



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIA DA COMPUTAÇÃO

Matheus Felipe Ferreira da Silva Lisboa Tigre

**Utilizando Modelos De Machine Learning Para Classificar Fake News De
COVID-19**

Recife

2023

Matheus Felipe Ferreira da Silva Lisboa Tigre

**Utilizando Modelos De Machine Learning Para Classificar Fake News De
COVID-19**

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre Profissional em Ciência da Computação.

Área de Concentração: Redes de Computadores e Sistemas Distribuídos

Orientador (a): Djamel Fawzi Hadj Sadok

Recife

2023

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

T568u Tigre, Matheus Felipe Ferreira da Silva Lisboa
Utilizando modelos de Machine Learning para classificar fake news de
COVID-19 / Matheus Felipe Ferreira da Silva Lisboa Tigre – 2023.
50 f.: il., fig., tab.

Orientador: Djamel Fawzi Hadj Sadok.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn,
Ciência da Computação, Recife, 2023.
Inclui referências.

1. Redes de computadores e Sistemas distribuídos. 2. Machine Learning. 3.
Fake News. 4. COVID-19. I. Sadok, Djamel Fawzi Hadj (orientador). II. Título

004.6

CDD (23. ed.)

UFPE - CCEN 2024 – 20

Matheus Felipe Ferreira da Silva Lisboa Tigre

“Utilizando modelos de Machine Learning para classificar Fake News de COVID-19”

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Redes de Computadores e Sistemas Distribuídos.

Aprovado em: 28/02/2023

BANCA EXAMINADORA

Prof.Dr. Nelson Souto Rosa
Centro de Informática/UFPE

Profa. Dra. Luciana Pereira Oliveira
Instituto Federal de Educação, Ciência e Tecnologia Paraíba

Prof. Dr. Djamel Fawzi Hadj Sadok
Centro de Informática / UFPE
(Orientador)

AGRADECIMENTOS

Queria agradecer primeiramente a minha família e minha namorada pelo grande apoio emocional e técnico que me deram durante o desenvolvimento desse trabalho.

Também queria agradecer pela orientação do professor Djamel Fawzi.

E também ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo financiamento do trabalho.

RESUMO

Com a chegada da pandemia de COVID-19 a internet se tornou um meio de comunicação muito utilizado para divulgação e consumo de informações sobre o que estava acontecendo no mundo e também de como se prevenir da contaminação do vírus. Acontece que ao mesmo tempo também eram espalhadas notícias que muitas vezes não condiziam com a realidade e eram utilizadas para confundir a população do Brasil, as chamadas *Fake News*. Devido ao seu grande volume, verificar todas de maneira manual pode ser uma tarefa quase impossível e com isso surge a necessidade da utilização de técnicas de Inteligência Artificial que consigam classificá-las de maneira automática. Para esse fim, modelos de *Machine Learning* podem vir a ser utilizados. Como não existiam bases de dados de texto em Português do Brasil, foi necessária a criação de um *dataset* colhendo textos em veículos de imprensa verificados e posteriormente ocorreu a análise dessas bases identificando diferenças em suas estruturas. Por fim, ao realizar experimentos com diferentes modelos de *Machine Learning* comparando diferentes combinações de *Features* foi identificado um desempenho superior quando utilizado como modelo o *Random Forest*.

Palavras-chave: machine learning; grid search; análise de tópicos; vetorização; técnicas de pré processamento.

ABSTRACT

With the arrival of the COVID-19 pandemic, the internet has become a means of communication widely used for the dissemination and consumption of information related to this illness that was occurring worldwide and how ways to prevent from virus contamination. However, at the same time, a great deal of news that often lacked any scientific basis was used to mislead Brazil's population in what is commonly known as Fake News. Due to its significant volume, manually verifying the soundness and veracity of each piece of news was practically impossible. As a result, there is a need for the use of verification strategies based on emerging technologies such as Artificial Intelligence capable of sifting among and classifying the news. This is yet a new area where machine Learning models can contribute a great deal in the combat of fake news. As there were no available text datasets in the Brazilian Portuguese language in this context, it was first necessary to create a new dataset by searching for texts in the verified press vehicles that later showed different structures after the text analysis. This work shows that when running Machine Learning models, an experiment comparing different features combinations, such as vectorization techniques and textual metrics, achieved good performance when utilizing Random Forest as a classification model.

Keywords: machine learning; grid search; topic analysis; vectorization; pre-processing techniques.

LISTA DE FIGURAS

Figura 1 – Bag Of Words com Frequency Vectors	18
Figura 2 – Fluxo da Elaboração dos Experimentos	30
Figura 3 – Cenário 1 com TF-IDF	40
Figura 4 – Cenário 1 com Frequency Vectors	41
Figura 5 – Cenário 1 com Word2Vec	42
Figura 6 – Cenário 2 com TF-IDF	43
Figura 7 – Cenário 2 com Frequency Vectors	44
Figura 8 – Cenário 2 com Word2Vec	45
Figura 9 – Cenário 3	46

LISTA DE TABELAS

Tabela 1 – 6 primeiras linhas do dataset de Fake News	26
Tabela 2 – Informações dos datasets	27
Tabela 3 – Conjunto de tópicos gerado analisando base de Fake News	34
Tabela 4 – Conjunto de tópicos gerado analisando base de Notícias Verdadeiras	35
Tabela 5 – Métricas textuais para a base de Fake News	36
Tabela 6 – Métricas textuais para a base de Notícias Verdadeiras	36
Tabela 7 – Resultados para o Cenário 1	37
Tabela 8 – Resultados para o Cenário 2	38
Tabela 9 – Resultados do GridSearch para o Cenário 2	38
Tabela 10 – Hiperparâmetros selecionados pelo Grid Search para o Cenário 3	39
Tabela 11 – Resultados para o Cenário 3	39

LISTA DE ABREVIATURAS E SIGLAS

FV	Frequency Vectors
SVM	Support Vector Machine
TF-IDF	Term Frequency–Inverse Document Frequency
W2V	Word2vec

SUMÁRIO

1	INTRODUÇÃO	11
1.1	MOTIVAÇÃO	12
1.2	OBJETIVOS	13
1.3	ORGANIZAÇÃO DA DISSERTAÇÃO	14
2	CONCEITOS BÁSICOS	15
2.1	FAKE NEWS	15
2.2	CLASSIFICAÇÃO DE TEXTO	16
2.3	MACHINE LEARNING	19
3	TRABALHOS RELACIONADOS	22
4	ELABORAÇÃO DOS EXPERIMENTOS	24
4.1	CRIAÇÃO DE DATASETS	24
4.2	CLASSIFICAÇÃO	27
5	RESULTADOS	31
5.1	ANÁLISE DAS BASES	31
5.2	MODELOS	36
	REFERÊNCIAS	47

1 INTRODUÇÃO

Em 26 fevereiro de 2020, o Ministério da Saúde confirmou o primeiro caso de COVID-19 no Brasil (UNA-SUS, 2020) e em 11 de março do mesmo ano, a OMS caracterizou a situação da doença como uma pandemia (WHO, 2020). De lá pra cá crescem o número de casos e o vírus se espalhou pelo mundo. Numa situação dessas, é necessário que a população siga as recomendações das autoridades de saúde de forma correta para amenizar os impactos causados na sociedade. Nos dias de hoje, um dos principais meios de comunicação para compartilhamento de informações é a internet, através, principalmente, de mídias sociais (SYN; OH, 2015) acontece que, muitas vezes, essas informações não são verificadas e podem acabar confundindo a população quando se trata de medidas sanitárias.

Como é gigantesco o volume de informações compartilhadas na internet, conferir todas, mesmo que existam portais confiáveis de checagem - como Boatos.org ¹, Agência Lupa ², G1 Fato ou Fake ³, entre outros - acaba se tornando um trabalho bastante complicado e com isso, as *Fake News* podem rapidamente se espalhar. A disseminação de *Fake News* pode trazer muitos impactos negativos para sociedade, principalmente quando se trata do comportamento de adoção de medidas protetivas de contenção do vírus da COVID-19 (LINDEN; ROOZENBEEK; COMPTON, 2020). Por exemplo, os autores em (FREEMAN et al., 2022) apresentam uma pesquisa sobre como as pessoas estão suscetíveis a acreditar em *Fake News* como teorias da conspiração tanto relacionadas a ser anti-vacinas, como à não adoção de medicações e até a não aceitação de orientações do governo em relação a medidas sanitárias.

No Brasil, isso não foi diferente. Durante o período pandêmico, um grande volume de desinformação foi espalhado e isso acabou gerando um impacto considerável na população. Uma quantidade considerável de brasileiros acredita fielmente em diferentes tipos de *Fake News*, desde coisas como "Água fervida serve para tratamento contra o corona vírus" até "Gargarejo com água morna ou com sal previne o vírus de chegar aos pulmões", entre outras, todas as informações que não possuem nenhuma comprovação científica de realmente funcionarem (GALHARDI et al., 2020). Além disso, o próprio governo federal, principalmente o Presidente da República, acabou contribuindo para a desinformação da população apresentando uma postura contrária à adoção de medidas sanitárias como o distanciamento social e o uso de máscaras,

¹ <https://www.boatos.org/>

² <https://piaui.folha.uol.com.br/lupa/>

³ <https://g1.globo.com/fato-ou-fake/>

através de falas e atitudes, fazendo com que uma parcela das pessoas acabasse não adotando estas práticas ou sendo contrárias a medidas sanitárias de controle do vírus da COVID-19 (GALHARDI et al., 2020)(CERON; LIMA-SANTOS; QUILES, 2021).

Uma estratégia que pode ser utilizada para controlar e fazer detecção automática de *Fake News* é a adoção de metodologias de mineração de dados, como por exemplo *Machine Learning* para auxiliar no processo da verificação de informações na web. Através de modelos matemáticos e data cêtricos e com a utilização de técnicas de pré processamento de texto, é possível treinar modelos que podem realizar a classificação de textos como sendo *fake news* ou informação real (MANZOOR; SINGLA et al., 2019) (AHMED et al., 2021) (APHIWONGSOPHON; CHONGSTITVATANA, 2018). Com a utilização dessas estratégias, é possível criar ferramentas que possam auxiliar a população quando receber notícias duvidosas através das redes sociais.

Para a criação destas ferramentas é necessário que existam *datasets* adequados para treinar os modelos sendo adotados. Acontece que, quando se trata de base de dados com textos específicos relacionados à COVID-19, existem poucas bases em Português do Brasil.

Isso cria boas oportunidades de pesquisa nessa área, onde podem ser criados novos *datasets* relacionados a COVID-19 colhendo textos de diferentes meios de comunicação confiáveis, assim como, também criar uma plataforma onde as pessoas consigam usar como auxílio para verificar se as informações que recebem na web são realmente verdadeiras.

1.1 MOTIVAÇÃO

Com a chegada da pandemia de COVID-19, muitas dúvidas surgiram em relação ao vírus e com o passar do tempo, muitas informações foram sendo compartilhadas na web. Acontece que tais informações eram difundidas de maneira muito rápida e em grande volume, principalmente através de redes sociais fazendo com que a verificação da veracidade de todas elas fosse quase impossível de ser efetuada manualmente (YANG; TIAN, 2021). A longo prazo, informações falsas podem ser muito prejudiciais à sociedade, principalmente no contexto pandêmico por se tratar de um problema que envolve a saúde da população. No Brasil, isso não foi diferente, foram inúmeras informações falsas compartilhadas por diferentes meios de comunicação.

Neste contexto é necessário que exista alguma estratégia onde essas informações possam ser verificadas e uma boa opção pode ser o uso de metodologias de classificação de texto como por exemplo, utilizando *Machine Learning*. Criando *datasets* bem classificados é possível identificar padrões tanto em notícias falsas como nas verdades, treinando modelos para que

os mesmos possuam altas taxas de acerto e consigam identificar a veracidade de informações compartilhadas de maneira textual (SHU et al., 2017).

Devido a isso, quando se tenta analisar um contexto nacional, surge a necessidade de encontrar *datasets* de *Fake News* relacionadas à COVID-19 em Português. Entretanto, existe uma carência de *datasets* nessa língua, por isso acaba sendo necessária a criação de uma base de dados confiável de notícias classificadas como falsas ou verdadeiras para treinar os modelos.

O foco do presente trabalho é treinar diversos modelos de *Machine Learning* para selecionar o modelo com melhor desempenho e usa-lo em uma plataforma que irá classificar automaticamente um texto como real ou *fake* e para isso, será criada uma base de dados contendo *Fake News* relacionadas à pandemia de COVID-19 em Português do Brasil. Através das análises, este trabalho tentará responder a seguinte pergunta de pesquisa "Como utilizar modelos de *Machine Learning* para classificar *Fake News* de COVID-19".

1.2 OBJETIVOS

Considerando as motivações apresentadas na seção anterior, o objetivo principal deste trabalho é aplicar técnicas de classificação de texto para comparar diferentes modelos *Machine Learning* de acordo com métricas pré estabelecidas, para escolher aquele com o melhor desempenho e utiliza-lo em uma ferramenta que irá avaliar a veracidade de informações que forem passadas. Para isso, será necessário também criar uma base de dados de *Fake News*, totalmente em português, contendo tanto textos classificados como notícias verdadeiras quanto notícias falsas Os objetivos específicos dessa dissertação são:

- Pesquisar e analisar estratégias para classificação automática de *Fake News*;
- Avaliar diferentes modelos de *Machine Learning* para classificação de texto.
- Analisar e comparar o desempenho dos modelos com diferentes configurações de características (*features*).
- Discutir os resultados obtidos com os experimentos comparando as diferentes configurações
- Através das análises, entender qual configuração dos modelos obteve o melhor desempenho e pode ser utilizada para classificar *Fake News*.

1.3 ORGANIZAÇÃO DA DISSERTAÇÃO

O trabalho é organizado da seguinte maneira: No Capítulo 2 serão apresentados os conceitos básicos necessário para o entendimento dos experimentos e análises que serão executadas. O Capítulo 3 da dissertação descreve a metodologia e técnicas utilizadas para a elaboração dos experimentos que terão os seus resultados demonstrados e discutidos no Capítulo 4. Por fim, no Capítulo 5, virão as conclusões que foram obtidas com as análises assim como os trabalhos futuros e dificuldades que foram encontradas na elaboração do trabalho.

2 CONCEITOS BÁSICOS

Neste Capítulo serão discutidos os principais conceitos utilizados para o desenvolvimento deste trabalho.

2.1 FAKE NEWS

Desde a época anterior à era da mídia impressa, existem evidências e registros de disseminação de rumores e notícias falsas (BURKHARDT, 2017). Com a chegada da internet e o desenvolvimento das redes sociais, nos dias de hoje, várias pessoas em várias partes do mundo tem acesso à informação que é disseminada de maneira rápida e com grande volume para todos. Observe que nem toda informação espalhada na internet realmente reflete a realidade dos fatos.

Um termo que é bastante popular para se referir a esse tipo de informação é "*Fake News*". Apesar de existirem registros de seu uso desde meados do século 19 (MERRIAM-WEBSTER-DICTIONARY, 2018), o termo ficou bastante popular nas eleições presidenciais de 2016 nos EUA, onde existiu uma grande quantidade de divulgações de informações imprecisas para confundir os eleitores. O termo é também utilizado como forma de ataque à imprensa, querendo descredibilizá-la (QUANDT et al., 2019).

Não existe uma definição absoluta para o termo *Fake News*. Logo, na literatura, os autores o definem de diversas maneiras. Em (GELFERT, 2018), por exemplo, *Fake News* é a divulgação, muitas vezes intencional, de falsas ou ilusórias informações que são apresentadas como notícias e que são construídas de maneira pensada para confundir o leitor. Já em (ZHANG; GHORBANI, 2020), *Fake News* representa qualquer tipo de notícia ou história falsa que majoritariamente distribuída pela internet para ludibriar os leitores e propiciar ganho financeiro ou político. Também pode se encontrar autores que citam *Fake News* como sendo postagens em redes sociais publicadas por contas não autênticas e que parecem reproduzir notícias (SOUZA et al., 2020).

Independente da definição que for utilizada, é possível perceber que a propagação de *Fake News* pode acabar sendo um problema para a sociedade. Como elas aparecem em grande volume (qualquer um pode criar e compartilhar notícias falsas), grande velocidade (com a internet esse compartilhamento é feito em tempo real), grande variedade (diversos assuntos

são abordados desde política à saúde) (ZHANG; GHORBANI, 2020), verificar a veracidade de notícias na internet é uma tarefa difícil porém bastante necessária.

Uma das técnicas adotadas é o *fact-checking*. Existem diversos portais e blogs online onde pessoas verificam, manualmente, a veracidade de notícias na internet. Só que, devido ao fato mencionando anteriormente do grande volume de *Fake News*, somente a análise manual não é suficiente. É necessária alguma estratégia para fazer a verificação automática dessas informações. Utilizar técnicas de classificação de texto podem auxiliar nesse processo, e este assunto será abordado na próxima seção.

2.2 CLASSIFICAÇÃO DE TEXTO

Grande parte do conteúdo que circula pela a internet é armazenado em texto e com isso, estudos relacionados a processamento e classificação de texto vem cada vez mais ganhando espaço e diversas pesquisas nesse campo foram executadas(KADHIM, 2019). O levantamento feito em (AGGARWAL; ZHAI, 2012) define o problema de classificação da seguinte maneira: Dado um conjunto de dados de treinamento $D = X_1, \dots, X_n$, em que cada registro possui uma etiqueta associada a uma classe representada por um valor pertencente a um conjunto $k = 1, \dots, k$ de valores discretos distintos.

Ou seja, existe um conjunto de texto (de treino) que será representado através de um formato numérico e cada elemento desse será associado a um valor categórico que será sua classe. O conjunto de treino será usado para a criação de um modelo de classificação que irá conseguir classificar novos registros submetidos a ele. Em outras palavras, é possível que um modelo bem treinado consiga, de maneira eficiente, classificar novos conjuntos de texto de acordo com as classes pré-definidas.

Para que um modelo de classificação seja bem construído é necessário a efetuação de alguns passos. Em (KORDE; MAHENDER, 2012) são apresentadas 6 etapas que podem ser executadas no processo de classificação de texto:

- **Coleta de documentos:** Esta é a etapa inicial onde os dados são colhidos em diversos formatos, como: HTML, PDF, Doc da Microsoft, etc;
- **Pré-processamento:** Aqui é necessário deixar o texto em um formato mais claro e padrão. Técnicas bastante utilizadas nesta etapa, são: A remoção de *stop-words* —que são palavras que podem ser retiradas do texto sem ocorrer alguma perda de sentido da

informação geral—, conversão de palavras para seu formato básico deixando somente o radical através de algoritmos de *stemming*, colocar as letras todas em minúsculo, entre outras coisas;

- **Vetorização:** Para que os algoritmos de classificação possam processar o texto, é necessário que este seja colocado em uma codificação específica. Normalmente, a representação mais utilizada é através de vetores, onde cada documento (ou cada conjunto de texto) é representado por uma matriz numérica onde cada posição do vetor pode apresentar um valor para alguma palavra, por exemplo;
- **Feature Selection:** Pode vir a ser necessário alterar o espaço do vetor e para isso, são utilizadas técnicas de *Feature Selection*. Através de algoritmos, é possível selecionar as melhores features (as palavras) de acordo com sua importância fazendo com que esse novo conjunto acabe melhorando o desempenho de um classificador de texto;
- **Classificação:** É o processo onde os documentos são analisados e classificados em categorias pré-definidas. Isso pode ser feito através da utilização de diversos modelos de classificação, como por exemplo: *Bayesian classifier*, *Decision Tree*, *K-nearest neighbor(KNN)*, *Support Vector Machines(SVMs)* e etc;
- **Avaliação de performance:** Ao executar os experimentos de classificação, é necessário utilizar alguma estratégia para analisar a eficiência dos modelos utilizados. Isso pode ser feito através da observação de métricas como *Accuracy*, *Precision*, *Recall*, *F1 Score* e *Specificity* que são descritas da seguinte maneira:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)}$$

$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$F1Score = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

$$Specificity = \frac{TN}{(TN + FP)}$$

Onde:

- **TP** = True Positive (Verdadeiro Positivo)

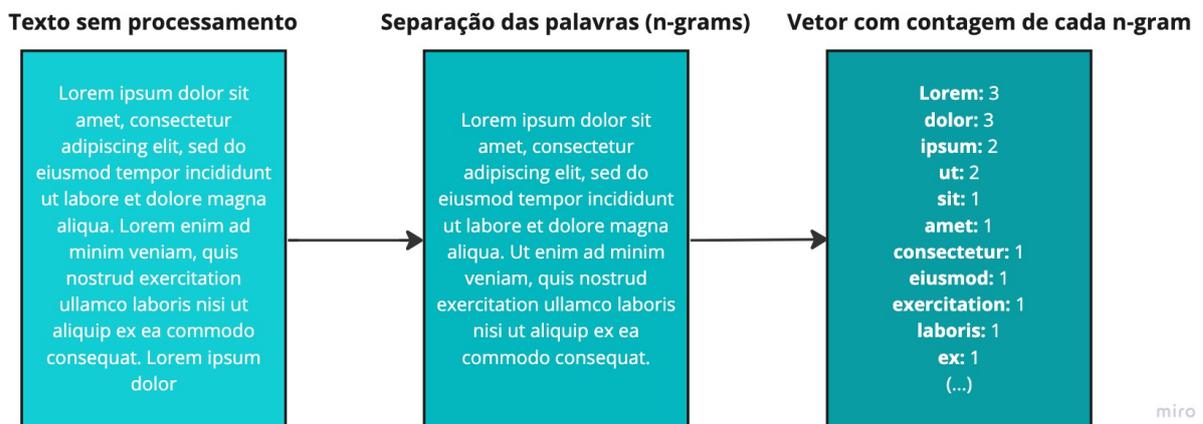
- **TN** = True Negative (Verdadeiro Negativo)
- **FP** = False Positive (Falso Positivo)
- **FN** = False Negative (Falso Negativo)

Através da análise dessas métricas é possível ter uma noção do quanto seu modelo consegue categorizar corretamente os documentos e fazer alterações que possam melhorar esse desempenho;

Quando se trata da etapa de vetorização, existem diversas técnicas que podem ser utilizadas de acordo com a preferência ou com o tipo de problema que será abordado. Exemplos bastante conhecidos são as técnicas baseadas em *Bag of Words* como *Frequency Vectors*, *Term Frequency Inverse Document Frequency* (ou simplesmente TF-IDF) e exemplos baseados no conceito de *Word Embeddings* como o *Word2Vec*.

Algumas das abordagens mais simples são as baseadas no conceito de *Bag-of-Words* onde basicamente existe um vetor de números, onde cada posição irá representar uma palavra, ou agrupamento de palavras (*N-grams*) e cada valor deste vetor irá representar sua presença ou ausência em determinado texto (CHORAŚ et al., 2021). Utilizando esse mesmo conceito também pode ser contabilizada a frequência de palavras ou conjuntos de palavras e essa abordagem pode ser chamada de *Frequency Vectors (FV)*. O fluxo do funcionamento dessa abordagem pode ser visualizado na figura 1

Figura 1 – Bag Of Words com Frequency Vectors



Fonte: Feito pelo autor.

Já o Term Frequency–Inverse Document Frequency (TF-IDF) é um algoritmo bastante utilizado quando se trata de mineração de texto e leva em conta a frequência que uma palavra aparece no documento (*Term Frequency*) e o log da probabilidade inversa daquele termo

aparecer nos demais documentos do corpus (*Inverse Document Frequency*) (LIU; YANG, 2012). Essa característica faz com que o TF-IDF consiga carregar uma maior quantidade de semântica da palavra já que também considera o contexto em que ela está inserida. (HAKIM et al., 2014) define da seguinte maneira o cálculo do TF-IDF:

$$tfidf(t, d, D) = tf(t, d)idf(t, D)$$

Onde t representa a quantidade de vezes que o termo aparece e d e D representam os documentos contidos no conjunto de texto, sendo:

- **$tfidf(t,d,D)$** = Peso calculado para a termo;
- **$tf(t,d)$ (frequência do termo)** = A quantidade de vezes que o termo é encontrado no texto
- **$idf(t,D)$ (frequência inversa dos documentos)** = Calculo do log multiplicado pela probabilidade inversa do termo ser encontrado em outros documentos do corpus

Outra forma de representar texto em formato numérico pode ser também aplicar técnicas de *Word Embeddings* que, basicamente, tem o objetivo de representar palavras através de vetores de tamanho fixo e esses vetores vão carregar, de certa forma, a semântica da palavra (ALMEIDA; XEXÉO, 2019). Uma das técnicas mais conhecidas é a W2V (CHURCH, 2017). Este modelo adota duas estratégias ambas utilizando técnicas de redes neurais: Continuous Bag-of-Words e Skip-gram Model (MIKOLOV et al., 2013). A primeira prevê a palavra do meio baseado nas palavras ao redor da mesma enquanto a segunda prevê as palavras ao redor da palavra observada.

2.3 MACHINE LEARNING

Devido ao grande volume de dados e *datasets* disponíveis surgiu a necessidade de retirar informações que podem vir a ser úteis deles. Nesse contexto, surge a aplicação de *Machine Learning* que é uma sub-área de Inteligência Artificial para auxiliar a tratar os dados com algoritmos que identificam padrões e ajudam em processos como tomadas de decisão, por exemplo (RAUT et al., 2017).

Em *Machine Learning* existem diversas categorias de aprendizado que podem ser utilizados e que possuem características diferentes entre si, como é possível citar: *Supervised Learning*,

Unsupervised Learning, Semi-Supervised Learning, Reinforcement learning, Transfer learning entre outras (ZHANG, 2020). Esta seção irá focar em *Supervised Learning* que representa o segmento utilizado neste trabalho.

Quando se trata de *Supervised Learning*, os algoritmos necessitam de algum tipo de auxílio externo, basicamente, são os casos onde é necessário dividir o *dataset* em teste e treino, onde existirão os valores para cada *feature* e os registros estarão associados a uma classe fazendo com que o algoritmo tenha então que aprender com os dados de treino e tente classificar ou prever corretamente os dados de teste (RAUT et al., 2017).

Abaixo serão apresentados alguns exemplos de algoritmos de *Supervised Learning*:

- **Decision Tree:** Em uma árvore de decisão, seus atributos são ordenados de acordo com seus valores e são divididas em nós e ramos, onde os nós representam um grupo que será classificado e os ramos representam os valores que o grupo pode assumir (KOTSIANTIS et al., 2007). Árvores de decisão são normalmente utilizadas para problemas de classificação.
- **Naive Bayes:** O princípio básico das Naive Bayes consiste no conceito de probabilidade condicional, onde, basicamente, são criadas árvores, também conhecidas como Redes Bayesianas, e o fluxo irá estar ligado com a probabilidade de cada coisa acontecer (LOWD; DOMINGOS, 2005)..
- **Suport Vector Machine:** Também usado para classificação, o *Suport Vector Machine* (Suport Vector Machine (SVM)) se baseia no conceito de cálculo de margem onde, basicamente são desenhadas margens entre as classes dispostas de maneira que a distância entre a margem e as classes é maximizada e ainda minimizando o erro de classificação (RAUT et al., 2017).
- **eXtreme Gradient Boosting:** É uma das implementações do conceito *Gradient Boosting* que é basicamente um conjunto de modelos de predição, normalmente árvores de decisão, só que possui uma série de otimizações que trazem uma melhor performance (CHEN; GUESTRIN, 2016).
- **Random Forest Classifier:** O *Random Forest* consiste em criar um conjunto de *Decision Trees* ao realizar o treinamento. Em problemas de classificação, por exemplo, o resultado é a classe escolhida pela maioria das árvores enquanto em problemas de regressão é apresentado a média do resultado delas (LIU; WANG; ZHANG, 2012).

Para que os modelos possam trabalhar com seu melhor desempenho possível é necessário que seja utilizada a melhor combinação de hiperparâmetros para o modelo utilizado. Hiperparâmetros são os valores passados para o modelo antes do treinamento e que irão definir o comportamento do mesmo na geração dos parâmetros (como pesos de um modelo de regressão linear) que ocorre nesta etapa de treino e que serão utilizados para processamento de dados de entrada (*input*) para gerar os valores das classificações do modelo (MACKAY, 1996).

Esses valores são externos ao modelo e que não podem ser estimados através dos dados inseridos no modelo. Uma técnica que é bastante utilizada para isso é o *Grid Search* que consiste, basicamente, em um algoritmo que através de uma busca exaustiva, utiliza diferentes configurações de hiperparâmetros (pré definidas) e, através de uma métrica especificada consegue definir qual a melhor combinação (LIASHCHYNSKYI; LIASHCHYNSKYI, 2019)

3 TRABALHOS RELACIONADOS

Em (PAIXAO; LIMA; ESPINASSE, 2020) são apresentados experimentos de classificação de texto utilizando tanto técnicas de *Supervised* e *Unsupervised Learning* com diferentes tipos de *features* como TF-IDF, POS Tag e *Word Embeddings*. Além disso, foi feita uma análise de tópicos na base de dados para detectar os padrões que diferem entre as notícias verdadeiras e as falsas. O trabalho obteve melhores resultados em comparação com outros do estado da arte que utilizavam do mesmo *dataset*.

Os autores em (ABONIZIO et al., 2020) apresentam uma solução que tenta conseguir classificar corretamente *fake news* independente do idioma em que o texto se encontra. Para isso são utilizadas *features* que não tenham relação única com algum idioma específico e sim que consigam abranger características gerais de qualquer que seja a língua em questão, que no trabalho eles dividem em três categorias: complexidade, estilo e características psicológicas. Para este trabalho foram utilizados *datasets* em três idiomas diferentes que são, português, inglês e espanhol. Os autores conseguiram uma média de acurácia de 85.3% utilizando com o classificador *Random Forest*.

Em (MONTEIRO; NOGUEIRA, 2019) os autores propõe metodologias para o carregamento de dados em um *Data Warehouse defake news*. Para isso, foi criado um *dataset* com notícias verdadeiras e falsas coletadas de portais na internet. A base foi testada com 6 modelos de *machine learning* diferentes e a pesquisa avaliou o desempenho de cada um. Também foi desenvolvida uma interface web onde seria possível colocar as notícias que serviriam como insumo para o *Data Warehouse* executar o processamento de texto e a classificação de *fake news*.

Uma abordagem um diferente foi utilizada em (MARTINS et al., 2021). Neste trabalho, os autores focaram em criar um *dataset* utilizando texto da ferramenta de troca de mensagens, Whatsapp. Eles colheram mensagens entrando em grupos públicos da ferramenta e filtraram aquelas que se referiam à pandemia de Covid-19. Após a montagem do *dataset* com essas mensagens, todas foram classificadas manualmente como mensagens com ou sem desinformação (que os autores classificam como o processo de criação e compartilhamento de mensagens, construídas propositalmente, contendo informações falsas, erradas ou tiradas de contexto para causar desordem). Em seguida foram feitos experimentos utilizando diferentes modelos de *Machine Learning*, configurações com dois tipos de vetorização (Bag-of-Words e

TF-IDF) e diferentes técnicas de pré-processamento. A melhor configuração obteve um F1 Score de 0.857.

Em (PATWA et al., 2021), os autores criam um *dataset* de postagens em redes sociais e artigos classificadas manualmente de maneira binária como real ou falso, onde as reais foram colhidas de portais verificados e as falsas em portais de *fact-checking*. Todo conteúdo foi agrupado e relacionado ao tópico de COVID-19. Após a criação e classificação do *dataset* foram utilizados alguns modelos de *Machine Learning* para comparação entre eles. Pra isso, foram retiradas as *stop-words*, links e caracteres especiais e para a vetorização, foi utilizado o *TF-IDF*. O modelo que obteve o melhor resultado entre as métricas analisadas foi o SVM.

Os trabalhos apresentados nessa seção buscam apresentar soluções para realização de classificações automáticas de *Fake News*, porém nenhum deles foca especificamente em notícias relacionadas a COVID-19 em Português o que acaba se tornando um diferencial do presente trabalho em relação aos demais. Para atacar esse contexto foi feita a primeira escolha de criar uma base de dados com textos relacionados à pandemia de COVID-19 totalmente em português brasileiro. Além disso, a utilização de diferentes configurações de *features* como a utilização de métricas textuais combinadas com métodos de vetorização também é feita de maneira diferente dos trabalhos apresentados acima, pois, são executadas diferentes combinações com ambos os tipos de *features*, assim como a utilização de mais métricas para avaliar os modelos.

4 ELABORAÇÃO DOS EXPERIMENTOS

Neste capítulo serão apresentadas todas as etapas da elaboração e execução dos experimentos conduzidos utilizando cinco modelos de *Machine Learning*, sendo eles: *Decision Tree*, *Naive Bayes*, *Support Vector Machine*, *eXtreme Gradient Boosting* e *Random Forest* (que são os modelos amplamente utilizados e citados na literatura como pode ser visto nos Capítulos 2). Além disso também será demonstrada como ocorreu a criação do *dataset* utilizado para alimentar os modelos assim como os resultados da execução dos mesmos com diferentes combinações de *Features* contando com a seleção da melhor configuração de hiper parâmetros feita pelo *Grid Search*.

4.1 CRIAÇÃO DE DATASETS

Inicialmente, devido ao fato de não existir uma boa quantidade de *datasets* em português que tratam de *Fake News* relacionadas a COVID-19, foi necessária a criação de uma base de dados para que os modelos pudessem ser alimentados com notícias falsas e verdadeiras. Para isso, foram escolhidos portais confiáveis de notícias e de *Fact-Checking*. Para colher as notícias verdadeiras o portal escolhido foi o G1 que é um portal bastante conhecido e importante do país, enquanto as notícias falsas foram adquiridas em dois sites de *Fact-Checking* que são a Agência Lupa e o Boatos.org, dois portais com conteúdo reconhecido e verificado por jornalistas profissionais.

Basicamente para conseguir os textos e agrupar as informações foram feitos alguns *scripts* utilizando a linguagem de programação Python ¹ junto com uma biblioteca chamada *Beautiful Soup* ², biblioteca essa que provê diversas funções para manipulação de dados no formato HTML que foram utilizadas para retirar os textos necessários das páginas web dos portais mencionados acima. Foram filtrados somente os textos relacionados à pandemia de COVID-19 fazendo *requests* utilizando o endereço web que trazia os textos relacionados a esse tópico. Após acessar estes endereços é possível ter acesso aos links das matérias de cada portal através de requisições do tipo *GET* e conseguir o html que contém as informações desejadas. Com isso, utilizando o *Beautiful Soup* é possível encontrar o conteúdo necessário para criação do *dataset* mapeando o HTML em objetos Python. A partir daí é que são agrupadas as informações em

¹ <https://www.python.org/>

² <https://pypi.org/project/beautifulsoup4/>

um formato de tabela utilizando a biblioteca pandas³ que também consegue gerar um csv que, neste caso, tem as seguintes colunas apresentadas na Tabela 1.

³ <https://pandas.pydata.org/>

Tabela 1 – 6 primeiras linhas do dataset de Fake News

link	date	title	text
https://www.boatos.org/tecnologia/governo-aluguel-social-teve-renda-afetada-compartilhou-link-whatsapp.html	2020-05-07T12:06:06-03:00	Governo está dando aluguel social para quem teve renda afetada e compartilhou link no WhatsApp #boato	Famílias que tiveram sua renda afetada pelas políticas de distanciamento social devido a pandemia do novo Coronavírus podem solicitar , a partir de hoje, o benefício do Aluguel Social. Acesse: [site]
https://www.boatos.org/tecnologia/governo-aluguel-social-teve-renda-afetada-compartilhou-link-whatsapp.html	2020-05-07T12:06:06-03:00	Governo está dando aluguel social para quem teve renda afetada e compartilhou link no WhatsApp #boato	Devido a Pandemia do novo Coronavírus (COVID-19) o Governo Federal, através do Programa Aluguel Social, em parceria com entidades filantrópicas, está cadastrando famílias que moram [...]
https://www.boatos.org/tecnologia/governo-aluguel-social-teve-renda-afetada-compartilhou-link-whatsapp.html	2020-05-07T12:06:06-03:00	Governo está dando aluguel social para quem teve renda afetada e compartilhou link no WhatsApp #boato	REQUISITOS: Para receber o benefício, o candidato(a) deve preencher os seguintes requisitos: – Ser Brasileiro(a); – Não estar trabalhando com carteira assinada; – [...]
https://www.boatos.org/tecnologia/governo-aluguel-social-teve-renda-afetada-compartilhou-link-whatsapp.html	2020-05-07T12:06:06-03:00	Governo está dando aluguel social para quem teve renda afetada e compartilhou link no W#boato	O governo federal está dando até três meses do benefício de aluguel social. Para ganhar, você precisa entrar em um site e compartilhar um link com amigos no WhatsApp [...]
https://www.boatos.org/mundo/china-cancelou-transacoes-dolar-bolsa-valores-abalou-mundo.html	2020-05-04T13:50:28-03:00	China cancelou transações em dólar na Bolsa de Valores e abalou o mundo #boato	“* Decisão súbita chinesa abala os cantos do mundo inteiro * A China, surpreendeu o mundo hoje e decidiu cancelar o pino do dólar nas transações na bolsa de valores e decidiu [...]
https://www.boatos.org/mundo/china-cancelou-transacoes-dolar-bolsa-valores-abalou-mundo.html	2020-05-04T13:50:28-03:00	China cancelou transações em dólar na Bolsa de Valores e abalou o mundo #boato	A China planeja um e-RMB em moeda digital separado para enfrentar brigas com US \$ e-RMB, que deve ser a primeira moeda digital operada por uma grande economia. [...]

Fonte: Feito pelo autor.

Na coluna *link* encontrasse a url de acesso referenciando onde foi tirado o texto, na coluna **data** sua data de publicação, na coluna **title** o título da mesma, em **text** o conteúdo que foi colhido e por fim, na coluna **classification** um número inteiro, que pode ser 0 ou 1 indicando se aquela informação se trata de algo verdadeiro (0) ou falso (1) que são valores atribuídos de acordo com a fonte em que foram encontrados.

Devido ao fato de que nem sempre os sites se encontram em uma estrutura padronizada, alguns erros podem ocorrer na criação da tabela, como por exemplo, textos que deveriam estar em uma linha acabam ficando divididos em várias linhas quando é executado o fluxo automático do *script*. Por causa disso, foi necessária uma checagem de toda a tabela criada para conferir as informações estavam em formato satisfatório e, quando não, foi executada uma correção manual.

Informações gerais sobre os *datasets* gerados podem ser vistos na Tabela 2:

Tabela 2 – Informações dos datasets

Classificação	Datasets	Quantidade de Textos	Período
Fake News	Boatos+Agência Lupa	2808	26 de Janeiro de 2020 à 05 de Julho de 2021
Notícias Verdadeiras	G1	10279	22 de Janeiro de 2020 à 28 de Fevereiro de 2021

Fonte: Feito pelo autor.

4.2 CLASSIFICAÇÃO

Com a base de dados pronta, primeiro foram utilizadas algumas técnicas de pré processamento visando padronizar os textos e prepara-los para serem inseridos e utilizados pelos modelos. Isso é necessário, primeiramente, para fazer uma adaptação dos textos retirando ruídos e trechos que não possuam valor semântico, por exemplo, e possam atrapalhar na classificação. Para isso, inicialmente são retiradas as urls, caracteres especiais, espaços em branco e todo o texto restante é convertido em letras minúsculas. Tudo isso foi feito utilizando um script em Python ⁴.

Após isso, é feita a retirada de *stop-words* que são palavras que são irrelevantes para o contexto semântico geral de um texto (SARICA; LUO, 2021), como por exemplo: "os", "as", "de", "para" e etc. Adicionalmente é feito o processo conhecido como *stemming* onde as palavras são transformadas para seu formato base ou raiz da palavra (ORENGO; HUYCK, 2001).

⁴ <https://github.com/Matheus-felipe/experimentos-mestrado>

Utilizando esta técnica, palavras que aparecem várias vezes devido a diferentes derivações das mesmas, faz com que todas sejam resumidas a um único termo.

Com o texto já tratado, agora é necessário que toda informação no formato textual seja convertida em um formato numérico que seja compreendido pelos modelos. Para isso, foram utilizadas as técnicas de vetorização citadas no Capítulo 2. Também foram adicionados alguns atributos textuais encontrados em (ZHOU; ZAFARANI, 2020), que descrevem, o conteúdo de notícias, sendo eles divididos em:

- Quantidade
 - Quantidade de Caracteres por texto
 - Quantidade de palavras por texto
 - Quantidade de sentenças por texto

- Complexidade
 - Média de caracteres por palavra
 - Média de palavras por sentença
 - Média de pontuação por sentença

Para gerar os conjuntos de dados que serviram como entrada de teste e de treino para os algoritmos avaliados foram utilizadas combinações destas *features* que geraram cenários específicos de acordo com a combinação. Isso foi feito para que pudesse ser mensurada e descoberta qual grupo de **features** reunido gera o melhor desempenho quando aplicado aos modelos, assim como, o que a adição ou remoção de features impacta na classificação. As combinações utilizadas são apresentadas a seguir:

- TF-IDF

- Bag-of-Words

- Word2Vec

- Atributos Textuais

- TF-IDF + Atributos Textuais

- Bag-of-Words + Atributos textuais

- W2V + Atributos textuais
- Bag-of-Words + Atributos textuais
- TF-IDF + Bag-of-Words + Word2Vec + Atributos Textuais

Para cada combinação de **features** foram utilizados os cinco modelos de classificação citados anteriormente, dividindo o conjunto de dados em 20% para testes e 80% para treino pois, para treinar os modelos, é necessário uma quantidade maior de dados. Para a etapa de treino dos modelos, foi utilizada a técnica de *Grid Search* para a otimização dos hiperparâmetros dos modelos.

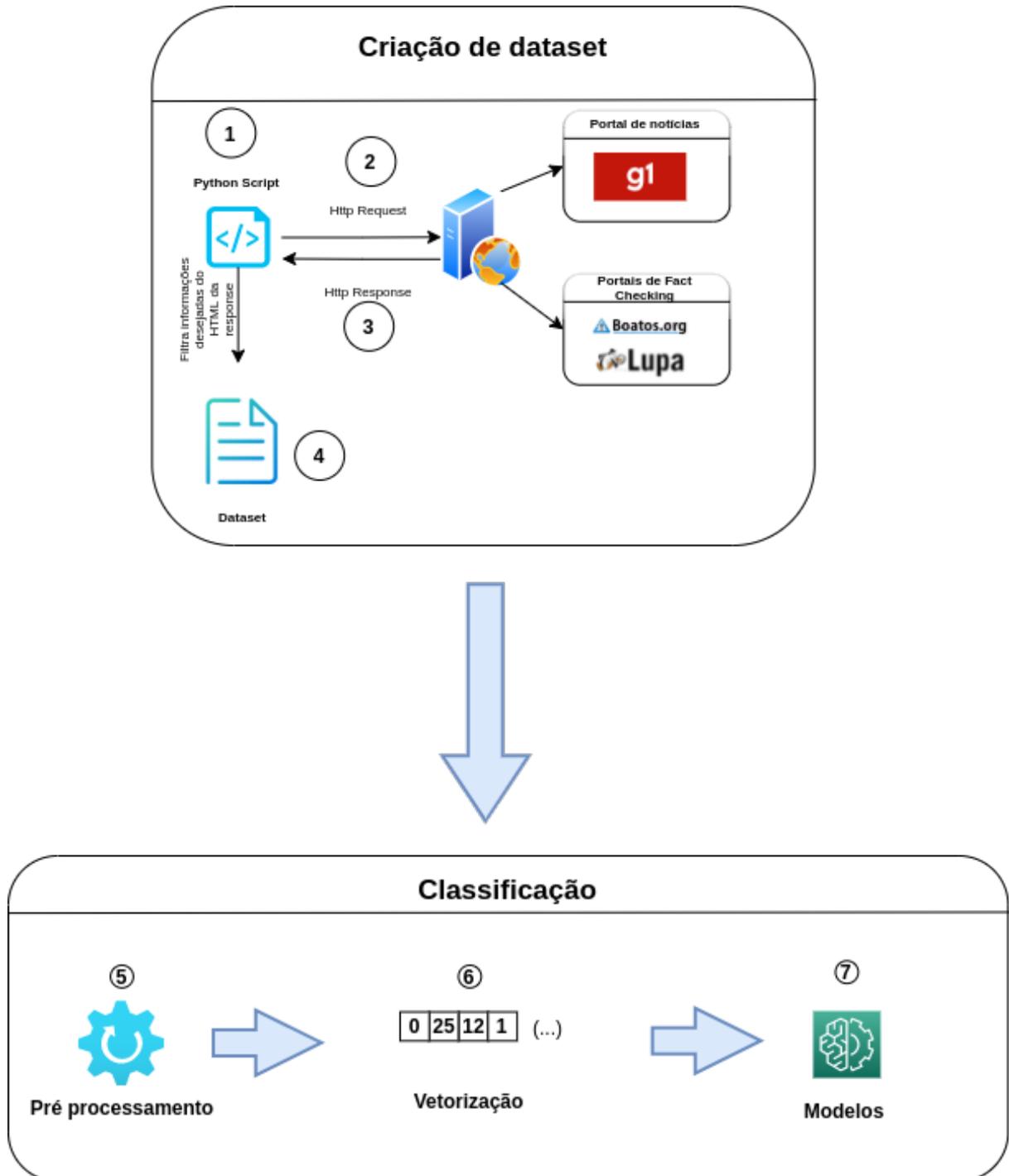
Após a execução do *Grid Search*, já com a melhor combinação de hiperparâmetros selecionada e o modelo treinado, foi possível chegar na etapa onde foi utilizado o conjunto de testes para executar a classificação e esta foi repetida trinta vezes pegando o resultado e calculando as métricas de avaliação de performance citadas na sub-seção 2.2 calculando a média e o desvio padrão de todas as execuções. Essa etapa foi feita porque os resultados podem variar em diferentes execuções do modelo então a média é calculada para ter um resultado mais preciso.

Na Figura 2 é possível encontrar uma visão geral da organização e fluxo de cada etapa que foi seguida na elaboração dos experimentos:

1. Script Python utilizado para obter e tratar as informações de websites.
2. Request HTTP feita para o servidor onde se encontram as informações desejadas (G1, boatos.org e Agência Lupa).
3. Response que trás o HTML que será tratado pelo script para geração do dataset.
4. Criação do dataset no formato csv utilizado para alimentar os modelos.
5. Etapa onde os textos do dataset serão tratados com as técnicas citadas anteriormente no Capítulo 2.
6. Transformação dos textos pré processados em formato numérico que é entendido pelos modelos.
7. Etapa onde os modelos são treinados e testados.

No próximo capítulo serão apresentados os resultados de experimentos assim como a discussão dos valores obtidos através das análises das métricas apresentadas no Capítulo 2 mais especificamente na Seção 4.2.

Figura 2 – Fluxo da Elaboração dos Experimentos



Fonte: Feito pelo autor.

5 RESULTADOS

Neste capítulo serão apresentados os resultados encontrados através dos experimentos e análises feitas no presente trabalho.

5.1 ANÁLISE DAS BASES

Antes da execução dos experimentos com os modelos de *Machine Learning*, foi executada uma análise de tópicos para identificar as principais características das bases de dados utilizadas tanto para *Fake News* como para *Notícias Verdadeiras*. O objetivo é detectar alguns padrões textuais em ambas. Para isso, foram utilizadas duas abordagens que são técnicas que auxiliam na análise da estrutura de textos:

- Métricas textuais analisadas encontradas em (ZHOU; ZAFARANI, 2020)
- Análise de tópicos como em (JACOBS; TSCHÖTSCHEL, 2019)

As métricas textuais são as citadas na Seção 4.2 que analisam **Quantidade** e **complexidade** dos textos. Na Tabela 5 é possível visualizar as métricas textuais para a base de *Fake News* e na Tabela 5.1 as mesmas métricas para a base de *Notícias Verdadeiras*. Quando observadas as métricas relacionadas a **Quantidade**, existem algumas diferenças entre as duas bases. Nos textos relacionados a notícias falsas é possível observar uma quantidade maior, quase o dobro, na média de quantidade de palavras (50.97 para *Fake News* e 25.67 para *Notícias Verdadeiras*), de caracteres (252.23 **Fake News** e 135.49 nas verdadeiras) e sentenças (3.82 e 2.71, respectivamente) mostrando que as notícias falsas da base acabam tendo uma quantidade maior de texto.

Outro ponto é quando analisado o desvio padrão das métricas é possível perceber um comportamento onde seus valores para a base de *Fake News* acabam sendo maiores quando observadas tanto as métricas de **Quantidade** como de **Complexidade**. Por exemplo, a quantidade de palavras por texto possui uma média de 50.97 e um desvio padrão de 58.90 (maior que a média) e isso também acontece com a quantidade de caracteres do texto (252.23 de média e 284.17 de desvio padrão) e quantidade de sentenças (3.82 de média e 4.09 de desvio padrão). Já na base de *Notícias Verdadeiras* o desvio padrão das métricas de *Quantidade* apresenta valores menores em relação a média, onde é possível ver a Quantidade de palavras

com média 25.67 e desvio padrão 8.77, de caracteres com média 135.49 e 44.90 de desvio padrão e sentenças com 2.71 de média e 0.88 de desvio padrão. Isso é explicado devido ao fato de que as *Fake News* não seguem um padrão, pois muitas vezes são feitas de maneira informal, compartilhadas em redes sociais, por exemplo, enquanto as *Notícias Verdadeiras* tem características técnicas com uma linguagem jornalística padronizada.

A dispersão dos dados também pode ser observada quando analisados os quartis, tendo em vista que, na base de *Fake News* os valores do primeiro quartil (25%) e do terceiro quartil (75%) tem uma diferença bastante significativa, olhando principalmente para Quantidade de palavras e caracteres, diferença essa que é bem menor quando analisadas essas mesmas métricas para a base de *Notícias Verdadeiras*. Já em relação a métricas de *Complexidade* as duas bases apresentam distribuições bem semelhantes.

Nas Tabelas 3 e 4 é possível observar os tópicos gerados para as bases, respectivamente de *Fake News* e de *Notícias verdadeiras*, assim como os *bi-grams* que estão associados a cada um. Na primeira, observa-se que existem muitos *bi-grams* que estão associados com notícias falsas sobre tratamentos relacionados a COVID-19, principalmente nos tópicos 6 e 10 que contam com conjuntos de palavras tais como: "cura covid", "uso cloroquina", "aprovar uso", "descobrir cura", "tratamento precoce", "adotar tratamento", "precoce ivermectina". Esses termos se relacionam com medicamentos que tiveram sua ineficácia comprovada cientificamente (GELLERIS et al., 2020). Além disso também existem textos relacionados com o uso da vacina que previne sintomas graves do vírus como nos tópicos 1,3 e 5, com termos como "vacina contra", "tomar vacina", "vacina pfizer", "vacina coronovac", "vacina chinês". Muitos textos dessa base questionam a eficácia da vacina argumentando sobre efeitos colaterais que ela pode causar, como pode ser visto nos Tópicos 1 e 3 com termos como "pessoa morrer" e "passar mal". Existem também textos onde são questionadas as medidas não farmacológicas relacionadas à prevenção da COVID-19 como nos Tópicos 7 e 8 onde se encontram termos como "isolamento social", "álcool gel", "usar álcool", "higienizar mão", "matar nada" e "ficar casa" que remetem a higienização pessoal e distanciamento social.

Já na tabela de tópicos gerados pela base de *Notícias Verdadeiras* é possível notar um comportamento diferente. Os Tópicos estão mais relacionados com a linguagem jornalística e com termos que remetem a informações técnicas. Como nos tópicos 1, 2, 5, 6 e 10 onde se encontram os termos "registro morte", "morte balanço", "total caso", "confirmar caso", "diagnóstico covid", "registrar diagnóstico", "confirmar doença", "país contabilizar" e "mil pessoa". Esses termos apresentam características informativas com dados sobre a situação da

doença no país como número de mortes e contagiados. No tópico 7 também é possível encontrar termos que remetem a outros veículos de imprensa como "extra estadão", "globo extra", "folha uol" que estão no mesmo tópico que "dados imunização", "levantamento junto", "saúde apontar". Isso pode demonstrar uma indicação de busca de referências para as informações passadas. Já os tópicos 8 e 9 contêm "ministério da saúde", "confirmação ministério", "rede social", "flavio dino", "jair bolsonaro", "enviar mensagem" que podem estar relacionados a declarações de órgãos públicos e personalidades políticas (que no caso são, respectivamente o Ministério da Saúde do Brasil e os políticos Flávio Dino e Jair Bolsonaro) que são divulgadas na mídia.

Tabela 3 – Conjunto de tópicos gerado analisando base de Fake News

Tópico		Bi-grams									
1	contra covid	vacina contra	vacinação contra	após tomar	vacinar contra	bill gates	peessoa morrer	reino unido	causa vacina	universidade oxford	
2	compartilhar link	link amigo	entrar site	amigo whatsapp	preciso entrar	cesta básico	ganhar preciso	dar cesta	governo federal	site compartilhar	
3	tomar vacina	vacina chines	após tomar	efeito colateral	chines contra	vacina covid	vacina pfizer	vacina covid	enfermeira tomar	passar mal	
4	novo coronavirus	quarentena novo	coronavirus covid	covid novo	isolamento social	durante quarentena	pandemia novo	durante pandemia	causa quarentena	combater novo	
5	joão doria	governador paulo	paulo joão	governador joão	presidente bolsonaro	vacina covid	número morto	lockdown paulo	decretar lockdown	pandemia covid	
6	tratamento precoce	morte covid	adotar tratamento	precoce ivermectina	uso hidroxícloroquina	precoce contra	maxaranguape rn	chapeco sc	caso morte	salvar vida	
7	vídeo mostrar	mostrar momento	pandemia covid	wuhan china	isolamento social	rio janeiro	hospital campanha	durante pandemia	covid coronavirus	protesto contra	
8	álcool gel	lei seca	usar álcool	higienizar mão	matar nada	chape capô	ficar casa	dentro capô	bem sincero	poder receber	
9	corona virus	contra corona	rede atacação	virus covid	novo corona	morte covid	ano comorbidade	dois rede	fernanda torres	cura contra	
10	cura covid	ministério saúde	estudo realizar	uso cloroquina	aprovar uso	governo federal	oms oferecer	falar sobre	coronavirus cura	descobrir cura	

Fonte: Feito pelo autor.

Tabela 4 – Conjunto de tópicos gerado analisando base de Notícias Verdadeiras

Tópico	Bi-gram									
	registro morte	incluir registro	morte balanço	hora incluir	incluir balanço	último hora	morte hora	hora confirmar	total caso	confirmar caso
1	registro morte	incluir registro	morte balanço	hora incluir	incluir balanço	último hora	morte hora	hora confirmar	total caso	confirmar caso
2	diagnóstico covid	país ter	ter óbito	registrar diagnóstico	óbito registrar	covid segundo	veículo imprensa	consórcio veículo	segundo levantamento	levantamento consórcio
3	novo coronavírus	pandemia novo	infectar novo	contra novo	causar novo	infecção novo	combate novo	caso novo	doença causar	morte novo
4	estadual saúde	secretaria estadual	sars cov	coronavirus	novo coronavirus	saúde confirmar	caso novo	cov morte	país caso	confirmar país
5	caso confirmar	ter caso	confirmar covid	novo caso	confirmar doença	estado ter	confirmar coronavirus	país ter	secretaria saúde	número caso
6	desde início	início pandemia	país contabilizar	balanço consórcio	segundo balanço	pandemia segundo	veículo imprensa	consórcio veículo	covid desde	contabilizar caso
7	extra estado	globo extra	folha uol	estado folha	imunização país	dado imunização	secretarias saúde	levantamento junto	junto secretarias	saúde apontar
8	ministério saúde	oficialmente reconhecer	saúde dizer	após confirmação	confirmação ministério	segundo ministério	saúde ainda	reconhecer suspeito	ainda oferecer	caso suspeito
9	rede social	fazer parte	social dizer	contra coronavirus	meio rede	flávio dino	enviar mensagem	jair bolsonaro	informação divulgar	saúde público
10	reino unido	segundo feira	estados unidos	contra covid	sexta feira	partir segundo	vacina contra	mil pessoa	novo caso	mil morte

Fonte: Feito pelo autor.

Tabela 5 – Métricas textuais para a base de Fake News

Característica	Métrica	Média	Desvio Padrão	Min	25%	50%	75%	Max
Quantidade	Quantidade de palavras	50.97	58.90	3.00	21.00	32.00	61.00	1153.00
	Quantidade de caracteres	252.23	284.17	21.00	108.00	158.00	306.75	5064.00
	Quantidade de sentenças	3.82	4.09	1.00	2.00	3.00	4.00	82.00
Complexidade	Média de caracteres por palavras	5.06	0.67	2.82	4.66	5.00	5.38	15.85
	Média de palavras por sentença	17.12	13.23	2.00	10.33	14.66	20.32	282.00
	Média de pontuação por sentença	1.73	2.11	0.00	0.50	1.18	2.20	48.50

Fonte: Feito pelo autor.

Tabela 6 – Métricas textuais para a base de Notícias Verdadeiras

Característica	Métrica	Média	Desvio Padrão	Min	25%	50%	75%	Max
Quantidade	Quantidade de palavras	25.67	8.77	5.0	20.0	25.0	30.0	116.0
	Quantidade de caracteres	135.49	44.90	30.0	106.00	134.00	159.00	606.0
	Quantidade de sentenças	2.71	0.88	1.0	2.00	3.00	3.00	9.0
Complexidade	Média de caracteres por palavras	5.32	0.59	3.5	4.92	5.28	5.70	9.25
	Média de palavras por sentença	16.98	7.34	3.0	12.00	15.50	21.00	65.0
	Média de pontuação por sentença	1.63	1.57	0.00	0.50	1.00	2.00	26.0

Fonte: Feito pelo autor.

5.2 MODELOS

Após a execução dos experimentos, os resultados foram agrupados em três diferentes cenários, para melhor análise dos impactos da mudança de configurações, assim como os modelos e suas respectivas configurações de hiper parâmetros que foram selecionadas pelo *Grid Search*. Os seguintes agrupamentos foram considerados:

- **Cenário 1:** Técnicas de vetorização isoladas
- **Cenário 2:** Técnicas de vetorização combinadas com *métricas* textuais
- **Cenário 3:** Todas as *features* combinadas

As configurações de hiperparâmetros para o cenário 1 podem ser vistas na Tabela ???. Já na tabela 7 é possível ver os resultados dos experimentos quando apenas as técnicas de vetorização foram utilizadas para geração de *features* onde estão destacados os melhores valores por métrica e tipo de vetorização. Como pode ser visto nos valores destacados em negrito, para o TF-IDF todas as métricas foram superiores as demais quando utilizado o SVM como modelo de classificação, alcançando 94.40% de *Precision*. Já com o *Frequency Vectors* e o W2V, o Random Forest foi o que obteve os melhores valores nas métricas.

O *Decision Tree* obteve os menores valores nas métricas com todas as técnicas de vetorização (exceto para o W2V) com destaque para o *TF-IDF* que obteve as menores porcentagens

com esse modelo em comparação com as demais. Já, apesar do desempenho superior com o *TF-IDF*, o SVM é o segundo modelo em termos de resultados de métricas com o *Frequency Vectors* e o *W2V* ambos com Accuracy, Recall, F1 Score e Specificity superiores em comparação à Decision Tree, XGBoost e Naive Bayes. Este último, por sua vez teve um desempenho bem próximo ao SVM olhando para a configuração que usou o *Frequency Vector* como técnica de vetorização com no máximo 2% de diferença nas métricas ainda tendo uma *Precision* maior, mas já utilizando o *W2V* teve um desempenho bem abaixo dos demais, tendo somente uma *Precision* maior (94.65%). Já em relação ao XGBoost, ele é o segundo pior em desempenho para todas as técnicas de vetorização

Tabela 7 – Resultados para o Cenário 1

Features	Modelo	Média				
		Accuracy	Precision	Recall	F1 Score	Specificity
TF-IDF	Random Forest	90.83%	90.29%	91.11%	90.70%	91.11%
	Decision Tree	78.36%	77.55%	77.98%	77.77%	77.98%
	SVM	93.77%	94.40%	92.67%	93.53%	92.67%
	Naive Bayes	89.85%	88.43%	90.72%	89.56%	90.72%
	XGBoost	90.03%	86.83%	93.30%	89.95%	93.30%
Frequency Vectors	Random Forest	92.43%	92.16%	93.28%	92.72%	93.28%
	Decision Tree	78.54%	77.52%	78.79%	78.15%	78.79%
	SVM	92.07%	90.69%	92.90%	91.78%	92.90%
	Naive Bayes	91.45%	91.33%	90.99%	91.16%	90.99%
	XGBoost	88.78%	90.09%	87.21%	88.63%	87.21%
W2V	Random Forest	92.52%	92.37%	93.32%	92.84%	93.32%
	Decision Tree	80.05%	79.79%	81.20%	80.49%	81.20%
	SVM	89.49%	90.60%	88.36%	89.46%	88.36%
	Naive Bayes	77.83%	94.65%	60.72%	73.98%	60.72%
	XGBoost	89.49%	90.65%	87.70%	89.15%	87.70%

Fonte: Feito pelo autor.

Já para o Cenário 2, combinando as técnicas de vetorização com as features textuais os resultados podem ser visualizados na Tabela 8. Nela, em todas as configurações o *Random Forest* foi o modelo com melhor desempenho das métricas com destaque para o *Frequency Vectors* que teve os maiores valores em todas as métricas comparado com as demais técnicas de vetorização, alcançando 95% de *Precision*. O SVM vem logo em seguida no quesito valores das métricas nas três configurações de Features. O Naive Bayes teve os menores valores na maioria das métricas comparado com os demais modelos em todas as três configurações de Features, principalmente quando é olhado o *W2V + Features textuais* onde ele obteve aproximadamente 43.15% de *Specificity* e *Recall*, 57% de *F1 Score* e 67.14% de *Accuracy*. O

Tabela 8 – Resultados para o Cenário 2

Features	Modelo	Média				
		Accuracy	Precision	Recall	F1 Score	Specificity
TF_IDF + Features Textuais	Random Forest	93.41%	94.22%	92.15%	93.17%	92.15%
	Decision Tree	81.03%	82.94%	79.02%	80.93%	79.02%
	SVM	92.88%	94.42%	91.85%	93.12%	91.85%
	Naive Bayes	69.55%	71.37%	66.61%	68.91%	66.61%
	XGBoost	91.18%	92.16%	89.66%	90.89%	89.66%
Frequency Vectors + Features Textuais	Random Forest	94.39%	95.81%	92.93%	94.35%	92.93%
	Decision Tree	82.72%	81.76%	84.37%	83.04%	84.37%
	SVM	91.36%	93.67%	88.87%	91.21%	88.87%
	Naive Bayes	85.13%	93.52%	74.40%	82.87%	74.40%
	XGBoost	91.27%	93.82%	88.04%	90.84%	88.04%
W2V + Features Textuais	Random Forest	93.41%	95.10%	91.77%	93.40%	91.77%
	Decision Tree	82.19%	83.39%	81.96%	82.67%	81.96%
	SVM	92.70%	92.44%	93.70%	93.06%	93.70%
	Naive Bayes	67.14%	85.86%	43.15%	57.44%	43.15%
	XGBoost	90.03%	94.84%	85.37%	89.86%	85.37%

Fonte: Feito pelo autor.

Tabela 9 – Resultados do GridSearch para o Cenário 2

Modelo	Hiperparâmetro	Valores		
		TF-IDF	Frequency Vectors	W2V
Random Forest	<i>Bootstrap</i>	false	false	false
	<i>Criterion</i>	gini	gini	entropy
	<i>Número máximo de features</i>	log2	log2	log2
	<i>Número de árvores na floresta</i>	200	200	200
Decision Tree	<i>Criterion</i>	gini	gini	gini
	<i>Profundidade máxima</i>	50	50	100
SVM	<i>Parâmetro de regularização</i>	1.0	1.0	1.0
	<i>Tipo de Kernel</i>	linear	linear	linear
Naive Bayes	<i>Alpha</i>	1.0	1.0	1.0
XGBoost	<i>Taxa de aprendizado</i>	0.4	0.4	0.4
	<i>Number of boosting stages</i>	1000	1000	1000

Fonte: Feito pelo autor.

XGBoost alcançou valores maiores para as métricas em comparação com *Decision Tree* que para todas as configurações de *Features* ficou com valores entre 74.02% (Recall TF-IDF + *Features Textuais*) e 84.37% (*Frequency Vectors + Features Textuais*). Os valores escolhidos pelo Grid Search podem ser encontrados na Tabela 9

Observando o Cenário 3 temos os seguintes valores para os hiperparâmetros apresentados na Tabela 5.2 e os resultados da execução dos experimentos na Tabela 5.2. Neste cenário o *Random Forest* também foi o modelo com maioria das métricas com os maiores valores com Accuracy de 92.34%, Precision de 95.08% seguido do SVM que teve Recall de 91.16% e

Specificity de 91.56%, maiores valores dessas métricas em comparação com os demais modelos. O *XGBoost* vem logo em seguida ordenando os resultados por valores de métricas variando de 88.25% (*Accuracy*) a 92.80% (*Precision*). O *Decision Tree* fica em penúltimo com valores variando de 80.87% em *Recall* a 83.18% em *Precision* ficando a frente somente do Naive Bayes teve os menores valores obtendo aproximadamente 46.17% em *Recall* e *Specificity*, por exemplo.

Tabela 10 – Hiperparâmetros selecionados pelo Grid Search para o Cenário 3

Modelo	Hiperparâmetro	Valores
Random Forest	<i>Bootstrap</i>	False
	<i>Criterion</i>	gini
	<i>Número máximo de features</i>	log2
	<i>Número de árvores na floresta</i>	200
Decision Tree	<i>Criterion</i>	gini
	<i>Profundidade máxima</i>	50
SVM	<i>Parâmetro de regularização</i>	1.0
	<i>Tipo de Kernel</i>	linear
Naive Bayes	<i>Alpha</i>	1.0
XGBoost	<i>Taxa de aprendizado</i>	0.4
	<i>Number of boosting stages</i>	1000

Fonte: Feito pelo autor.

Tabela 11 – Resultados para o Cenário 3

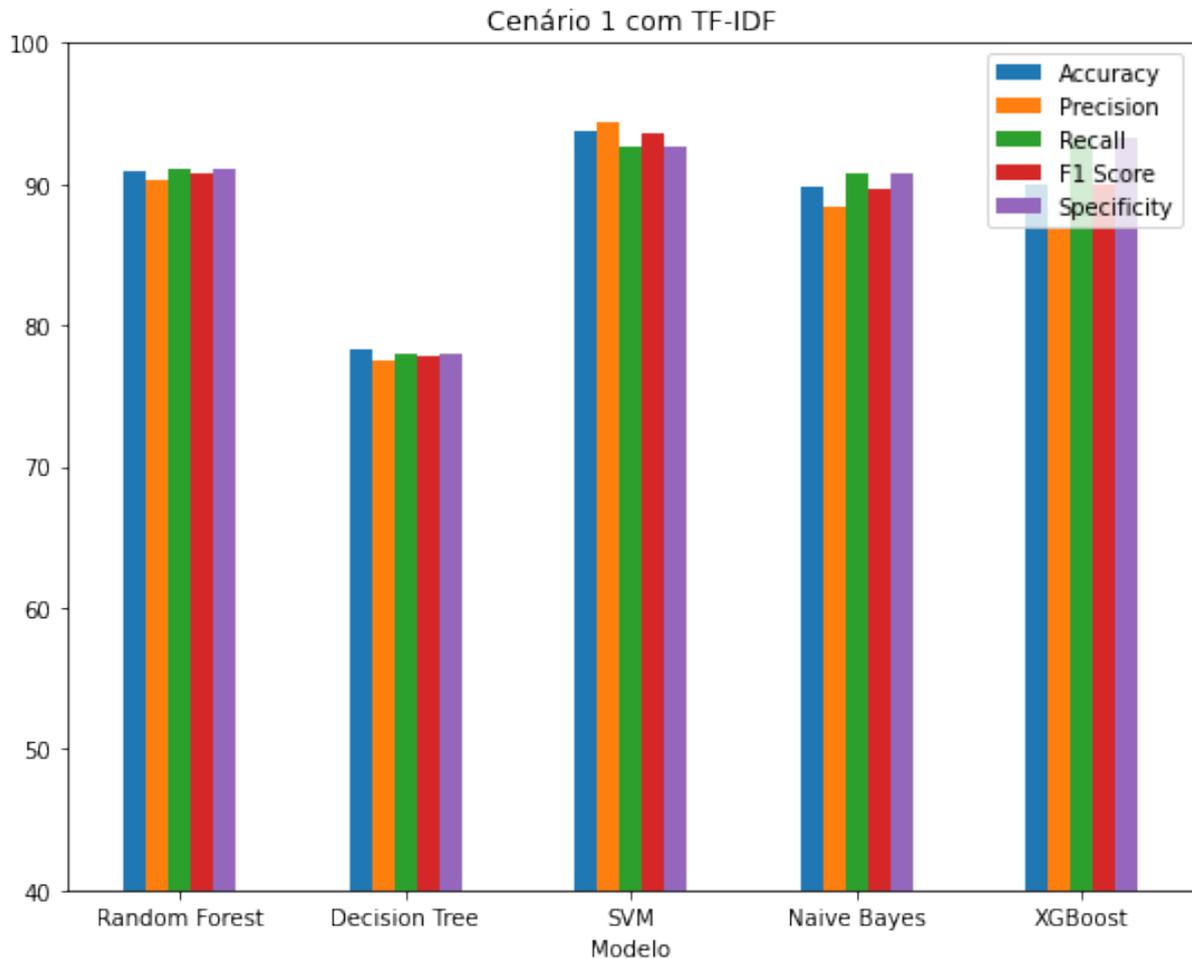
Features	Modelo	Média				
		Accuracy	Precision	Recall	F1 Score	Specificity
ALL	Random Forest	92.34%	95.08%	89.85%	92.39%	89.85%
	Decision Tree	81.83%	83.18%	80.87%	82.01%	80.87%
	SVM	91.36%	91.96%	91.16%	91.16%	91.56%
	Naive Bayes	70.88%	88.77%	46.17%	60.74%	46.17%
	XGBoost	90.56%	92.80%	88.25%	90.47%	88.25%

Fonte: Feito pelo autor.

Em seguida serão apresentados alguns gráficos que ilustram mais detalhadamente os resultados apresentados anteriormente comparando os modelos de *Machine Learning* utilizados. Na Figura 3 é possível observar que o SVM foi o modelo com os melhores resultados das métricas seguido do *Random Forest*, *XGBoost* e *Naive Bayes* que alcaçaram valores bem próximos en-

tre si. Nessa configuração, o *Decision Tree* teve um desempenho bem inferior em comparação com os demais modelos.

Figura 3 – Cenário 1 com TF-IDF



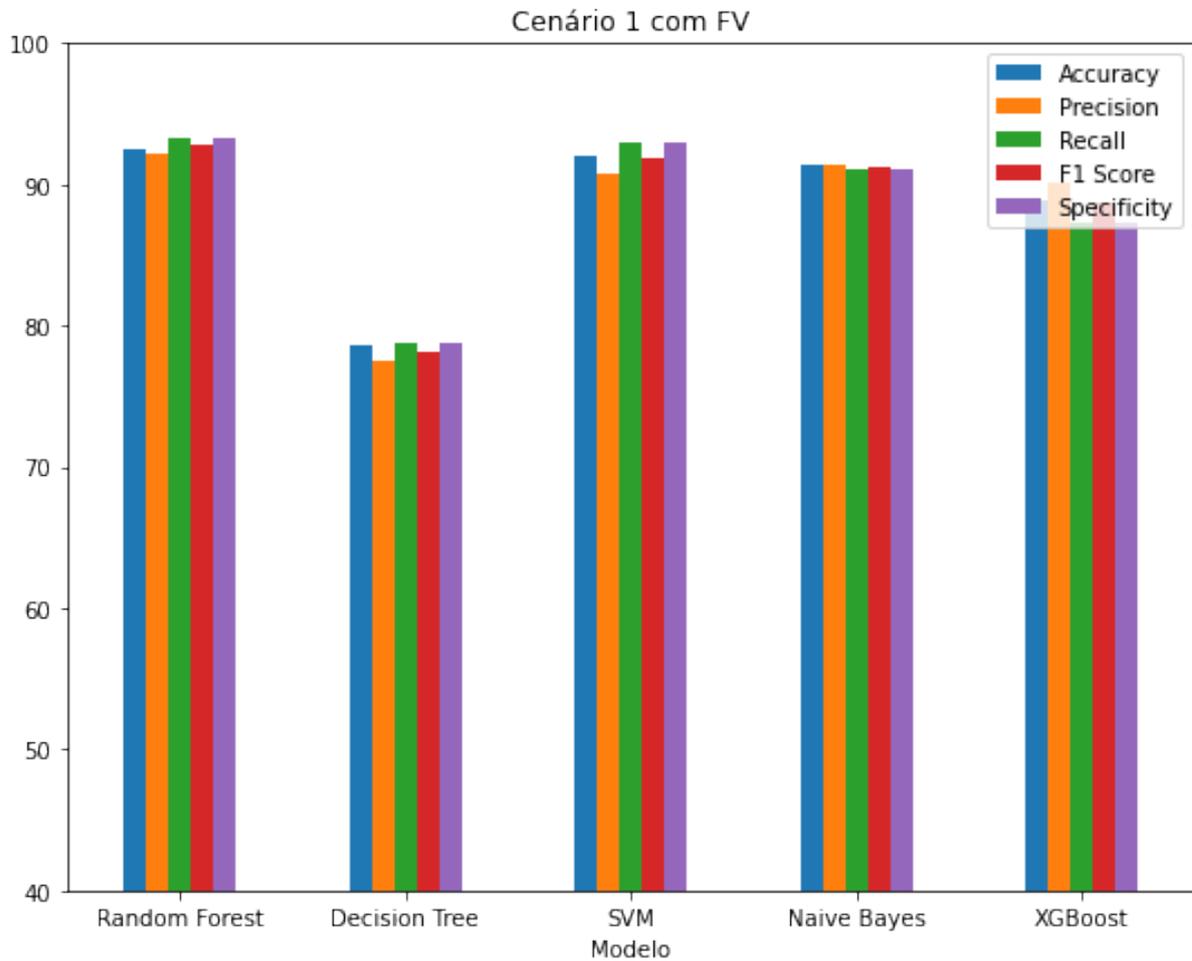
Fonte: Feito pelo autor.

Já na Figura 4, quando utilizado o *Frequency Vectors* como técnica de vetorização, é possível observar um comportamento bem similar ao anterior com *Random Forest*, *SVM*, *Naive Bayes* e *XGboost* bem próximos (mas nesse caso, o *Random Forest* se sobressai) e novamente o *Decision Tree* com os menores valores bem abaixo dos outros.

Quando observado o Cenário 1 com o *Word2Vec* é possível notar uma diferença no comportamento dos resultados. O *Random Forest* continua sendo o modelo com valores maiores, próximos do *XGBoost* e do *SVM*, porém, nessa configuração, o *Naive Bayes* apresentou os piores resultados em todas as métricas com exceção da *Precision* que obteve um valor maior do que os demais modelos.

Quando observado o Cenário 2 na configuração presente na Figura 6 o *Naive Bayes* teve o pior desempenho enquanto *Random Forest* e o *SVM* tiveram resultados muito próximos, com

Figura 4 – Cenário 1 com Frequency Vectors



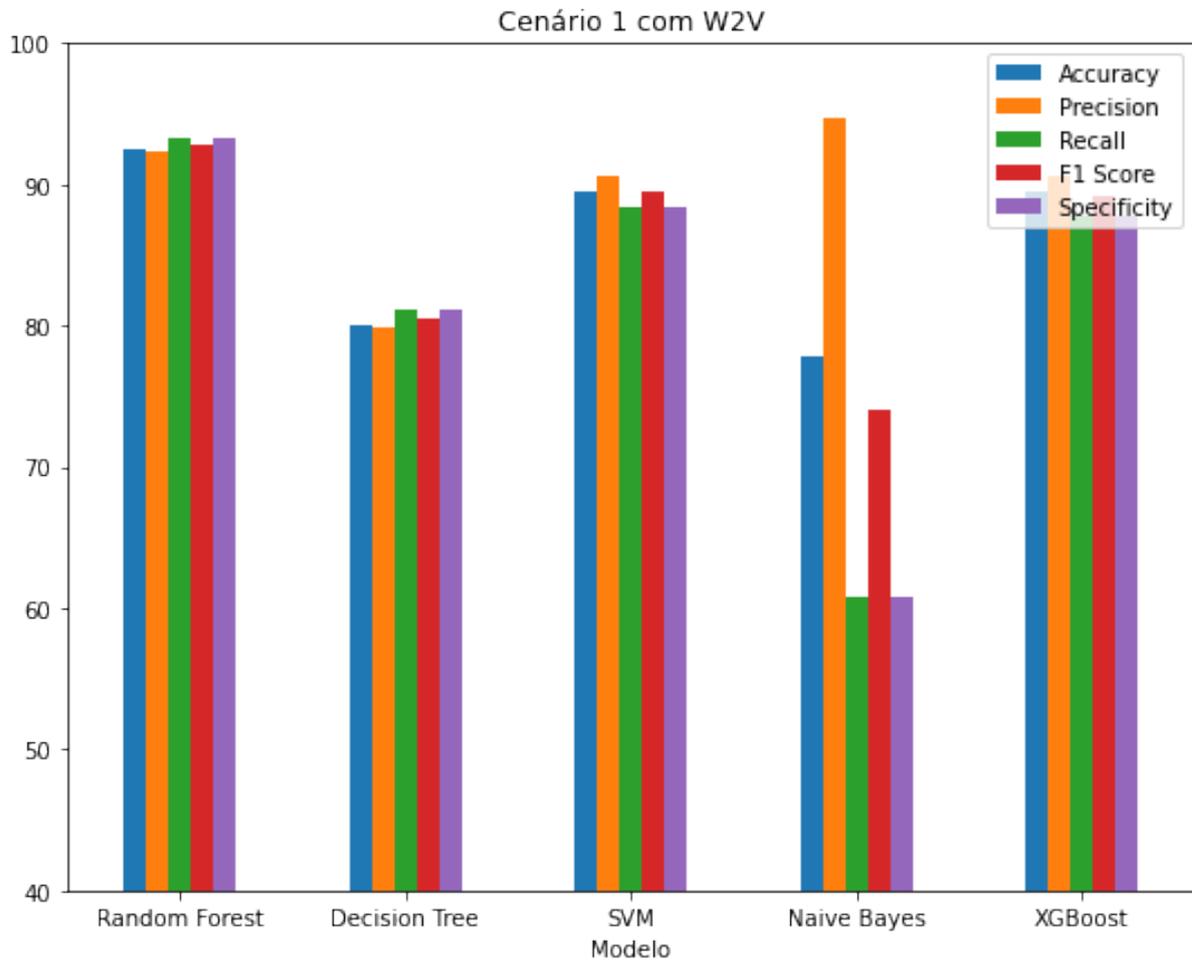
pequena vantagem do primeiro.

Na Figura 7 quando utilizado o *Frequency Vectors* combinado com as métricas textuais, *Random Forest*, *Decision Tree* e *XGBoost* seguem o mesmo comportamento dos cenários anteriores, porém, *Naive Bayes* e *Decision Tree*, estão com os menores valores, sendo que o segundo apresenta somente *Accuracy* e *Precision* maiores que o primeiro.

Observando o Cenário 2 utilizando W2V como estratégia de vetorização (Figura 8) o *Naive Bayes* teve o pior desempenho em comparação com diferentes modelos em diferentes cenários em quatro das cinco métricas utilizadas (*Accuracy*, *Precision*, *Recall* e *F1 Score*). Neste cenário, o *Random Forest* e *SVM* obtiveram os melhores resultados, sendo o primeiro melhor em algumas métricas com o *XGBoost* sendo o terceiro melhor. Este comportamento é bastante semelhante quando comparado com o Cenário 3 (Figura 9).

No geral, as bases de *Fake News* e de notícias verdadeiras apresentaram estruturas peculiares quando analisadas, tanto em sua estrutura (métricas textuais) como no conteúdo que

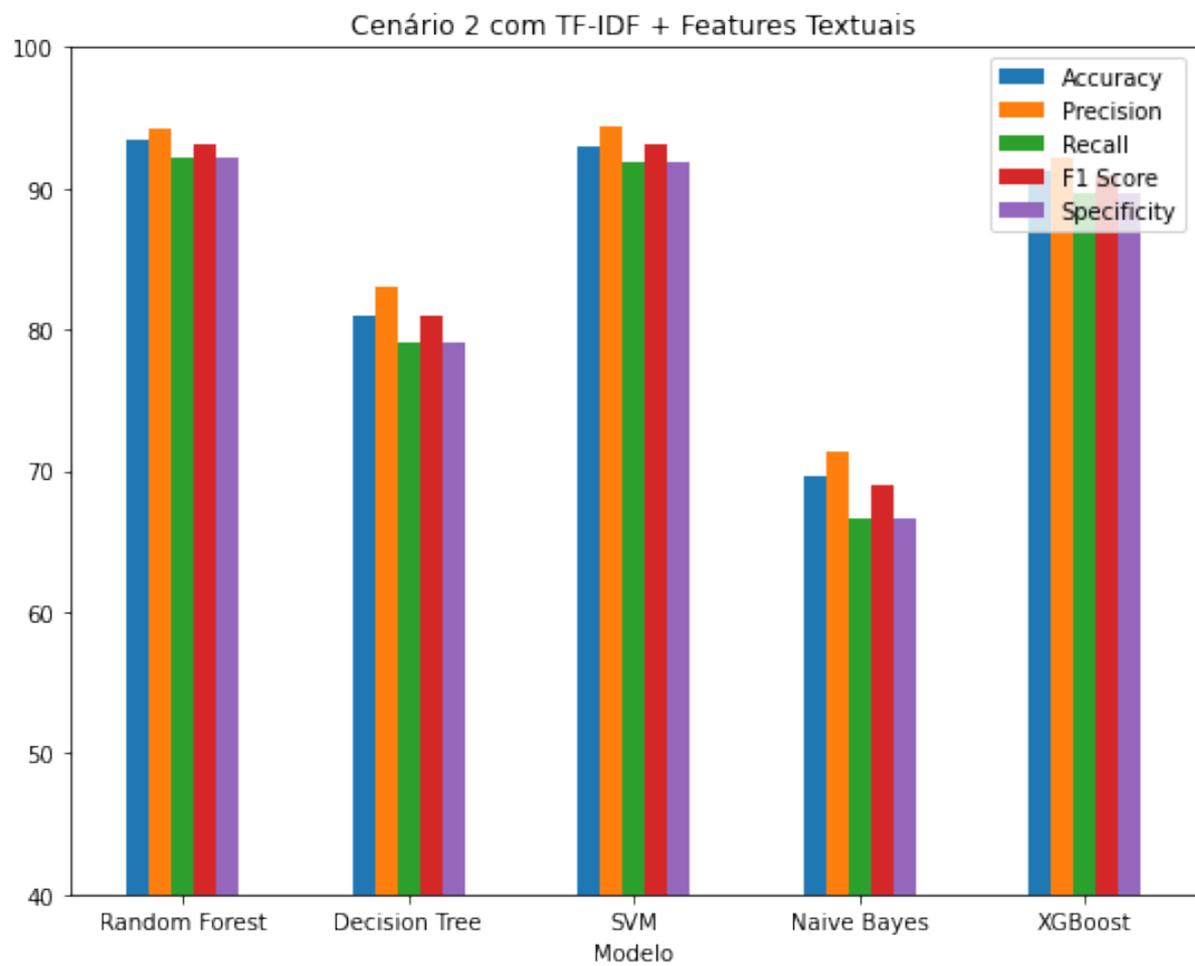
Figura 5 – Cenário 1 com Word2Vec



Fonte: Feito pelo autor.

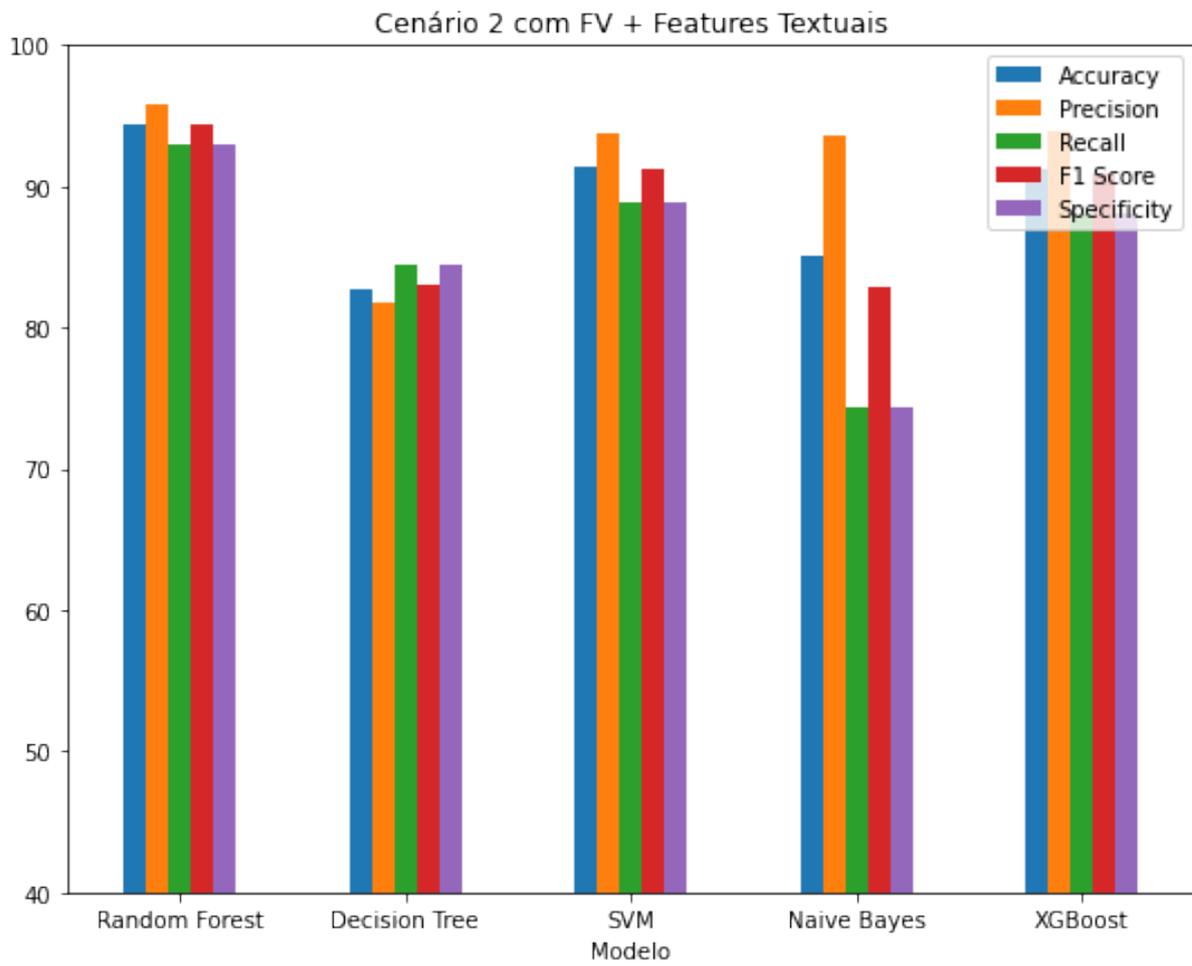
possuem (análise de tópicos e nuvem de palavras). Quando observados os modelos, o *Random Forest* foi o que obteve o melhor desempenho no geral comparando todos os cenários. Em relação as configurações de *features* houve uma melhora no desempenho dos modelos quando combinadas as técnicas de vetorização com as métricas textuais.

Figura 6 – Cenário 2 com TF-IDF



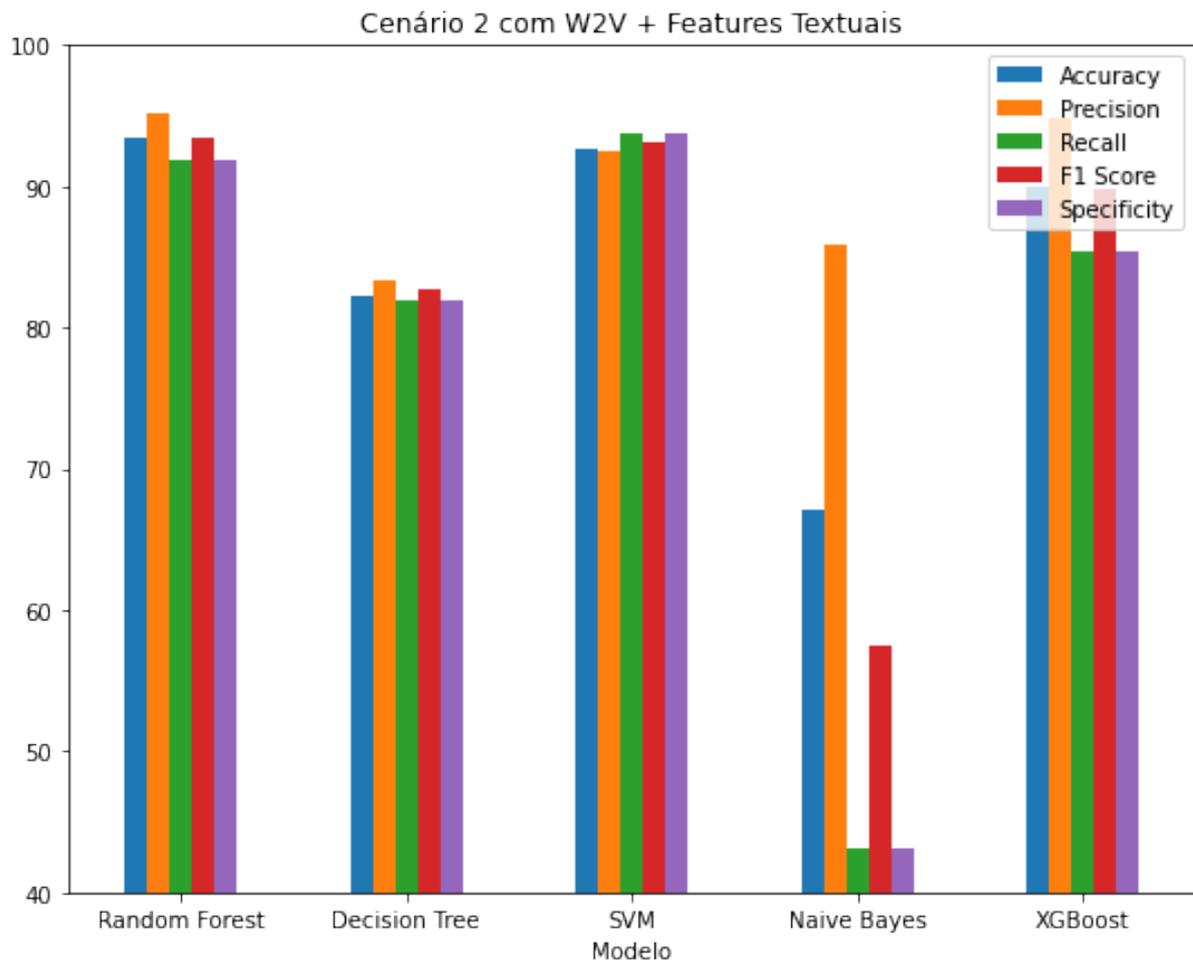
Fonte: Feito pelo autor.

Figura 7 – Cenário 2 com Frequency Vectors



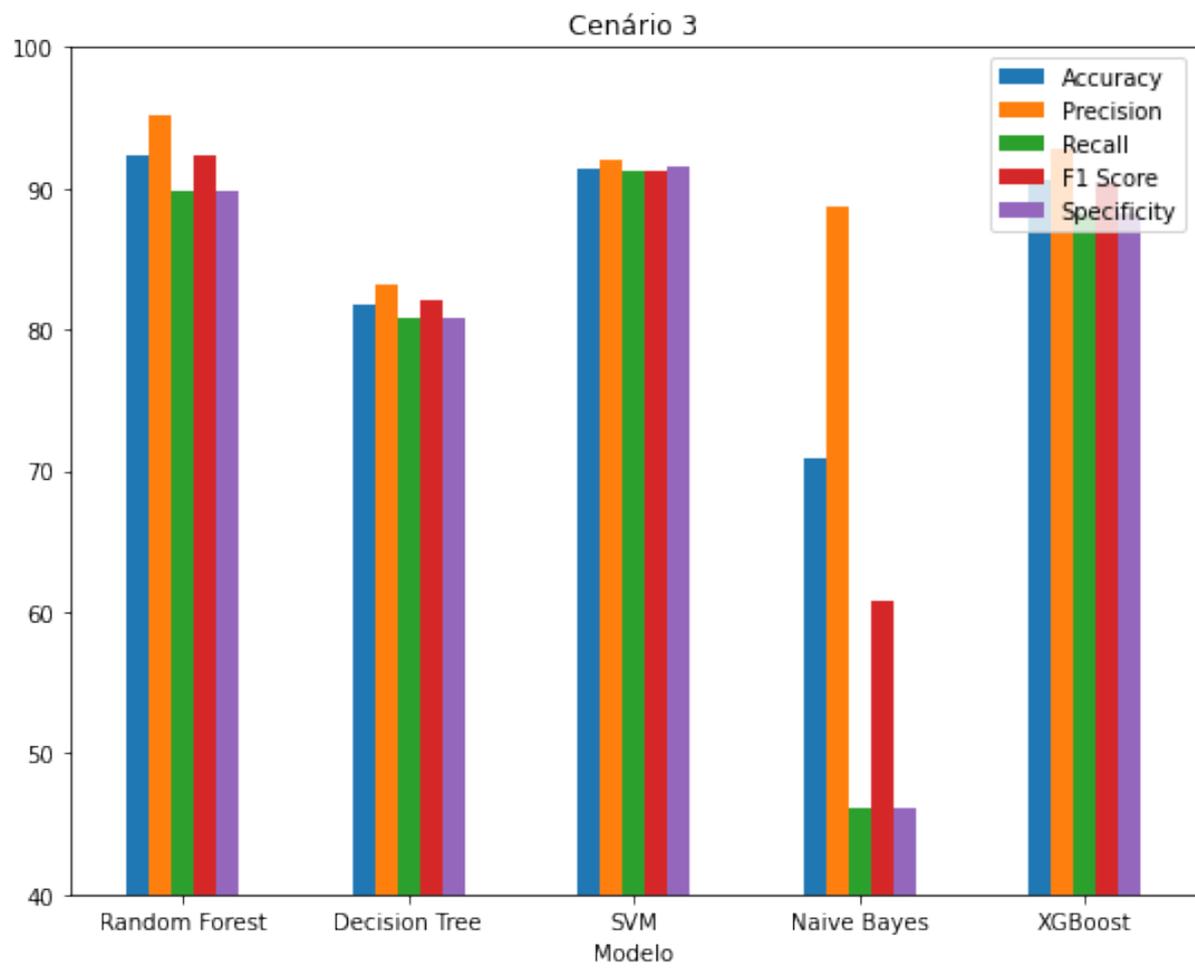
Fonte: Feito pelo autor.

Figura 8 – Cenário 2 com Word2Vec



Fonte: Feito pelo autor.

Figura 9 – Cenário 3



Fonte: Feito pelo autor.

REFERÊNCIAS

- ABONIZIO, H. Q.; MORAIS, J. I. de; TAVARES, G. M.; JUNIOR, S. B. Language-independent fake news detection: English, portuguese, and spanish mutual features. *Future Internet*, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 87, 2020.
- AGGARWAL, C. C.; ZHAI, C. A survey of text classification algorithms. In: *Mining text data*. [S.l.]: Springer, 2012. p. 163–222.
- AHMED, A. A. A.; ALJABOUH, A.; DONEPUDI, P. K.; CHOI, M. S. Detecting fake news using machine learning: A systematic literature review. *arXiv preprint arXiv:2102.04458*, 2021.
- ALMEIDA, F.; XEXÉO, G. Word embeddings: A survey. *arXiv preprint arXiv:1901.09069*, 2019.
- APHIWONGSOPHON, S.; CHONGSTITVATANA, P. Detecting fake news with machine learning method. In: IEEE. *2018 15th international conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON)*. [S.l.], 2018. p. 528–531.
- BURKHARDT, J. M. History of fake news. *Library Technology Reports*, v. 53, n. 8, p. 5–9, 2017.
- CERON, W.; LIMA-SANTOS, M.-F. de; QUILES, M. G. Fake news agenda in the era of covid-19: Identifying trends through fact-checking content. *Online Social Networks and Media*, Elsevier, v. 21, p. 100116, 2021.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2016. p. 785–794.
- CHORAŚ, M.; DEMESTICHAS, K.; GIEŁCZYK, A.; HERRERO, Á.; KSIENIEWICZ, P.; REMOUNDOU, K.; URDA, D.; WOŹNIAK, M. Advanced machine learning techniques for fake news (online disinformation) detection: a systematic mapping study. *Applied Soft Computing*, Elsevier, v. 101, p. 107050, 2021.
- CHURCH, K. W. Word2vec. *Natural Language Engineering*, Cambridge University Press, v. 23, n. 1, p. 155–162, 2017.
- FREEMAN, D.; WAITE, F.; ROSEBROCK, L.; PETIT, A.; CAUSIER, C.; EAST, A.; JENNER, L.; TEALE, A.-L.; CARR, L.; MULHALL, S. et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in england. *Psychological medicine*, Cambridge University Press, v. 52, n. 2, p. 251–263, 2022.
- GALHARDI, C. P.; FREIRE, N. P.; MINAYO, M. C. d. S.; FAGUNDES, M. C. M. Fact or fake? an analysis of disinformation regarding the covid-19 pandemic in brazil. *Ciencia & saude coletiva*, SciELO Public Health, v. 25, p. 4201–4210, 2020.
- GELERIS, J.; SUN, Y.; PLATT, J.; ZUCKER, J.; BALDWIN, M.; HRIPCSAK, G.; LABELLA, A.; MANSON, D. K.; KUBIN, C.; BARR, R. G. et al. Observational study of hydroxychloroquine in hospitalized patients with covid-19. *New England Journal of Medicine*, Mass Medical Soc, v. 382, n. 25, p. 2411–2418, 2020.

- GELFERT, A. Fake news: A definition. *Informal Logic*, Informal Logic, v. 38, n. 1, p. 84–117, 2018.
- HAKIM, A. A.; ERWIN, A.; ENG, K. I.; GALINIUM, M.; MULIADY, W. Automated document classification for news article in bahasa indonesia based on term frequency inverse document frequency (tf-idf) approach. In: IEEE. *2014 6th international conference on information technology and electrical engineering (ICITEE)*. [S.l.], 2014. p. 1–4.
- JACOBS, T.; TSCHÖTSCHHEL, R. Topic models meet discourse analysis: a quantitative tool for a qualitative approach. *International Journal of Social Research Methodology*, Taylor & Francis, v. 22, n. 5, p. 469–485, 2019.
- KADHIM, A. I. Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, Springer, v. 52, n. 1, p. 273–292, 2019.
- KORDE, V.; MAHENDER, C. N. Text classification and classifiers: A survey. *International Journal of Artificial Intelligence & Applications*, v. 3, n. 2, p. 85, 2012.
- KOTSIANTIS, S. B.; ZAHARAKIS, I.; PINTELAS, P. et al. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, Amsterdam, v. 160, n. 1, p. 3–24, 2007.
- LIASHCHYNSKYI, P.; LIASHCHYNSKYI, P. Grid search, random search, genetic algorithm: a big comparison for nas. *arXiv preprint arXiv:1912.06059*, 2019.
- LINDEN, S. van D.; ROOZENBEEK, J.; COMPTON, J. Inoculating against fake news about covid-19. *Frontiers in psychology*, Frontiers, v. 11, p. 2928, 2020.
- LIU, M.; YANG, J. An improvement of tfidf weighting in text categorization. *International proceedings of computer science and information technology*, v. 47, p. 44–47, 2012.
- LIU, Y.; WANG, Y.; ZHANG, J. New machine learning algorithm: Random forest. In: SPRINGER. *International Conference on Information Computing and Applications*. [S.l.], 2012. p. 246–252.
- LOWD, D.; DOMINGOS, P. Naive bayes models for probability estimation. In: *Proceedings of the 22nd international conference on Machine learning*. [S.l.: s.n.], 2005. p. 529–536.
- MACKAY, D. J. Hyperparameters: optimize, or integrate out? In: *Maximum entropy and bayesian methods*. [S.l.]: Springer, 1996. p. 43–59.
- MANZOOR, S. I.; SINGLA, J. et al. Fake news detection using machine learning approaches: A systematic review. In: IEEE. *2019 3rd international conference on trends in electronics and informatics (ICOEI)*. [S.l.], 2019. p. 230–234.
- MARTINS, A. D. F.; CABRAL, L.; MOURÃO, P. J. C.; MONTEIRO, J. M.; MACHADO, J. Detection of misinformation about covid-19 in brazilian portuguese whatsapp messages. In: SPRINGER. *International Conference on Applications of Natural Language to Information Systems*. [S.l.], 2021. p. 199–206.
- MERRIAM-WEBSTER-DICTIONARY. *The Real Story of 'Fake News'*. 2018. Disponível em: <<https://www.merriam-webster.com/words-at-play/the-real-story-of-fake-news>>.

- MIKOLOV, T.; CHEN, K.; CORRADO, G.; DEAN, J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- MONTEIRO, R. O.; NOGUEIRA, R. R. Development of a system for automatic classification of fake news in portuguese language. 2019.
- ORENGO, V. M.; HUYCK, C. R. A stemming algorithm for the portuguese language. In: *spire*. [S.l.: s.n.], 2001. v. 8, p. 186–193.
- PAIXAO, M.; LIMA, R.; ESPINASSE, B. Fake news classification and topic modeling in brazilian portuguese. In: IEEE. *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. [S.l.], 2020. p. 427–432.
- PATWA, P.; SHARMA, S.; PYKL, S.; GUPHA, V.; KUMARI, G.; AKHTAR, M. S.; EKBAL, A.; DAS, A.; CHAKRABORTY, T. Fighting an infodemic: Covid-19 fake news dataset. In: SPRINGER. *International Workshop on Combating On line Ho st ile Posts in Regional Languages dur ing Emerge ncy Si tuation*. [S.l.], 2021. p. 21–29.
- QUANDT, T.; FRISCHLICH, L.; BOBERG, S.; SCHATTO-ECKRODT, T. Fake news. *The international encyclopedia of journalism studies*, John Wiley & Sons, Inc. Hoboken, NJ, USA, p. 1–6, 2019.
- RAUT, P. P.; BORKAR, N. R.; STUDENT, M.; PROFESSOR, A.; KAMLATAI, S. Machine learning algorithms: Trends, perspectives and prospects. *International Journal of Engineering Science*, v. 4884, 2017.
- SARICA, S.; LUO, J. Stopwords in technical language processing. *Plos one*, Public Library of Science San Francisco, CA USA, v. 16, n. 8, p. e0254937, 2021.
- SHU, K.; SLIVA, A.; WANG, S.; TANG, J.; LIU, H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, ACM New York, NY, USA, v. 19, n. 1, p. 22–36, 2017.
- SOUZA, J. V. de; JR, J. G.; FILHO, F. M. de S.; JULIO, A. M. de O.; SOUZA, J. F. de. A systematic mapping on automatic classification of fake news in social media. *Social Network Analysis and Mining*, Springer, v. 10, n. 1, p. 1–21, 2020.
- SYN, S. Y.; OH, S. Why do social network site users share information on facebook and twitter? *Journal of Information Science*, SAGE Publications Sage UK: London, England, v. 41, n. 5, p. 553–569, 2015.
- UNA-SUS. *Coronavírus: Brasil confirma primeiro caso da doença*. 2020. Disponível em: <<https://www.unasus.gov.br/noticia/coronavirus-brasil-confirma-primeiro-caso-da-doenca>>.
- WHO. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. 2020. Disponível em: <<https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>>.
- YANG, J.; TIAN, Y. “others are more vulnerable to fake news than i am”: Third-person effect of covid-19 fake news on social media users. *Computers in Human Behavior*, Elsevier, v. 125, p. 106950, 2021.
- ZHANG, X.; GHORBANI, A. A. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, Elsevier, v. 57, n. 2, p. 102025, 2020.

ZHANG, X.-D. Machine learning. In: *A Matrix Algebra Approach to Artificial Intelligence*. [S.l.]: Springer, 2020. p. 223–440.

ZHOU, X.; ZAFARANI, R. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 53, n. 5, p. 1–40, 2020.