

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

IVAN TAKESHI OSHITA

CLASSIFICAÇÃO DE FAKE NEWS POR MINERAÇÃO DE TEXTO

LONDRINA

2021

IVAN TAKESHI OSHITA

CLASSIFICAÇÃO DE FAKE NEWS POR MINERAÇÃO DE TEXTO

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Engenharia de Produção, do Departamento de Engenharia de Produção, da Universidade Tecnológica Federal do Paraná.

Orientador: Profa. Dr. Bruno Samways dos Santos

LONDRINA

2021

IVAN TAKESHI OSHITA

CLASSIFICAÇÃO DE FAKE NEWS POR MINERAÇÃO DE TEXTO

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia de Produção da Universidade
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 20/agosto/2021

Bruno Samways dos Santos
Doutor
Universidade Tecnológica Federal do Paraná

Rafael Henrique Palma Lima
Doutor
Universidade Tecnológica Federal do Paraná

Rogério Tondato
Doutor
Universidade Tecnológica Federal do Paraná

LONDRINA

2021

AGRADECIMENTOS

Agradeço, primeiramente ao criador de tudo que é (Deus), por todas as experiências sagradas vividas, pelas oportunidades e tornar tudo possível em minha vida.

Agradeço aos meus pais por sempre me apoiar, incentivar, possibilitar ao acesso aos conhecimentos e por acreditar em meus sonhos e objetivos. Todo esse apoio que me foi dado, se tornou em força para vencer os obstáculos enfrentados durante toda minha trajetória.

Um agradecimento especial ao Professor Bruno Samways dos Santos pelos ensinamentos, pela dedicação, orientação, paciência e pela compreensão durante todo desenvolvimento deste trabalho.

E por fim gostaria de agradecer aos meus amigos e familiares pelo incentivo e apoio, que também tornaram tudo isso possível.

RESUMO

As notícias falsas (conhecidas por *fake news*) estão sendo intensificadas nas mídias sociais e fóruns online, com o papel fundamental de propagar a desinformação, influenciar e distorcer a realidade. A desinformação está cada vez mais presente em uma sociedade devido à rapidez das transmissões de falsas alegações nas redes sociais, falta de averiguação dos fatos e o aumento do acesso da população às mídias sociais. Além disso, as *fake news* têm chamado muita atenção do mundo em função de seus impactos negativos na política, economia ou até mesmo pessoais. Portanto, este estudo teve como objetivo classificar notícias em verdadeiras ou em *fake news* por meio de técnicas de mineração de textos. Para classificação foi utilizada a ferramenta de mineração de dados WEKA (*Waikato Environment for Knowledge Analysis*) para aplicação da técnica de árvore de decisão e os recursos *Bag-of-Words* (BOW) e *Term Frequency-Inverse Document Frequency* (TF-IDF). O experimento foi avaliado pelas métricas de acurácia, precisão, sensibilidade (ou *recall*) e *f-score*, e os resultados obtidos mostraram-se promissores. Os modelos foram validados com o conjunto *Fake.br Corpus*, o qual possui 7.200 textos de *websites* que já estão pré-processados no idioma português. O resultado deste experimento mostrou que o modelo TF-IDF obteve o melhor desempenho em relação ao BOW, com uma acurácia de 89,82%.

Palavras-chave: *Fake News*. KDT. Mineração de Texto. Classificação. Árvore de Decisão.

ABSTRACT

False news (known as fakes news) is being intensified in social media and online forums, with the fundamental role of spreading misinformation, influencing, and distorting reality. Disinformation is increasingly present in a society due to the rapid transmission of false allegations on social networks, lack of fact-finding and the population's increased access to social media. In addition, fake news has drawn a lot of attention in the world due to its negative impacts on politics, economy, or even personal ones. Therefore, this study aimed to classify news as true or fake news through text mining techniques. For classification, the data mining tool WEKA (Waikato Environment for Knowledge Analysis) was used to apply the decision tree technique and the resources Bag-of-Words (BOW) and Term Frequency-Inverse Document Frequency (TF-IDF). The experiment was evaluated by accuracy, precision, sensitivity (or recall) and f-measure metrics, and the results sought to be promising. The models were validated with the Fake.br Corpus set, which has 7.200 websites texts that are already pre-processed in Portuguese. The result of this experiment showed that the TF-IDF model obtained the best performance in relation to the BOW, with an accuracy of 89.82%.

Keywords: *Fake News. KDT. Text Mining. Classification. Decision tree.*

LISTA DE ILUSTRAÇÕES

Figura 1 - Os Estágios de Análise no PLN	22
Figura 2 - Representação de um modelo de <i>Bag-of-words</i>	24
Figura 3 -Representação de <i>Bag-of-Words</i> base baseados nas sentenças...24	
Figura 4 - Representa uma árvore de decisão e seus componentes	27
Figura 5 - Fluxograma das Etapas do Processo.....	33
Figura 6 - Média das métricas obtidas.....	36

LISTA DE TABELAS

Tabela 1 - Resultados do Experimento.....	35
Tabela 2 – Matriz de Confusão.....	37
Tabela 3 - Termos relevantes.....	37

LISTA DE ABREVIATURAS

AD	Árvores de Decisão
BoW	<i>Bag-of-Words</i>
CVPS	<i>Cross Validation Parameter Selection</i>
DCD	<i>Document-Class Distance</i>
DM	<i>Data Mining</i>
MIT	<i>Massachusetts Institute of Technology</i>
FEBRABAN	Federação Brasileira de Bancos
FEND	<i>Fake News Detection</i>
GWO	<i>Grey Wolf Optimization</i>
KDD	<i>Knowledge Discovery in Databases</i>
KDT	<i>Knowledge Discovery from Text</i>
LWL	<i>Locally Weighted Learning</i>
MD	Mineração de Dados
MT	Mineração de Textos
NB	<i>Naive Bayes</i>
PNL	Processamento de Linguagem Natural
POS	<i>Part of Speech</i>
RF	<i>Random Forest</i>
RFC	<i>Randomized Filtered Classifier</i>
RNC	<i>Redes Neurais Convolucionais</i>
RNN	<i>Recurrent Neural Networks</i>
RNR	Redes Neurais Recorrentes
SECOM	Secretaria Especial de Comunicação
SSO	<i>Salp Swarm Optimization</i>

SVM	<i>Support Vector Machine</i>
TF	<i>Term Frequency</i>
TF-IDF	<i>Term Frequency - Inverse Document Frequency</i>
WIHW	<i>Weighted Instances Handler Wrapper</i>

Sumário

1. INTRODUÇÃO.....	12
1.1. Objetivo Geral.....	14
1.2. Objetivos Específicos	14
1.3. Justificativa	14
1.4. Estrutura do Trabalho.....	17
2. REFERENCIAL TEÓRICO	18
2.1. Fake News	18
2.2. Descoberta de Conhecimento em Bases de Dados – KDD	19
2.3. Mineração de Texto.....	20
2.4. Processamento de Linguagem Natural (PLN).....	22
2.5. Tokenização	23
2.6. Bag-of-Words (Bow).....	23
2.7. Term Frequency - Inverse Document Frequency (TF-IDF)	25
2.8. Árvores de Decisão	26
2.9. Métricas de Avaliação	27
2.10. Trabalhos Correlatos.....	29
3. MÉTODOS E TÉCNICAS	33
3.1. Descrição Do Conjunto De Dados.....	33
4. RESULTADOS E DISCUSSÃO	35
5. CONCLUSÃO.....	39
6. REFERÊNCIAS BIBLIOGRÁFICAS	40

1. INTRODUÇÃO

A exposição de notícias e informações ocorre cada vez mais por meio das redes sociais, e isso tem impactado na veracidade das informações contidas nesse meio de comunicação devido ao surgimento de notícias falsas. Popularmente essas notícias são chamadas de *fake news*, que se referem às desinformações que são disseminadas no mundo virtual (LIMA *et al.* 2020).

As mídias sociais conseguem propagar falsas alegações rapidamente (VOSOUGHI *et al.* 2018) e aumentar a exposição seletiva a notícias e informações consistentes em atitudes (BAKSHY *et al.* 2015), sendo esse padrão intensificado pelos algoritmos de informações (HELMSLEY, 2018). Diante deste cenário as redes sociais se tornaram um veículo ideal para transmissão das *fake news*, isto ocorre devido à velocidade de alcance das mídias sociais e a falta de averiguação dos fatos (SWIRE, ECKER, 2018).

De acordo com a pesquisa realizada por *Pew Research Center* (2017) 67% da população norte-americana recebem notícias nas mídias sociais, sendo que estas notícias desempenham comportamento variável nas redes sociais. O Facebook lidera os sites de mídias sociais, isso se dá devido aos 66% da população americana que utilizam a rede social e a maioria dos seus usuários recebem notícias pelo site, o que representa 45% da população adulta. O Youtube está em segundo lugar de maior alcance em termos de uso geral, com uma base de usuários que representam 58% da população norte-americana, porém uma parcela de 18% que recebem notícias, que é inverso do Twitter em que 74% dos seus usuários recebem notícias sendo que sua audiência é significativamente menor no geral. Visto que, no geral, menos americanos recebem notícias no Twitter (11% dos adultos no EUA).

O cenário brasileiro não é muito diferente do norte-americano, pois uma pesquisa realizada em 2016 pela Secretaria Especial de Comunicação (SECOM) evidencia que 49% da população se informam pela internet. O relatório Digital in 2019 publicado em janeiro de 2019, por *We Are Social* e *Hootsuite* mostra que 81% dos brasileiros estão ativos nas redes sociais, considerando a população com 13 anos ou mais, a penetração das redes sociais é de 66%. O Brasil ocupa o segundo lugar em termos de horas gastas em plataformas de mídia social, consumindo em média 3 horas e 34 minutos por dia. O Facebook possui 130 milhões de usuários brasileiros,

sendo que as postagens nesta página têm uma taxa de engajamento de 4,22% no Brasil, acima da média mundial que equivale 3,75%.

Perante este contexto brasileiro as redes sociais têm se tornado cada vez mais um meio importante para os cidadãos compartilharem notícias e busca de informação. Os tipos de notícias e informações que os usuários estão expostos nas redes sociais diferem dos noticiários motivados emocionalmente e polarizador de conteúdo (MACHADO *et al.* 2018). O Brasil vem enfrentando uma crise da desinformação devido às crescentes fontes de propagação de informações enganosas disseminadas por contas ou falsos, que muitas vezes imitam as notícias e utilizam a desinformação para chamar atenção do espectador (GARCIA, 2020; MACHADO, KIRA 2018; HOWARD, 2018).

As divulgações das *fake news* sucedem em épocas significativas do interesse público. Os algoritmos das mídias sociais são capazes de impulsionar e espalhar material conspiratório sobre informações precisas (HOWARD, WOOLEY, CALO, 2018). Essas manipulações de informações vêm gerando uma onda de preocupações aos órgãos públicos brasileiros, pois criam uma falsa percepção do âmbito político a certa proposta, ideia ou figura pública, afetando o mercado de ações, propagando rumores, notícias falsas e teorias conspiratórias, causando as desinformações. Além disso, atrai os usuários para *links* maliciosos que roubam dados pessoais e entre outros riscos (RUEDIGER, 2018, p.9).

Segundo uma pesquisa realizada pelo Instituto de Tecnologia de Massachusetts (MIT), às *fake news* se difundem significativamente mais longe e mais rápidas do que as notícias verdadeiras, em diversas categorias e em muitos casos por uma ordem de magnitude. As notícias falsas têm 70% mais chances de serem divulgadas e notícias verdadeiras levam cerca de seis vezes mais tempo para atingir 1.500 pessoas quando comparado a notícias falsas. Entretanto as disseminações das *fake news* não ocorrem principalmente por *bots*, robôs que são programados para divulgar histórias imprecisas. Visto que as notícias falsas tendem a se espalhar rapidamente por pessoas reais.

O KDD é uma metodologia composta por diversas etapas, que podem ser classificadas em grupos: pela seleção de dados, pré-processamento, transformação dos dados e estabelecimento de padrões úteis na extração de conhecimento (VIANA *et al.*, 2010). Em algumas pesquisas da área da computação, esta metodologia é bastante utilizada para descoberta de identificação de padrões válidos e não triviais

(VIANA *et al.*, 2010; CARVALHO *et al.*, 2012), o qual permite encontrar vínculos entre os dados armazenados mais facilmente do que com as técnicas tradicionalmente utilizadas (CARVALHO *et al.*, 2012).

Uma das etapas que está no centro do processo é a Mineração de Dados (*Data Mining* - DM), que consiste na interpretação de padrões e na criação de conhecimento após a verificação dos resultados obtidos (FAYYAD *et al.*, 1996). A mineração de dados apresenta quatro tarefas principais: classificação, regressão, descoberta de regras de associação e agrupamento, que podem ser utilizadas em diversas situações que dependem do objetivo da aplicação. O presente estudo tem o intuito de detectar as *fake news* nos *sites* brasileiro. Para tal, será aplicada a técnica de Árvores de decisão.

1.1. Objetivo Geral

O objetivo deste trabalho é aplicar a técnica de Mineração de Textos (MT) para classificar as *fake news* partir da ferramenta de Árvore de Decisão (AD).

1.2. Objetivos Específicos

- Pré-processar dados não estruturados utilizando a abordagem BOW.
- Transformar e selecionar os atributos por meio do TF-IDF.
- Aplicar a técnica de classificação (árvores de decisão) no ambiente WEKA.
- Comparar os resultados obtidos por meio da acurácia, sensibilidade e precisão.

1.3. Justificativa

O ano de 2020 foi propício para circulação de notícias falsas e desinformação sobre os acontecimentos no cenário brasileiro devido às inúmeras notícias falsas sobre pandemia do Covid-19. Além disso, as eleições para eleger um novo prefeito para governar os municípios brasileiros em 2020, o que intensifica a manipulação de informações e rumores sobre o âmbito político.

O governo brasileiro está atento a essas ondas de notícias e tomando medidas para orientar a população sobre as *fake news*, criando canais de comunicação e divulgando propagandas na TV para combater as propagações das notícias falsas que são promovidas através das redes sociais. Segundo o Ministério da Saúde (2020), a saúde é o tema mais abordado pelas *fake news*, exemplo disso é a circulação de notícias de alimentos que curam ou previnem o coronavírus, sendo que não existe nenhuma base científica que garanta a cura ou prevenção do Covid-19. Para combater as *fake news* sobre a saúde, o Ministério da Saúde criou o canal “SAÚDE SEM FAKE NEWS” de recebimento e envio de mensagens instantâneas, via WhatsApp. Por meio deste canal, o cidadão pode enviar uma mensagem com imagens ou textos que tenha recebido nas redes sociais para averiguar se a informação é verdadeira ou falsa.

De acordo com o relatório emitido pela CPMI das *Fake News*, o Governo Federal exibiu 653.378 vezes anúncios em 47 canais de notícias falsas, nas quais foram detectadas dois milhões de propagações de anúncios do governo em canais de conteúdo inadequado (REZENDE, 2020). Este número representa 4,37% das 47.188.047 divulgações de anúncios analisados, que é o resultado para contar quantas vezes um anúncio foi espalhado em campanhas patrocinadas pelo governo. A investigação engloba canais que mais veicularam inserções da campanha “Nova Previdência” do governo (REZENDE, 2020).

A Federação Brasileira de Bancos (Febraban) divulgou um observatório em setembro de 2020, o qual mostra que os internautas estão desconfiados em relação aos conteúdos divulgados na internet. Embora hoje existem muitos avanços da tecnologia em relação à segurança de dados e de informações, grande parte dos usuários da internet mostra-se descontente com o baixo desempenho do poder público no Brasil em relação ao combate da *fake news*, sendo que 66% dos entrevistados acreditam que a autoridades não estão tomando ações eficientes (FEBRABAN, 2020). Ainda segundo a Febraban, 86% dos usuários demonstram preocupações em relação às notícias falsas. Ainda que haja esta desconfiança em relação às informações recebidas pelas redes sociais, uma pesquisa realizada pelo DataSenado 2019 aponta que 79% dos usuários leem informações e veem noticiários pelo *WhatsApp* na maior parte do tempo, uma quantia que supera outros meios de informação.

O Senado brasileiro aprovou em junho de 2020 o projeto de lei de combate às notícias falsas, a qual se chama Lei Brasileira de Liberdade, Responsabilidade e Transparência na internet, como normas para redes sociais e serviços de mensagem como *WhatsApp* e *Telegram*. A intenção é evitar a propagação de notícias falsas que possam causar danos individuais ou coletivos e à democracia.

Segundo Garcia (2020) o governo deveria investir em educação e responsabilizar os financiadores das redes de notícias falsas, pois a lei promulgada é a pior forma de combater o problema por ser uma das leis de internet que visa a maior restrição de liberdade no mundo dos usuários. Em 2014, o Brasil foi precursor dos direitos digitais e, quando aprovou o Marco Civil da Internet, criou uma ampla garantia de liberdade de expressão online. Contudo, a Lei de notícias falsas circundaria a estrutura, resultando na restrição de liberdade de expressão de milhares de brasileiros.

As *fakes news* tornaram-se o centro de atenção e preocupação da sociedade em geral devido aos seus aspectos negativos que geram impactos na economia, no âmbito político e área de informações, sendo assim diversos estudos vem se desenvolvendo por meio de ferramentas e técnicas computacionais que são eficaz na identificação e na classificação automática de notícias falsas ou manipulações de informações. Uma das abordagens utilizadas para a análise de informações contidas no banco de dados gerados nas redes sociais é a Mineração de Texto abordagem na qual se extrai um grande volume de informações de bases textuais de modo claro.

1.4. Estrutura do Trabalho

O presente trabalho é composto por cinco sessões, sendo que a primeira sessão é constituída pela introdução, objetivo geral, objetivos específicos e a justificativa do trabalho. Na segunda seção foi desenvolvido o referencial teórico, a qual é abordado conceitos de *fake news*, dos métodos de KDT para dados não estruturados utilizados no trabalho e os trabalhos correlacionados ao tema.

A terceira sessão descreve o conjunto de dados e é explicado como o conjunto de dados foi desenvolvido pelos seus criadores. E também na mesma sessão é apresentado os métodos e técnicas aplicadas neste trabalho, descrevendo o passo a passo de como foi elaborada a pesquisa, desde a informação do conjunto de dados até a técnica aplicada.

Na quarta sessão são apresentados os resultados dos experimentos e também é feita uma análise dos resultados obtidos. E, por último, a quinta sessão é composta pela conclusão e as considerações finais, além de sugestões de exploração de outras técnicas.

2. REFERENCIAL TEÓRICO

O presente capítulo irá apresentar os conceitos sobre fake news, KDT e as técnicas utilizadas, bem como os trabalhos correlatos que nortearam os métodos adotados neste trabalho.

2.1. Fake News

A palavra *Fake News* começou a ser comumente utilizada após as eleições presidenciais de 2016 nos Estados Unidos da América. No Brasil, a palavra se instituiu em momentos-chave do cenário político brasileiro, tais como: a aprovação da Reforma Trabalhista, a greve geral de 2017, as eleições de 2014 e o debate sobre o Impeachment da então Presidenta Dilma Rousseff (RUEDIGER, 2017).

O termo *fake news* não é novo. Tornou-se uma palavra popular mundialmente, mas as referências atuais definem de forma diferente das definições anteriores. No presente, especificamente a mídia, assemelham as notícias falsas como postagens virais baseadas em relatos fictícios feitos para parecer como notícias (TANDOC et al., 2018). Segundo Allcott e Gentzkow (2017) a expressão *fake news* é definida como artigos de notícias que são intencionalmente e comprovadamente falsos a fim de manipular e enganar os leitores.

De acordo com Abreu (2019) as *Fake News* simulam as informações reais com o intuito de ocasionar polêmicas, ou difamar a imagem de uma pessoa ou empresa. Isto ocorre devido ao seu potencial de viralização na internet, o que costuma chamar bastante a atenção em massa das pessoas. Desta forma, as notícias falsas ganham grande força de persuasão sobre a população.

As *Fake News* são disseminadas por meios digitais, através de redes sociais como Facebook, Youtube, Twitter, Whatsapp e sites ou blogs que utilizam nomes semelhantes àqueles que são autorais. Além destes meios mencionados acima, a disseminação de *Fake News* pode ser feita por *bots*, o qual tem o objetivo à monetização mediante a publicidade e propaganda, obtidas por visitas do usuário a determinados sites (SHAO et al,2017).

Segundo Markines *et al.* (2009), para atrair os usuários a estas notícias falsas são criados textos atraentes que possuem palavras-chave populares e que a informação pareça genuína mesmo que os seus conteúdos não façam sentido. Os programas de *bots* procuram por palavras de maior relevância, os nomes de domínios

baseados nessas palavras e geram contas de hospedagem para esses domínios. Depois de concluída essa etapa, são criados os sites com notícias falsas integrados com palavras-chave e uma série de anúncios publicitários e finalizam carregando os sites para as contas de hospedagem nos seus respectivos domínios, e são referenciadas com links, para impulsionar a chance de *sites* serem buscados no Google, através do Algoritmo de PageRank (MARUMO, 2018).

2.2. Descoberta de Conhecimento em Bases de Dados – KDD

Descoberta de Conhecimentos em Bases de Dados (*Knowledge Discovery in Databases* — KDD) é uma metodologia que utiliza métodos e ferramentas que visam transformar os dados armazenados em dados de conhecimento, ou seja, tange a todo processo de descoberta de conhecimento útil de dados (Lemos et al., 2005).

Segundo Fayyad *et al.* (1996) o KDD é um processo não trivial que objetiva encontrar novos padrões compreensíveis que possam ser interpretados como conhecimento útil, buscando tornar os dados de baixo nível em conhecimento de alto nível.

O KDD é composto por um conjunto de atividades contínuas que compartilham o conhecimento descoberto através de bases de dados. Esse conjunto é constituído por cinco etapas: seleção dos dados; pré-processamento e limpeza dos dados; transformação dos dados; mineração de dados; interpretação e avaliação dos resultados (FAYYAD *et al.*, 1996).

O KDD inicia-se com a compreensão do domínio das aplicações e do conhecimento inicial, e define os objetivos finais a serem obtidos. Posteriormente, é criado um conjunto de dados selecionados com o intuito de averiguação dos mesmos.

A seguir, a etapa de limpeza dos dados (*Data Cleaning*) é constituída por um processamento de dados que visa a apropriar-se dos algoritmos, através da incorporação de dados heterogêneos, com remoção de dados incompletos e outros procedimentos. Os dados pré-processados passam por uma modificação para o seu armazenamento apropriado, simplificando para o uso das técnicas de mineração de dados.

A mineração de dados (MD) faz parte do processo do KDD, que consiste na execução de análises de dados e na aplicação de algoritmos de descoberta que produzem um conjunto de padrões, ou seja, baseia-se na interpretação destes

padrões e da concepção de conhecimentos após a análise dos resultados obtidos (FAYYAD *et al.*, 1996).

Segundo Berry e Linoff (1997) a “mineração de dados é a exploração e a análise, por meio automático ou semi-automático, de grandes quantidades de dados, a fim de descobrir padrões e regras significativas”. A MD tem como seu principal objetivo descobrir relacionamentos entre dados e fornecer recursos que sejam capazes de fazer uma previsão de tendências futuras baseadas no passado (DIAS, 2002). Os resultados atingidos com a MD podem ser utilizados no gerenciamento de informação, processamento de pedidos de informação, tomada de decisão, controle de processo e muitas outras aplicações (DIAS, 2002).

Seguindo no processo, a fase de MD inicia-se com a escolha das técnicas e ferramentas a serem utilizadas, que dependem essencialmente do objetivo do processo de KDD: classificação, estimação, agrupamento ou associação. De modo geral, o processo utiliza técnicas especializadas que buscam padrões nos dados (WITTEN; FRANK, 2005).

Dentre as técnicas utilizadas na mineração de dados, este estudo aplicou as Árvores de Decisão (AD), que são utilizadas para a tarefa de classificação de dados, e também usadas em conjunto com a abordagem de indução de regras, porém são as únicas que contêm resultados de forma hierárquica (com priorização) (LEMOS *et al.*, 2005), auxiliando na sua interpretação.

2.3. Mineração de Texto

A Mineração de Textos (MT) é um processo de descoberta de conhecimento a partir de textos (*Knowledge Discovery from Text – KDT*), e identifica automaticamente padrões relevantes em dados textuais (Feldman; Dagan, 1995). Segundo Tan (1999) a MT é definida como um processo de extração de conhecimentos e padrões não triviais de documentos textuais não estruturados.

A MT é similar à MD, com diferença das ferramentas de MD (WAEGEL, 2006) que são designados para lidar com dados estruturados de bancos de dados, enquanto que a MT também pode executar os dados não estruturados ou conjuntos de dados semiestruturados, como e-mails, documentos de texto e arquivos HTML, etc. (KUMAR; BHATIA, 2013). Além disso, a MT é utilizada para analisar as tendências de pesquisa em diversas áreas de estudo, como sistemas de informação, gestão de

tecnologia, educação, biblioteca e ciência da informação, psicologia, sociologia, entre outros (JUNG; LEE, 2020).

Segundo CHATTERJEE *et al.* (2020) os dados estruturados são mais compreensíveis e úteis, porém ao trabalhar com dados não estruturados é possível obter mais informações em consequência das análises por meio da combinação de técnicas qualitativas e quantitativas. O único problema de MT é a extração de conteúdos significativos, útil e preliminarmente informações desconhecidas de dados textuais, pois quando comparado com os dados armazenados em bancos de dados, o texto não estruturado apresenta ambiguidade e o que dificulta processar as informações textuais (KUMAR; BHATIA, 2013).

No sistema de aprendizagem as representações de dados influenciam na precisão dos resultados, especialmente em MT, pois os problemas de análise de texto precisam ser transformados em uma representação apropriada para que o método seja aplicado (OZBAY; ALATAS, 2019). Portanto, na aplicação dos métodos de processamento na MT, é importante que os dados de textos não estruturados sejam limpos (CHATTERJEE *et al.*, 2020), ou seja, o processo de análise de MT inicia-se com o pré-processamento dos dados de texto.

O método de pré-processamento exerce uma função muito importante nas técnicas e aplicações de MT, sendo a primeira etapa do processo (VIJAYARANI; ILAMATHI; NITHYA, 2015). Nessa fase ocorre uma análise morfológica a qual classifica as frases em classe gramaticais. É nesta etapa que são determinadas as principais palavras, as quais são extraídas com base em tópicos e palavras-chave que aparecem no mesmo parágrafo ou sentenças (JUNG; LEE, 2020). Desta forma as palavras que não agregam significado (pronomes, preposições, etc.) são removidas devido à necessidade de reduzir o tamanho dos dados de texto e auxilia na eficiência do processamento (CHATTERJEE *et al.*, 2020).

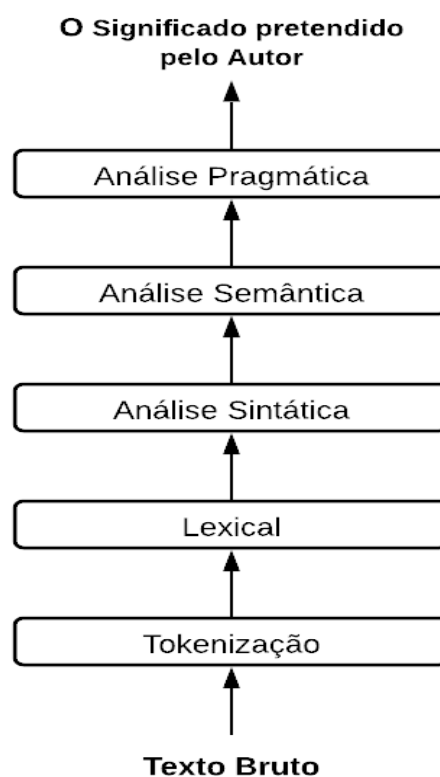
As seções a seguir discutem as principais etapas do método de pré-processamento, tais como: Tokenização; Processamento de Linguagem Natural (PNL); *Bag-of-Words* (BOW) e *Term Frequency - Inverse Document Frequency* (TF-IDF).

2.4. Processamento de Linguagem Natural (PLN)

O Processamento de Linguagem Natural (PLN) é uma técnica aplicada nos métodos de MT. A técnica de PLN é relacionada à interação homem-máquina (KAUR; MADA, 2015), visando analisar e representar textos de ocorrências naturais em vários níveis de análise linguística com a finalidade de alcançar o processamento de linguagem semelhante aos humanos para diversas aplicações e tarefas (LIDDY, 2001).

Segundo Medeiros (2010) técnicas de PLN são utilizadas para aperfeiçoar os resultados das etapas de MT, por meio da modelagem de textos, considerando as funções gramaticais e sintáticas de palavras, estruturas das frases, parágrafos e textos inteiros. De acordo com Indurkha e Damerau (2010) geralmente as técnicas de PLN se iniciam através do processo de separação de palavras, denominado de “tokenização” a etapa de pré-processamento, seguido de análises léxicas, sintáticas e semânticas.

Figura 1 - Os Estágios de Análise no PLN



Fonte: Indurkha e Damerau (2010, tradução livre)

2.5. Tokenização

A técnica de tokenização é comum em modelagem de textos com PLN, sendo assim, é a primeira etapa do processamento de texto, constituindo-se uma fase essencial de segmentação de informação (INDURKHYA; DAMERAU, 2010). O processo de tokenização consiste em quebrar as palavras em unidades, delineadas por caracteres pré-estabelecidos, na qual as unidades são denominadas de token (MEDEIROS,2010; ALLAHYARI *et al.*,2017).

Segundo Medeiros (2010) os tokens variam dependendo de sua granularidade, isto significa que os textos podem ser fracionados em capítulos, seções, parágrafos, frases, palavras, entres outros, de acordo com sua aplicação, sendo que o mais comum é a quebra de texto em frases e palavras. Em alguns idiomas (na maioria ocidentais) o caractere é delimitado por espaço, tendo uma segmentação de informação, contribuindo com o método de tokenização (INDURKHYA; DAMERAU, 2010). A tokenização apresenta desvantagem pois é difícil de utilizar a técnica em idiomas tais como chinês, japonês e tailandês, que utilizam a escrita em 'scriptio continua' que é uma escrita sem espaços ou outras marcas entre palavras ou frases (VIJAYARANI; JANANI, 2016).

2.6. Bag-of-Words (Bow)

A metodologia *Bag-of-Words* ("saco de palavras", em português) normalmente é utilizada para modelar o contexto léxico combinado de quaisquer usos de uma determinada palavra ou conjunção de palavras (INDURKHYA; DAMERAU, 2010). Nesta abordagem os dados não estruturados são transformados em um formato estruturado (MARTINS, 2003), sendo representados por um modelo vetorial (atributo-valores) para representação de documentos textuais (ROSSI, 2011). Ao considerar este procedimento como um vetor, é viável fazer comparações de contextos (WIDDOWS; DOROW, 2005). Ou seja, cada palavra detectada no conjunto textual pode virar um atributo de representação de documentos textuais. A figura 2 representa um modelo de BOW:

Figura 2 - Representação de um modelo de *Bag-of-words*

	w_1	w_2	w_3	...	w_n
d_1	a_{11}	a_{12}	a_{12}	...	a_{1n}
d_2	a_{21}	a_{22}	a_{23}	...	a_{2n}
...	\vdots	\ddots	\vdots
d_n	a_{n1}	a_{n2}	a_{n3}	...	a_{nn}

Fonte: Alves (2010).

De acordo com Alves (2010), w_i representa uma palavra, d_j representa um documento e a_{ij} é considerado o peso de cada palavra no documento, visto que os valores dos pesos de a_{ij} podem ser calculados por várias medidas. Essas medidas podem ser classificadas em dois tipos: binárias (0 se o termo não ocorre no documento, e 1 se o ele ocorre) e a frequência, como a frequência do atributo no documento (TF) que visa contabilizar o número de ocorrências de um dado termo (Nogueira,2009). A figura 3 é uma representação de BOW, a partir das sentenças dadas por Ribeiro e Silva (2018):

- Gostei do celular, mesmo sendo caro;
- Não gostei desse celular, estraga.

Figura 3 - Uma representação de *Bag-of-Words* baseados nas sentenças

	caro	celular	desse	estraga	fácil	gostei	mesmo	não	sendo
tweet1	1	1	0	0	0	1	1	0	1
tweet2	0	1	1	1	1	1	0	1	0

Fonte: Ribeiro e Silva (2010).

No entanto, o atributo de representação textual na técnica BOW tem como principal característica a alta dimensionalidade devido ao grande número de palavras diferentes contidas em uma coleção de texto (ROSSI, 2011). Isto significa que determinadas palavras não são representativas para discriminar o conteúdo dos textos. Apesar disso, a abordagem mais utilizada na estruturação de textos é a

representação BOW, sendo uma das alternativas para trabalhar a representação de textos (MARTINS, 2003).

2.7. Term Frequency - Inverse Document Frequency (TF-IDF)

O TF-IDF (*Term Frequency - Inverse Document Frequency*) visa determinar quais são as palavras mais relevantes em um conjunto textual (SHOUZHONG; MINLIE, 2016), ou seja, busca determinar a relevância de uma palavra em relação ao texto que pertence através de cálculos que atribuem valores para cada palavra do documento, mediante a uma proporção inversa da frequência da palavra em um banco de textos para determinar a proporção que a palavra aparece no documento (RAMOS, 2003).

O algoritmo é utilizado na fase de pré-processamento, especialmente na etapa de ranqueamento das palavras. Portanto o TF-IDF é influenciado pela quantidade de vezes que um termo aparece no texto, isto quer dizer que é determinado pela frequência da palavra (CAPOBIANCO, 2016). O TF-IDF é muito utilizado na abordagem estatística que retrata a importância de um termo para determinado documento em relação ao corpus (SOUMYA; PRAMOD, 2020).

No TF-IDF, a frequência do termo (TF) é definida pela probabilidade de ocorrência de um termo no documento, enquanto a frequência no documento inverso (IDF) indica a particularidade do termo, sendo uma operação de multiplicação entre essas duas estatísticas (Eqs. (1) e (2)) que fornece TF-IDF conforme fornecido pela Eq. (3), (THAKKAR; CHAUDHARI, 2020).

$$TF(t, d) = \frac{f_{t,d}}{|\{f_{t',d}: t' \in d\}|} \quad (1)$$

onde, $f_{t,d}$ indica a frequência do termo t ocorrendo no documento d ; $f_{t',d}$ apresenta a frequência de qualquer termo t de tal modo que $|\{f_{t',d}: t' \in d\}|$ dá o número total de termos no documento d .

$$DF(t, D) = \frac{|D|}{|\{d' \in D: t \in d'\}|} \quad (2)$$

onde, D é o corpus do documento e $|D|$ indica o número total de documentos no corpus dado; para o termo t , $|\{d' \in D: t \in d'\}|$ que fornece o número total de documentos no corpus D que contém o termo t .

$$TF-IDF(t, d, D) = TF(t, d) * IDF(t, D) \quad (3)$$

Com base na importância derivada de um termo em relação ao dado documento em um corpus, os termos podem ser fornecidos por pontuações do TF-IDF, com a sua relevância para a entrada da consulta podendo também ser obtida (THAKKAR; CHAUDHARI, 2020). O algoritmo é adotado em diversos estudos para detectar *fake news* através do banco de dados textuais espalhados por várias notícias, mídias sociais, blogs, entre outros.

2.8. Árvores de Decisão

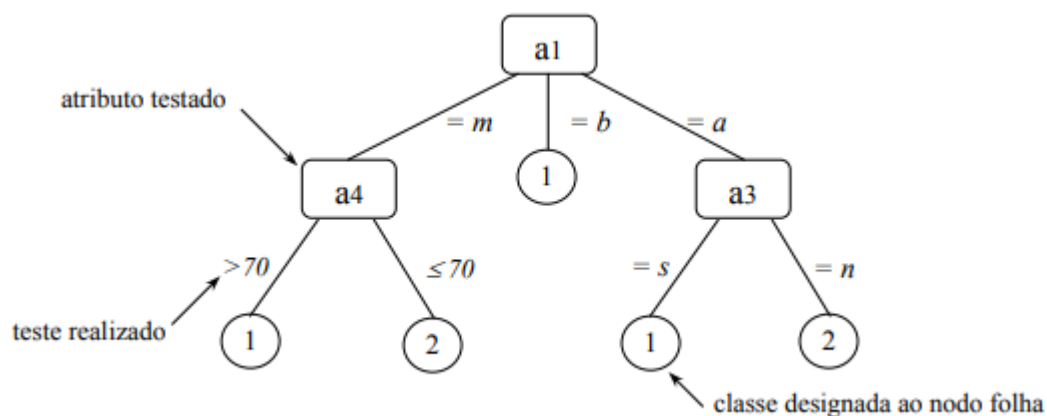
A AD é integrada com uma cadeia de nós de decisão, interligados através de ramificações, amplificando-se desde o nó raiz até os nós folhas, abrangendo os requisitos básicos a existência de um atributo alvo (LAST *et al.*, 2016; LAROSE; LAROSE, 2014). O primeiro nó apresentando na árvore é o atributo mais importante e os outros menos relevantes são apresentados nos nós subsequentes (Lemos *et al.*, 2005).

A sua vantagem está na tomada de decisões tendo em conta os atributos mais pertinentes, sendo compreensíveis para a maioria das pessoas. Sendo assim, ao definir e identificar os atributos em ordem de importâncias, possibilita aos usuários a compreensão de quais fatores mais influenciam os seus trabalhos (Lemos *et al.*, 2005).

Segundo Garcia (2003) e Meira *et al.* (2008), as AD são compostas de:

- **Nó raiz:** é o primeiro nó da árvore, no topo da estrutura, que representa os atributos.
- **Nós internos ou ramos:** originários dos nós, os quais são atribuídos valores para os atributos, os que definem nós de tomadas de decisão.
- **Nós folhas:** São os nós que ficam nas extremidades da árvore, que apresenta diferentes classes de um conjunto de treinamento, sendo assim cada folha faz associação a uma classe, que corresponde a uma regra de classificação. Ou seja, representam valores de predição para distribuição de probabilidade. A figura 4 mostra a representação de um classificador baseado em árvore de decisão:

Figura 4 - Representa uma árvore de decisão e seus componentes



Fonte: Garcia (2003).

A AD simboliza o conhecimento de forma simples e eficiente, delineando os classificadores que preveem classes apoiadas nos valores atribuídos de um conjunto de dados (GARCIA, 2003). Ela é baseada na tática de “dividir-para-conquistar”, ou seja, um problema complexo é deteriorado subdividido em subproblemas mais simples, respectivamente a mesma tática é aplicada a cada subproblema (GAMA, 2002).

Quinlan (1993) contribuiu para o desenvolvimento do algoritmo ID3 e suas evoluções (ID4, ID6, C 4.5, See 5), sendo eles adequados para a construção de AD.

2.9. Métricas de Avaliação

As métricas de avaliação são constantemente utilizadas em algoritmos supervisionados e estas permitem testar a eficácia do algoritmo (OBZAY; ALATAS, 2019). Neste estudo, será utilizada a Matriz de Confusão para avaliar o desempenho da detecção de *Fake News* por diferentes métricas derivadas.

A Matriz de Confusão apresenta informações sobre a frequência com que um determinado comportamento é detectado corretamente e com qual frequência é classificado com outro comportamento (SALLA *et al.*, 2018). Esta matriz possui duas dimensões, sendo que uma é indexada pela classe real de um objeto, e a outra é

indexada pela classe que o classificador prevê (DENG *et al.*, 2016). O desempenho de um classificador é frequentemente resumido por indicadores como precisão, sensibilidade (ou recall) e especificidade (SALLA *et al.*,2018), podendo também incluir a acurácia, pontuação *F-score* e área sob a curva da *receiver operating characteristic* (auc-ROC).

De acordo com Wang, Zuo e Dong (2020) a classificação dos resultados são expressos por uma matriz de confusão que contém quatro categorias: Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN). Sendo que um resultado positivo verdadeiro na matriz, denota o número de previsões corretamente classificadas como positivos, já um falso positivo demonstra o número de previsões incorretamente classificado como positivo. No caso de um verdadeiro negativo indica o número de previsões corretamente presumida como negativo. Logo um falso negativo demonstra o número de previsões incorretamente classificadas como negativas (KOK; AZWEEN; JHANJHI, 2020).

Com base nessas frequências é possível calcular os indicadores de desempenho de classificação que manifestam o desempenho do classificador ao detectar a classe dada (SALLA *et al.*,2018). Utilizando a matriz de confusão, serão aplicados os seguintes critérios de avaliação de desempenho: Acuracidade, Precisão, Recall e *F-score*.

A métrica de acuracidade visa determinar o quão bem o modelo preditivo é apto de fazer predição. Sendo assim, a classificação de acuracidade é definida como a proporção de previsões corretas (TP + TN) para o número total de previsões (TP + FP + FN + TN), a qual a precisão correta expressa a probabilidade prevista maior que 0,5 (; ZUO; DONG, 2020).

$$Acuracidade = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (3)$$

A métrica de precisão indica a proporção de previsões corretamente presumidas para ser positiva sobre o total de previsões positivas, ou seja, define quanto o modelo preditivo pode ser confiável quando a previsão é positiva (KOK; AZWEEN; JHANJHI, 2020).

$$Precisão = \frac{TP}{(FP+TP)} \quad (4)$$

O recall é uma métrica que designa a capacidade de um modelo de previsão de selecionar as instâncias de uma determinada classe de um conjunto de dados, isto significa que o modelo preditivo consegue prever corretamente os valores positivos (DENG et al., 2016).

$$Recall = \frac{TP}{(FN+TP)} \quad (5)$$

O *F-score*, que também é conhecida como a pontuação F, é na verdade um valor médio harmônico do valor de recall e o valor de precisão obtido para banco de dados, sendo assim, esta métrica determina o quanto o modelo de previsão consegue prever corretamente os valores positivos (KOK; AZWEEN, JHANJHI, 2020; OBZAY; ALATAS, 2019).

$$F - score = 2 * \frac{(Recall \times Precisão)}{(Recall + Precisão)} \quad (6)$$

2.10. Trabalhos Correlatos

Bharadwaj e Shao (2019), apresentaram uma comparação entre as técnicas de *Recurrent Neural Networks* (RNN), *Naïve Bayes* (NB), *Random Forest* (RF), com a utilização de recursos linguísticos, incluindo TF, TF-IDF, BIGRAM, TRIGRAM, QUADGRAM e GloVe para detecção de notícias falsas. O conjunto de dados utilizado para este artigo foi extraído da plataforma kaggle.com, a qual contém 6.256 artigos, incluindo seus títulos. O banco de dados foi dividido em 80% para treinamento e 20% para teste. Após a comparação das técnicas, concluíram que o resultado é promissor com *BIGRAM* e RF que alcançaram uma precisão de 95,66%.

Ozby e Alatas (2019) propuseram um modelo de detecção de *fake news* nas redes sociais, contendo duas etapas diferentes. Na primeira etapa, foram aplicados os métodos de mineração de texto para um conjunto de dados de notícias online, visto que o objetivo dos métodos e técnicas de textos era obter dados

estruturados de notícias de artigos a partir de dados não estruturados. Já na segunda etapa, algoritmos de inteligência artificial de aprendizagem supervisionada (*BayesNet*, *JRip*, *OneR*, *Decision Stump*, *ZeroR*, *SGD*, *CV Parameter Selection* - CVPS, *Randomized Filtered Classifier* - RFC, *Logistic Model Tree*, *Locally weighted learning* - LWL, CvC, WIHW, MLP, OLM, ASC, J48, SMO, *Bagging*, Árvores de decisão, k-vizinhos mais próximos - IBk, KLR) foram aplicados às notícias estruturadas do conjunto de dados. O modelo foi testado em três conjuntos de dados e avaliado através das métricas de acurácia, recall, precisão e medida F. De acordo com os resultados atingidos, os melhores valores médios em termos de acurácia, precisão e medida F foram obtidos a partir dos algoritmos de árvore de decisão (AD), *ZeroR*, *Cross Validation Parameter Selection* (CVPS), e algoritmos do tipo *Weighted Instances Handler Wrapper* (WIHW), com valor de 1.000, apresentando serem os melhores algoritmos em termos de *recall*.

Nasir, Khan e Valarmis (2020), implementam um novo modelo híbrido de *deep learning* que combina as *Redes Neurais Convolucionais* (RNC) e Recorrentes (RNR) para a classificação de notícias falsas. Foram utilizados dois conjuntos de dados (FA-KES e ISOT) para avaliação de desempenho do modelo implementado, alcançando resultados (100% de precisão no conjunto de dados ISOT, 45.000 artigos, 60% de precisão no conjunto de dados FA-KES, 804 artigos) de detecção que são significativamente melhores do que outros métodos não híbridos.

O estudo de Zhang *et al.* (2019) teve como objetivo garantir a redução de riscos representados pelo consumo de *fake news*, sendo assim, propuseram uma nova abordagem de *Fake News Detection* (FEND) orientada para a detecção de notícias falsas. Para avaliação quantitativa da eficácia da abordagem, foi desenvolvido um repositório de notícias reais e falsas. Para a construção do conjunto de dados foram coletadas notícias reais disponíveis na CNN e The New York Times para o treinamento do modelo. Além disso, para avaliar a precisão da detecção foram utilizados conjuntos testes de sites classificados como domínios de notícias falsas (www.greenvillegazette.com, www.politicot.com, www.advocate.com e www.naturalnews.com). Esta nova abordagem apresentou 92,49% de precisão de classificação e 94,16% de recall com base no especificado do valor limite de 0,6.

Agarwal *et al.* (2019), pesquisaram sobre o conjunto de dados LIAR, que contém 12.836 declarações curtas rotuladas para veracidade, assuntos, contexto, local, partido e entre outros (WANG,2017). Os autores simplificaram a base de dados

em apenas duas variáveis para a tarefa de classificação. Técnicas de pré-processamento como a tokenização, *stemming* (derivação), BOW e *N-grams* foram utilizadas para seleção de recursos, e o TF como a ponderação TF-IDF foi aplicado. Para a classificação, as técnicas *Naïve Bayes*, Regressão Logística, Linear SVM, *Stochastic Gradient Classifier* e RF foram usadas. Os resultados foram medidos pelas métricas de precisão, recall e medida F, concluindo que SVM e Regressão Logística obtiveram o melhor desempenho neste conjunto de dados.

Segundo Faustini e Covões (2020) propuseram em sua pesquisa a detecção de notícias falsas em três idiomas distintos e utilizando apenas recursos de textos que podem ser gerados independentemente da plataforma de origem. Portanto, foram realizados experimentos a partir de cinco conjuntos de dados: FakeBrCorpus, TwitterBR, FakeOrRealNews, Fakenewsdata1 e btvlifestyle, sendo de três línguas diferentes, tais como: Português (latina), Inglês (germânica) e Búlgaro (eslavo). As técnicas de classificação KNN, *Gaussian Naïve Bayes*, RF e SVM, e de pré-processamento (*Customised*, *Word2Vec*, *Document-Class Distance* - DCD e BOW) foram aplicados e testados nos conjuntos de dados. A partir do experimento, concluiu-se que as técnicas SVM e RF superaram as demais.

Krishnan e Chen (2018) tiveram o intuito de criar um *Framework* avançado para identificar publicações no Twitter, com conteúdo de notícias falsas. Para a sua validação, foi testado um conjunto de dados fornecido por Boididou *et al.* (2014). O conjunto de dados é composto por tweets sobre o episódio do furacão Sandy e outros acontecimentos, incluindo MH370, maratona de Boston, ataque de Paris e ataque aéreo da Rússia à Síria, entre outros. As técnicas de classificação AD (com o algoritmo) J48 e SVM foram aplicadas e avaliadas por métricas como precisão, recall e medida F. Ao ser medida a eficiência do Framework proposto, o mesmo mostrou excelente desempenho ao prever tweets falsos com recall significativo, sem sacrificar o valor da precisão.

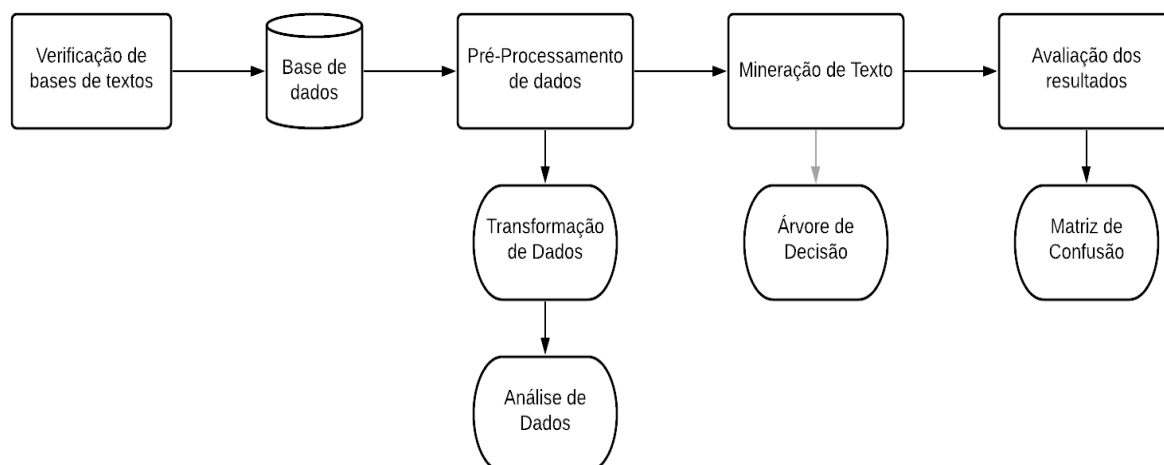
Ozbay e Alatas (2019) publicaram uma nova abordagem para detecção de notícias falsas nas redes sociais usando algoritmos de otimização de metaheurísticas. Neste novo modelo é proposto uma adaptação dos algoritmos *Grey Wolf Optimization* (GWO) e *Salp Swarm Optimization* (SSO). Foram utilizados três conjuntos de dados (Results for BuzzFeed Political News, Liar Benchmark e Random Political New) para avaliação da nova abordagem proposta. Os resultados foram comparados com sete técnicas de classificação (AD, *Naïves Bayes*, SVM, *Gradient Boosted Trees*, *Ridor*,

J48 e SMO). Após aplicação, concluiu-se que a melhor precisão obtida foi do GWO em todos os conjuntos de dados. O GWO também conseguiu a melhor precisão e valores da medida F em dois dos três conjuntos de dados. Já o SSO superou todos os outros algoritmos (exceto GWO) em termos de precisão dentro de dois conjuntos de dados.

3. MÉTODOS E TÉCNICAS

A Figura 5 mostra o fluxograma com todas as etapas dos métodos, técnicas aplicadas e avaliação dos resultados.

Figura 5 - Fluxograma das Etapas do Processo



Fonte: Autor (2021)

Metodologicamente, nesta pesquisa serão executados os seguintes passos:

- a. Os procedimentos para obtenção dos dados serão realizados via download no sítio eletrônico do Fake BR Corpus, fundamentados anteriormente.
- b. Após a etapa 1, os dados foram tratados, treinados pela técnica AD, classificados e avaliados no *software Waikato Environment of Knowledge Analysis (WEKA)*, disponível em: <https://www.cs.waikato.ac.nz/ml/weka/>. Este *software* é uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. Ele contém ferramentas para preparação de dados, classificação, regressão, agrupamento, mineração de regras de associação e visualização.

3.1. Descrição Do Conjunto De Dados

Para o presente estudo foi utilizado o conjunto de dados público denominado Fake.br Corpus (Monteiro *et.al*, 2018), disponibilizado em

<https://github.com/roneysco/Fake.br-Corpus>. O conjunto de dados possui 7.200 textos de websites, sendo que 3.600 são notícias verídicas e outros 3.600 são *fake news*, todos no idioma português. O conjunto de dados possui uma pasta pré-processada, a qual está no formato CSV contendo etiqueta de notícias e texto de notícias pré-processada, como palavras irrelevantes (*StopWords*) em português removidas, sotaque e diacrítico.

4. RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados do experimento realizado para a classificação de *fake news*, desta forma, foi utilizado a ferramenta WEKA para aplicação da técnica de AD com os dois recursos: BOW e TF-IDF, com a finalidade de comparar seus desempenhos. No experimento, foi usado o esquema de validação cruzada k -fold, o qual o conjunto de dados é dividido em k subconjuntos, e o método de validação é repetido k vezes, sendo que um dos k subconjuntos é utilizado como o conjunto teste e os outros $k - 1$ subconjuntos são agrupados para formar um conjunto de treinamento. Em seguida, o erro médio em todas as k tentativas é calculado. Segundo Krishnan e Chen (2018) a vantagem deste método é que não é considerado como os dados são divididos, pois cada ponto de dados entra em um conjunto de teste exatamente uma vez e em um conjunto de treinamento $k - 1$ vezes, com a variância da estimativa resultante é reduzida à medida que k aumenta. A tabela 1 mostrar os resultados obtidos através da matriz de confusão:

Tabela 1 - Resultados do Experimento

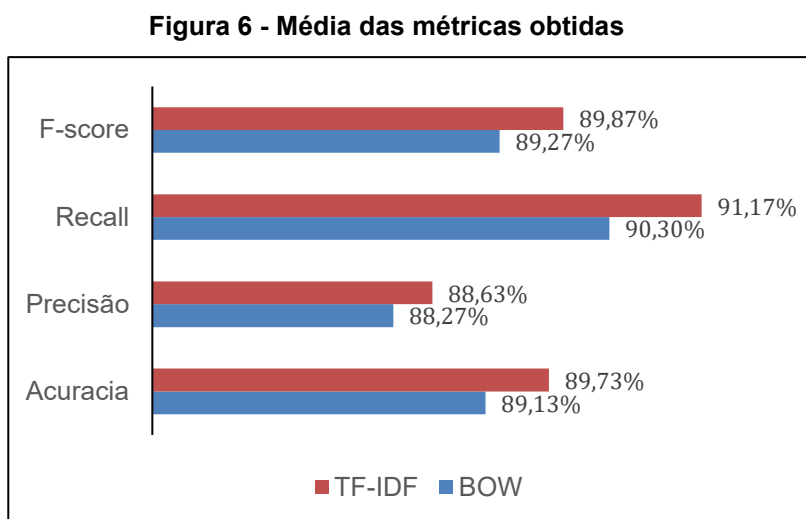
Recurso	K -fold	Acurácia	Precisão	Recall	F-score
BOW	10	88,79%	88,00%	89,90%	88,90%
BOW	20	89,21%	88,20%	90,60%	89,40%
BOW	50	89,40%	88,60%	90,40%	89,50%
TF-IDF	10	89,79%	88,60%	91,30%	89,90%
TF-IDF	20	89,58%	88,50%	91,00%	89,70%
TF-IDF	50	89,82%	88,80%	91,20%	90,00%

Fonte: Elaborado pelo autor (2021)

A partir dos resultados obtidos, foi possível observar que ambos os recursos resultaram em desempenho promissor ao classificar *fake news*. Porém, em termos de todas as métricas de avaliação, a técnica de AD teve um melhor desempenho ao utilizar o recurso TF-IDF, sendo que ao utilizar o TF-IDF ($k = 10$) obteve uma acurácia de 89,79 %, mas teve uma precisão 88,60%, ou seja, o algoritmo é um bom modelo de classificação, porém de todas as classificações que o modelo classificou como *fake news*, 88,60% pertenciam realmente a esta classe. O recurso TF-IDF ($k = 50$) obteve uma acurácia melhor, 89,82 %, precisão 88,80%, no entanto

teve um *recall* (sensibilidade) de 91,20%. Isto significa que de todas as *fake news*, o algoritmo conseguiu classificar 91,20% das notícias que, era de fato, *fake news*.

A figura 6 é a média do desempenho obtidos dos dois modelos em relação a acurácia, precisão, recall e *F-score* são demonstrados abaixo:



Fonte: Elaborado pelo autor (2021).

A partir da média calculada das métricas, nota-se que o modelo TF-IDF teve alguns ganhos em relação ao recurso BOW, sendo que alcançou 0,6% a mais de acurácia, 0,37% a mais na precisão, na sensibilidade teve uma vantagem de 0,87% e no *F-score* o ganho foi de 0,6%. Por mais que o pré-processamento TF-IDF tenha apresentado um resultado levemente superior, não se pode afirmar que essas médias dos resultados são diferentes. Apesar do modelo TF-IDF apresentar ganhos sobre o BOW, pode-se dizer que os resultados do modelo BOW também é eficaz na classificação de *fake news*. Além disso, a matriz de confusão do modelo TF-IDF, mostra que o experimento não é apenas promissor na classificação de notícias falsas, ou seja, o modelo apresenta um desempenho eficaz ao prever notícias verdadeiras, isto significa que o experimento é um modelo generalizado. A tabela 2 mostrar a matriz de confusão do modelo TF-IDF:

Tabela 2 – Matriz de Confusão

Classificação	Notícias Falsas (<i>Fake</i>)	Notícias Verídicas
<i>Fake</i>	3282	318
Verdadeira	415	3185

Fonte: Elaborado pelo autor (2021).

De acordo com o experimento, os termos transformados a partir do modelo TF-IDF apresentam conexão com a base do conjunto de dados, sendo que os termos mais relevantes estão relacionados ao âmbito político. A tabela 3 mostra os termos relevantes do conjunto de dados:

Tabela 3 - Termos relevantes

Ganho de informação	Frequência	Termo	Ganho de informação	Frequência	Termo
0.034206	1003	zona	0.079194	911	supremo
0.031402	1002	votos	0.042981	907	stf
0.027209	1001	voto	0.008857	875	senadores
0.017951	1000	votar	0.014945	874	senador
0.024238	999	votacao	0.019518	873	senado
0.024216	992	vitimas	0.341362	868	segundo
0.018102	991	vitima	0.078168	867	segunda
0.035223	989	violencia	0.034951	673	partidos
0.139025	947	tribunal	0.118985	566	ministerio
0.014727	913	suspeito	0.032985	569	ministros
0.026998	912	suspeita	0.059038	568	ministro

Fonte: Elaborado pelo autor (2021).

Segundo Junior (2007), a importância de um determinado termo para discriminação entre classes de documentos previamente conhecidos, é definido como ganho de informação, o qual tem o intuito de verificar o quanto cada termo está correlacionado com cada classe. Um dos princípios da construção de AD, é maximizar o ganho de informação, pois sabe-se que a propriedade da entropia é baseada no valor máximo alcançado quando todas as mensagens são igualmente prováveis (YASHKOV, 2014). Diante dos resultados da tabela 3, nota-se que os termos como: segundo, segunda, supremo, tribunal, voto, votação, stf e entre outros, provavelmente foram significativas para determinar e classificar as notícias como *fake news* ou verdadeira. Visto que as *fake news* geralmente estão interligadas ao contexto político ou de interesse público.

No experimento Faustini e Covões (2020) foi utilizado o conjunto de dados *FakeBr Corpus*, o qual foi avaliado em quatro conjuntos testes Customised, Word2Vec, DCD e BOW e em cada conjunto foram aplicadas as técnicas de classificação KNN, NB, SVM e RF. Neste experimento, a técnica de classificação que teve um desempenho promissor foi SVM e o conjunto de teste BOW. Esse modelo proposto atingiu uma acurácia de 91% e a F1 alcançou 91 %.

Ao comparar-se o desempenho do experimento de Faustini e Covões (2020) com o presente estudo, nota-se que o algoritmo SVM proposto foi mais promissor do que a técnica de AD. Outro ponto interessante do modelo de Faustini e Covões (2020) é que o recurso BOW teve um ganho significativo com a técnica SVM. Além disso, o modelo proposto no experimento atual possui uma limitação devido o conjunto de dados ser pré-processado, implicando na exploração de outros recursos para extração de informações de dados textuais. Apesar dessas limitações, o modelo proposto é considerado promissor em predição de *fake news*, sendo que o modelo atingiu um recall 91,30% na identificação de notícias falsas, onde a métrica *recall* costuma ser considerada mais importante do que a precisão (KRISHNAN; CHEN, 2018).

5. CONCLUSÃO

As *fake news* têm chamado a atenção em nível global e vem ganhando espaço em diversas áreas de pesquisa, principalmente na área da computação, a qual geralmente se baseia nos estudos com o foco em um tipo de método para detecção e verificação da autenticidade de notícias publicadas em meio de mídias sociais. Neste contexto, o presente estudo teve como objetivo principal extrair informações não triviais sobre estas *fake news* por meio da metodologia KDT.

O método proposto foi baseado em MT, utilizando a técnica de AD e os recursos TF-IDF e BOW para criação de um modelo capaz de extrair informações de textos de um conjunto de dados previamente pré-processado (*FakeBr Corpus*), ou seja, com a lematização e *stemming* realizados. O experimento obteve resultados promissores com os ambos os recursos, sendo que o TF-IDF alcançou a maior acurácia 89,72%, ou seja, o modelo proposto é eficaz em classificar estas notícias. Além disso, o experimento evidencia que o modelo também é interessante em predição de notícias verdadeiras, ou seja, com uma alta especificidade.

Apesar do conjunto de dados apresentar uma limitação devido ser um banco de dados já pré-processado, a técnica de MT aplicada neste estudo mostrou-se promissora e pode ser aplicada para extração de informações não triviais. Contudo, ressalta-se a importância de se explorar conjuntos de dados (Dados Brutos) e outras técnicas da literatura que têm disso apontadas por pesquisadores como boas classificadoras (SVM, RF e *Deep Learning*, por exemplo) para alcançar resultados mais aprimorados e melhorar facilitar a extração de informações de dados textuais.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, L. R. **Mineração de Texto Aplicada à Identificação de Fake News**. 2019. 46 p. Monografia (Bacharelado em Ciência da Computação) - Universidade Federal do Maranhão, São Luís, 2019. Disponível em: <<https://monografias.ufma.br/jspui/handle/123456789/4253>>.

AGARWAL, V., SULTANA, H.P., MALHOTRA, S., SARKAR, A. Analysis of Classifiers for Fake News Detection. **Procedia Computer Science**, v. 165, n. 2019, p. 377–383, 2019. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050920300430>>.

ALLAHYARI, M., POURIYEH, S., ASSEFI, M., SAFAEI, S., TRIPPE, E. D., GUTIERREZ, J. B., KOCHUT, K. A Brief Survey of Text Mining: Classification, Clustering and Extraction. **Techniques In Proceedings of KDD Bigdas**, Halifax, Canada, V.2,2017. <<http://arxiv.org/abs/1707.02919>>.

ALLCOTT, H., GENTZKOW, M. Social Media and Fake News in the 2016 Election. **Journal of Economic Perspectives**, v. 31, n. 2, p. 211-236, 2017. DOI 10.1257/jep.31.2.211. Disponível em: <<https://www.aeaweb.org/articles?id=10.1257/jep.31.2.211>>.

ALVES, A.I. M. Modelo de representação de texto mais adequado à classificação. 2010. 92 f. DISSERTAÇÃO (Mestre em Engenharia Informática) - Instituto Superior de Engenharia do Porto, Porto, Portugal, 2010. Disponível em: http://alvarestech.com/temp/Rapidminer/DM_AlexandraAlves_2010_MEI.pdf.

BAKSHY, E., MESSING, S., ADAMIC, L. A. Exposure to ideologically diverse news and opinion on Facebook. **Science**, v. 348, n. 6239, p. 1130–1133, 2015. Disponível em <<https://science.sciencemag.org/content/348/6239/1130.abstract>>.

BAPTISTA, R. Redes sociais influenciam voto de 45% da população, indica pesquisa do DataSenado. **Brasil: Senado Notícias**, 12 dez. 2019. Disponível em: <<https://www12.senado.leg.br/noticias/materias/2019/12/12/redes-sociais-influenciam-voto-de-45-da-populacao-indica-pesquisa-do-datasenado>>.

BERRY, M.J.A., LINOFF, G. Data mining techniques. **New York: John Wiley & Sons**, Inc. 1997.

BHARADWAJ, P., SHAO, Z. Fake News Detection with Semantic Features and Text Mining. **International Journal on Natural Language Computing**, v. 8, n. 3, p. 17–22, 2019. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3425828>.

CAPOBIANO, K. R. **AVALIAÇÃO DA ETAPA DE PRÉ-PROCESSAMENTO NA MINERAÇÃO DE TEXTO EM REDES SOCIAIS DIGITAIS**. 2016. 57 f. TCC (Bacharelado em Ciência da Computação) - Universidade Estadual de Londrina - UEL,

Londrina, 2016. Disponível em: <http://www.uel.br/cce/dc/wp-content/uploads/TCC-KELVIN_RAMIRES_CAPOBIANCO-BCC-UEL-2015.pdf>.

CARVALHO, D, R., MOSER, A; D.; SILVA, V; A.; DALLAGASSA, M; R. **Mineração de Dados Aplicada à Fisioterapia**. *Fisioter. Mov.*, Curitiba, v. 25, n. 3, p. 595-605, jul./set. 2012. Disponível em: < <http://www.scielo.br/pdf/fm/v25n3/15.pdf> >.

CHATTERJEE, S., GOYAL, D., PRAKASH, A., SHARMA, J. Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. **Journal of Business Research**, p.815–825,12 oct. 2020. <<https://doi.org/10.1016/j.jbusres.2020.10.043>.Disponível em: <https://www.elsevier.com/pt-br>>.

DENG, X., LIU, Q., DENG, Y., MAHADEVAN, S. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. **Information Sciences**, v. 340–341, p. 250–261, 2016. Disponível em:<<https://www.sciencedirect.com/science/article/abs/pii/S002002551600044X?via%3Dihub>>.

DIAS, M.M. Parâmetros na Escolha de Técnicas e Ferramentas de Mineração de Dados. **Acta Scientiarum: Technology**, Maringá, Paraná, v. 24, ed. 6, p. 1715-1725, 2002. Disponível em: <https://www.researchgate.net/publication/267770511_Parametros_na_escolha_de_tecnicas_e_ferramentas_de_mineracao_de_dados>.

Fake News: Saúde Sem Fake News. **Brasil: Ministério da Saúde**, 2020. Disponível em: <<https://antigo.saude.gov.br/fakenews/>>.

FAUSTINI, P. H. A., COVÕES, T. F. Fake news detection in multiple platforms and languages. **Expert Systems with Applications**, v. 158, p. 113503, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0957417420303274>>.

FAYYAD, U., PIATETSKY-SHAPIRO, G., SMYTH, P. Knowledge Discovery and Data Mining: Towards a Unifying Framework. **AAAI.ORG**, p. 88–82, 1996b.

FAYYAD, U., PIATETSKY-SHAPIRO, G., SMYTH, P. Knowledge Discovery in From Data Mining to Databases. **AAAI.ORG**, v. 17, n. 3, p. 637–648, 1996a.

FEBRABAN.**OBSERVATÓRIO FEBRABAN**. IV. ed.:FEBRABAN, set 2020. Disponível em: <https://cmsportal.febraban.org.br/Arquivos/documentos/PDF/200926_iD_%20OBSERVAT%3%93RIO%20FEBRABAN%20IV_%20SETEMBRO%202020%20%23BRASILONLINE_final.pdf>.

FELDMAN, R., DAGAN, I. Knowledge Discovery in Textual Databases (KDT). **International Conference on Knowledge Discovery and Data Mining (KDD)**, p. 112–117, 1995.

GAMA, J. **Árvores de Decisão**. p. 1–22, 2002. Disponível em: <https://www.dcc.fc.up.pt/~ines/aulas/MIM/arvores_de_decisao.pdf>.

GARCIA, R. T. Brazil's “fake news” bill won't solve its misinformation problem. **MIT Technology Review**, 10 de set. de 2020. Disponível em <<https://www.technologyreview.com/2020/09/10/1008254/brazil-fake-news-bill-misinformation-opinion>>.

GARCIA, S. C. **O Uso de Árvores de Decisão na Descoberta de Conhecimento na Área da Saúde**. p. 1–88, 2003. Dissertação (Mestre em Ciência da Computação) – Universidade Federal do Rio Grande do Sul - UFRGS, Porto Alegre, 2003.

HELMSLEY, J. **The role of middle-level gatekeepers in the propagation and longevity of misinformation**. In Southwell, B. G., Thorson, E. A., Sheble, L. *Misinformation and Mass Audiences*, p.263–273 (University of Texas Press, 2018).

HOWARD, P. N., WOOLLEY, S., CALO, R. Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. **Journal of Information Technology & Politics**.11 de abr. de 2018. Disponível em <<https://www.tandfonline.com/doi/pdf/10.1080/19331681.2018.1448735?needAccess=true>>.

INDURKHYA, N., DAMERAU, F. J. **NATURAL LANGUAGE PROCESSING**. 2. ed. Boca Raton, FL: Taylor & Francis Group, 2010. 676 p. ISBN 13: 978-1-4200-8593-8.

JUNG, H., LEE, B. G. Research trends in text mining: Semantic network and main path analysis of selected journals. **ELSEVIER**, South Korea, p. 1-12, 14 jul. 2020. DOI <https://doi.org/10.1016/j.eswa.2020.113851>. Disponível em:<<https://www.elsevier.com/pt-br>>.

JUNIOR, J. R. C. **Desenvolvimento de uma Metodologia para Mineração de Textos**. 2007. TESE (Mestrado em Engenharia Elétrica) - Departamento de Engenharia Elétrica da PUC-Rio, Rio de Janeiro, 2007. Disponível em: <<https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=11675@1>>.

KAUR, P., MADAN, N. Technique of Text Mining: A Survey. **International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)**, p.513–519, 2015. Disponível em: <<http://www.iasir.net/IJETCASpapers/IJETCAS14-409.pdf>>.

KEMP, S. DIGITAL 2019: GLOBAL DIGITAL OVERVIEW. **DATAREPORTAL**,2019. Disponível em <<https://datareportal.com/reports/digital-2019-global-digital-overview>>.

KOK, S. H., AZWEEN, A., JHANJHI, N. Z. Evaluation metric for crypto-ransomware detection using machine learning. **Journal of Information Security and Applications**, v. 55, p. 102646, 2020. Disponível em:<<https://www.sciencedirect.com/science/article/abs/pii/S221421262030805X>>.

KRISHNAN, S., CHEN, M. Identifying tweets with fake news. **Proceedings - 2018 IEEE 19th International Conference on Information Reuse and Integration for Data Science**, IRI 2018, v. 67, p. 460–464, 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8424744>>.

KUMAR, L., BHATIA, P. K. TEXT MINING: CONCEPTS, PROCESS AND APPLICATIONS. **Journal of Global Research in Computer Science**, v. 4, ed. 3, p. 36-39, 3 mar. 2013. Disponível em: <<https://www.rroij.com/open-access/text-mining-concepts-process-and-applications-36-39.php?aid=38178>>.

LAROSE, D. T. (2014). **Discovering knowledge in data: an introduction to data mining**, John Wiley & Sons. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=9hOpAWAAQBAJ&oi=fnd&pg=PR11&dq=Discovering+knowledge+in+data:+an+introduction+to+data+mining+LAROSE+2014&ots=9Q4B7VdQWb&sig=PaHi7LHow4j38cd37Ut7cnS0OZg#v=onepage&q=Discovering%20knowledge%20in%20data%3A%20an%20introduction%20to%20data%20mining%20LAROSE%202014&f=false>>.

LAST, M., TOSAS, O., CASSARINO, T. G., KOZLAKIDIS, Z. Evolving classification of intensive care patients from event data. **Artificial Intelligence in Medicine**, v. 69, p. 22–32, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S093336571530083X?via%3Dihub>>.

LEMOS, E. P., STEINER, M. T. A., NIEVOLA, J. C. Análise de crédito bancário por meio de redes neurais e árvores de decisão: uma aplicação simples de data mining. **Revista de Administração - RAUSP**, v. 40, n. 3, p. 225–234, 2005.

LIDDY, E.D. Natural Language Processing. In **Encyclopedia of Library and Information Science**, 2nd Ed. NY. Marcel Deck, Inc.2001 Disponível em: <<https://surface.syr.edu/cnlp/11/>>.

LIMA, G. E., SANTOS, M. J., MOREIRA, D. R. R. O combate às fake news por meio da checagem de informação. **Mackenzie**, 2020. Disponível em: <<https://www.mackenzie.br/fakenews/noticias/arquivo/n/a/i/o-combate-as-fake-news-por-meio-da-checagem-de-informacao/>>.

MACHADO, C., KIRA, B., HIRSCH, G., MARCHAL, N., KOLLANYI, B., HOWARD, P. N., GRAPHIKA, L. T., GRAPHIKA, V. B. News and Political Information Consumption in Brazil: Mapping the First Round of the 2018 Brazilian Presidential Election on Twitter. **The Computational Propaganda Project Algorithms, Automation and Digital Politics**. 5 de out. de 2018. Disponível em: <https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/10/machado_et_al.pdf>.

MARKINES, B., CATTUTO, C., MENCZER, F. Social spam detection. **ACM International Conference Proceeding Series**, n. January 2009, p. 41–48, 2009.

MARTINS, C. A. **Uma abordagem para pré-processamento de dados textuais em algoritmos de aprendizado**. 2003. 208 f. Tese (Ciências - Ciências de Computação e Matemática Computacional) - Instituto de Ciências de Matemática e de Computação - ICMC-USP, São Carlos, SP., 2003.

MARUMO, F. S. **Deep Learning para Classificação de Fake News por Sumarização de Texto**. 2018 54 p. TCC (Bacharelado em Ciência da Computação) - Universidade Estadual de Londrina, Londrina, PR, 2018. Disponível em: <http://www.uel.br/cce/dc/wpcontent/uploads/Fabiano_Preliminar.pdf>.

MEDEIROS, D. M. R. **Análise de dados por meio de agrupamento fuzzy semi-supervisionado e mineração de textos**. 2010. 111 p. TESE (Doutorado em Ciências - Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemática e de Computação - ICMC - USP, São Carlos, SP, 2010. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/55/55134/tde-15022011-092004/publico/deboramedeiros.pdf>>.

MEIRA, C. A. A., RODRIGUES, L. H. A., MORAES, S. A. Análise da epidemia da ferrugem do cafeeiro com árvore de decisão. **Tropical Plant Pathology**, 33(2), p.114–124., 2008. <https://doi.org/10.1590/S1982-56762008000200005>. Disponível em: <<https://sbfitepatologia.org.br/>>.

MONTEIRO R.A., SANTOS R.L.S., PARDO T.A.S., de ALMEIDA T.A., RUIZ E.E.S., VALE O.A. (2018) **Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection Results**. In: Villavicencio A. et al. (eds) Computational Processing of the Portuguese Language. PROPOR 2018. Lecture Notes in Computer Science, vol 11122.

NASIR, J. A., KHAN, O. S., VARLAMIS, I. Fake news detection: A hybrid CNN-RNN based deep learning approach. **International Journal of Information Management Data Insights**, v. 1, n. 1, p. 100007, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2667096820300070?via%3Dihub>>.

NOGUEIRA, B. M. **Avaliação de métodos não-supervisionados de texto de seleção de atributos para Mineração de Textos**. 2009. 104 f. DISSERTAÇÃO (Mestre em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação - ICMC - USP, São Carlos, SP, 2009. DOI 10.11606/D.55.2009.tde-06052009-154832. Disponível em: <<https://teses.usp.br/teses/disponiveis/55/55134/tde-06052009-154832/pt-br.php>>.

OZBAY, F. A., ALATAS, B. A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. **Elektronika ir Elektrotechnika**, v. 25, n. 4, p. 62–67, 2019. Disponível em: <https://www.researchgate.net/publication/335133816_A_Novel_Approach_for_Detection_of_Fake_News_on_Social_Media_Using_Metaheuristic_Optimization_Algorithms>.

OZBAY, F. A., ALATAS, B. Fake news detection within online social media using supervised artificial intelligence algorithms. **Physica A: Statistical Mechanics and its**

Applications, v. 540, p. 123174, 2020. <https://doi.org/10.1016/j.physa.2019.123174>. Disponível em: <<https://www.elsevier.com/pt-br>>.

QUINLAN, J. R. (1993). C4.5: Programming for machine learning, **Morgan Kauffmann**. RAMOS, J. Using TF-IDF to Determine Word Relevance in Document Queries. **Proceedings of the first instructional conference on machine learning**, v. 242, n. 1, p. 29–48, 2003.

REZENDE, C. Governo anunciou em 47 sites de notícias falsas em 38 dias, diz CPMI: CPMI das Fake New. **Notícias UOL**, BRASIL, 3 jun. 2020. Política. Disponível em: <<https://noticias.uol.com.br/colunas/constanca-rezende/2020/06/03/governo-publicou-653-mil-anuncios-em-canais-de-noticias-falsas-avalia-cpmi.htm>>.

RIBEIRO, A. P; SILVA, N.F.F. Um estudo comparativo sobre métodos de análise de sentimentos em tweets. **Revista de Sistemas de Informação da FSMA**, Goiás, ano 2018, n. 22, p. 35-48, 1 jan. 2018. Disponível em: <http://www.fsma.edu.br/si/edicao22/FSMA_SI_2018_2_Principal_01.pdf>.

ROSSI, R. G. **Representação de coleções de documentos textuais por meio de regras de associação**. 2011. 133 p. Dissertação (Mestrado em Ciências - Ciências de Computação e Matemática Computacional) - Instituto de Ciências de Matemática e de Computação - ICMC-USP, São Carlos, SP, 2011.

RUEDIGER, M. A. COORD. **Robôs, redes sociais e política no Brasil: estudo sobre interferências ilegítimas no debate público na web, riscos à democracia e processo eleitoral de 2018**. p. 1–11, 2017. Disponível em <http://bibliotecadigital.tse.jus.br/xmlui/bitstream/handle/bdtse/4433/2017_ruediger_robos_redes_sociais_.pdf?sequence=1>.

SALLA, R., WILHELMIINA, H., SARI, k., MIKAELA, M., PEKKA, M., JAAKKO, M. Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. **Behavioural Processes**, v. 148, p. 56–62, 2018. <https://doi.org/10.1016/j.beproc.2018.01.004>. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0376635717301146?via%3Dihub>>.

Secretaria Especial de Comunicação Social da Presidência da República (Secom/PR). **Pesquisa Brasileira de Mídia 2016 - Hábitos de consumo de mídia pela população brasileira**. Disponível em: <<http://www.secom.gov.br/atuacao/pesquisa/lista-de-pesquisas-quantitativas-e-qualitativas-de-contratos-atuais/pesquisa-brasileira-de-midia-pbm-2016-1.pdf/view>>.

SHAO, C., LUCA, G., VAROL, O., FLAMMINI, A., MENCZER, F. The spread of low-credibility content by social bots. **Nature Communications**, v. 9, n. 1, p. 1–16, 2018. Disponível em: <<https://www.andyblackassociates.co.uk/wp-content/uploads/2015/06/fakenewsbots.pdf>>.

SHEARER, E., GOTTFRIED, J. News Use Across Social media platforms 2017. **Pew Research Center**, p. 17, 2017. Disponível em <

<https://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/> >.

SHOUZHONG, T., MINLIE, H. Mining microblog user interests based on TextRank with TF-IDF factor. **Journal of China Universities of Posts and Telecommunications**, 23(5), p.40–46,2016. [https://doi.org/10.1016/S1005-8885\(16\)60056-0](https://doi.org/10.1016/S1005-8885(16)60056-0).

SOUMYA, S., PRAMOD, K.V. Sentiment analysis of malayalam tweets using machine learning techniques. **ICT Express**, v. 6, n. 4, p. 300–305, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405959520300382>>.

SWIRE, B., ECKER, U. K. **Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication**. In Southwell, B. G., Thorson, E. A., Sheble, L.(eds.) *Misinformation and Mass Audiences*, p.195–211 (University of Texas Press, 2018).

TAN, A. Text Mining: The state of the art and the challenges Concept-based. **Proceedings of the PAKDD 1999 Workshop on**, n. March, p. 65–70, 2011.

TANDOC, E. C., LIM, Z. W., LING, R. Defining “Fake News”: A typology of scholarly definitions. **Digital Journalism**, v. 6, n. 2, p. 137–153, 2018. Disponível em <https://edisciplinas.usp.br/pluginfile.php/4948550/mod_resource/content/1/Fake%20News%20Digital%20Journalism%20-%20Tandoc.pdf >.

THAKKAR, A., CHAUDHARI, K. Predicting stock trend using an integrated term frequency–inverse document frequency-based feature weight matrix with neural networks. **Applied Soft Computing Journal**, v. 96, p. 106684, 2020.

VIANNA, R. C. X. F., MORO, C. M. C. de B.; MOYSES, S. J., CARVALHO, D., NIEVOLA, J. C. Mineração de Dados e Características da Mortalidade Infantil. **Cadernos de Saúde Pública**, Rio de Janeiro, 26(3), p.535–542,2010. <https://doi.org/10.1590/s0102-311x201000030001>. Disponível em: <<https://www.scielo.br/j/csp/a/cpJFDz8tCwHgXLJRvqPRzJz/?lang=pt> >.

VIJAYARAN, S., JANANI, R. Text Mining: Open Source Tokenization Tools – An Analysis. **Advanced Computational Intelligence: An International Journal (ACIJ)**, 3(1), p.37–47,2016. <https://doi.org/10.5121/acij.2016.3104>.

VIJAYARANI, S., ILAMATHI, Ms. J., NITHYA, Ms. Preprocessing Techniques for Text Mining Preprocessing Techniques for Text Mining. **International Journal of Computer Science & Communication Networks**, v. 5, n. October 2014, p. 7–16, 2015.

VOSOUGHI, S., DEB, R., ARAL, S. THE SPREAD OF TRUE AND FALSE NEWS ONLINE. **MIT Initiative on the Digital Economy Research Brief**, v. 359, n. 6380, p. 1146–1151, 2018. Disponível em <<https://science.sciencemag.org/content/359/6380/1146> >.

WAEGEL, D. (2006). The Development of Text Mining Tools and Algorithms. **Ursinus College**, 2006. Disponível em: <<http://webpages.ursinus.edu/akontostathis/WaegelHonorsThesis.pdf>>.

WANG, Z., ZUO, R., DONG, Y. Mapping Himalayan leucogranites using a hybrid method of metric learning and support vector machine. **Computers and Geosciences**, v. 138, n. May 2019, p. 104455, 2020.

WEKA. **The workbench for machine learning**. Disponível em: <https://www.cs.waikato.ac.nz/ml/weka/>.

WIDDOWS, D., DOROW, B. Automatic extraction of idioms using graph analysis and asymmetric lexicosyntactic patterns. **Proceedings of the ACL-SIGLEX Workshop on Deep Lexical Acquisition**, June, p.48–56,2005. <https://doi.org/10.3115/1631850.1631856>.

WITTEN, I. H., FRANK, E. Data Mining: Pratical Machine Learning Tools and Techniques. 2 ed,2005. San Francisco, CA: **Elsevier**.

YASHKOV, I. B. Feature selection using decision trees in the problem of JSM classification. **Automatic Documentation and Mathematical Linguistics**, v. 48, n. 1, p. 6–11, 2014.

ZHANG, C., GUPTA, A., KAUTEN, C., DEOKAR, A.V., Qin,X. Detecting fake news for reducing misinformation risks using analytics approaches. **European Journal of Operational Research**, v. 279, n. 3, p. 1036–1052, 2019. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0377221719304977?via%3Dihub>>.