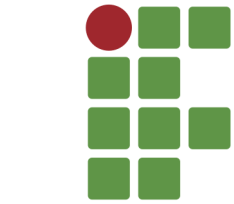


Instituto Federal de Brasília
Campus Brasília
Curso Superior de Tecnologia em Sistemas para Internet

DETECÇÃO DE *FAKE NEWS* RELACIONADAS À COVID-19 NO BRASIL

Por

FRANCISCA DE FÁTIMA DE ARAÚJO LUCENA
MARIANA GUEDES DA SILVA

Tecnólogo em Sistemas para Internet

BRASÍLIA
2023

Francisca de Fátima de Araújo Lucena
Mariana Guedes da Silva
Detecção de *fake news* relacionadas à COVID-19 no Brasil/ Francisca de Fátima de Araújo Lucena
Mariana Guedes da Silva . – BRASÍLIA, 2023-
44 p. : il. (algumas color.) ; 30 cm.

Orientador MsC. Diego Cesar Florencio de Queiroz

Tecnólogo em Sistemas para Internet – Instituto Federal de Brasília, 2023.

1. COVID-19. 2. *Fake news*. I. Prof. MsC Diego Cesar Florencio de Queiroz. II. Instituto Federal de Brasilia. III. Curso Superior de Tecnologia em Sistemas para internet. IV. Detecação de *fake news* relacionadas à COVID-19 no Brasil.

CDU 004

Francisca de Fátima de Araújo Lucena
Mariana Guedes da Silva

Detecção de *fake news* relacionadas à COVID-19 no Brasil

Trabalho de conclusão de curso de graduação apresentado a Coordenação do Curso Superior de Tecnologia em Sistemas para Internet do Instituto Federal de Brasília – Campus Brasília, como requisito parcial para a obtenção do título de Tecnólogo em Sistemas para Internet.

Aprovado em: _____ de _____ de _____.

BANCA EXAMINADORA

Prof. MsC. Diego Cesar Florencio de Queiroz
Computação/IFB

Prof. Dr. Caio Moura Daoud
Computação/IFB

Prof. Dr. Daniel Sundfeld
Engenharia de Software/UNB

BRASÍLIA
2023

Dedicamos este trabalho aos mais de 690 mil brasileiros que morreram por COVID-19 e suas famílias. A história não esquecerá e poderá ser contada como um momento de cegueira, descaso pelo elemento humano e ganância dos que foram eleitos para trabalhar pela sociedade.

Este trabalho surgiu da revolta, mas também da esperança de podermos contribuir para escrever uma nova história, um novo país.

A história fará justiça e nós nunca esqueceremos.

Agradecimentos

Agradecemos aos professores Daniel Sundfeld e Diego Queiroz por acolherem e incentivarem nossas ideias. Temos muito orgulho de tê-los em nossa trajetória acadêmica.

A todas e todos do corpo operacional, técnico e acadêmico do IFB, campus Brasília. O apoio e mobilização de vocês foi fundamental. Sabemos dos desafios e agradecemos todos os esforços para adaptarem os processos e ambiente acadêmico para o cenário da pandemia da COVID-19.

Às nossas famílias pelo amor e por nos permitirem sonhar e realizar mais essa etapa de formação. Aos colegas de trabalho por nos inspirarem no aprendizado dos conteúdos do curso e aos colegas de turma pela parceria e apoio mútuo. Juntos somos mais fortes.

Agradecemos aos que, em 2008, reestruturam o sistema educacional brasileiro e permitiram o retorno das atividades dos institutos federais, com formação técnica de excelência. Naqueles tempos nosso país sabia sonhar. Pela oportunidade de estudar numa instituição de ensino pública, gratuita e de qualidade.

Em tempos de obscurantismo e retrocessos, oferecemos esse trabalho àqueles que negam a importância do conhecimento e da ciência. Àqueles que ainda não compreendem a importância da educação para a transformação individual, coletiva e de país.

Seguiremos lutando juntos, em uníssono com outros bravos educadores e cientistas, na reconstrução do nosso país, quando esse momento desafiador passar.

Resumo

LUCENA, Francisca de Fátima de Araújo . SILVA, Mariana Guedes. Detecção de *fake news* relacionadas à COVID-19 no Brasil. 2023. Trabalho de Conclusão de curso Tecnólogo em Sistemas para Internet. Instituto Federal de Brasília – Campus Brasília. Brasília/DF, 2023.

A pandemia de coronavírus foi declara pela Organização Mundial da Saúde (OMS) em março de 2022. A veiculação de grande volume de *fake news* nas mídias sociais nos últimos anos, trouxe visibilidade global para a importância da observação da veracidade das informações trafegadas. O uso de ferramentas de inteligência artificial possibilita a construção de algoritmos capazes de automatizar diversos processos. O processamento de linguagem natural (PLN) é uma subárea de Ciência da Computação, inteligência artificial (IA) e linguística, que estuda os problemas da geração e compreensão da linguagem humana e que são tipicamente baseados em algoritmos de aprendizado de máquina. A aplicação de técnicas de PLN permite o processamento automático para analisar e decidir sobre o conteúdo de textos, a partir de regras de classificação estabelecidas. Este trabalho teve como objetivo desenvolver e aplicar uma ferramenta de IA capaz de analisar e classificar o conteúdo de notícias sobre COVID-19 e identificar sua veracidade. Para isso, foi construído um *corpus* através de uma coleta semiautomática de notícias, por meio de raspagem de dados implementada em linguagem de programação Python. Em seguida foi realizada uma análise manual para classificação das notícias, a fim de garantir que todas as notícias identificadas como falsas recebessem a marcação FAKE, bem como aquelas com conteúdo verdadeiro fosse marcadas como REAL, para treinar o algoritmo. A etapa seguinte foi o pré-processamento dos dados e extração dos atributos, por meio da biblioteca Natural Language Toolkit (NLTK) da linguagem Python 3.6. A partir da construção do *corpus* e do processamento dos atributos, foram construídos os conjuntos de aplicação do algoritmo, com a proporção 80/20, ou seja, 20% da amostra retida para teste e 80% dedicada ao treinamento. O processamento e classificação foram realizados utilizando a associação de TF-IDF com 3 métodos de aprendizado de maquina: *Naive Bayes* (NB), *Random Forest* (RF) e *Support Vector Machine* (SMV). Os resultados mostraram que o método *Naive Bayes* se destacou, entre os demais métodos, obtendo uma acurácia de 70%, enquanto RF e SVM obtiveram 35% e 45%, respectivamente.

Palavras-chave: Notícias falsas. COVID-19. Inteligência artificial. Aprendizagem de máquina. Algoritmo. Acurácia.

Abstract

LUCENA, Francisca de Fátima de Araújo. SILVA, Mariana Guedes. Detection of fake news related to Covid-19 in Brazil. 2023. Trabalho de Conclusão de Curso – Tecnólogo em Sistemas para Internet. Instituto Federal de Brasília – Campus Brasília. Brasília/DF, 2023.

The coronavirus pandemic was declared by the World Health Organization (WHO) in March 2022. The dissemination of a large volume of fake news on social media in recent years has brought global visibility to the importance of observing the veracity of the information transmitted. The use of artificial intelligence tools enables the construction of algorithms capable of automating various processes. Natural language processing (NLP) is a subarea of Computer Science, artificial intelligence (AI) and linguistics, which studies the problems of generating and understanding human language and which are typically based on machine learning algorithms. The application of NLP techniques allows automatic processing to analyze and decide on the content of texts, based on established classification rules. This work aimed to develop and apply an AI tool capable of analyzing and classifying news content about COVID-19 and identifying its veracity. For this, a corpus was built through a semi-automatic collection of news, through data scraping implemented in the Python programming language. Then, a manual analysis was performed to classify the news, in order to ensure that all news identified as false received the FALSE mark, as well as those with true content were marked as REAL, to train the algorithm. The next step was data preprocessing and attribute extraction, using the Natural Language Toolkit (NLTK) library of the Python 3.6 language. From the construction of the corpus and the processing of the attributes, the application sets were built with an 80/20 ratio, that is, 20% of the sample retained for testing and 80% dedicated to training. Processing and classification were performed using the association of TF-IDF with 3 machine learning methods: Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SMV). The results showed that the Naive Bayes method stood out among the other methods, obtaining an accuracy of 70% while RF and SVM obtained 35% and 45%, respectively.

Keywords: Fake news. COVID-19. Artificial inteligence. Marchine learning. Algorithm. Accuracy.

Lista de Figuras

Lista de Tabelas

Sumário

1

Introdução

A autenticidade da informação é um problema emergente que afeta a sociedade e os indivíduos, e que pode ter um impacto negativo na capacidade de tomada de decisão das pessoas (HABIB, 2019). O fenômeno de propagação de notícias falsas, aqui denominado de *fake news*, movimenta também a comunidade científica no sentido de buscar estratégias e ferramentas efetivas de identificação da veracidade do conteúdo das notícias, contribuindo com a sociedade no processo de busca por informações confiáveis. O volume crescente de notícias disseminadas nas redes sociais e aplicativos tornou impraticável a tarefa de verificação manual dos fatos veiculados (YANG et al., 2018).

O processo de utilização dos algoritmos de inteligência artificial (IA) representa um importante aliado para verificação da veracidade de informações (LIVINGSTON, 2020). No contexto de inteligência artificial, o processamento de linguagem natural (PLN) permite o processamento automático de linguagens faladas e escritas (PIERRI; PICCARDI; CERI, 2020). Adicionalmente, técnicas de aprendizagem de máquina (AM) são importantes, por exemplo, na produção e implementação de um classificador automatizado para avaliação de informações de interesse (SGANTZOS; GRIGG, 2019).

O grande volume de notícias disseminadas nas redes sociais e aplicativos de mensagens, ou seja, o aumento na dimensionalidade das informações trafegadas, lançam ainda mais destaque para a importância do uso dos algoritmos de IA, a partir de técnicas de PLN e implementação em AM (HABIB, 2019). Aplicações dessas técnicas são encontradas, por exemplo, em temas políticos (NELSON; TANEJA, 2018), acadêmicos (PUERTA-DIAZ et al., 2021) e de saúde pública (KOENIGKAM-SANTOS et al., 2019).

Em 2019, o mundo vivenciou o surgimento da pandemia do COVID-19, conforme denominação da Organização Mundial da Saúde (OMS), causada pelo coronavírus (Sars-CoV-2). No final de 2022 foram mais de 700 milhões de casos confirmados e mais de 6,5 milhões de óbitos em todo o mundo. No mesmo período, o Brasil registrou mais de 35 milhões de casos confirmados e 690 mil óbitos (WHO, 2021).

No contexto da pandemia de COVID-19, o volume de informações veiculadas sobre o assunto abordava questões desde transmissão do vírus, quantidade de casos e óbitos registrados, efetividade da medidas de contenção do vírus (uso de máscaras, distanciamento e isolamento

social), uso de medicamentos e vacinas. Com a massiva produção de informações sobre a pandemia, a capacidade de dicernimento sobre notícias confiáveis e falsas representou uma desafio a mais no enfrentamento da pandemia.

Considerando a dimensionalidade do volume de informações sobre a COVID-19 e a importância de acessar informações confiáveis para o enfrentamento de questões de saúde pública trazidas pela pandemia, este trabalho tem como proposta a utilização de ferramentas computacionais de inteligência artificial para o desenvolvimento de uma aplicação para detecção de informações falsas (*fake news*) sobre COVID-19.

1.1 Tema

Utilização de ferramentas de inteligência artificial no desenvolvimento de uma aplicação para detecção de *fake news* relacionadas à pandemia da COVID-19.

1.2 Problema

Desenvolvimento e aplicação de uma ferramenta de inteligência artificial capaz de analisar e classificar o conteúdo de notícias e identificar sua veracidade.

1.2.1 Objetivo geral

Desenvolver e aplicar uma ferramenta de inteligência artificial para detecção de *fake news* sobre COVID-19, utilizando processamento de linguagem natural (PLN), por aprendizagem de máquina (AM).

1.2.2 Objetivos específicos

1. Construir um *corpus* de notícias sobre COVID-19;
2. Criar um classificador de *fake news*;
3. Treinar o classificador com o *corpus* de notícias sobre COVID-19;
4. Avaliar o método com melhor desempenho na classificação de *fake news* sobre COVID-19;
5. Disponibilizar a metodologia desenvolvida em *Git hub*, para aprimoramento em novas pesquisas.

1.3 Estrutura do TCC

O presente trabalho é organizado da seguinte forma: O capitulo 1 apresenta introdução, tema e objetivo deste trabalho e o capitulo 2 é composto por conceitos gerais, com revisão de literatura. O capítulo 3 mostra a metodologia de implementação das ferramentas propostas. O

capítulo 4 apresenta os resultados da aplicação e principais achados do trabalho e, por último, o capítulo 5 descreve a conclusão do estudo, bem como sugestões para sua continuidade.

2

Conceitos gerais e revisão da literatura

2.1 Conceitos gerais

O presente estudo foi estruturado a partir elementos de diversas áreas do conhecimento em sua estruturação. A pandemia de coronavírus (COVID-19) aparece com o elemento de interesse de análise, fornecendo o contexto do uso de ferramentas de inteligência artificial e seus diversos métodos, no processo de análise e classficação sobre a veracidade das notícias veiculas na grande mídia e redes sociais. Nos itens subsequentes, estão apresentados os diversos conceitos necessários ao entendimento do tema de estudo e aplicação das ferramentas computacionais propostas.

2.1.1 COVID - 19

A pandemia de coronavirus foi declara pela Organização Mundial da Saúde (OMS) em março de 2020 (DUCHARME, 2020), após o registro dos primeiros casos identificados em dezembro de 2019, na província de Wuhan, na China (WU et al., 2020).

A COVID-19 é uma doença cujo diagnóstico é realizado por meio de exame molecular específico, a partir da coleta de amostras respiratórias. Os exames são realizados em pacientes com suspeita de infecção e que apresentam sinais e sintomas como febre, tosse seca e falta de ar (ZHANG et al., 2020).

Adicionalmente, o mundo enfrentou um desafio em relação às informações sobre a pandemia, pela veiculação de um grande volume de *fake news* sobre o diagnóstico, as manifestações clínicas, o tratamento e, mais recentemente, a vacinação contra a COVID-19.

Dessa maneira, uma importante contribuição científica é a proposição de estatégias de busca, análise e identificação de *fake news* sobre COVID-19. Essas estratégias representam uma importante ferramenta de enfrentamento dessa pandemia, em meio a tantas incertezas.

2.1.2 Notícias falsas (*fake news*)

A autenticidade da informação é um problema emergente que afeta a sociedade e os indivíduos e que pode ter um impacto negativo nas capacidades de tomada de decisão dos indivíduos (HABIB, 2019).

De acordo com TANDOC JR; LIM; LING (2018), as *fake news* se caracterizam por serem criadas e compartilhadas na internet e possuírem conteúdo falso, em função da falta de embasamento técnico ou científico. Adicionalmente, são notícias criadas com o objetivo de enganar e manipular pessoas e situações.

A veiculação de grande volume de *fake news* nas mídias sociais nos últimos anos trouxe visibilidade global para a importância da observação da veracidade das informações. A ampla veiculação de *fake news* mostrou sua importância em grandes eventos globais, com eleições e questões sanitárias (pandemias, tratamento, vacinação, entre outros).

Uma avaliação sobre conteúdo e veracidade de notícias é realizada a partir da observação de suas propriedades e atributos. A variedade de notícias e suas diferentes estruturas podem tornar difícil a identificação das *fake news*. (WARDLE, 2017) as dividiu em sete categorias diferentes, sendo elas: sátira ou paródia, falsa conexão, conteúdo enganoso, falso contexto, conteúdo impostor, conteúdo manipulado e conteúdo fabricado.

No escopo das informações em saúde, as *fake news* apresentam impacto importante na identificação, diagnóstico e tratamento dos pacientes e na condiçao de saúde e doença da população. Por isso a importância do desenvolvimento de ferramentas de classificação de notícias e classificação de *fake news* e sua aplicação no escopo de saúde pública.

A escolha de grafia da expressão *fake news* em língua inglesa, no presente trabalho, se deve ao fato das principais publicações utilizarem essa expressão, adotada em todo o documento e nas produções posteriores.

2.1.3 Inteligência artificial (IA)

A inteligência artificial (IA), resumidamente, é a possibilidade de uma máquina, através da implementação de algoritmos, automatizar e repetir diversos processos de maneira otimizada, com melhor desempenho e com maior velocidade do que se fosse realizado por humanos (SILVA; MAIRINK, 2019). Por exemplo, pode ser aplicada na área do direito, onde as máquinas poderiam realizar todo o trabalho repetitivo e maçante de forma ininterrupta, economizando tempo, recursos humanos e otimizando processos.

SILVA; MAIRINK (2019) mostram que por se tratar de um meio de facilitação do cotidiano, a inteligência artificial traz diversas vantagens quanto ao seu uso. Sendo uma inovação tecnológica, que são modernizações apresentadas para solucionar algum tipo de problema, e para o caso da inteligência artificial o maior deles: tempo.

Para que um sistema seja capaz de usar as combinações de habilidades supracitadas para demonstrar algum tipo de comportamento, é necessário que esse sistema aprenda como fazê-lo. Para tal, o estudo do aprendizado de máquina deve ser explorado. Para NASCIMENTO JR; YONEYAMA (2000) “o aprendizado é a aquisição de conceitos e de conhecimentos estruturados”.

2.1.4 Processamento de linguagem natural (PLN)

O processamento de linguagem natural (PLN) é 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 de os computadores analisarem, compreenderem e derivarem o significado da linguagem humana de uma maneira inteligente e útil. Utilizando a PLN, os desenvolvedores podem organizar e estruturar o conhecimento para executar tarefas como resumo automático, tradução, análise de sentimentos e outros. (PEREIRA, 2011).

Os algoritmos de PLN são tipicamente baseados em algoritmos de aprendizado de máquina. Em vez de codificar manualmente grandes conjuntos de regras, a PLN pode confiar na aprendizagem automática para aprender automaticamente essas regras, analisando um conjunto de exemplos (ou seja, um *corpus* grande, como um livro, até uma coleção de sentenças) e fazendo uma inferência estática. Em geral, quanto mais dados forem analisados, mais preciso será o modelo (PEREIRA, 2011).

2.1.5 Aprendizado de máquina (AM)

O aprendizado de máquina (AM) apresenta-se como um importante aliado na classificação de textos, tratando-se de um processo de criação de um classificador por aprendizagem para máquinas, implementado de forma indutiva, a partir de um conjunto específico de características.

O campo do aprendizado de máquina está preocupado com a questão de como construir programas de computador que melhoram automaticamente com a experiência (MITCHELL, 1997).

Essa ferramenta mostra grande precisão e adaptabilidade às mudanças de características para julgamento no processamento, aplicável às mais diversas áreas do conhecimento. De um modo geral, a implementação ocorre em duas fases: fase de treinamento, na qual o sistema recebe informações necessárias para o processamento, ou seja, ele aprende a partir dos exemplos apresentados para que, na segunda fase, a de aplicação, sejam utilizados os aprendizados no processamento e classificação dos documentos (SEBASTIANI, 2002).

2.1.6 Técnicas para pré-processamento

2.1.6.1 N-*grams*

O algoritmo subdivide uma *string* em *subtrings* de tamanho *n* (a aplicação mais conhecida é o bi-*gram*, que divide em *subtrings* de tamanho 2). O algoritmo pode ser aplicado para uma filtragem inicial na captura de palavras semelhantes numa busca textual ou para identificar sequência de texto com palavras comuns (GANDRABUR; FOSTER, 2003).

Ao comparar duas *strings* S e T, são gerados todos os n-*grams* possíveis para cada uma delas e comparados para definir quantos são os comuns. O cálculo final da distância n-*grams* deve tomar o número de possíveis n-*grams* formados pela maior *string* e subtrair o número de

n-*grams* comuns às duas *strings*.

2.1.6.2 *Term frequency* (TF)

Suponha que um conjunto de elementos (e) é formado por um conjunto finito de n palavras (p), sendo que, em cada elemento do conjunto, podem aparecer quaisquer palavras. A medida TF representa o número de vezes que a palavra p aparece em um elemento e.

2.1.6.3 *Inverse document frequency* (IDF)

O peso das palavras (p) no conjunto de elementos (e), determinado pelo inverso de sua frequência entre os elementos (IDF).

O cálculo do IDF é importante para que palavras (p) de maior frequência no conjunto de elementos (e) tenham sua participação ponderada. Seja E o conjunto de todos os elementos de um arquivo R e Ep o conjunto dos elementos de E que contém determinada palavra p.

O cálculo do IDF de p é dado por:

$$W_p = 1 + log\frac{|E|}{|E_p|} \tag{2.1}$$

Nomes grandes tendem a ser favorecidos por conterem maior número de palavras diferentes. Com isso, é necessário realizar uma normalização em função do tamanho do elemento, que é determinada pela fórmula:

$$W_e = \sqrt{\sum {W_{e,p}}^2} \tag{2.2}$$

Onde W(e,p) corresponde ao peso das palavras em relação ao elemento e. O cálculo é realizado por meio da regra TF × IDF:

$$W_{e,p} = (1 + log(f_{e,p})) * W_p \tag{2.3}$$

Sendo f(e,p) o número de ocorrências da palavra p no elemento e.

2.1.6.4 *Bag of words*

Considerada uma das medidas de similaridade mais simples (CHAKRABARTI, 2003). Suponha um conjunto de elementos composto por um grupo finito de n palavras. Cada elemento será representado por um vetor n-dimensional, em que cada eixo desse vetor representa uma palavra p e recebe o valor 1, caso p seja observado em e, e 0, caso contrário.

O cálculo da similaridade entre dois elementos e_1 e e_2 será realizado por meio do produto interno entre os vetores e_1 e e_2, ou seja, a multiplicação direta dos elementos das posições $i = 1,2,...,n$ do vetor e_1 pelos elementos das posições $j = 1,2,...,n$ do vetor e_2, conforme proposto a seguir:

$$bow(e_1, e_2) = \overrightarrow{e_1}\overrightarrow{e_2} \tag{2.4}$$

Nesse caso, também é necessário definir experimentalmente um número mínimo de ocorrências, μ, que deve ser satisfeito para que dois elementos sejam considerados similares.

2.1.7 Algoritmos para categorização automática de textos

A categorização automática de textos é um processo que permite a classificar um texto em categorias maiores ou grandes temas. As técnicas de categorização foram impulsionadas pela expansão da dimensionalidade, volume de dados, a partir do surgimento da internet, bem como, pelo desenvolvimento de computadores com maior capacidade de processamento e facilidade de publicação eletrônica de documentos (SEBASTIANI, 2002).

A principal via de implementação de estratégias de categorização automática de texto é por meio da aprendizagem de máquina, que consiste de um processo de criação de um classificador por aprendizagem, implementado de forma indutiva, a partir de um conjunto específico de características.

No escopo das diversas técnicas de aprendizagem de máquina, destacam-se em eficácia, a aprendizagem baseada em instâncias (KNN) e aprendizagem bayesiana do tipo *Naive* (WANG, 2006; MCCALLUM; NIGAM, 1998).

2.1.7.1 *Naive bayes* (NB)

O classificador bayesiano *Naive bayes* (NB) supõe que os elementos que se deseja comparar são independentes, indiferente de sua ordem ou posição de entrada. Apesar de facilmente verificar que essa suposição nem sempre ocorre na prática, esse método realiza a classificação de modo muito satisfatório (DOMINGOS; PAZZANI, 1997), sem prejuízo para a eficiência do classificador.

Existem dois tipos de modelos estatísticos para os classificadores NB:

a) O multinomial, representado por um vetor de frequência dos nomes (JOACHIMS, 1998);

b) O binário, que considera apenas a ocorrência dos nomes (ROBERTSON; JONES, 1976).

MCCALLUM; NIGAM (1998) realizaram simulações e compararam os modelos multinomial e binário, concluindo que o multinomial apresenta melhores resultados. O cálculo da probabilidade, pelo modelo multinomial, de um nome pertencer a uma sequência de caracteres será:

$$P(C|D) = \frac{P(C)}{P(D)} P(D|C) \tag{2.5}$$

Sendo:

P(C) = probabilidade de a sequência de caracteres ocorrer.

P(D) = probabilidade de o nome ocorrer.

P(D|C) = probabilidade de a sequência de caracteres ocorrer, dado que o nome ocorreu.

P(C|D) = probabilidade de, dado que a sequência de caracteres ocorreu, de o nome estar contido nessa sequência.

2.1.7.2 *Random forest* (RF)

Segundo PICCHI NETTO (2013) árvores de decisão (AD) são ferramentas de suporte à decisão que mostram as possíveis consequências, incluindo as chances de sucesso em cada evento. Uma AD pode mostrar o caminho a ser percorrido para que um determinado evento ocorra.

Uma extensão da aplicação foi desenvolvida e registrada sob a denominação de *random forest*, que se refere ao algoritmo que cria, de forma aleatória, várias árvores de decisão (*decision trees*), e combina seus resultados para chegar no resultado final. De um modo geral, é uma ferramenta de suporte à decisão que mostra as possíveis consequências, incluindo as chances de sucesso em cada evento.

Em termos de aplicação, após a construção da árvore de decisão, é possível que o classificador construído seja muito específico para a base de treinamento. Dessa maneira, a árvore de decisão pode ser utilizada para classificar novos exemplos, iniciando pela raiz da árvore e caminhando, através de cada um dos nós de decisão, até que uma folha seja encontrada. Quando uma folha é encontrada, a classe do novo exemplo é dada pela classe daquela folha.

2.1.7.3 *Support vector machine* (SVM)

O SVM, também conhecido como máquina de suporte vetorial, teve seu processo de aprimoramento registrado no estudo proposto por BOSER; GUYON; VAPNIK (1992). É um algoritmo de aprendizado supervisionado, cujo objetivo é classificar determinado conjunto de dados, mapeados para um espaço de características multidimensional, usando uma função Kernel para classificação de problemas. Nessa função, o limite de decisão no espaço de entrada é representado por um hiperplano em dimensão superior à do espaço. No caso do Kernel linear, um conjunto de dados de treinamento de n pontos é construído da seguinte forma:

$$(x1, y1), ..., (xn, yn) \tag{2.6}$$

Onde os *yi* são 1 ou -1, cada um indicando a classe à qual o ponto *xi* pertence. Cada *xi* é um p-vetor real tridimensional. O objetivo é encontrar o hiperplano de margem máxima que divide o grupo de pontos *xi* para qual *yi* = 1, do grupo de pontos para os quais *yi* = -1, definido de modo que a distância entre o hiperplano e o ponto mais próximo *yi* de qualquer um dos grupos é maximizado (MONTEIRO; NOGUEIRA, 2019).

2.2 Revisão de literatura

As subseções a seguir contam com uma breve revisão da literatura, na qual estão destacados três trabalhos que foram importantes na motivação e escolha dos métodos utilizados na elaboração do estudo proposto.

2.2.1 Detecção automática de notícias falsas para o português, 2018

MONTEIRO et al. (2018) apresentou os resultados da investigação de métodos de detecção automática de informações falsas na *web*, escritas em português. As atividades proostas foram:

(i) Compilar um *corpus* de notícias falsas e verdadeiras para subsidiar o treinamento e teste dos métodos de aprendizagem de máquina (AM);

(ii) Adaptar e criar os recursos e ferramentas linguístico-computacionais necessários para os métodos estudados;

(iii) Implementar os métodos utilizados em Pérez-Rosa e Mihalcea (2015) e outros mais comumente utlizados na literatura, a fim de demonstrar e comparar a eficácia de cada método;

(iv) Avaliar os métodos implementados sobre o *corpus* compilado, destacando suas vantagens e limitações.

MONTEIRO et al. (2018) utilizou ferramentas e recursos necessários para implementar os métodos estudados que já existiam, como o etiquetador morfossintático NLPNet e o léxico LIWC, o qual lista as palavras do português e suas possíveis associações com classes semânticas e com sentimentos, além de outras ferramentas básicas de pré-processamento de texto, como o *stemmer*, do projeto Snowball, e lista de *stopwords*, entre outras.

Os resultados obtidos nos experimentos realizados durante o projeto foram positivos, com destaque para a acurácia de 89% obtida com o uso de *bag of words* e SVM. Ainda que sejam resultados impressionantes, o foco da abordagem foram as notícias com conteúdo falso, sendo a detecção de meias verdades, boatos e textos satíricos desafios a serem resolvidos em trabalhos futuros. MONTEIRO et al. (2018) destaca que é interessante que os modelos treinados em seu projeto sejam avaliados em um *corpus* de validação, independente do *corpus* apresentado, com o intuito de estudar a generalização dos modelos desenvolvidos pela abordagem proposta.

2.2.2 Inteligência artificial aplicada à detecção de *fake news*, 2019

OLIVEIRA (2019) mostra que informações são veiculadas ao público através da internet, televisão, jornais, revistas e outros meios de comunicação a todo instante, mas nem sempre se apresentam com conteúdo confiável. O acesso a esse tipo de conteúdo e a propagação de notícias que não são verdadeiras torna-se um fato alarmante por suas intenções. As *fake news* são conhecidas por essas características, o uso mal intencionado de informações pode, entre outros, comprometer a democracia e manipular a opinião do leitor.

Preocupado com a gravidade que o tema apresenta à sociedade, OLIVEIRA (2019) propôs em seu trabalho uma aplicação que permitisse a detecção de *fake news*, em meio a notícias verdadeiras, utilizando algoritmos de inteligência artificial, a partir da aplicação de técnicas como árvore de decisão, *naive bayes*, *support vector machine (SVM)* e aprendizagem profunda. Para analisar esses padrões, o autor utilizou duas abordagens, A e B.

A abordagem A utilizou algoritmos clássicos de aprendizagem de máquina: árvore de decisão (AD), *naive bayes (NB)* e *support vector machine (SVM)*. Na abordagem B foram aplicadas técnicas de redes neurais, a partir da escolha de três redes neurais distintas, as quais mostraram melhores resultados e desempenho no conjunto de testes. Os resultados para abordagem A mostraram que o algoritmo SVM teve maior acurácia, atingindo 90,13%, e da abordagem B a rede 1 foi a melhor, atingindo a acurácia de 92,36%.

2.2.3 Classificação de *fake news* utilizando algoritmos de aprendizado de máquina e aprendizado profundo, 2019

Com a popularização da internet e a facilidade de acesso às informações possibilitada pelas mídias sociais, redes sociais e aplicativos de mensagens, a rápida disseminação de notícias falsas vem se tornando uma preocupação. Essa disseminação vem ocasionando problemas sérios e a utilização de soluções de inteligência artificial vem ganhando espaço e obtendo ótimos resultados na classificação de textos e identificação de notícias falsas.

O objetivo do trabalho foi propor modelos preditivos utilizando técnicas de *word embedding*, em conjunto com os algoritmos de aprendizado de máquina e aprendizado profundo, para classificação de notícias falsas.

Foram escolhidos cinco algoritmos de aprendizado de máquina: (i) *logistic regression*, (ii) *random forest*, (iii) *naive bayes*, (iv) SVM e (v) xgboost. Os algoritmos de aprendizado profundo escolhidos foram: (i) redes neurais *feedforward* e (ii) redes neurais convolucionais.

Os algoritmos conseguiram resultados satisfatórios classificando as notícias. Dentre os algoritmos testados, os melhores desempenhos ficaram com redes neurais convolucionais, em associação com a técnica de *word embedding* (*Global Vectors for Word Representation*) - GloVe - com uma acurácia de 97,50%. Esse modelo utilizou um vetor de *embeddings* de 300 dimensão.

O segundo melhor resultado foi o algoritmo redes neurais *feedforward*, juntamente com a técnica de *word embedding* - word2vec - com uma acurácia de 91,37%. Com essa combinação de técnicas, os modelos desenvolvidos no referido trabalho conseguiram lidar muito bem na classificação das notícias falsas.

3

Metodologia

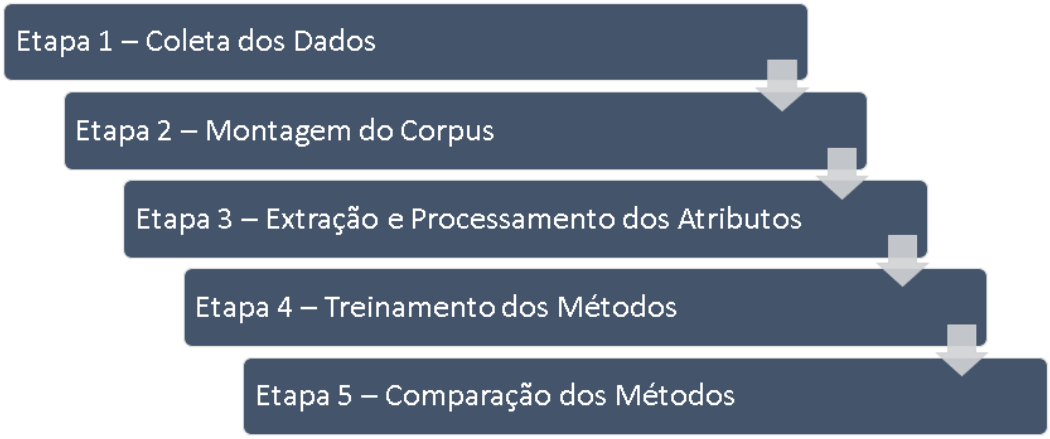
3.1 Organização do trabalho

Para uma implementação organizada, este trabalho foi divido em 5 etapas:

- Etapa 1: Coleta de dados (notícias).
- Etapa 2: Montagem e balanceamento do *corpus*.
- Etapa 3: Extração e processamento dos atributos.
- Etapa 4: Treinamento dos métodos.
- Etapa 5: Comparação dos métodos.

Todas as etapas foram executadas de maneira sequencial, conforme ilustra a Figura 3.1.

Figura 3.1 Etapas de projeto



3.2 Coleta e criação do *corpus*

Inicialmente, com o intuito de capturar instâncias representativas do problema abordado, definiu-se que a coleta seria de mensagens e notícias com conteúdo enganoso. A coleta desse

grupo de notícias foi realizada em *sites* de checagem de fatos, *sites* de notícias e páginas em redes sociais.

Para atender ao escopo do projeto, decidiu-se pela exclusão de qualquer material que com recursos de mídia, tais como áudios, vídeos ou imagens. Essas exclusões permitiram a manutemção da homogeneidade do tipo de notícia coletada, de forma a facilitar a classificação em conteúdo enganoso (FAKE) ou não-enganoso (REAL).

Em relação à estratégia para a coleta das notícias, optou-se por utilizar uma coleta semi-automática de raspagem de dados, ou *web scraping*, realizada utilizando a linguagem de programação Python. As fontes de notícias foram *sites* oficiais do governo federal brasileiro, criado para o combate à desinformação sobre a COVID-19, disponível no endereço eletrônico (https://antigo.saude.gov.br/fakenews/) e *sites* conhecidos como propagadores de *fake news*.

Figura 3.2 Balanceamento do *corpus*

text	label
EUA encontraram a cura para o novo coronavírus após teste no Hospital de Stanford. EUA desenvolveram a cura para a Covid-19 e conseguiram tratar, de forma definitiva, 40 pacientes em estado grave durante um teste do Hospital de Stanford. "Senhores, acabou de ser divulgado em Stanford que 100% dos 40 casos testados com hidroxicloroquina e azitromicina associados zeraram o PCR para Covid-19, ou seja, eles curaram a doença. Então, esse protocolo já tá sendo usado no Hospital do Coração e na Prevent Sênior a partir de hoje. Isso, provavelmente, vai mudar completamente o curso da doença".	FAKE

text	label
Ao contrário do que aponta o texto, o CEO da Novartis não afirmou que a doroquina ou hidroxidoroquina é a cura para a Covid-19. No dia 27 de março de 2020, em uma entrevista publicada pela CNBC, o CEO da Novartis Vas Narasimhan disse ser "muito cedo" para afirmar se a hidroxicloroquina poderia funcionar contra a Covid-19. Depois de meses e diversas pesquisas, ficou comprovado que a hidroxidoroquina, de fato, não funciona para esse fim	REAL

Para garantir o balanceamento do *corpus* foi realizada, após a coleta inicial, uma análise manual das notícias, a fim de garantir que todas as notícias marcadas como FAKE continham conteúdo falso e todas aquelas marcadas como REAL eram formadas por conteúdos verdadeiros. Assim, o *corpus* do trabalho foi composto por 100 notícias, sendo 50 notícias com marcação FAKE e 50 notícias com marcação REAL.

3.3 Extração de atributos

Na elaboração de um classificador automático utilizando aprendizado de máquina, se faz necessário que atributos textuais das notícias sejam extraídos e representados em um formato numérico. Neste caso foi utilizado os mesmo que MONTEIRO et al. (2018). São eles:

- Unigramas / *bag of words*: representação em *bag of words* da ocorrência (ou não) das

palavras presentes nos textos, utilizando valores booleanos. A extração ocorre após um tratamento das palavras, realizando (i) conversão das palavras do texto para caixa baixa, (ii) remoção de numerais, (iii) pontuação, (iv) *stopwords* (palavras de classe fechada, como preposições, pronomes e artigos), e (v) radicalização das palavras.

- POS Tags (etiquetas morfossintáticas, ou classes gramaticais): número normalizado de ocorrências de cada etiqueta morfossintática no texto, indicado pelo etiquetador/ *tagger* NLPNet.

Para que os algoritimos de aprendizado de máquina (AM) funcionem adequadamente, as normalizações apresentadas foram realizadas dividindo-se os números calculados pela contagem total de palavras presentes no texto. Dessa maneira, foram obtidos valores compreendidos no intervalo entre 0 e 1 e passíveis de comparação livre.

Todo o processamento dos atributos foi realizado utilizando a linguagem de programação Python 3.6, por meio da biblioteca *Natural Language Toolkit* (NLTK), que possui uma grande variedade de técnicas de processamento de linguagem natural.

3.4 Aprendizado de máquina

A partir da construção do *corpus* e do processamento dos atributos, modelo foi ajustado no conjunto destinado ao treinamento, para fazer previsões sobre dados que não foram treinados. Assim, dividiu-se os dados em duas partes: dados de treinamento e dados de teste. O conjunto de treinamento contém uma saída conhecida e o modelo aprende com esses dados, para ser generalizado para outros dados posteriormente. Dessa forma, obteve-se o conjunto de dados de teste (ou subconjunto) para testar a previsão do modelo neste subconjunto.

A proporção utilizada foi 80/20, em que 20% da amostra (*corpus*) foi retida para teste e o restante dedicada ao treino.

Figura 3.3 Proporção do *corpus* para treino e teste

As bibliotecas utilizadas no processamento do *corpus* foram:

- Pandas: Utilizada para realizar o carregamento do arquivo como um quadro de dados do Pandas e permirtir, então, a análise dos dados.

- Sklearn (subbiblioteca model_selection): Importação da train_test_split para dividir o conjunto de dados em treinamento e teste.
- Matplotlib : Importação do pyplot para construção dos gráficos de análise dos dados.

Foram utilizados os atributos TF-IDF, *bag of words* e POS *tags*, bem como os algoritmos de classificação *naive bayes*, SVM e *random forest*, no processamento dos dados. Esse atributos e algoritmos de classificação são amplamente utilizados em tarefas de classificação de textos, e permitem visualizar de forma compreensível a estrutura de decisão gerada.

A escolha de combinação de atributos e classificadores foi realizada a fim de identificar qual combinação obteve o melhor desempenho na tarefa de classificação.

4

Apresentação e análise dos resultados

4.1 *Corpus* e pré-processamento

Ao final de sua montagem, o *corpus*, contava com 100 notícias no total, sendo 50 notícias com conteúdo falso (FAKE) e 50 notícias com conteúdo verdadeiro (REAL).

O passo seguinte foi realizar a etapa de análise dos dados integrantes do *corpus*, a partir da aplicação de atributos e algoritmos de classificação, disponíveis na biblioteca NLTK, da linguagem Python.

Nesta etapa, a ação de extrair os atributos e realizar o preparo dos dados teve a finalidade de melhorar o *corpus*, para minimizar interferências nas fases de teste e treinamento.

Inicialmente, foi realizado o processo de limpeza da codificação, considerando aqui a retirada de acentos e pontuações. Assim, passou-se para a retirada das *stopwords*.

Foram analisadas as palavras que mais apareciam no *corpus* de duas formas: por quantidade de repetição e por frequência relativa.

Na Figura 4.1 é possível observar o resultado da análise de uma parte do *corpus* com a contagem dos termos mais frequentes.

```
[('covid', 130),
 ('nao', 113),
 ('vacinas', 52),
 ('pessoas', 47),
 ('vacina', 44),
 ('contra', 41),
 ('ivermectina', 39),
 ('vai', 32),
 ('sao', 28),
 ('coronavirus', 28),
 ('vacinados', 27),
 ('todos', 27),
 ('voce', 26),
 ('ja', 26),
 ('pra', 24)]
```

Figura 4.1 Palavras que mais se repetem no *corpus*

Na Figura 4.2 é possível observar a frequência relativa das palavras. A partir dessas informações, foram identificadas as caracteristicas do *corpus*.

```
[('nao', 117),
 ('covid', 54),
 ('vacinas', 30),
 ('vacina', 30),
 ('casos', 28),
 ('ja', 27),
 ('contra', 26),
 ('sao', 25),
 ('mortes', 21),
 ('saude', 20),
 ('pessoas', 19),
 ('disso', 18),
 ('alem', 18),
 ('pessoa', 17),
 ('doenca', 17)]
```

```
[('nao', 0.028903162055335968),
 ('covid', 0.0133399209486166),
 ('vacinas', 0.007411067193675889),
 ('vacina', 0.007411067193675889),
 ('casos', 0.00691699604743083),
 ('ja', 0.006669604743083),
 ('contra', 0.006422924901185771),
 ('sao', 0.006175889328063241),
 ('mortes', 0.005187747035573123),
 ('saude', 0.004940711462450593),
 ('pessoas', 0.004693675889328063),
 ('disso', 0.004446640316205534),
 ('alem', 0.004446640316205534),
 ('pessoa', 0.00419960474308300 4),
 ('doenca', 0.00419960474308300 4),
```

Figura 4.2 Frequência relativa de palavras do *corpus*

Os resultados aqui apresentados foram padronizados através do relatório de classificação, resultante da biblioteca sklearn.

4.2 Comparação dos resultados

Usualmente, precisão e acurácia são métricas utilizadas para avaliar modelos de classificação. O resultado que se procura na aplicação realizada neste trabalho é aquele que combina maior precisão e acurácia, conforme ilustrado na Figura 4.3.

Figura 4.3 Ilustração das medidas de precisão e acurácia

Neste relatório são utilizados para a interpretação dos resultados os seguintes parâmetros da biblioteca sklearn: acurácia (*accuracy*), precisão (*precision*), revogação (*recall*) e escore padronizado (f1-*score*).

4.2.1 TF-IDF e *naive bayes*

Utilizando a metodologia descrita e aplicando TF-IDF combinado com método de classificação *naive bayes*, os resultados encontrados são apresentados na tabela 4.1.

A partir desses resultados, os parâmetros mostram que o algoritimo obteve desempenho de acurácia na ordem de 70%. Considerando agora a precisão para a classificação como REAL, o algoritmo alcançou 60% e, para a classificação como FAKE, 80%.

Outra ferramenta para demonstrar e comparar os resultados é a matriz de confusão, que auxilia na medição de desempenho dos modelos na classificação das notícias do *corpus*, no teste,

Tabela 4.1 Resultados do relatório de classificação - TF-IDF com *naive bayes*

	Precisão (Precision)	Revogação (Recall)	Escore padronizado (F1-score)
FAKE	80%	67%	73%
REAL	60%	75%	67%
Acurácia (Accuracy)			70%

após a etapa de treinamento. Dessa maneira, pode-se avaliar o desempenho na classificação das notícias pela combinação de TF-IDF com *naive bayes*, conforme ilustra a Figura 4.4.

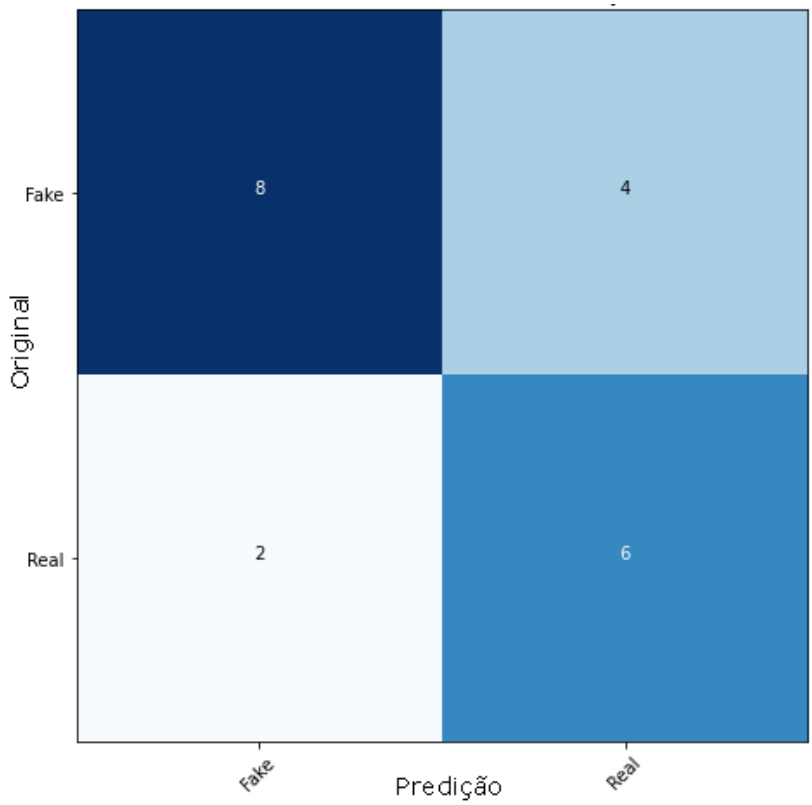


Figura 4.4 Matriz de confusão TF-IDF com *naive bayes*

Analisando a matriz de confusão em questão, o resultado do modelo previu corretamente 8 vezes mais notícias FAKE e 6 vezes notícias REAL, que eram originalmente dessas categorias. Em contrapartida, previu erroneamente 4 vezes para FAKE e 2 vezes para REAL.

4.2.2 TF-IDF e *random forest*

Agora, aplicando TF-IDF combinado com método de classificação *random forest*, os resultados encontrados são apresentados na tabela 4.2.

Os resultados obtidos na aplicação combinada do algoritimo apresentou parâmetro global

Tabela 4.2 Resultados do relatório de classificação - TF-IDF com *random forest*

	Precisão (Precision)	Revogação (Recall)	Escore padronizado (F1-score)
FAKE	100%	8%	15%
REAL	42%	100%	59%
Acurácia (Accuracy)			45%

de acurácia na ordem de 45%. Considerando a classificação REAL, o algoritmo alcançou 42% de precisão e para FAKE 100%, apresentando melhor resultado de precisão para classificação de FAKE.

A matriz de confusão da aplicação pode ser visualizada na Figura 4.5.

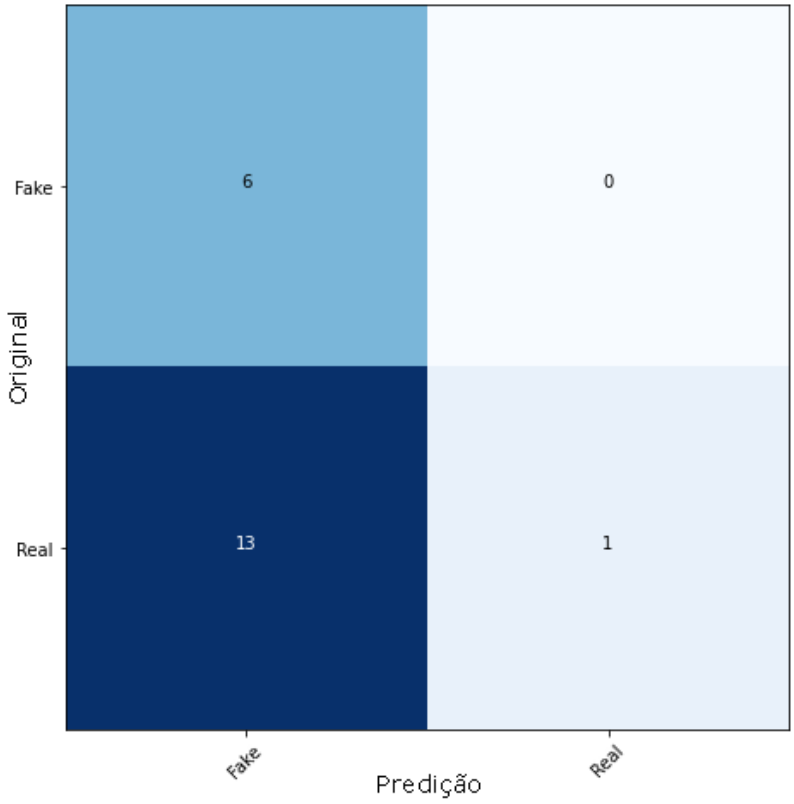


Figura 4.5 Matriz de confusão TF-IDF com *random forest*

Analisando a matriz de confusão em questão, tem-se como resultado que o modelo previu corretamente 6 vezes para notícias FAKE e 1 vez para notícias REAL. Em contrapartida, previu erroneamente 13 vezes para notícias FAKE e nenhuma para REAL, levando ao achado de que este modelo não foi o que apresentou resultado mais relevantes e adequados para o estudo.

4.2.3 TF-IDF e *support vector machine*

Utilizando a metodologia descrita e aplicando TF-IDF com o o método *support vector machine*, os resultados encontrados são apresentados na tabela 4.3.

O resultado da aplicação combinada apresentou o resultado de 35% de acurácia. Considerando a precisão para a classificação REAL, o método alcançou 100% e, para FAKE 32%, ressaltando que o modelo apresenta melhor resultado de precisão para classificação REAL.

Tabela 4.3 Resultados do relatório de classificação - TF-IDF com *support vector machine*

	Precisão (Precision)	Revogação (Recall)	Escore padronizado (F1-score)
FAKE	32%	100%	48%
REAL	100%	7%	13%
Acurácia (Accuracy)			35%

Analisando os resultados apresentados, realizou-se a contrução da matriz de confusão (Figura 4.6).

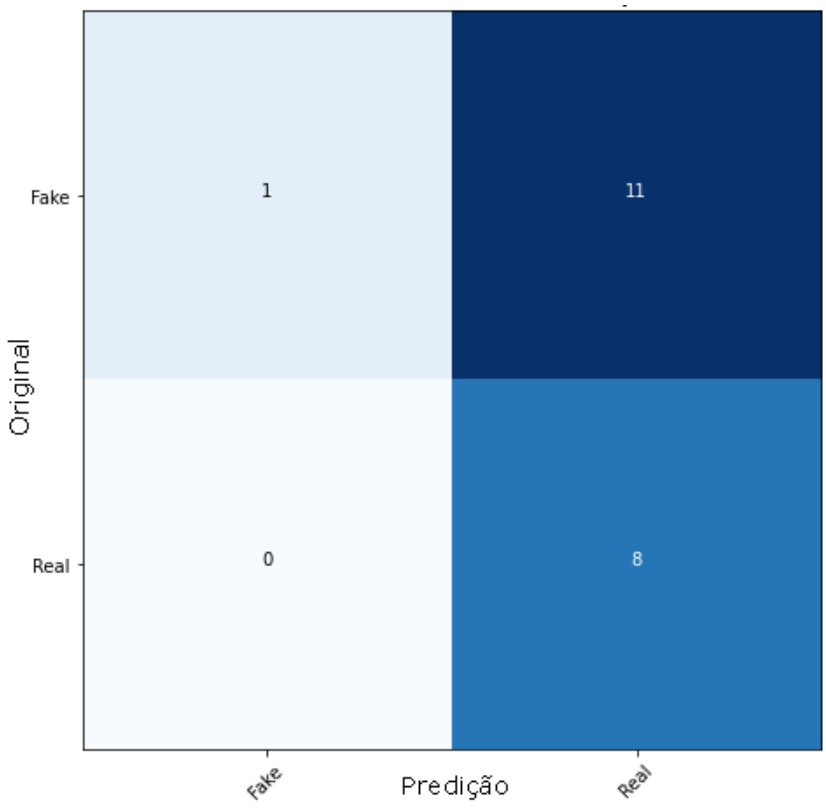


Figura 4.6 Matriz de confusão TF-IDF com *support vector machine*

Analisando a matriz de confusão em questão, temos como resultado que este modelo previu corretamente 1 vez para notícias FAKE e 8 vezes para noticias identificadas como REAL.

Em contrapartida, previu erroneamente 11 vezes para REAL e nenhuma vez para FAKE.

5

Conclusões e trabalhos futuros

Este trabalho contou com a construção de um conjunto de notícias (*corpus*), relacionadas à COVID-19, formado por 100 notícias sendo 50 com conteúdos verdadeiros (REAL) e 50 com conteúdos enganosos (FAKE). Esse conjunto de informações foi submetido inicialmente às etapas de pré-processamento e padronização e, então, aplicados a três modelos de classificação de aprendizagem de máquina.

A montagem do *corpus* com notícias em língua portuguesa, representou um importante e valioso recurso para academia, uma vez que, no momento do início deste trablho, não foram encontrados repositórios qualificados com informações sobre COVID-19 no Brasil.

As abordagens utilizadas nas etapa de qualificação e padronização das notícias contemplou a criação de uma *bag of words* com todas as palavras que ocorreram no conjunto de treinamento, bem como a aplicação da técnica de TF-IDF para avaliar a importância das palavras contidas nas notícias.

Posteriormente, na etapa de aplicação, foi utilizada a associação de TF-IDF com 3 métodos de aprendizado de maquina para classificação: *naive bayes*, *random forest* e SVM.

Os resultados mostraram que o método *naive bayes* se sobressaiu entre os demais obtendo uma acurácia de 70%, enquanto o desempenho dos demais métodos, random forest e SVM foram de 35% e 45%, respectivamente. Os resultados evidenciam que os métodos *random forest* e SVM tiveram dificuldade em predizer a classificação das notícias em FAKE e REAL, em relação ao algoritmo *naive bayes*.

Dentre as causas possíveis para a diferença de resultados está o volume ainda pequeno de notícias contidas no *corpus*, o que contribui com o desempenho do modelo em ajustar os dados de treinamento e, portanto, perder as tendências nos dados na fase de testes.

Como contribuições à academia, além no ineditismo do *corpus* na temática de COVID-19, em lígua portuguesa, fica a sugestão da utilização do conjunto de informações produzida em trabalhos futuros, por parte de estudantes e grupos de pesquisa. Indica-se a atualização e aumento no número de notícias, para tornar o *corpus* mais robusto e melhorar o desempenho dos modelos. Outra recomendação é a criação de uma interfaçe *web* e disponibilização desses modelos na rede, para que este se torne um aliado a população contra a desinformação.

Referências

BOSER; GUYON; VAPNIK. A Training Algorithm for Optimal Margin Classifiers. Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT'92). **Pittsburgh, 27-29 July 1992,**, [S.l.], p.144–152, 1992.

CHAKRABARTI, S. Chapter 1 - Introduction. In: CHAKRABARTI, S. (Ed.). **Mining the Web**. San Francisco: Morgan Kaufmann, 2003. p.1–13. (The Morgan Kaufmann Series in Data Management Systems).

DOMINGOS, P.; PAZZANI, M. On the optimality of the simple Bayesian classifier under zero-one loss. In: MACHINE LEARNING. **Anais...** [S.l.: s.n.], 1997. p.103–137.

DUCHARME, J. **World Health Organization Declares COVID-19 a Pandemic.** 2020.

GANDRABUR, S.; FOSTER, G. Confidence Estimation for Translation Prediction. In: SEVENTH CONFERENCE ON NATURAL LANGUAGE LEARNING AT HLT-NAACL 2003 - VOLUME 4, USA. **Proceedings...** Association for Computational Linguistics, 2003. p.95–102. (CONLL '03).

HABIB, A. e. a. False information detection in online content and its role in decision making: a systematic literature review. **Social Network Analysis and Mining (2019) 9:50 https://doi.org/10.1007/s13278-019-0595-5**, [S.l.], 2019.

JOACHIMS, T. Text categorization with support vector machines: learning with many relevant features. In: ECML-98, 10TH EUROPEAN CONFERENCE ON MACHINE LEARNING, Chemnitz, DE. **Proceedings...** Springer Verlag: Heidelberg: DE, 1998. n.1398, p.137–142.

KOENIGKAM-SANTOS, M. et al. Inteligência artificial, aprendizado de maquina, diagnostico auxiliado por computador e radiômica: avanços da imagem rumo à medicina de precisão. **Radiol Bras, 2019 Nov/Dez; 52(6):387–396.**, [S.l.], 2019.

MCCALLUM, A.; NIGAM, K. A comparison of event models for naive bayes text classification. In: AAAI 1998. **Anais...** [S.l.: s.n.], 1998.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997.

MONTEIRO, R. A. et al. Contributions to the study of fake news in portuguese: new corpus and automatic detection results. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL PROCESSING OF THE PORTUGUESE LANGUAGE. **Anais...** [S.l.: s.n.], 2018. p.324–334.

MONTEIRO, R. O.; NOGUEIRA, R. R. Projeto de um Sistema Web a Classificação de Fake News. In: XVI CONGRESSO LATINO-AMERICANO DE SOFTWARE LIVRE E TECNOLOGIAS ABERTAS. **Anais...** [S.l.: s.n.], 2019. p.120–123.

NASCIMENTO JR, C. L.; YONEYAMA, T. Inteligência artificial em controle e automação. **Editora Edgard Blücher Ltda**, [S.l.], 2000.

NELSON, J. L.; TANEJA, H. The small, disloyal fake news audience: the role of audience availability in fake news consumption. **New Media & Society**, [S.l.], v.20, n.10, p.3720–3737, 2018.

OLIVEIRA, L. M. R. Inteligência Artificial Aplicada a Detecção de Fake News. , [S.l.], 2019.

PEREIRA, S. d. L. **Processamento de linguagem natural**. [S.l.]: São Paulo: Universidade de São Paulo, 2011.

PICCHI NETTO, O. **Um filtro iterativo utilizando árvores de decisão**. 2013. Tese (Doutorado em Ciência da Computação) — Universidade de São Paulo.

PIERRI, F.; PICCARDI, C.; CERI, S. A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter. **EPJ Data Science**, [S.l.], v.9, 2020.

PUERTA-DIAZ, M. et al. O Processamento de Linguagem Natural nos Estudos Metricos da Informação: uma análise dos artigos indexados pela web of science (2000- 2019). **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S.l.], v.26, 2021.

ROBERTSON, S.; JONES, K. Relevance weighting of search terms. **J. Am. Soc. Inf. Sci.**, [S.l.], v.27, p.129–146, 1976.

SEBASTIANI, F. Machine Learning in Automated Text Categorization. **ACM Comput. Surv.**, New York, NY, USA, v.34, n.1, p.1–47, Mar. 2002.

SGANTZOS, K.; GRIGG, I. Artificial Intelligence Implementations on the Blockchain. Use Cases and Future Applications. **Future Internet**, [S.l.], v.11, n.8, 2019.

SILVA, J. A. S. da; MAIRINK, C. H. P. Inteligência artificial. **LIBERTAS: Revista de Ciênciais Sociais Aplicadas**, [S.l.], v.9, n.2, p.64–85, 2019.

TANDOC JR, E. C.; LIM, Z. W.; LING, R. Defining “fake news” A typology of scholarly definitions. **Digital journalism**, [S.l.], v.6, n.2, p.137–153, 2018.

WANG, H. Nearest neighbors by neighborhood counting. **IEEE transactions on pattern analysis and machine intelligence**, [S.l.], v.28(6), p.942–953, 2006.

WARDLE, C. Fake news. It’s complicated. , [S.l.], 2017.

WHO. **WHO Coronavirus (COVID-19) Dashboard**. [S.l.]: https://covid19.who.int/region/amro/country/br, 2021.

WU, F. et al. Author Correction: a new coronavirus associated with human respiratory disease in china. **Nature**, [S.l.], v.580, 2020.

YANG, Y. et al. TI-CNN: convolutional neural networks for fake news detection. **arXiv preprint arXiv:1806.00749**, [S.l.], 2018.

ZHANG, N. et al. Recent advances in the detection of respiratory virus infection in humans. **Journal of medical virology**, [S.l.], v.92(4), p.408–417, 2020.