (TANH). Isso retorna a tangente hiperbólica do ângulo x. O argumento x deve ser expresso em radianos (MEDCALC, 2022c). A tangente hiperbólica pode ser definida conforme equação (1) abaixo:

$$tanh(x) = \frac{sinh(x)}{cosh(x)} \qquad (1)$$

Outra função de ativação muito utilizada é a função sigmóide. Esta função recebe qualquer número real como entrada e retorna um valor entre 0 e 1 (BAHETI, 2022).
Matematicamente, pode ser expresso (equação 2) como:

$$f(x) = \frac{1}{1 + e^{-x}} \qquad (2)$$

2.6 Métricas de avaliação

Ao desenvolver projetos de aprendizado de máquina e ciência de dados, é importante usar as métricas certas para cada problema. Seus valores refletem a qualidade do modelo, portanto, se você fizer uma escolha errada, não poderá avaliar se o modelo realmente atende aos requisitos exigidos.

Por exemplo, erros diferentes podem ter custos diferentes e o cálculo da métrica pode precisar refletir essa diferença. Outra situação é quando modelos de qualidade diferente possuem o mesmo valor para uma métrica, refletindo uma má

escolha ou a necessidade de usar mais de uma. Também é importante considerar o quão interpretável é para as pessoas que trabalham no projeto.

Um modelo de classificação binária visa determinar, entre duas classes possíveis, a qual classe pertence uma nova observação. Duas classes, comumente chamadas de positiva (P) e negativa (N), indicam a ocorrência ou não de um determinado evento. Um exemplo é classificar se um determinado paciente tem uma determinada doença (positivo) ou não (negativo).

2.6.1 Matriz de confusão

Uma maneira muito fácil de visualizar o desempenho de um modelo de classificação é a matriz de confusão. Esta matriz mostra quantos exemplos existem em cada grupo: falsos positivos (FP), falsos negativos (FN), verdadeiros positivos (TP) e verdadeiros negativos (TN). Como o número de exemplos em cada classe pode variar, é interessante mostrar os números desses grupos tanto em números absolutos quanto em percentuais das classes reais. A figura 04 demonstra visualmente a matriz de confusão.

Figura 4 — Exemplo de uma matriz de confusão



Fonte: KUNUMI (2020)

2.6.2 Acurácia

A acurácia, ou *accuracy em inglês,* dá o número de exemplos que são realmente classificados corretamente, independentemente da classe.

Por exemplo, se você tiver 100 observações e 90 delas forem classificadas corretamente, seu modelo terá 90% de acurácia (KUNUMI, 2020). A acurácia é definida pela fórmula da Figura 05.

Figura 5 — Fórmula da acurácia

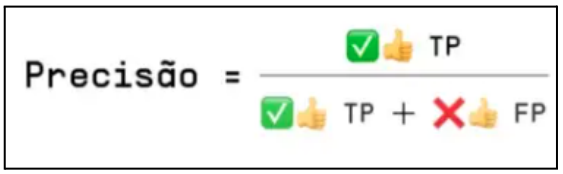


Fonte: KUNUMI (2020)

2.6.3 Precisão

A precisão também é uma das métricas mais populares para avaliar modelos de classificação. Essa métrica é definida pela razão entre o número de exemplos positivos avaliados corretamente e o número total de exemplos avaliados positivamente de acordo com a fórmula da Figura 06.

Figura 6 — Fórmula da precisão

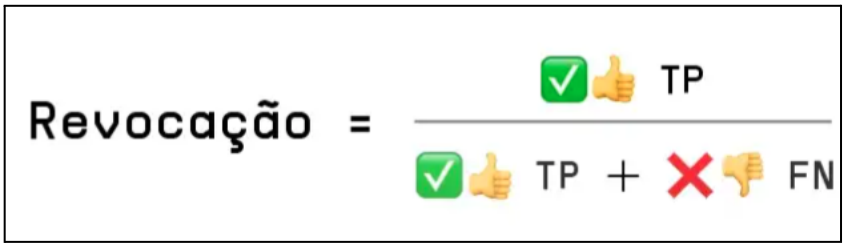


Fonte: KUNUMI (2020)

2.6.4 Recall

A revocação, ou recall do inglês, também chamado de sensibilidade ou taxa de verdadeiros positivos (TPR), concentra-se em erros falsos-negativos. Essa métrica é definida pela razão entre o número de exemplos corretamente classificados como positivos e o número de exemplos realmente positivos, conforme a fórmula da Figura 07.

Figura 7 — Fórmula da revocação, ou recall do inglês



Fonte: KUNUMI (2020)

3 METODOLOGIA

Trata-se de uma pesquisa aplicada com o objetivo de criar um classificador com a capacidade de determinar se uma notícia é verdadeira ou falsa, além de explicar por que a informação pode ser falsa através dos possíveis padrões que tais notícias costumam ter.

Para realizar este trabalho, foi utilizado métodos de pesquisa exploratório, com o objetivo de desenvolver um classificador de notícias. A pesquisa exploratória é uma ótima alternativa para compreender situações específicas e ser capaz de usar métodos qualitativos e quantitativos.

Quanto à abordagem, utilizou-se a metodologia triangular, que consiste na combinação de duas ou mais abordagens metodológicas, o primeiro método utilizado é um método qualitativo referente à classificação dos textos, e o segundo método é um método quantitativo, para análise dos resultados, que são apresentados na forma de tabelas e gráficos.

Quanto às técnicas de investigação e procedimentos, esta pesquisa se classifica em bibliográfica e experimental, visto que os recursos técnicos aplicados foram providos de bibliografias existentes, pois se fundamenta em técnicas de reconhecimento e classificação de textos bastantes estudadas quando se trata de *machine learning*, e de um banco de dados já existente com os dados necessários para o desenvolvimento.

É experimental porque utiliza algoritmos e bibliotecas escritas na linguagem de programação Python, e um classificador de notícias que pode analisar o impacto de cada palavra em um texto, usando conceitos de rede neural, padrões para detectar mensagens, incluindo mensagens falsas, e por fim, classificar o texto analisado como verdadeiro ou falso.

4 CLASSIFICADOR DE FAKE NEWS

4.1 Conjunto de Dados

Uma busca em bases de dados que tratam da questão das notícias falsas revelou que, embora haja muito trabalho de base em inglês, tópicos em português ainda não foram pesquisados. Felizmente, foi possível encontrar uma base que atende aos requisitos propostos neste trabalho.

O corpus Fake.br, produzido pelo projeto Opinando dentro do Núcleo Interinstitucional de Linguística Computacional, NILC[2], agrega 7.200 notícias em português, 3.600 falsas e 3.600 verdadeiras em diversos gêneros (MONTEIRO et al., 2018). O corpus também fornece meta-informações sobre cada item, como: autor, link, fonte, gênero e outros indicadores semânticos e sintáticos. O NILC é administrado por pesquisadores da área de ciência da computação do Instituto de Ciências Matemática e de Computação de São Carlos (ICMC) da USP, da Universidade Federal de São Carlos (UFSCar), da Universidade Estadual de São Paulo (UNESP) e da Universidade Estadual Maringá (UEM).

Infelizmente, não existe um formato uniforme para componentes de banco de dados. Consiste em mensagens de texto simples e arquivos contendo meta-informações para cada mensagem, todos os arquivos são arquivos de texto com a extensão ".txt". A Figura 8 mostra um arquivo de exemplo.

Figura 8 — Exemplo de notícia falsa presente na base de dados

```
Kátia Abreu diz que vai colocar sua expulsão em uma moldura, mas não para de reclamar.
A senadora Kátia Abreu (sem partido-TO) disse que sua expulsão do PMDB foi resultado de
uma ação da cúpula atual da legenda que, segundo ela, é oportunista. “Amanhã eu vou botar
numa moldura dourada a minha expulsão, porque das mãos de onde veio, é um atestado de boa
conduta para o meu currículo. Essas pessoas que me expulsaram não servem ao país. Eles se
servem do país em seus benefícios próprios”, disse Kátia Abreu. Ué, mas se a expulsão é
algo tão bom para seu currículo, por que tanta choradeira, Kátia? Sabemos o motivo.
Provavelmente Kátia não tem valor para o PT, partido que já deveria tê-la absorvido.
Ao que parece o PT gostava de Kátia somente se ela ficasse entrincheirada dentro do PMDB.
Ou seja, isso é se rebaixar demais. Resta a Kátia ficar chorando as pitangas por todos os
cantos. Em tempo: até o momento o PT não cadastrou Kátia Abreu em suas fileiras. Que situação
patética para a ex-ministra da Agricultura de Dilma.
```

Fonte: Elaborada pelo autor (2022).

[2] Disponível em: https://sites.google.com/view/nilc-usp Acesso em: 4 nov 2022.

4.2 Pré-processamento

Devido à natureza do conjunto de dados, contem-se muitos ruídos por se tratar de uma linguagem natural. Vários cálculos são realizados para tornar os dados adequados para processamento por algoritmos. A normalização de dados é uma etapa de pré-processamento necessária (AGRAWAL, 2021).

Nela é realizada a conversão para letras minúsculas, remoção dos *stopwords* e remoção de acentos e pontuação. Este processo visa encurtar o conjunto de dados, removendo informações desnecessárias.

Além de ajustar os dados para que o modelo seja renderizado corretamente, normalizamos os dados para remover informações desnecessárias, mas também pré-processamos todo o conjunto de dados para reduzir a quantidade de informações a serem processadas.

Para tarefas de processamento de linguagem natural, as técnicas de pré-processamento de texto melhoram o desempenho e a precisão dos dados.

4.2.1 Conversão para letras minúsculas

Ao processar o mesmo texto, as máquinas tratam letras minúsculas e maiúsculas de forma diferente. Por exemplo, palavras como "Bola" e "bola" são tratadas de forma diferente na mesma máquina. Para que as máquinas possam interpretar palavras facilmente o texto deve ser escrito igualmente em maiúsculas ou em minúsculas, mas recomenda-se usar minúsculas para evitar tais problemas (AGRAWAL, 2021).

A Figura 9 apresenta a função responsável por fazer a limpeza dos textos. O primeiro comando é responsável por converter as sentenças para letras minúsculas. O segundo comando é responsável por remover acentos e pontuações.

Figura 9 — Código responsável pela limpeza dos textos

```
def cleaning_sentence(sentence):
    sentence = sentence.lower()
    sentence = sentence.translate(str.maketrans('', '', string.punctuation))
    return sentence
```

Fonte: Elaborada pelo autor (2022).

4.2.2 Stopwords

Basicamente, *stopwords* (ou palavras de parada - tradução livre) são palavras que não adicionam valor semântico a uma frase. Alguns exemplos são os adjetivos, advérbios, preposições e conjunções. Como o conjunto de dados consiste em várias palavras de parada, é imperativo remover essas palavras irrelevantes antes de alimentar os dados como entrada para o classificador.

A Figura 10 mostra um método para remover palavras de parada. A biblioteca NLTK (abreviação de *Natural Language Toolkit)* é usada.

Figura 10 — Código responsável por remover *stopwords*

```
def removeStopWords(sentence):
    stopwords_nltk = nltk.corpus.stopwords.words('portuguese')
    sentence = ' '.join([word for word in sentence.split() if word not in stopwords_nltk])
    return sentence
```

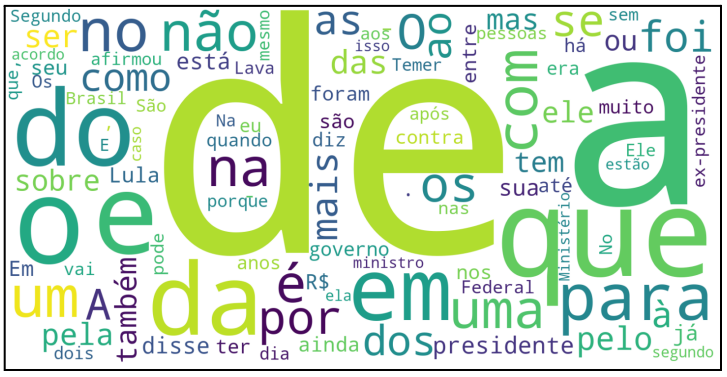
Fonte: Elaborada pelo autor (2022).

As *stopwords* usadas estão listadas no Apêndice. Após removê-las do conjunto de dados, a saída é reduzida a palavras únicas (AGRAWAL, 2021).

Nas Figuras 11 e 12, temos nuvens de palavras que representam as palavras mais repetidas na base de texto classificadas por tamanho, com as palavras mais repetidas sendo maiores que as menos repetidas.

Na Figura 11, o conjunto de dados ainda não foi pré-processado. Isso significa que ainda temos *stopwords*, acentos e sinais de pontuação. Palavras como 'a', 'da', 'de' e 'do' não têm valor semântico, mas são frequentes no conjunto de dados.

Figura 11 — Nuvem de palavras das notícias verdadeiras sem a remoção dos *stopwords*



Fonte: Elaborada pelo autor (2022).

Na Figura 12, o conjunto de dados já foi pré-processado, e pode-se ver que a predominância de palavras com alto poder semântico é maior do que na Figura 11.

Figura 12 — Nuvem de palavras das notícias verdadeiras com a remoção dos *stopwords*



Fonte: Elaborada pelo autor (2022).

4.3 Treinamento do modelo

A validação cruzada, ou *cross validation*, é uma técnica amplamente utilizada para avaliar o desempenho de modelos de aprendizado de máquina. validação cruzada consiste em dividir os conjuntos de dados em partes, usando um conjunto para treinamento e outro conjunto para testar e avaliar o desempenho do modelo.

Antes de iniciar o treinamento, os dados do conjunto de dados foram misturados e separados de maneira balanceada, com 80% usados para treinamento e 20% para teste e desempenho do modelo de aprendizado, que respectivamente são 5760 e 1440. Foram 20 épocas no total.

Ou seja, todos os dados de treinamento foram processados 20 vezes para determinar quais épocas deram os melhores resultados.

A Figura 13 mostra um fragmento de código que divide o conjunto de dados em 80% de treinamento e 20% de teste. O método *train_test_split* inclui *dataset* e *test_size* como parâmetros. Este último refere-se à porcentagem determinada para o teste (20% neste caso).

Figura 13 — Código responsável por separar o conjunto em teste e treinamento

```
x_train, x_test, y_train, y_test = train_test_split(dataset, test_size=0.33)
```

Fonte: Elaborada pelo autor (2022).

A Figura 14 mostra o trecho de código responsável por treinar o modelo. Foi utilizado um *perceptron* multicamada, uma rede neural multicamada com cinco camadas ocultas e como função de ativação a tangente hiperbólica (TANH).

Figura 14 — Código responsável por separar o *conjunto de dados* em teste e treinamento

```
neural_classifier = MLPClassifier(hidden_layer_sizes=5, activation='tanh')
```

Fonte: Elaborada pelo autor (2022).

4.4 Desempenho do classificador

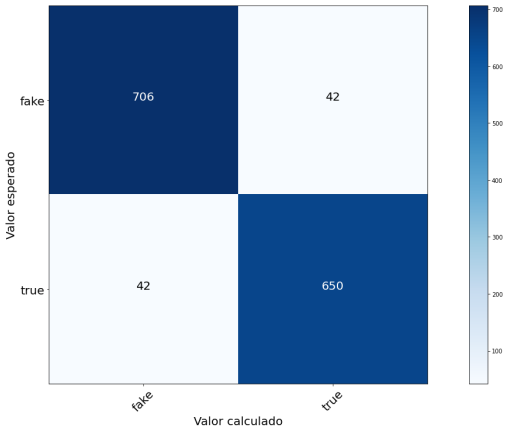
Agora que o classificador foi treinado, é hora de testá-lo e ver como ele se comporta. Para quantificar o desempenho do classificador, usamos três parâmetros: acurácia, precisão e *recall*.

A acurácia é quantas de suas predições foram corretas, considerando notícias falsas e verdadeiras. Precisão, refere-se aos números de notícias que foram classificadas como verdadeiras e que realmente eram verdadeiras. *Recall* indica quantas notícias verdadeiras foram detectadas considerando o máximo de notícias verdadeiras que existiam no conjunto de teste, que eram 720 notícias verdadeiras.

Considerando a média de 12 testes que foram realizados o classificador conseguiu obter uma acurácia de 94% e uma precisão em torno de 95% e um *recall* também de 95%.

A Figura 15 mostra a matriz de confusão representando o teste do modelo. Observe que das 720 notícias falsas, 706 (cerca de 98%) foram classificadas corretamente como falsas e as demais foram classificadas incorretamente como verdadeiras. Para mensagens verdadeiras, de 720 notícias verdadeiras, 650 (aproximadamente 90%) foram classificadas corretamente como verdadeiras, e as demais foram classificadas incorretamente como falsas.

Figura 15 — Matriz de confusão



Fonte: Elaborada pelo autor (2022).

5 CONCLUSÃO

Este trabalho foi desenvolvido com o objetivo de criar um classificador de notícias que as define como verdadeiras ou falsas utilizando técnicas de aprendizado de máquina. Durante o desenvolvimento, foi difícil encontrar conjuntos de dados públicos contendo grandes quantidades do tipo de dados que o projeto requer, e é difícil encontrar dados confiáveis, o que afeta os resultados mostrados no trabalho.

Espera-se que novos dados sejam aplicados no futuro para alcançar possíveis melhorias no desempenho dos classificadores desenvolvidos. Um estudo sobre os efeitos negativos das *fake news* se mostrou importante para motivar o desenvolvimento deste trabalho.

Foi alcançado uma precisão de 94%, mas esperava-se resultados ainda maiores. No entanto, acredita-se que tal resultado é decorrente do uso de um conjunto de dados com poucas notícias, o que limitou o potencial que o modelo poderia alcançar, baseando-se em outros classificadores que já foram desenvolvidos e tiveram resultados levemente superiores.

O estudo de determinados algoritmos de *machine learning* se mostrou de grande valia, pois as técnicas aplicadas para o desenvolvimento do modelo contribuíram bastante para que seu desempenho fosse rápido e eficiente mesmo com o uso de pouco processamento computacional. Mesmo com resultados bons e que podem ser utilizados para contribuir no combate a *fake news*, acredita-se que

ainda há bastante espaço para melhorias, visto que se trata de uma área da tecnologia ainda nova.

A utilização da matriz de confusão tem se mostrado suficiente para mostrar os dados que podem ser utilizados para determinar métricas importantes relacionadas ao desempenho do modelo, mostrando a quantidade de dados classificados corretamente e incorretamente.

No futuro, espera-se que este trabalho, e outros nesta área, contribuam para o desenvolvimento de aplicações que possam ser utilizadas por todos, e que sejam mais fiáveis no que diz respeito aos materiais que consomem. Isso nos permite tomar melhores decisões quando necessário, criando um ambiente social mais saudável.

- Testar outros algoritmos.
- Utilização de outras bases de dados para coleta de registros.
- Desenvolvimento de aplicativos utilizando o modelo criado para ajudar aos usuários na detecção de *fake news*.
- Analisar a influência nos resultados da eleição.

REFERÊNCIAS

AGRAWAL, R. **Must Known Techniques for text preprocessing in NLP.** 2021. Disponível em: https://www.analyticsvidhya.com/blog/2021/06/must-known-techniques-for-text-preprocessing-in-nlp/. Acesso em: 3 Nov. 2022.

ALLCOTT, H; GENTZKOW, M. **Social Media and Fake News in the 2016 Election**. 2017. Disponível em: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.31.2.211. Acesso em: 28 abr. 2022.

BAHETI, P. **Activation Functions in Neural Networks [12 Types & Use Cases].** 2022. Disponível em: https://www.v7labs.com/blog/neural-networks-activation-functions. Acesso em: 18 Nov. 2022.

BARANAUSKAS, J; MONARD, M. **Conceitos sobre Aprendizado de Máquina**. 2003. Disponível em: https://dcm.ffclrp.usp.br/~augusto/publications/2003-sistemas-inteligentes-cap4.pdf. Acesso em: 01 jun. 2022.

KUMAR, U. **Single Layered Neural Networks in R Programming.** 2020. Disponível em: https://www.geeksforgeeks.org/single-layered-neural-networks-in-r-programming. Acesso em: 29 mai. 2022.

MEDCALC. **TANH function.** 2022c. Disponível em: https://www.medcalc.org/manual/tanh-function.php. Acesso em: 18 Nov. 2022.

MITTAL, S. **Deep Learning Techniques for Text Classification**. 2019. Disponível em: https://medium.datadriveninvestor.com/deep-learning-techniques-for-text-classification-9392ca9492c7. Acesso em: 29 abr. 2022.

MONTEIRO, R. A.; SANTOS, R. L. S.; PARDO, T. A. S.; ALMEIDA, T. A.; RUIZ, E. E. S.; VALE, O. A. **Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection Results.** 2018. Disponível em: https://sites.icmc.usp.br/taspardo/PROPOR2018-MonteiroEtAl.pdf. Acesso em: 27 Out. 2022.

OLMEDO, M. T. C.; PAGELOW, M; MAS, J; ESCOBAR, F. **Geomatic Approaches for Modeling Land Change Scenarios.** 2017. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-60801-3_27. Acesso em: 15 Nov. 2022.

ORACLE. **O que é inteligência artificial – IA?**. c2022. Disponível em: https://www.oracle.com/br/artificial-intelligence/what-is-ai. Acesso em: 02 jun. 2022.

PRAKASH, S. **Neural network: Is your brain like a computer?.** 2017. Disponível em: https://medium.com/towards-data-science/neural-networks-is-your-brain-like-a-computer-d76fb65824bf. Acesso em: 29 abr. 2022.

PEREIRA, P; TOLEDO, L. F; MONNERAT, A. **Disseminação de "fake news" para atacar candidatos marca eleição.** 2018. Disponível em: https://exame.com/brasil/disseminacao-de-fake-news-para-atacar-candidatos-marca-eleicao/. Acesso em: 18 Nov. 2022.

RAMCHOUN, H; IDRISSI, M. A. J; GHANOU, Y. ETTAOUIL, M. **Multilayer Perceptron: Architecture Optimization and Training.** 2016. Disponível em: https://reunir.unir.net/bitstream/handle/123456789/11569/ijimai20164_1_5_pdf_30533.pdf?sequence=1. Acesso em: 18 Nov. 2022.

SHARIATMADARI, D. **Could language be the key to detecting fake news?.** 2019. Disponível em: https://www.theguardian.com/commentisfree/2019/sep/02/language-fake-news-linguistic-research. Acesso em: 18 Nov. 2022.

SHARMA, S; SHARMA, S; ATHAIYA, A. **ACTIVATION FUNCTIONS IN NEURAL NETWORKS.** 2020. Disponível em: https://www.ijeast.com/papers/310-316,Tesma412,IJEAST.pdf. Acesso em: 18 Nov. 2022.

SHEA, K. O; NASH, R. **An Introduction to Convolutional Neural Networks**. 2015. Disponível em: https://arxiv.org/pdf/1511.08458.pdf. Acesso em: 01 dez. 2022.

SILVERMAN, C. **This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook**. 2016. Disponível em: https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook#.uc9gevywE. Acesso em: 19 mar. 2022.

TIBCO. **O que é aprendizagem supervisionada?**. c2022. Disponível em: https://www.tibco.com/pt-br/reference-center/what-is-supervised-learning. Acesso em: 02 jun. 2022.

VICTOR, F. **Notícias falsas existem desde o século 6, afirma o historiador Robert Darnton**. 2017. Disponível em: https://www1.folha.uol.com.br/ilustrissima/2017/02/1859726-noticias-falsas-existem-desde-o-seculo-6-afirma-historiador-robert-darnton.shtml. Acesso em: 19 mar. 2022.

APÊNDICE — LISTA DE STOPWORDS

Figura 16 — Lista de stopwords

a, agora, ainda, alguém, algum, alguma, algumas, alguns, ampla, amplas, amplo, amplos, ante, antes, ao, aos, após, aquela, aquelas, aquele, aqueles, aquilo, as, até, atraves, cada, coisa, coisas, com, como, contra, contudo, da, daquele, daqueles, das, de, dela, delas, dele, deles, depois, dessa, dessas, desse, desses, desta, destas, deste, deste, destes, deve, devem, devendo, dever, deverá, deverão, deveria, deveriam, devia, deviam, disse, disso, disto, dito, diz, dizem, do, dos, e, é, ela, elas, ele, eles, em, enquanto, entre, era, essa, essas, esse, esses, esta, está, estamos, estão, estas, estava, estavam, estávamos, este, estes, estou, eu, fazendo, fazer, feita, feitas, feito, feitos, foi, for, foram, fosse, fossem, grande, grandes, há, isso, isto, já, la, lá, lhe, lhes, lo, mas, me, mesma, mesmas, mesmo, mesmos, meu, meus, minha, minhas, muita, muitas, muito, muitos, na, não, nas, nem, nenhum, nessa, nessas, nesta, nestas, ninguém, no, nos, nós, nossa, nossas, nosso, nossos, num, numa, nunca, o, os, ou, outra, outras, outro, outros, para, pela, pelas, pelo, pelos, pequena, pequenas, pequeno, pequenos, per, perante, pode, pude, podendo, poder, poderia, poderiam, podia, podiam, pois, por, porém, porque, posso, pouca, poucas, pouco, poucos, primeiro, primeiros, própria, próprias, próprio, próprios, quais, qual, quando, quanto, quantos, que, quem, são, se, seja, sejam, sem, sempre, sendo, será, serão, seu, seus, si, sido, só, sob, sobre, sua, suas, talvez, também, tampouco, te, tem, tendo, tenha, ter, teu, teus, ti, tido, tinha, tinham, toda, todas, todavia, todo, todos, tu, tua, tuas, tudo, última, últimas, último, últimos, um, uma, umas, uns, vendo, ver, vez, vindo, vir, vos, vós, ;, ., +, -, *, /, \, !, ?, :

Fonte: Elaborada pelo autor (2022).