

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COORDENADORIA DO CURSO DE ENGENHARIA DE SOFTWARE

VINICIUS CASANI

**USO DE APRENDIZADO DE MÁQUINA PARA
IDENTIFICAÇÃO DE PERFIS DEPRESSIVOS EM REDES
SOCIAIS**

TRABALHO DE CONCLUSÃO DE CURSO

DOIS VIZINHOS

2019

VINICIUS CASANI

**USO DE APRENDIZADO DE MÁQUINA PARA
IDENTIFICAÇÃO DE PERFIS DEPRESSIVOS EM REDES
SOCIAIS**

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

Orientador: Prof. Dr. Francisco Carlos
Monteiro Souza

Coorientador: Prof. Dr. Rafael Gomes
Mantovani

DOIS VIZINHOS

2019



TERMO DE APROVAÇÃO

Uso de Aprendizado de Máquina para Identificação de Perfis Depressivos em Redes Sociais

por

Vinicius Casani

Este Trabalho de Conclusão de Curso foi apresentado em 25 de Novembro de 2019 como requisito parcial para a obtenção do título de Bacharel em Engenharia de Software. O(a) candidato(a) foi arguido(a) pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Francisco Carlos Monteiro Souza
Presidente da Banca

Alinne Cristinne Correa Souza
Membro Titular

Dionatan Augusto Guimaraes Cieslak
Membro Titular

* A Folha de Aprovação assinada encontra-se na Coordenação do Curso

AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo dom da vida, por ter me concedido sabedoria e força para seguir neste caminho em busca de conhecimento.

A meus pais Remoaldo Vilmar Casani e Janete Soares Casani, meu irmão João Vitor Casani e toda a minha família, por sempre acreditar nos meus sonhos e pelo incentivo durante todo o curso. Também agradeço imensamente a minha namorada Larissa Quadros da Silva por todo o apoio, auxílio, incentivo e compreensão em todos os momentos.

Ao meu orientador Prof. Dr. Francisco Carlos Monteiro Souza e ao meu coorientador Prof. Dr. Rafael Gomes Mantovani, pela paciência e ensinamentos, suas contribuições foram de suma importância para que o desenvolvimento deste trabalho fosse realizado. Agradeço também pela amizade que construímos durante os últimos meses. Também agradeço a psicóloga Jéssica Aparecida Rodrigues dos Santos pelo auxílio durante todo o processo. A contribuição de vocês foi fundamental para a conclusão deste trabalho.

Também agradeço a todos os professores da UTFPR, por seus ensinamentos durante todo o curso, em especial a professora Alinne Cristinne Corrêa Souza por todos os auxílios fornecidos em diversos momentos.

Por fim, agradeço aos meus amigos por todos os momentos de apoio e descontração. Ainda agradeço a todos que de alguma forma contribuíram para o desenvolvimento deste trabalho.

RESUMO

Casani, V. USO DE APRENDIZADO DE MÁQUINA PARA IDENTIFICAÇÃO DE PERFIS DEPRESSIVOS EM REDES SOCIAIS. 81 f. Trabalho de Conclusão de Curso – Coordenadoria do Curso de Engenharia de Software, Universidade Tecnológica Federal do Paraná. Dois Vizinhos, 2019.

A depressão é uma doença que atinge milhões de pessoas por todo o mundo, independente de sua idade, classe social ou nacionalidade. Análise de Sentimentos vem sendo utilizada para extrair informações relevantes de postagens dos usuários em redes sociais, a fim de identificar o sentimento desses referente aos mais variados assuntos. O presente trabalho tem como objetivo utilizar técnicas de análise de sentimentos para detectar possíveis perfis depressivos no *Twitter*. Para tal, foi criada uma estratégia intitulada *DP-Symptom-Identifier* a qual identifica possíveis sintomas depressivos presentes nos *tweets* e em seguida possibilita a análise do perfil como um todo, de forma temporal, facilitando o processo de identificação de traços depressivos nas *timelines*. Foram conduzidos experimentos para identificar o melhor algoritmo de aprendizado de máquina para compor a estratégia e também para avaliar a eficiência da estratégia *DP-Symptom-Identifier*. O algoritmo selecionado foi o *Multilayer Perceptron*, o qual obteve uma taxa de assertividade de 90% com um tempo de processamento inferior a 5 segundos. A estratégia proposta mostrou-se eficaz para identificar possíveis perfis depressivos analisando *timelines* extraídas do *Twitter*. Portanto, os resultados indicam que a estratégia consegue identificar sintomas de depressão e possibilita a identificação de perfis depressivos de forma correta.

Palavras-chave: Análise de Sentimentos, Aprendizado de Máquina, Processamento de Linguagem Natural, Depressão, Redes Sociais

ABSTRACT

Casani, V. USING MACHINE LEARNING FOR DETECTING DEPRESSIVE PROFILES IN SOCIAL MEDIA. 81 f. Trabalho de Conclusão de Curso – Coordenadoria do Curso de Engenharia de Software, Universidade Tecnológica Federal do Paraná. Dois Vizinhos, 2019.

Regardless of age, social class or nationality, depression is a psychic disorder that affects millions of people in the world. Sentiment Analysis has been used to extract relevant information from social media users' posts in order to identify their feelings regarding the most varied subjects. The present study aims to use Sentiment Analysis techniques to detect depressive profiles on Twitter. To this end, a strategy entitled DP-Symptom-Identifier was created which identifies possible depressive symptoms present in tweets and then enables profile analysis as a whole, temporally, facilitating the process of identifying depressive profiles in timelines. Experiments were conducted to identify the best machine learning algorithm to compose the strategy and also to evaluate the efficiency of the DP-Symptom-Identifier strategy. The algorithm selected was the Multilayer Perceptron, which obtained a 90% assertiveness rate with a processing time of fewer than 5 seconds. The proposed strategy proved effective in identifying possible depressive profiles analyzing timelines extracted from Twitter. Therefore, the results indicate that the strategy can identify symptoms of depression and allows the identification of depressive profiles correctly.

Keywords: Sentiment Analysis, Machine Learning, Natural Language Processing, Depression, Social Media

LISTA DE FIGURAS

FIGURA 1	– Hierarquia do Aprendizado Indutivo.	16
FIGURA 2	– Exemplo de tarefa de Regressão.	18
FIGURA 3	– Exemplo de tarefa de aprendizado não supervisionado onde grupos (<i>clusters</i>) são criados com base nas características similares dos objetos.	19
FIGURA 4	– Estágios da análise no Processamento de Linguagem Natural.	21
FIGURA 5	– Modelo de neurônio artificial.	24
FIGURA 6	– Topologia de uma rede Perceptron Multicamadas (MLP).	25
FIGURA 7	– String de Busca da Questão de Pesquisa.	30
FIGURA 8	– Etapas do Processo de Seleção dos Estudos.	32
FIGURA 9	– Estudos por Ano de Publicação.	34
FIGURA 10	– Estudos por País de Publicação.	37
FIGURA 11	– Técnicas identificadas nos estudos primários.	40
FIGURA 12	– Tipos de rotulação identificadas nos estudos primários.	41
FIGURA 13	– <i>DP-System-Identifier</i> : uma estratégia para identificação de sintomas depressivos.	53
FIGURA 14	– Captura da Tela de Listagem dos Dados.	55
FIGURA 15	– Captura da Tela de Análise dos Dados.	56
FIGURA 16	– Captura da Tela de Rotulação dos Dados.	57
FIGURA 17	– Exemplificação da Base de Dados de Treinamento.	58
FIGURA 18	– Distribuição das classes presentes no conjunto de treinamento.	58
FIGURA 19	– Exemplo de curva AUC.	63
FIGURA 20	– AUC obtida por cada estratégia de aprendizado com cada um dos algoritmos de AM.	67
FIGURA 21	– Tempo de Processamento de cada Estratégia com cada um dos Algoritmos.	68
FIGURA 22	– Frequência de Predições corretas em cada categoria sintomática.	70
FIGURA 23	– Síntese das predições realizadas pelo modelo para cada um dos usuários.	71

LISTA DE TABELAS

TABELA 1	–	Critérios DSM-V	28
TABELA 2	–	<i>Overview</i> dos Estudos Primários selecionados durante o MS.	33
TABELA 3	–	Análise dos Estudos I	34
TABELA 4	–	Análise dos Estudos II	37
TABELA 5	–	Sintomas presentes em sentenças rotuladas com ajuda de profissional da área de psicologia.	54
TABELA 6	–	Tecnologias e ferramentas utilizadas para o desenvolvimento do trabalho.	61
TABELA 7	–	Resultados do <i>Nemenyi test</i>	68
TABELA 8	–	Postagens do usuário de id 1	72

LISTA DE SIGLAS

ACM	<i>Association for Computing Machinery</i>
AM	Aprendizado de Máquina
AS	Análise de Sentimentos
AUC	<i>Area Under The Curve</i>
BLSTM	<i>Bi-directional Long Short-Term Memory</i>
BoW	<i>Bag-of-Words</i>
CES-D	<i>Center for Epidemiological Scale - Depression</i>
CNN	<i>Convolutional Neural Network</i>
CV	<i>Cross-Validation</i>
DT	<i>Decision Trees</i>
IA	Inteligência Artificial
IEEE	<i>Institute of Electrical and Electronic Engineers</i>
KNN	<i>K Nearest Neighbors</i>
LDA	<i>Latent Dirichlet Allocation</i>
MANCOVA	<i>Multivariate Analysis of Covariance</i>
ME	<i>Maximum Entropy</i>
MLP	<i>Multilayer Perceptron</i>
MS	Mapeamento Sistemático
MSE	<i>Mean Squared Error</i>
NB	<i>Naïve Bayes</i>
OMS	Organização Mundial da Saúde
PCA	<i>Principal Component Analysis</i>
PLN	Processamento de Linguagem Natural
QP	Questão de Pesquisa
RF	<i>Random Forest</i>
RNAs	Redes Neurais Artificiais
RNN	<i>Recurrent Neural Network</i>
ROC	<i>Receiver Operating Characteristic</i>
SVM	<i>Support Vector Machine</i>
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>
UTFPR	Universidade Tecnológica Federal do Paraná

SUMÁRIO

1 INTRODUÇÃO	11
1.1 MOTIVAÇÃO	12
1.2 OBJETIVOS	12
1.3 ESTRUTURA DA MONOGRAFIA	13
2 ASPECTOS CONCEITUAIS	14
2.1 APRENDIZADO DE MÁQUINA	15
2.1.1 Aprendizado Supervisionado	16
2.1.2 Aprendizado Não Supervisionado	18
2.1.3 Algoritmos de AM	18
2.2 PROCESSAMENTO DE LINGUAGEM NATURAL	20
2.3 REDES NEURAIS ARTIFICIAIS	23
2.4 DEPRESSÃO	25
2.4.1 Sintomas	26
2.4.2 Diagnóstico	27
2.5 CONSIDERAÇÕES FINAIS	28
3 MAPEAMENTO SISTEMÁTICO	29
3.1 PLANEJAMENTO	29
3.2 CONDUÇÃO	31
3.3 RESULTADOS E DISCUSSÕES	31
3.3.1 QP ₁ : Quais técnicas têm sido utilizadas para análise de sentimentos e identificação de perfis depressivos em redes sociais?	33
3.3.2 QP ₂ : Quais são os tipos de rotulação das bases de dados que têm sido utilizadas?	36
3.3.3 Análise e Discussão	45
3.4 AMEAÇAS À VALIDADE	49
3.5 CONSIDERAÇÕES FINAIS	50
4 PROPOSTA	51
4.1 PROPOSTA DE PESQUISA	51
4.2 ESTRATÉGIA <i>DP-SYMP-TOM-IDENTIFIER</i>	52
4.2.1 Coleta de dados	53
4.2.2 Pré-Processamento	57
4.2.3 Extração de Características e Treinamento	58
4.2.4 Estratégias de classificação exploradas	59
4.2.5 Avaliação	60
4.2.6 Tecnologias e Ferramentas	60
4.3 CONSIDERAÇÕES FINAIS	61
5 AVALIAÇÃO EXPERIMENTAL	62
5.1 DEFINIÇÃO DO EXPERIMENTO	62
5.2 DESIGN DO EXPERIMENTO	65
5.3 RESULTADOS E DISCUSSÕES	66
5.3.1 Desempenho dos algoritmos de classificação (<i>QP₁</i>)	66
5.3.2 Frequência de identificação das categorias sintomáticas (<i>QP₂</i>)	69

5.3.3 Identificação de perfil com tendências depressivas (<i>QP3</i>)	69
5.4 AMEAÇAS À VALIDADE	73
5.5 CONSIDERAÇÕES FINAIS	74
6 CONSIDERAÇÕES FINAIS	75
6.1 CONTRIBUIÇÕES	76
6.2 LIMITAÇÕES E DIFICULDADES ENCONTRADAS	76
6.3 TRABALHOS FUTUROS	77
REFERÊNCIAS	78

1 INTRODUÇÃO

Considerada o mal do século pela Organização Mundial da Saúde (OMS) (ORGANIZATION et al., 2017), a depressão é um transtorno psiquiátrico que afeta o emocional da pessoa. Uma pessoa com depressão apresenta alguns sintomas, como: desânimo; baixa autoestima; impossibilidade de sentir alegria; sentimento de inutilidade ou culpa; e em casos graves, desejo de morrer. Muitas vezes, esses sintomas modificam a maneira como as pessoas usam e processam a linguagem.

A depressão atinge pessoas de todas as idades, classes sociais e nacionalidades. Segundo dados da OMS, o número de casos de depressão aumentou em 18,4% entre 2005 e 2015, afetando 4,4% da população mundial. No Brasil, a parcela da população afetada é de 5,8%, o que coloca o país como o maior detentor do transtorno na América Latina, e o segundo nas Américas, ficando atrás apenas dos Estados Unidos, que tem 5,9% da sua população afetada.

Atualmente, técnicas de Análise de Sentimentos (AS) (LIU, 2012) tem sido amplamente utilizadas para detecção e extração de sentimentos a partir de dados textuais. Historicamente a análise de textos é uma tarefa complexa, custosa e tediosa de se fazer manualmente. No entanto, com o avanço da tecnologia e o desenvolvimento de algoritmos de Aprendizado de Máquina (AM) (MITCHELL, 1997); técnicas de Processamento de Linguagem Natural (PLN) (LIDDY, 2001); e a quantidade exorbitante de dados disponibilizados diariamente em redes sociais, fóruns e páginas na Web; atualmente é possível extrair informações riquíssimas e relevantes de dados textuais.

Dentre as várias vertentes de trabalhos em AS, alguns em específico tem explorado a detecção automática e textual de distúrbios mentais (Hassan et al., 2017; Deshpande; Rao, 2017; Islam et al., 2018). Entretanto, esta é ainda uma área de pesquisa emergente, principalmente no Brasil, onde foi encontrado apenas um estudo relacionado (Rosa et al., 2019).

1.1 MOTIVAÇÃO

Quando uma pessoa escreve um texto em alguma rede social, fórum ou ambiente virtual, sem a intenção específica de relatar seu estado emocional atual, ela pode relatar involuntariamente a polaridade do mesmo. Segundo estudos do departamento de psicologia da *University of Reading* (AL-MOSAIWI; JOHNSTONE, 2018), que analisaram mais de 6,400 usuários em 63 fóruns e *blogs* relacionados à depressão, suicídio e ansiedade, pessoas com perfil depressivo tendem a utilizar um padrão específico de palavras e sentenças, tornando-o bem característico. No padrão observado pelos pesquisadores, a frequência que palavras e sentenças absolutistas aparecem no texto é maior que o convencional utilizado por usuários saudáveis. Palavras e sentenças absolutistas transmitem magnitudes absolutas, como por exemplo: "sempre", "nada", "completamente", entre outras. Ainda segundo o estudo estas foram consideradas melhores marcadores para fóruns de saúde mental do que pronomes ou palavras de emoção negativa.

Entretanto, um dos principais problemas relacionados ao diagnóstico da depressão é a desinformação. Não saber identificar indícios sintomáticos dificultam à procura por profissionais que possa fazer um tratamento adequado, antes que se atinga quadros mais graves. Essa desinformação prejudica o diagnóstico como um todo, pois alguns dos sintomas da depressão são problemas rotineiros como: falta de apetite, insônia, desânimo e desinteresse. Além disso, a literatura também salienta que um quadro depressivo é caracterizado pela presença desses e demais sintomas por um período de duas semanas (American Psychiatric Association , 2013).

1.2 OBJETIVOS

Dessa forma, o objetivo geral deste trabalho é realizar a análise de postagens coletadas do *Twitter* a fim de identificar automaticamente a ocorrência de sintomas de depressão. Essa estratégia automatizada, mesclando técnicas de PLN e algoritmos de AM, foi batizada de *DP-Symptom-Identifier*.

Além disso, espera-se que a solução proposta possa auxiliar no processo de identificação de perfis com depressão presentes no *Twitter*. Para atingir tal objetivo, os seguintes objetivos específicos foram criados:

- desenvolver uma aplicação para coleta de postagens do *Twitter*;
- identificar e selecionar palavras e sentenças que apresentam indícios depressivos.

Além disso, cruzar tal informação com a literatura específica da área, visando direcionar a coleta das postagens;

- criar uma base de dados para treinamento dos algoritmos de AM, e determinar a melhor estratégia para rotulação da base;
- desenvolver uma aplicação *web* para auxiliar no processo de rotulação e análise dos dados coletados;
- identificar e selecionar algoritmos de AM e PLN mais adequados para compor a estratégia proposta; e
- realizar uma avaliação experimental e estatística das estratégias propostas.

Com o presente projeto interdisciplinar: há uma junção entre conhecimentos de Computação e Psicologia; foi necessário o auxílio de uma profissional da psicologia para construir o embasamento teórico e experimental. A psicóloga Jéssica Aparecida Rodrigues dos Santos (CRP: 08/26401), que acompanhou o desenvolvimento deste projeto e: guiou as pesquisas referentes à depressão; sanou dúvidas que surgiram durante a elaboração do método; avaliou os conceitos de depressão apresentados no Capítulo 2; auxiliou também no processo de criação e rotulação da base de dados de treinamento; e pode avaliar a efetividade da estratégia proposta.

1.3 ESTRUTURA DA MONOGRAFIA

O presente trabalho é composto por cinco capítulos, cujos conteúdos são apresentados da seguinte forma. O Capítulo 2 apresenta definições conceituais de Aprendizado de Máquina, Processamento de Linguagem Natural, Redes Neurais Artificiais, e também a formalização da Depressão, tema central do trabalho. No Capítulo 3 é apresentado como foi realizado o mapeamento sistemático de literatura, seguido da apresentação e discussão das evidências encontradas. O Capítulo 4 apresenta a solução criada neste trabalho, juntamente com a descrição detalhada de seu desenvolvimento. No Capítulo 5 é apresentada a metodologia experimental e a avaliação dos resultados, descrevendo o protocolo de pesquisa, as questões de pesquisa criadas, e os resultados propriamente ditos, seguidos de suas implicações. Por fim, o Capítulo 6 apresenta as considerações finais do trabalho.

2 ASPECTOS CONCEITUAIS

A Análise de Sentimentos (AS), também conhecida como “mineração de opiniões”, é uma área de pesquisa que tem como objetivo analisar opiniões, sentimentos, avaliações, atitudes e emoções das pessoas, referente à empresas, produtos, serviços, pessoas, figuras públicas, eventos ou a si próprio (LIU, 2012). Atualmente a AS vem sendo amplamente utilizada por: empresas para identificar a opinião dos clientes sobre seus produtos; figuras públicas (políticos, influenciadores, criadores de conteúdo, etc) para identificar qual foi a reação/repercussão de uma postagem; entre outros contextos.

A AS pode ser dividida em três níveis principais de análise:

- i)* **análise de documentos:** consiste em identificar se a opinião geral expressada sobre um determinado assunto presente em um documento/texto é positiva ou negativa;
- ii)* **análise de sentenças:** verifica se uma sentença é negativa, neutra ou positiva, pois entende que cada sentença/frase pode expressar uma opinião diferente sobre um mesmo tópico; e por fim
- iii)* **análise de aspectos e entidades:** considera documentos e sentenças com uma maior granularidade, a fim de extrair o maior número possível de detalhes sobre a opinião do usuário referente a um determinado assunto. Além da polaridade (negativo, positivo e neutro), existe ainda a possibilidade de se extrair emoções como felicidade, tristeza, raiva, entre outras.

A AS, na maioria das vezes, é voltada para a análise de publicações em redes sociais e microblogs, como: *Facebook*¹, *Twitter*², *Instagram*³ e *Reddit*⁴. Plataformas como *Youtube*⁵ também podem ser exploradas, pois os usuários produzem diariamente uma quantidade exorbitante de dados sobre os mais variados assuntos.

¹<http://www.facebook.com/>

²<http://www.twitter.com/>

³<http://www.instagram.com/>

⁴<https://www.reddit.com/>

⁵<https://www.youtube.com/>

Atualmente, o acesso às redes sociais é extramente simples e pode ser realizado por meio de *smartphones*, *laptops* e computadores pessoais. Em qualquer instante de tempo, dada uma conexão com a internet, é possível gerar conteúdo 24 horas por dia. A maioria das redes sociais disponibiliza APIs públicas para a coleta de seus dados, facilitando assim a criação de conjuntos de dados para pesquisas científicas e estudos direcionados a um determinado problema.

Nesse ponto, as técnicas de PLN auxiliam o processo de AS, realizando o pré-processamento e normalização dos dados textuais que são utilizados como entrada por algoritmos de detecção de padrões (AM). Tais técnicas podem auxiliar também usando dicionários léxicos, como o LIWC⁶. Baseando-se no uso de dicionários contendo suas palavras e categorias, estes estimam componentes emocionais, cognitivos e estruturais de um dados texto. Além disso, a AS utiliza ainda algoritmos de Aprendizado de Máquina para realizar a classificação dos textos.

Desta forma, este capítulo visa apresentar os conceitos fundamentais explorados por este trabalho, e está dividido da seguinte maneira: Na Seção 2.1 é apresentada uma breve introdução referente a AM e alguns algoritmos frequentemente utilizados na literatura relacionada. A descrição do que é PLN e suas técnicas pode ser encontrada na Seção 2.2. Na Seção 2.3 são explorados um pouco além conceitos e definições de Redes Neurais Artificiais. Por fim, é também apresentado na Seção 2.4, de acordo com a Psicologia, a definição de Depressão, seus sintomas e diagnóstico.

2.1 APRENDIZADO DE MÁQUINA

AM é uma subárea da Inteligência Artificial (IA), cujo objetivo é estudar e desenvolver técnicas/algoritmos que permitam sistemas aprenderem automaticamente por meio de experiência. Tal tarefa é realizada explorando-se conjuntos de dados, e visa melhorar o desempenho de uma determinada tarefa ao longo do tempo.

Algoritmos de AM utilizam um viés indutivo para derivar novos conhecimentos e prever eventos futuros (MITCHELL, 1997). O viés é uma inferência lógica que permite criar hipóteses e conclusões generalizadas com base no conjunto de dados utilizado como exemplo para treinamento. Indo além, esse viés indutivo é utilizado para restringir o espaço de busca, auxiliando assim na generalização (MONARD; BARANAUSKAS, 2003). Assim, diferentes algoritmos de AM apresentam diferentes vieses indutivos, ou seja, aprendem de

⁶<http://www.liwc.net/>

maneiras distintas.

Uma vez selecionado um algoritmo, ele passa por uma etapa de treinamento com os dados, a fim de se induzir um modelo preditivo. É o modelo o responsável por fazer previsões manipulando novos dados, não explorados em etapas prévias de treinamento.

É importante salientar que um modelo já induzido pode apresentar dois problemas referentes a sua generalização: *overfitting* e *underfitting* (ALPAYDIN, 2009). O primeiro ocorre quando o algoritmo se adapta muito bem aos dados de treinamento, mas não consegue generalizar quando realiza previsões com dados não vistos anteriormente. Esse problema ocorre principalmente quando o modelo é extremamente complexo. Já o *underfitting* se dá quando o modelo não consegue se ajustar aos dados, realizando previsões ruins até mesmo com os dados de treinamento. Ele ocorre principalmente quando o modelo obtido é excessivamente simples.

Existem diferentes tarefas de aprendizado de máquina. A Figura 1 mostra uma hierarquia simples de tarefas do aprendizado indutivo em AM. Na figura, são destacadas as tarefas de aprendizado supervisionado e não supervisionado; descritas respectivamente nas subseções 2.1.1 e 2.1.2 deste capítulo.

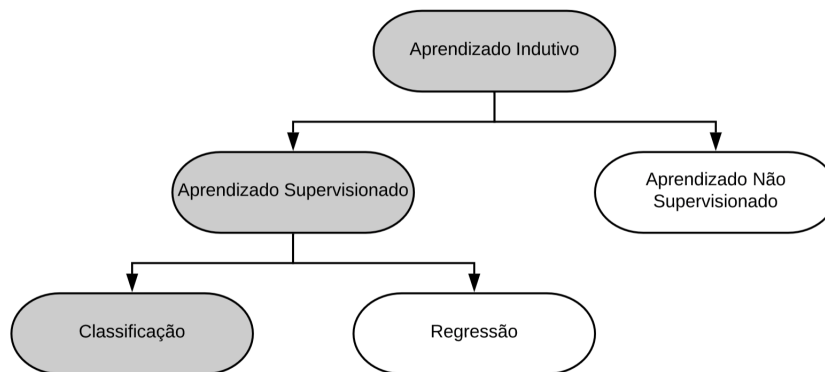


Figura 1: Hierarquia do Aprendizado Indutivo.

Fonte: Adaptada de Monard e Baranauskas (2003)

2.1.1 APRENDIZADO SUPERVISIONADO

No aprendizado supervisionado os algoritmos são treinados por meio de um conjunto de exemplos, onde cada item do conjunto está associado a um rótulo. Segundo Haykin (1999), no aprendizado supervisionado existe a figura de um professor externo, o qual possui o conhecimento do ambiente e com base neste conhecimento insere a saída (classe) desejada

para cada entrada (atributo) efetuada. Esse professor então rotula todo o conjunto de exemplos que será utilizado para treinamento do algoritmo. O objetivo é que o algoritmo extraia o conhecimento a partir do conjunto de exemplos utilizado, com isso conseguindo identificar saídas corretas para cada nova entrada, mesmo que essa não seja conhecida previamente.

O desempenho dos algoritmos de aprendizagem supervisionada está diretamente relacionado com a qualidade do conjunto de exemplos utilizado para o treinamento dos mesmos. Isso ocorre, pois caso os exemplos não sejam relevantes o suficiente, os algoritmos não serão capazes de generalizar os dados de entrada e seu desempenho será afetado.

Em geral, os algoritmos de AM podem ser classificados de acordo com o problema que os mesmos se propõe a resolver e o formato do respectivo resultado, podendo ser esses, problemas de regressão ou classificação.

Em problemas de regressão tem-se como objetivo estimar resultados *numéricos* ou *contínuos* em dados ainda desconhecidos, tendo como base dados conhecidos. Como pode ser visto na Figura 2, os itens não estão perfeitamente alinhados, mas ainda é possível identificar uma “*tendência*” de crescimento. O eixo Y pode representar, por exemplo, os valores das casas de uma determinada cidade, e o eixo X representar a quantidade de cômodos presente nas casas. Dessa forma, pode-se estimar o valor de uma casa sabendo a quantidade de cômodos que a mesma possui ($Y \sim X$).

Já a classificação visa identificar uma nova saída *categórica* a partir de um novo dado fornecido como entrada. O valor de saída é comumente chamado de “*rótulo*” ou “*classe*”. Os algoritmos de classificação usam como base exemplos já previamente rotulados. A quantidade de classes possíveis de se identificar vai depender do algoritmo/técnica utilizado para realizar a mesma.

Para se calcular o desempenho de um algoritmo em tarefas supervisionadas, compara-se a saída obtida (O), com a saída esperada (Y). Na regressão, são usadas medidas que tentam minimizar a diferença entre duas distribuições numéricas. Exemplos de medidas são medidas de erro, como o Erro Quadrado Médio (*Mean Squared Error (MSE)*), ou medidas de correlação, como o coeficiente de correlação de Pearson (PROVOST; KOHAVI, 1998). Para tarefas de classificação, a estratégia mais simples é contabilizar a porcentagem de dados classificados corretamente entre todos os dados fornecidos. Tal medida é conhecida por “*Acurácia*”, entretanto muitas outras podem ser usadas para avaliar um modelo supervisionado já induzido.

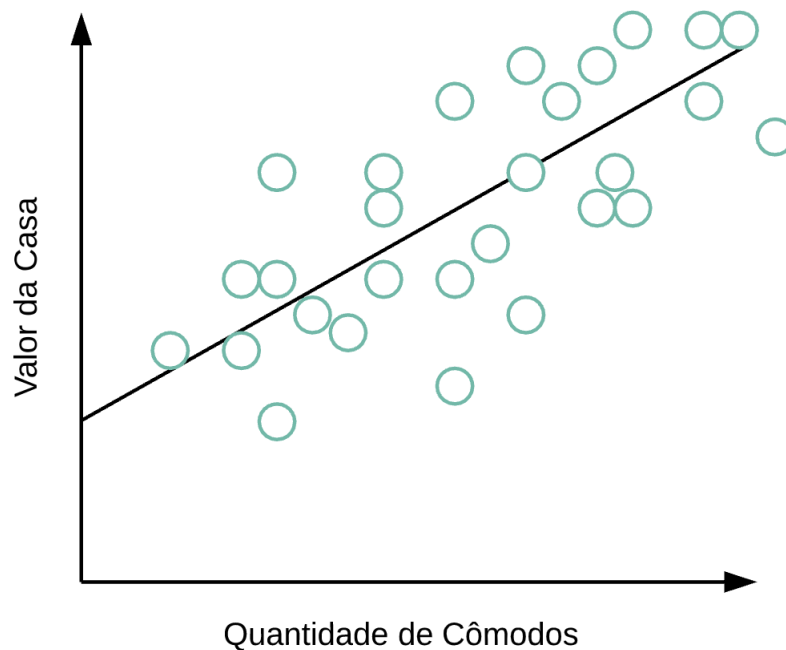


Figura 2: Exemplo de tarefa de Regressão.

Fonte: Autoria própria

2.1.2 APRENDIZADO NÃO SUPERVISIONADO

Já no aprendizado não supervisionado, os dados para treinamento não são previamente rotulados. Ou seja, não existe a figura do professor para rotular qual deve ser a saída correta para cada entrada, como ocorre no aprendizado supervisionado. Nesta categoria de AM o processo é guiado apenas pelas relações existentes entre os dados, ou seja, o aprendizado ocorre por meio da identificação de padrões e grupos (*clusters*) presentes nos dados, permitindo assim um melhor entendimento desses dados. A Figura 3, mostra um exemplo de aprendizado não-supervisionado demonstrado a divisão de exemplos em grupos distintos.

2.1.3 ALGORITMOS DE AM

Existem diversos algoritmos para reconhecimento de padrões em tarefas de aprendizado supervisionado, dentre eles a literatura específica do trabalho cita: *Support Vector Machines* (SVMs), o *Naive Bayes* (NB) e o *Random Forest* (RF), que serão descritos a seguir.

- ***Support Vector Machines***: as Máquinas de Suporte Vetorial (SVMs) são algoritmos

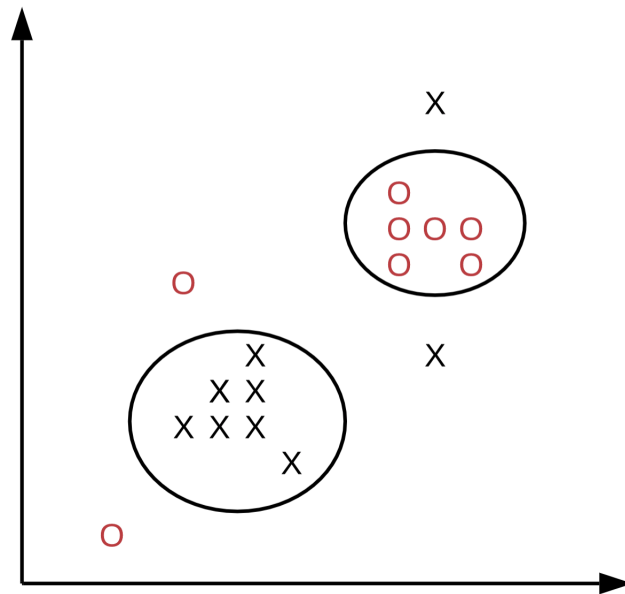


Figura 3: Exemplo de tarefa de aprendizado não supervisionado onde grupos (*clusters*) são criados com base nas características similares dos objetos.

Fonte: Autoria própria

baseados na teoria do aprendizado estatístico, proposta por Boser, Guyon e Vapnik (1992). Essa teoria estabelece princípios que visam obter uma boa generalização de classificadores, em geral.

As SVMs separam as classes com superfícies que maximizam a margem entre as mesmas, assim encontrando uma separação ótima. As entradas são convertidas para um vetor de características de alta dimensionalidade, visando separar padrões que sejam linearmente separáveis. Após o espaço ser definido, a SVM seleciona o hiperplano especial, chamado de hiperplano de margem máxima, que corresponde à maior distância de seus padrões no conjunto de treinamento, os quais são chamados de vetores de suporte.

- **Náive Bayes:** é um algoritmo que realiza aprendizado probabilístico baseado no teorema de *Bayes*. Esse teorema realiza um cálculo individual para cada variável, desconsiderando completamente a relação de cada variável com as demais, por isso leva *Náive* (ingênuo) no nome. Quando se trata da identificação de polaridade em textos, o NB considera uma sentença positiva quando a maior parte das palavras tem uma probabilidade maior de serem positivas. Por ser um algoritmo de fácil implementação e com um bom desempenho, se popularizou para classificação de textos.

- **Random Forest:** é um algoritmo de aprendizado supervisionado que cria um comitê de várias árvores de decisão (*Decision Trees* - (DTs)). O princípio fundamental por trás do RF é que um grande número de modelos não correlacionados operando de forma conjunta superam qualquer um dos modelos que o compõem de forma individual. Ou seja, os possíveis erros individuais de cada DT acontecem de forma isolada e o comitê como um todo não é influenciado por este erro, já que o resultado final é direcionado com base na maioria dos resultados individuais de cada modelo. É atualmente um dos algoritmos mais robustos para qualquer tipo de problema em AM (BREIMAN, 2001).

2.2 PROCESSAMENTO DE LINGUAGEM NATURAL

PLN é também uma subárea da IA, cujo objetivo é estudar a capacidade, formas e limitações que uma máquina tem de entender, interpretar e manipular a linguagem humana. Liddy (2001) define a PLN como teorias produzidas por diversas técnicas computacionais, que são utilizadas para analisar e representar textos providos da linguagem natural. Essas técnicas tem objetivos de processar a linguagem humana para diversas aplicações. Alguns exemplos de aplicações de PLN são: sumarização automática de textos em linguagem natural, análise morfológica e segmentação de palavras, tradução automática, *chatbots*, reconhecimento de fala, entre outros.

O PLN é subdividido em duas categorias: i) processamento e ii) geração de linguagem. A primeira tem seu foco voltado para produzir representações úteis e significativas a partir da análise da linguagem. A segunda, consiste em gerar linguagem natural a partir de uma representação. Como o foco deste trabalho é na análise textual, apenas a categoria de processamento de linguagem será discutida no texto.

Com o objetivo de facilitar a compreensão do processamento de linguagem natural, Indurkha e Damerau (2010) divide o processamento em si em cinco etapas, apresentadas na Figura 4. Cada uma destas etapas explicadas detalhadamente a seguir.

1. **Tokenização das Palavras:** consiste no processo de identificação do começo e final de uma palavra por meio da identificação dos pontos de quebra da sentença. Essas quebras são utilizadas como parâmetros de entrada de uma função, assim obtendo as palavras separadas para poder realizar as análises posteriores. Esses pontos de quebra normalmente são espaços, ou caracteres de pontuações em geral (ponto final, ponto e vírgula, vírgula, dois pontos, etc).

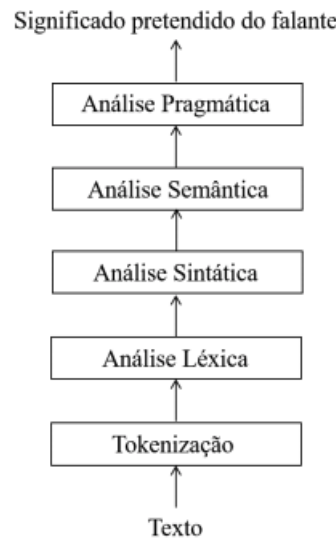


Figura 4: Estágios da análise no Processamento de Linguagem Natural.

Fonte: Adaptada de Indurkha e Damerau (2010)

2. **Análise Léxica:** realiza a análise das palavras de uma frase ou sentença de forma isolada, considerando-se especificamente sua estrutura, assim extraindo informações úteis para o restante do processamento. Decompor as palavras e detectar regras de formatação possibilita economizar espaço de armazenamento, também ajuda a aumentar a velocidade de processamento. Na execução deste processo existe a possibilidade do programa se deparar com uma palavra nova, não conseguindo reconhecer a mesma. Nessa situação, o processamento morfológico pode oferecer algumas estratégias para o tratamento da mesma (INDURKHYA; DAMERAU, 2010).

Um exemplo é o processo de decomposição, conhecido como *stemming*, que consiste basicamente em reduzir as palavras complexas morfológicamente em seus respectivos radicais. Frequentemente esta redução é realizada removendo os afixos derivacionais. Por exemplo, realizando processo de *stemming* na palavra "amigo", obtém-se o *stemm* "amig". Vale a pena ressaltar que o mesmo *stemm* é obtido com as palavras "amiga", "amigão". Este processo ajuda a reduzir o vocabulário e também a abstrair os seus significados.

3. **Análise Sintática:** tem como objetivo principal analisar a sentença de acordo com sua estrutura, verificar as formas de combinação das regras gramaticais e identificar a função de cada palavra na frase. Esta etapa visa transformar a sentença em uma estrutura hierárquica de unidades de significado da sentença, ou seja cada ponto

da estrutura é gerado a partir da aplicação de regras gramaticais. Esta fase é de fundamental importância, pois analisa a sentença como um todo, já que uma sentença não é apenas um conjunto linear de palavras.

4. **Análise Semântica:** realizada com o produto gerado pela fase anterior, esta fase tem como objetivo identificar e extrair o significado de uma sentença. Um dos principais problemas desta etapa está em como lidar com sentidos ambíguos encontrados. Essas ambiguidades podem ser: (i) léxicas - quando as palavras tem mais de um significado; (ii) referenciais - pois em alguns casos pode ser complicado até mesmo para os seres humanos entender a referência de pronomes nas sentenças; ou ainda (iii) de escopo - quando algumas palavras ou operadores de negação são aplicados em locais diferentes dentro da sentença.
5. **Análise Pragmática:** tem como objetivo verificar se o significado associado pela análise semântica é o mais apropriado no contexto geral. Usado para determinar significados que não estejam claros o suficiente. Os sistemas de PLN nunca serão completamente efetivos caso considere o texto como conjunto de sentenças independentes entre si.

Além disso, diversas técnicas são utilizadas durante o processamento de linguagem natural. A seguir são listadas e discutidas as principais técnicas de PLN.

- ***Bag-Of-Words (BoW)*:** esta técnica tem como objetivo extrair recursos dos textos utilizados como entrada para preparar a entrada dos algoritmos de AM. A técnica consiste em realizar a separação e contagem das palavras presentes em um texto;
- ***Term Frequency-Inverse Document Frequency (TF-IDF)*:** tem como objetivo expressar a relevância de uma palavra em um dado documento textual. Para isso são realizadas duas etapas, a primeira (TF) calcula a frequência que a palavra aparece em um texto. Calcula-se o número de vezes que uma palavra aparece no texto, e divide-se pela quantidade total de palavras do texto. A segunda etapa (IDF), frequência de documento inversa, realiza o cálculo por meio de um logaritmo do número de documentos na base, dividido pelo número de documentos em que a palavra em questão aparece.
- ***Pré-Processamento*:** consiste em analisar a base de dados coletada, procurando itens que devam ser normalizados ou removidos, visando um melhor entendimento

para o modelo computacional que irá processar a base. A seguir são listados algumas etapas presentes no pré-processamento.

- Remoção de *Stopwords*: *stopwords* são palavras que devem ser desconsideradas na análise pois não apresentam informações relevantes para a construção do modelo que será analisado, alguns exemplos de *stopwords* são: "o", "e", "a", "de", "que", entre outras. A remoção destas palavras diminui o processamento e melhora o entendimento da máquina.
- *Stemming*: reduz palavras ou termos ao seu radical, na maioria das vezes removendo afixos. Reduzindo assim a quantidade de palavras presentes no vocabulário, aumentando o processamento e reduzindo o número de erros, e auxiliando na abstração dos significados.

2.3 REDES NEURAIIS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) são algoritmos computacionais com processamento distribuído, criados a partir de modelos matemáticos inspirados nas estruturas neurais de organismos inteligentes. Como um algoritmo de AM, também aprendem por experiência, são capazes de realizar o processamento de uma grande quantidade de dados, e realizar reconhecimento de padrões. Uma RNA é formada a partir da interconexão de unidades computacionais, chamados neurônios. Os neurônios realizam o processamento básico da informação, semelhante ao que ocorre no sistema nervoso dos seres vivos (HAYKIN, 2009).

A Figura 5 mostra uma representação do neurônio computacional de Haykin (1999). No modelo visual, x_i representam os valores de entrada para o neurônio em questão, que podem ser providos da entrada da rede ou da saída de outros neurônios. w_{ki} representam os pesos sinápticos que estão relacionados a cada uma das conexões (sinapses). Cada peso tem como objetivo determinar qual a importância que a sua entrada referente tem para o neurônio, podendo até mesmo ser um valor negativo. A junção aditiva, somatório ou combinador linear, tem a função de realizar a somatória dos valores de entrada ponderados pelos pesos em suas respectivas sinapses. Além disso, há ainda o bias, que na figura é representado por b_k . O objetivo do bias é aumentar o valor da liquidez da função caso o mesmo seja positivo, ou diminuir a liquidez da função de ativação, quando possuir um valor negativo. Por fim, há a função de ativação, que procura restringir a amplitude do sinal de saída gerado pelo neurônio.

Existem inúmeras topologias e algoritmos de RNAs, as quais pode-se citar: os

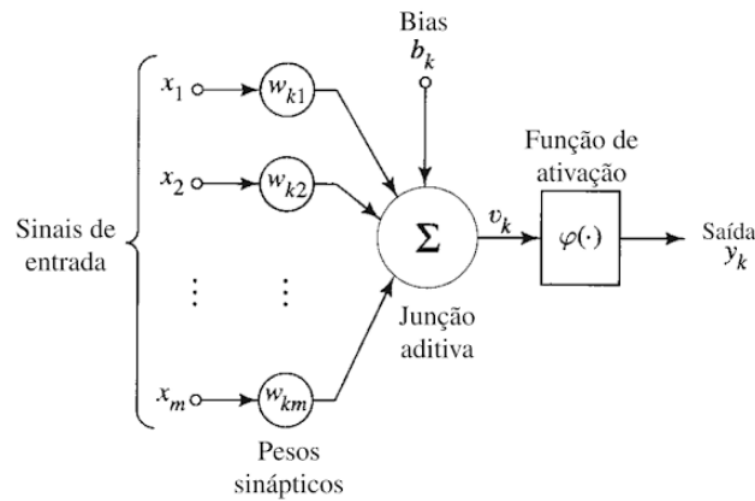


Figura 5: Modelo de neurônio artificial.

Fonte: Haykin (2009)

Perceptrons, Multilayer Perceptron (MLP), Redes Neurais Recorrentes, Redes Convolucionais, Redes de Hopfield, entre outras.

O *Perceptron* é um dos primeiros modelos de rede neural, proposto por Rosenblatt (1958). É também uma das representações mais simples de rede neural. É utilizado para realizar a classificação de padrões separando-os por meio de um hiperplano. O *Perceptron* utiliza o conceito de alimentação para frente (*feedforward*) no seu funcionamento: o sinal é transmitido apenas em uma direção, das entradas para a saída. Sua etapa de treinamento é feita por meio da apresentação de dados de treinamento para a rede, atualizando o valor dos pesos das sinapses por meio de regras de aprendizado, até que a saída realizada pela rede esteja correta.

O MLP é considerado uma extensão do *Perceptron*, mas que possui algumas características específicas. É utilizado para resolver problemas mais complexos que podem exigir a identificação de diversas classes. O MLP é formado por uma camada de entrada, uma ou mais camadas ocultas, e uma camada de saída, como ilustrado na Figura 6. Na figura podemos também verificar que as camadas são interconectadas entre si, onde cada neurônio propaga sua saída para todos os neurônios presentes na camada posterior. Isso faz com que as informações relevantes dos dados fornecidos como entrada sejam extraídas progressivamente conforme é realizado o processamento de cada uma das camadas. A função de ativação dos neurônios utilizados para compor uma rede MLP deve ser contínua e derivável (HAYKIN, 1999).

O treinamento de uma rede MLP é frequentemente realizado por meio do algoritmo

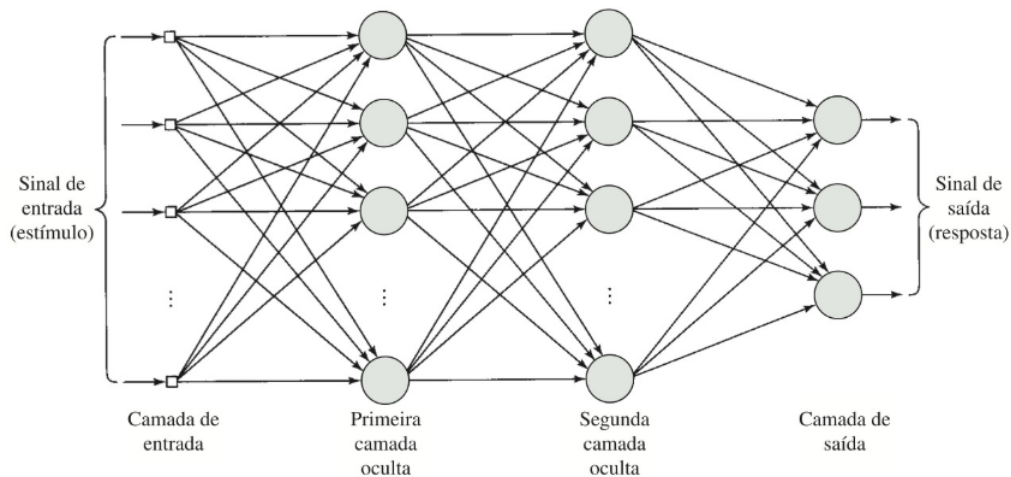


Figura 6: Topologia de uma rede Perceptron Multicamadas (MLP).

Fonte: Haykin (2009)

de Retropropagação (*Backpropagation*). Esse processo é executado em duas etapas: (i) *forward pass* - que consiste em realizar o processamento dos dados em sequência entre todas as camadas até que a camada de saída produza o resultado do processamento; e (ii) *backward pass* - onde o resultado da etapa anterior é comparado com o resultado desejado, o erro é calculado com base nessa comparação, e propagado da camada de saída até a camada de entrada, modificando os pesos das sinapses de cada um dos neurônios. O *backpropagation* utiliza a Regra Delta Generalizada que calcula os ajustes dos pesos das sinapses utilizando derivadas e noções de gradiente descendente. Nos últimos anos, as RNAs tem sido utilizadas com sucesso para tarefas de reconhecimento de diversos tipos de padrões, inclusive os textuais.

2.4 DEPRESSÃO

A depressão é o principal motivo de incapacitação mental e física da população mundial, e situa-se em quarto lugar entre as dez principais causas da carga mundial de doenças (ORGANIZATION et al., 2017). De acordo com a OMS, em 2017, 300 milhões de pessoas em todo o mundo sofriam de depressão, e embora muitos estudos estão direcionados a esta doença, as causas do seu desenvolvimento ainda não são bem compreendidas pela comunidade científica.

2.4.1 SINTOMAS

A depressão caracteriza-se principalmente por sintomas psíquicos, fisiológicos e evidências comportamentais, que são apresentados a seguir (American Psychiatric Association , 2013).

- **Sintomas Psíquicos:**

- **Humor deprimido:** forte sensação de tristeza profunda, desesperança, falta de ânimo para realização de atividades em geral. Alguns indivíduos relatam um profundo sentimento de "vazio", e de que tudo ao seu redor perdeu a graça.
- **Perda de interesse:** mesmo as atividades prediletas e frequentemente realizadas anteriormente com alegria, se tornam desinteressantes. Atividades sociais são tratadas com indiferença e passam a ter um sentido de obrigação. Em alguns casos o interesse sexual também é afetado.
- **Fatigabilidade:** cansaço constante, mesmo sem realizar atividades que exijam grande esforço ou até sem realizar nenhum esforço físico. Aumento do tempo para realização das tarefas em geral.
- **Redução da capacidade de concentração, pensamento e tomada de decisões:** decisões simples como a de escolher palavras adequadas para uma mensagem ou relatório tornam-se extremamente complexas, afetando diretamente o desempenho de uma pessoa deprimida. Frequentemente as escolhas realizadas em decisões mais complexas são prejudiciais, pois são fortemente influenciadas por expectativas negativas referentes às consequências das decisões tomadas. Lentidão no pensamento e raciocínio e *deficit* de atenção estão diretamente relacionados a falta de interesse geral.
- **Sentimento de culpa e inutilidade e redução da auto-estima:** o indivíduo deprimido constantemente se avalia de forma negativa. Se cobra excessivamente sobre eventos do dia a dia, mesmo que não possua responsabilidade sobre tais eventos. Sente-se como um peso para as demais pessoas presentes em seus vínculos inter-pessoais.

- **Sintomas Fisiológicos:**

- **Sono perturbado:** na maioria das vezes referente a insônia, que caracteriza-se pela dificuldade de dormir, ou para permanecer dormindo. Podendo ser também

caracterizado pela hipersonia, quando existe uma sonolência excessiva durante grande parte do dia.

- **Alteração do apetite:** perda ou aumento de apetite, influenciando diretamente em perda ou ganho significativos de peso. O aumento de apetite é muitas vezes em relação a carboidratos e doces.

- **Evidências comportamentais:**

- **Retraimento social:** atividades sociais são frequentemente afetadas, pois passam a parecer uma obrigação. Esse retraimento esta diretamente ligado a perda de interesse generalizado.
- **Retardo ou agitação psicomotora:** sensação de peso nos membros. Lentidão na fala, atividade corporal e pensamento. Diminuição do volume da voz, e dificuldade para diversificar assuntos falados. Existe também a possibilidade dos indivíduos sofrerem de agitação psicomotora, onde os mesmos apresentam movimentos involuntários e/ou não conseguem se manter parados em um mesmo lugar.
- **Pensamentos de morte ou ideação suicida:** o indivíduo deprimido pode vir a ter desejos de morrer, podendo chegar até mesmo a ideações suicidas, definindo um plano elaborado contendo forma, data e local da morte. Essa situação esta diretamente ligada ao desejo de por fim ao sofrimento que o transtorno causa, tanto no próprio indivíduo, bem como ao peso que o indivíduo julga exercer sobre as pessoas que fazem parte de seus círculos inter-pessoais. Muitas vezes é vista como a única solução para acabar com o sofrimento, visto que existe uma enorme desesperança de que este sofrimento possa ser extinguido ou até mesmo amenizado.

2.4.2 DIAGNÓSTICO

Os sintomas da depressão muitas vezes estão relacionados com outras doenças, e também são confundidos como problemas cotidianos e passageiros, que na maioria das vezes, acabam passando despercebidos, principalmente pela falta de informação da população em geral.

O DSM-V considera nove sintomas para diagnosticar um indivíduo com depressão, esse diagnóstico pode ser classificados em três grupos: (*i*) depressão menor; (*ii*) distímia; e (*iii*) depressão maior. Os sintomas são apresentados na Tabela 1.

Tabela 1: Critérios DSM-V

Sintomas Primários	Sintomas Secundários
Humor deprimido	Sentimento de culpa ou inutilidade
Perda de interesse	Problemas de concentração
	Fadiga
	Sono perturbado
	Retardo ou agitação psicomotora
	Aumento ou redução significativa de peso
	Recorrentes ideias suicida e de morte

Fonte: Autoria própria

Caracteriza-se **depressão menor** quando o quadro depressivo apresentar por duas ou mais semanas, de dois a quatro sintomas, contanto que entre esses sintomas esteja incluído pelo menos um sintoma primário. A **distímia** é caracterizada quando o quadro depressivo apresenta de três a quatro sintomas, incluindo humor deprimido, por no mínimo dois anos. Já a **depressão maior** caracteriza-se quando o quadro depressivo apresenta por duas semanas ou mais, cinco sintomas estando incluso nesses, pelo menos um sintoma primário American Psychiatric Association (2013).

2.5 CONSIDERAÇÕES FINAIS

No decorrer deste capítulo foram apresentados conceitos e técnicas relevantes para a execução do presente trabalho, já que segundo a literatura encontrada no contexto do projeto, as técnicas de PLN são utilizadas para normalizar e pré-processar os dados textuais, e a classificação desses dados é realizada por algoritmos de AM ou RNAs. Os conceitos referentes à depressão demonstram quais são os seus sintomas, como são categorizados e ainda como esses apresentam-se nos indivíduos.

3 MAPEAMENTO SISTEMÁTICO

Esse capítulo apresenta um Mapeamento Sistemático (MS) conduzido para a elaboração deste trabalho, e está organizado da seguinte forma: o planejamento juntamente com a estratégia de busca e critérios de seleção dos estudos são apresentados na Seção 3.1. Na Seção 3.2 é descrito como foi a condução do processo de seleção dos estudos. Na seção 3.3 são discutidos os resultados alcançados. As ameaças à validade referentes a execução deste MS são demonstradas na Seção 3.4. Por fim, na Seção 3.5 são apresentadas as considerações finais deste Capítulo.

Um MS, tem como objetivo identificar, categorizar e avaliar estudos, assim reunindo o máximo de informações possíveis de uma determinada área de pesquisa, possibilitando a identificação de lacunas que indicam possíveis áreas de pesquisa para serem focadas em futuras Revisões Sistemáticas da Literatura (BUDGEN et al., 2008).

O objetivo deste MS é identificar as principais evidências disponíveis na literatura referentes a classificar palavras, sentenças ou textos depressivos utilizando AS. O processo utilizado para a realização deste MS é o proposto por Kitchenham et al. (2010) que propõe que sejam executadas três fases, são elas: (i) Planejamento; (ii) Condução; e (iii) Análise;

3.1 PLANEJAMENTO

Esta fase do MS consiste na formulação das questões de pesquisa e a definição da estratégia de busca, dos critérios de inclusão e exclusão. Para a condução deste MS as seguintes Questões de Pesquisa (QPs) foram desenvolvidas:

- **QP₁: Quais técnicas têm sido utilizadas para análise de sentimentos e identificação de perfis depressivos em redes sociais?**
- **QP₂: Quais tipos de rotulação de base de dados têm sido utilizadas para análise de sentimentos e identificação de perfis depressivos em redes sociais?**

Com a QP₁ espera-se identificar diferentes tipos de técnicas relacionadas à AS que têm sido utilizadas para identificar perfis depressivos em textos em redes sociais. A QP₂ visa identificar de onde são extraídas as informações textuais para criação das bases de dados utilizadas para análise de sentimentos e identificação de perfis depressivos, e como estas bases estão sendo rotuladas, por exemplo rotulação manual, rotulação automática ou outros tipos.

A partir das QPs, foi gerada a *string* de busca que consistiu na combinação das palavras “*sentiment analysis*” e “*depression*” e seus respectivos sinônimos, todas no idioma inglês. Portanto, a seguinte *string* de busca foi gerada: (“*sentiment analysis*” OR “*data analysis*” OR “*data mining*”) AND (*depression* OR “*depressive mood*”). Este processo está representado na Figura 7.

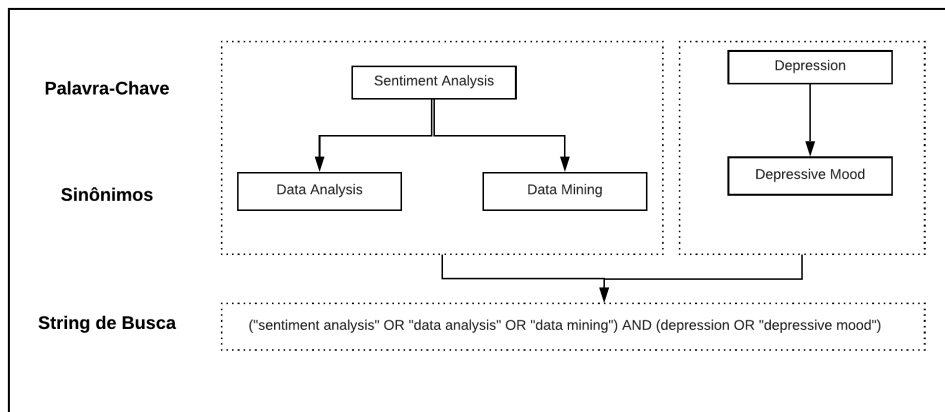


Figura 7: String de Busca da Questão de Pesquisa.

Fonte: Autoria própria

Com a *string* definida, foi realizada a busca automática em quatro bases de dados eletrônicas (*ACM Digital Liabrary, IEEE Xplore, Scopus e Springer*) de acordo com as diretrizes propostas por Brereton et al. (2007). É importante destacar que nesta etapa foram selecionadas apenas quatro bases de busca, devido a quantidade de tempo limitada para a elaboração do mesmo.

Para identificar a relevância dos estudos encontrados, foram definidos os seguintes Critérios de Inclusão e Exclusão:

- **Critérios de Inclusão (CI):**

- **CI₁:** estudos primários que apresentem técnicas para auxiliar a análise de sentimentos e a identificação de perfis depressivos em redes sociais;

- **CI₂**: estudos primários que apresente pelo menos uma forma de categorizar sentimentos a partir de dados textuais.

- **Critérios de Exclusão (CE)**

- **CE₁**: estudos primários que não sejam *full paper* ou *short paper* (pôsteres, tutoriais, dissertações e teses);
- **CE₂**: estudos primários que não estejam escritos em inglês ou português;
- **CE₃**: estudos primários que não estejam disponíveis; e
- **CE₄**: estudos primários que seja uma versão anterior de um estudo mais completo sobre a mesma investigação.

3.2 CONDUÇÃO

O processo de condução do MS apresentado na Figura 8 consistiu na execução da *string* de busca nas bases selecionadas. É importante ressaltar que as buscas foram conduzidas no período de março até setembro de 2019.

Na fase 1 foram retornados 124 estudos na *IEEE*, 31 estudos na *ACM*, 88 estudos na *Springer* e 229 no *Scopus*. A partir dos 472 estudos retornados das fontes de busca, os estudos repetidos foram identificados e excluídos, os títulos e resumos dos estudos restantes (402) foram lidos e neles aplicados os critérios de inclusão e exclusão (Fase 2). Na Fase 3, foi realizada a leitura da introdução e conclusão dos 77 estudos que foram incluídos durante a fase 2 e neles foram aplicados os critérios de exclusão e inclusão. Na fase 4, os 28 estudos selecionados na etapa anterior foram lidos na íntegra, acarretando na exclusão de 3 estudos, totalizando 25 estudos primários incluídos.

3.3 RESULTADOS E DISCUSSÕES

Esta seção visa apresentar e discutir os resultados alcançados por meio do processo de classificação dos estudos. Os 25 estudos selecionados são apresentados na Tabela 2, onde na primeira coluna tem-se a referência do estudo. Na segunda coluna a base de dados em que o estudo foi identificado é exibida. Por fim, na terceira coluna encontra-se o tipo de publicação do estudo, que pode ser Conferência ou Periódico.

De acordo com os resultados, é notável que a maior concentração dos estudos incluídos foi por meio da base *IEEE* com 44% (11/25), seguida da *Scopus* com 32%

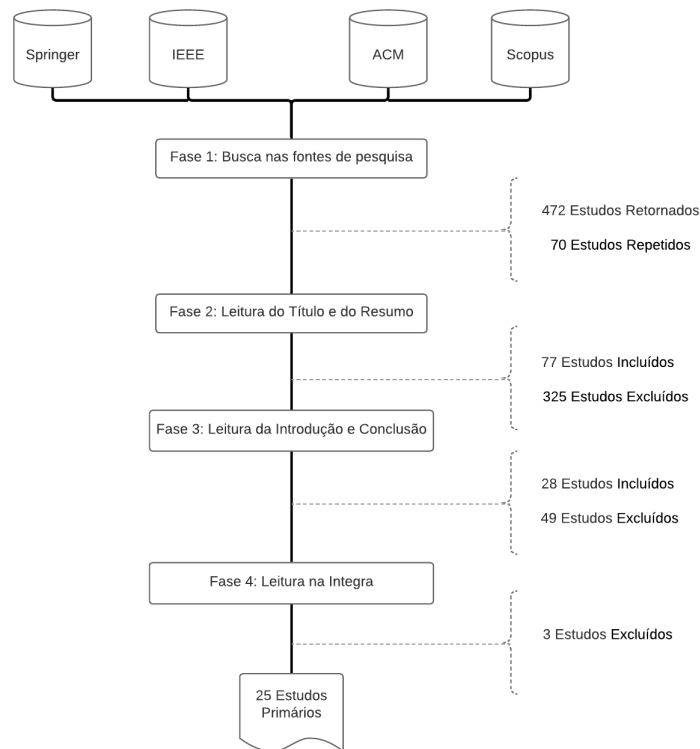


Figura 8: Etapas do Processo de Seleção dos Estudos.

Fonte: Autoria própria

(8/25), *Springer* com 16% (4/25) e *ACM* com 8% (2/25). Outra informação extraída foi o percentual de estudos primários selecionados de acordo com o tipo de publicação. 68% (17/25) dos estudos selecionados foram publicados em , conferências enquanto que 32% (8/25) foram em periódicos.

A partir dos 25 estudos primários analisados é possível observar que o primeiro estudo foi publicado no ano de 2013. A partir deste pode-se observar uma crescente em relação aos estudos publicados em cada ano, este efeito pode estar atribuído a popularização do tema, bem como uma maior preocupação da população com relação a depressão e suas consequências na sociedade global. O ano de 2017 é o que mais se destaca com seis publicações realizadas. Nos anos de 2018 e 2019 houve uma queda no número de estudos publicados, conforme pode ser visto na Figura 9.

Na Figura 10 são apresentados o número de estudos por países. Para classificar o país do estudo, foi utilizado o país de origem do primeiro autor do estudo. A Austrália, por exemplo, se destaca por conter cinco estudos publicados, seguida pelo Estados Unidos com três estudos. Bangladesh, Coreia do Sul, Marrocos e Taiwan possuem duas publicações cada. Os demais países contem apenas um estudo publicado. Também pode-se observar

Tabela 2: Overview dos Estudos Primários selecionados durante o MS.

Referência	Base	Tipo
(Hassan et al., 2017)	IEEE	Conferência
(Deshpande; Rao, 2017)	IEEE	Conferência
(Islam et al., 2018)	IEEE	Conferência
(Wang; Zhang; Sun, 2013)	IEEE	Conferência
(Seah; Jin Shim, 2018)	IEEE	Conferência
(Aldarwish; Ahmad, 2017)	IEEE	Conferência
(Nguyen et al., 2014)	IEEE	Periódico
(Rosa et al., 2019)	IEEE	Periódico
(Noureen; Qamar; Ali, 2017)	IEEE	Conferência
(Larsen et al., 2015)	IEEE	Periódico
(Dao et al., 2015)	IEEE	Conferência
(ZUORBA; OLAN; CANTARA, 2017)	ACM	Conferência
(VEDULA; PARTHASARATHY, 2017)	ACM	Conferência
(ISLAM et al., 2018)	Springer	Periódico
(WU et al., 2018)	Springer	Periódico
(HASAN; RUNDENSTEINER; AGU, 2019)	Springer	Periódico
(NGUYEN et al., 2017)	Springer	Periódico
(TAO et al., 2016)	Scopus	Conferência
(LEIVA; FREIRE, 2017)	Scopus	Conferência
(BIRJALI; BENI-HSSANE; ERRITALI, 2017)	Scopus	Conferência
(BRIAND; ALMEIDA; MEURS, 2018)	Scopus	Conferência
(TUNG; LU, 2016)	Scopus	Periódico
(BOUKIL et al., 2019)	Scopus	Conferência
(MA; WANG; ZHANG, 2017)	Scopus	Conferência
(Keumhee Kang; Chanhee Yoon; Eun Yi Kim, 2016)	Scopus	Conferência

Fonte: Autoria própria

que o Brasil possui apenas um estudo publicado, reiterando que pouco tem sido investigado nesta área no cenário brasileiro utilizando bases de dados na língua portuguesa. Dentre os autores dos estudos, apenas Nguyen et. al possuem mais de uma publicação (Nguyen et al., 2014; NGUYEN et al., 2017).

3.3.1 QP₁: QUAIS TÉCNICAS TÊM SIDO UTILIZADAS PARA ANÁLISE DE SENTIMENTOS E IDENTIFICAÇÃO DE PERFIS DEPRESSIVOS EM REDES SOCIAIS?

Para responder a QP₁, na Figura 11 são sumarizadas 39 técnicas diferentes utilizadas para AS no contexto deste trabalho. Dentre tais técnicas é importante destacar o algoritmo de classificação SVM que foi utilizado em 11 estudos, seguido pelo algoritmo NB, utilizado em nove estudos. Outros algoritmos que se destacam são o algoritmo

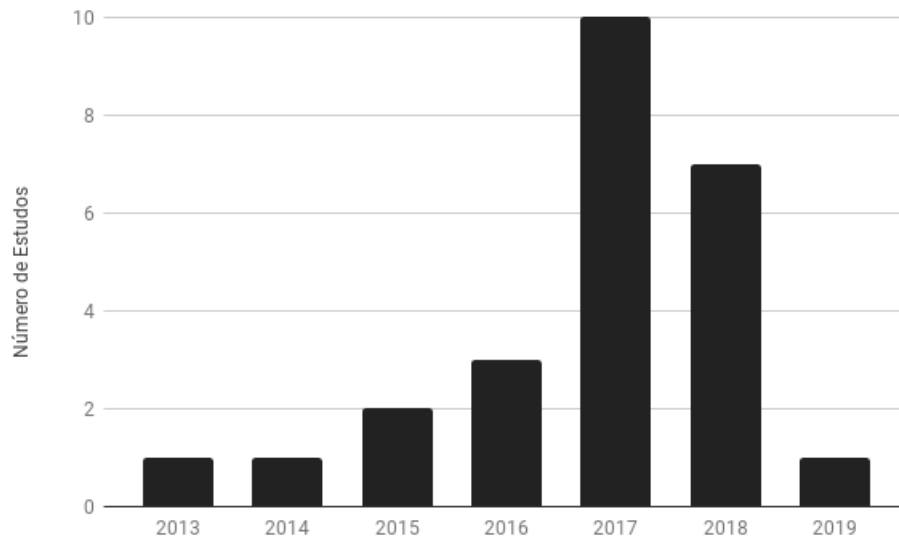


Figura 9: Estudos por Ano de Publicação.

Fonte: Autoria própria

K Nearest Neighbor (KNN) que aparece sendo utilizado por cinco estudos e o de DTs, utilizado por quatro estudos.

Ainda pode-se observar que diversos estudos utilizaram técnicas de PLN para realizar o pré-processamento dos dados de extraídos das fontes de dados, como remoção de *stopwords*, *stemming* e entre outras. Mas alguns estudos (Larsen et al., 2015; ZUORBA; OLAN; CANTARA, 2017; TUNG; LU, 2016) utilizaram dicionários léxicos como principal forma de identificar os sentimentos nos textos, ou ainda com o objetivo de auxiliar na vetorização dos dados visando facilitar o processamento realizado por técnicas de AM.

Na Tabela 3 são apresentadas as técnicas utilizadas por cada estudo na segunda coluna. Os estudos que realizaram experimentos, tem os mesmos exibidos na terceira coluna. Por fim, na ultima coluna é apresentado o desempenho que os estudos alcançaram com cada técnica utilizada.

Tabela 3: Análise dos Estudos I

Estudo	Técnicas	Experimento	Acurácia
E1	SVM; NB; ME	A performance foi avaliada em duas bases de dados, <i>twitter</i> e <i>20newsgroups</i>	SVM: 91%; NB: 83%; ME: 80%

Continuação na próxima página

Tabela 3 – Continuação da página anterior

Estudo	Técnicas	Experimento	Acurácia
E2	NB; SVM	-	NB: 83%; SVM: 79%
E3	KNN	Foi realizado utilizando <i>10-fold cross-validation</i> nas bases de teste	60 - 70%
E4	<i>Linear Regression</i>	Foram analisados dados de 86 usuários do Sina Micro Blog	95%
E5	<i>Latent Dirichlet Allocation (LDA)</i>	-	-
E6	SVM; NB	O modelo é testado com conjuntos de dados obtidos das redes sociais dos pacientes	SVM: 67%; NB: 63%
E7	Lasso	-	-
E8	NB; SVM; RF; CNN BLSTM-RNN -Softmax; CNN BLSTM-RNN -SVM;	146 pessoas participaram da fase de testes, totalizando 25.192 sentenças coletadas e avaliadas	NB: 69%; SVM: 84%; RF: 81%; CNN BLSTM-RNN - Softmax: 89%; CNN BLSTM-RNN - SVM: 87%
E9	NB; SVM; Lasso; <i>Bayes Net</i> ; DT; KNN; <i>Regression Analysis</i> ; <i>Binary Logistic Regression</i> ; <i>Rocchio Classifier</i> ; <i>Perceptron</i> ; <i>Bagging</i> ; MANCOVA; <i>K-means Algorithm</i> ; <i>Latent Semantic Analysis</i>	-	-
E10	PCA juntamente com ANEW e LIWC	-	-

Continuação na próxima página

Tabela 3 – Continuação da página anterior

Estudo	Técnicas	Experimento	Acurácia
E11	<i>Hierarchical Dirichlet Process</i>	-	-
E12	<i>Logistic Regression</i> juntamente com Emolex/IsaWika Lexicon; LDA	127 alunos tiveram suas contas de redes sociais avaliadas, após isso, foram questionados se os resultados foram precisos. Dos 127, 116 concordaram que o modelo foi assertivo.	99.92%
E13	DT	-	90%
E14	SVM; KNN; DT	-	60% a 80%
E15	RNN	-	-
E16	SVM; NB; DT	-	90%
E17	Lasso; NB; SVM; Logistic Regression	-	-
E18	PLN	-	-
E19	NB; SVM; <i>Logistic Regression</i> ; KNN	-	-
E20	IBK, J48, SMO; PART	-	-
E21	PLN; SVM; RF	-	-
E22	PLN juntamente com Lexicos	-	-
E23	CNN; NB; KNN	-	-
E24	PLN juntamente com LDA	-	-
E25	SVM; LIWC	-	-

Fonte: Autoria própria

3.3.2 QP₂: QUAIS SÃO OS TIPOS DE ROTULAÇÃO DAS BASES DE DADOS QUE TÊM SIDO UTILIZADAS?

Para a QP₂, é importante destacar que em alguns casos, o tipo de rotulação escolhida pelos autores dos estudos estava relacionada às fontes de dados utilizadas para extração das informações. A seguir é realizada uma análise dos dados identificados, bem

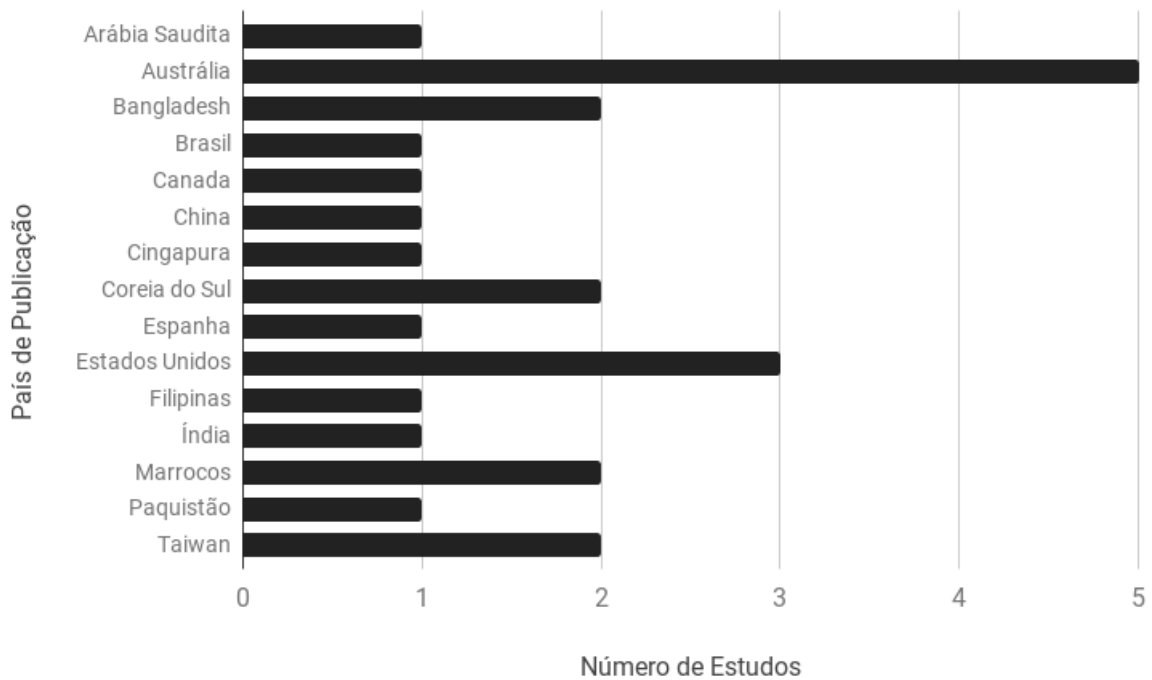


Figura 10: Estudos por País de Publicação.

Fonte: Autoria própria

como uma breve descrição das fontes de dados. Por fim, são apresentadas as formas de rotulação utilizadas por cada um dos estudos.

Na Tabela 4 são exibidas a origem dos dados utilizados para criação das bases de dados (coluna 2). Na terceira coluna é demonstrado o tamanho da base. Já na quarta coluna é representado como foi realizada a rotulação da base de dados. Por fim, na quinta coluna é informado se a base de dados está disponível

Tabela 4: Análise dos Estudos II

Estudo	Base	Tamanho	Rotulação	Disponível
E1	<i>Twitter;</i> <i>20newsgroups;</i>	Não informado	Não informado	<i>Twitter:</i> Não; <i>20newsgroups:</i> Sim;
E2	<i>Twitter</i>	10000 postagens	Automática	Não
E3	<i>Facebook</i>	7145 comentários	Manual	Não
E4	Sina Micro-Blog	27518 postagens	Manual	Não
E5	<i>Reddit</i>	21000 comentários	Manual	Não

Continuação na próxima página

Tabela 4 – Continuação da página anterior

Estudo	Base	Tamanho	Rotulação	Disponível
E6	<i>Twitter;</i> <i>LiveJournal;</i> <i>Facebook;</i>	<i>Twitter:</i> 2354; <i>LiveJournal:</i> 2132; <i>Facebook:</i> 2287;	Manual	Não
E7	<i>LiveJournal</i>	Base clínica: 38401; Base de Controle: 229563;	Automática	Não
E8	<i>Facebook</i>	27308 postagens	Automática	Não
E9	<i>Facebook</i>	Não informado	Não informado	Não
E10	<i>Twitter</i>	2,73 x 10 ⁹	Automática	Sim
E11	<i>LiveJournal</i>	268400	Automática	Não
E12	<i>Twitter;</i> <i>Facebook</i>	<i>Twitter:</i> 215100; <i>Facebook:</i> 51500;	Automática	Não
E13	<i>Twitter</i>	Não informado	Automática	Não
E14	<i>Facebook</i>	7145 comentários, sendo que 58% indicam depressão e 42% não indicam depressão	Automática	Não
E15	<i>Facebook</i> , dados do CES-D e dados de ambiente	873524 postagens	Automática	Não
E16	<i>Twitter</i>	Não informado	Automática	Não
E17	<i>LiveJournal</i>	Base clínica: 25012; Base de Controle: 25012;	Automática	Não
E18	<i>Twitter;</i> Vocabulários <i>online</i>	Não informado	Manual	Não
E19	<i>Reddit</i>	Não informado	Manual	Sim

Continuação na próxima página

Tabela 4 – Continuação da página anterior

Estudo	Base	Tamanho	Rotulação	Disponível
E20	<i>Twitter</i>	Não informado	Automática	Não
E21	<i>Reddit</i>	Risco de depressão: 49557; Sem risco: 481837	Manual	Sim
E22	Fórum de discussão PTT	18000 postagens	Manual	Não
E23	Notas escritas em Árabe	Indicação suicida: 2000 palavras; Sem indicação: 3550 palavras	Manual	Não
E24	<i>Twitter</i> e <i>blogs</i> de depressão	54 milhões de postagens	Automática	Não
E25	<i>Twitter</i>	Não informado	Manual	Não

Fonte: Autoria própria

Com base na análise realizada e nas informações apresentadas na Tabela 4, observa-se que postagens feitas no *Twitter* são utilizadas para classificação de depressão em 35,5% (11/25) dos estudos. Postagens realizadas no *Facebook* são utilizadas em 22,6% (7/25) dos estudos.

A comunidade *LiveJournal* é utilizada em 12,9% (4/25) dos estudos. Por outro lado, as postagens realizadas no *Reddit* compõem a base de dados de 9,7% (3/25) dos estudos analisados. Por fim, postagens realizadas no Sina Micro-Blog, no *blog* PTT, *blogs* de depressão em geral, o *dataset 20newsgroups*, coleções de notas em árabe e informações de vocabulários *online* são utilizadas em 3,2% (1/25) dos estudos.

A utilização do *Twitter* e *Facebook* em 58,1% (18/25) dos estudos, pode estar ligada ao fato de que ambas as redes sociais, possuem APIs públicas para coleta de dados, facilitando assim a construção das bases de dados tanto de treinamento com as de testes.

Postagens do *LiveJournal* também são frequentemente utilizadas, pois o *blog* possui um conjunto predefinido de 132 humores disponíveis para os usuários marcarem em suas postagens, auxiliando assim na classificação das bases. *LiveJournal* é uma comunidade virtual onde os usuários da Internet podem manter um blog, um jornal ou um diário. Cada usuário tem um perfil com o qual pode interagir com comunidades e outros usuários.

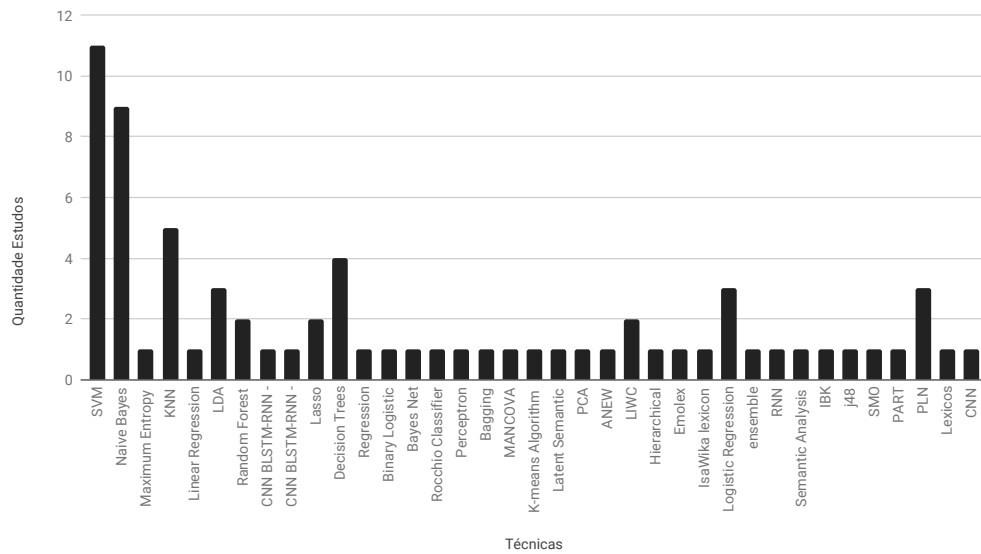


Figura 11: Técnicas identificadas nos estudos primários.

Fonte: Autoria própria

Reddit é ma comunidade de fóruns no qual os usuários podem criar tópicos para outros usuários interagirem. Sina Micro-Blog é um micro *blog* similar ao *Twitter* existente na China. PTT é um fórum de discussões chinês. Por fim, o *Center for Epidemiologic Studies Depression Scale* (CES-D) é um questionário de auto-relatos, que tem como finalidade medir a gravidade de sintomas depressivos na população em geral.

A análise referente as bases de dados citadas pelos estudos possibilitou identificar que apenas o estudo de Rosa et al. (2019) teve sua base criada em português. Isto indica a defasagem de estudos focados em analisar sentimentos no contexto de doenças mentais em português. Isso pode estar ligado ao fato da complexidade de realizar análises neste idioma, principalmente em publicações de redes sociais, já que o número de publicações com ambiguidade, figuras de linguagem, abreviações, entre outros aspectos é elevado neste contexto.

Para responder a QP₂, na Figura 12 são sumarizadas os dois tipos de rotulação das bases de dados. É importante destacar que somente dois estudos (Hassan et al., 2017; Noureen; Qamar; Ali, 2017) não informaram como realizaram a rotulação das bases de dados. Os demais estudos foram divididos nas seguintes categorias: Rotulação Manual e Rotulação Automática, e são descritos a seguir.

- **Rotulação Manual:**

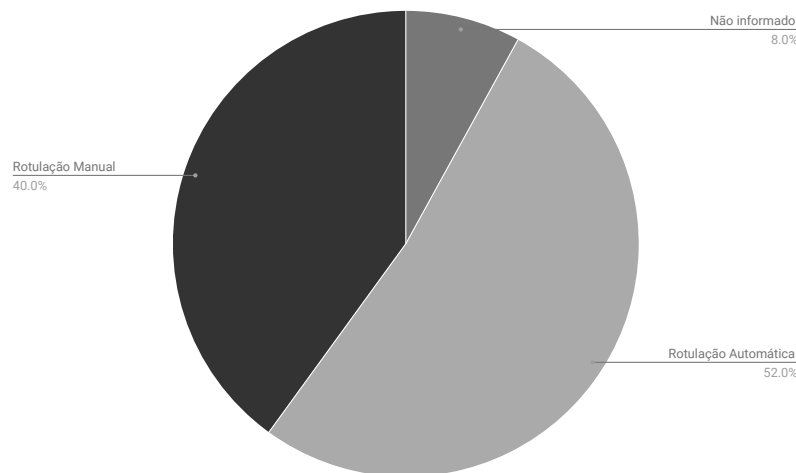


Figura 12: Tipos de rotulação identificadas nos estudos primários.

Fonte: A autoria própria

O estudo de Islam et al. (2018) rotulou a base de dados com o auxílio de dois especialistas, que rotularam a base entre comentários que indicavam depressão e que não indicavam depressão, ao final, 58% dos comentários indicavam depressão e 42% não indicavam.

Já o estudo de Wang, Zhang e Sun (2013) utilizou postagens de 50 usuários depressivos, e 50 de usuários não depressivos. Os usuários foram extraídos do conjunto de 180 usuários (90 depressivos e 90 não depressivos) que foram detectados por um grupo de psicólogos em (WANG et al., 2013).

Em Seah e Jin Shim (2018) foram coletados comentários do *Reddit* que continham termos como "*feel useless*", "*hang myself*", métodos de suicídio, nomes de medicamentos, entre outros. Após isso os pesquisadores inspecionaram manualmente os dados para garantir a relevância dos mesmos.

Aldarwish e Ahmad (2017) classificaram 2073 postagens como deprimidas e 4700 como não deprimidas, após isso as postagens deprimidas foram categorizadas entre os nove sintomas definidos pelo DSM-V.

Para criar a base de dados, Tao et al. (2016) realizaram buscas manuais de termos referentes a depressão no *Twitter* e em vocabulários *online*. Após a etapa de coletar dados, a base gerada foi analisada por vários psicólogos.

Leiva e Freire (2017) utilizaram a base de dados fornecida por Losada e Crestani (2016), onde estes realizaram a coleta de postagens de diferentes usuários do *Reddit*. Após

a coleta, os dados foram rotulados de forma manual onde os possíveis rótulos eram: se um usuário estava ou não com risco de possuir depressão.

O estudo de Briand, Almeida e Meurs (2018) utilizou a base de dados disponibilizada por Losada, Crestani e Parapar (2017), a qual foi criada recuperando informações de usuários do *Reddit*, onde parte do conteúdo da base foi rotulado manualmente por psicólogos, entre usuário com risco (deprimido) ou sem risco (não deprimido).

No estudo de Tung e Lu (2016) foi coletado um total de 18.000 postagens em chinses do fórum PTT. Dessas, 724 foram selecionados para treinamento e teste, essas postagens foram segmentadas e marcadas. Em seguida, foram rotuladas por três alunos de mestrado do *Institute of Behavioral Medicine College of Medicine, National Cheng Kung University*, como verdadeiras ou falsas referente a possuírem tendência à depressão.

A base de dados do estudo de Boukil et al. (2019) foi criada a partir de postagens escritas em árabe, coletadas de diversos *sites* (*Facebbok*, *Twitter*, *blogs* e fóruns). Todos os dados coletados foram manualmente analisados e rotulados nas seguintes classes: notas com indicação suicida e notas sem indicação suicida.

Keumhee Kang, Chanhee Yoon e Eun Yi Kim (2016) coletaram postagens do *Twitter* que continham as seguintes palavras ou sentenças: "*Christmas*", "*Suicide*", "*I feel relaxed*", "*I feel good*", "*want to die*", "*I feel stressed*", "*I feel sad*", "*kill myself*" e "*want to commit suicide*"). Após a coleta os dados foram rotulados por três pessoas com experiência no uso de redes sociais. Para cada postagem era possível aplicar os seguintes rótulos: negativo, neutro ou positivo. Foram incorporados a base de dados somente postagens com os rótulos designados por duas pessoas no mínimo, correspondendo a 89,9%.

- **Rotulação Automática:**

Deshpande e Rao (2017) separaram 80% do *dataset* para treinamento, e 20% para a realização dos testes. Os *tweets* coletados para a base de treinamento deveriam conter palavras como "*depressed*", "*suicide*" e "*hopeless*". Já para a base de testes os *tweets* foram coletados de forma aleatória.

Nguyen et al. (2014) criaram uma base clínica, utilizando a "busca de comunidades por interesse" disponível no *LiveJournal*, utilizando o termo "*depression*" como entrada da busca. O resultado foram 24 comunidades, que somadas, continham 38,401 postagens. Também criou uma base de controle utilizando a mesma busca supracitada, utilizando

como entrada as cinco categorias mais populares do blog, o resultado desta foi busca foi de 229,563 postagens contidas em 23 comunidades.

O estudo de Rosa et al. (2019) filtrou sentenças providas de postagens do *Facebook* que continham expressões como "odeio minha vida", "me sentindo triste", "estou estressado", entre outras. Também utilizou um conjunto de sentenças positivas para classificar como "não depressão". A partir disso, 27,308 mensagens foram coletadas, onde 23,7% é referente à depressão, 26,18% é referente à estresse, por fim 50.1% são mensagens que não tem relação com estresse e nem com depressão.

Para identificar os potenciais *tweets* emocionais, Larsen et al. (2015) utilizaram os vocabulários léxicos ANEW e o LIWC, onde o primeiro fornece um conjunto de avaliações emocionais para uma grande quantidade de palavras em inglês, e o segundo é uma ferramenta proprietária que auxilia na extração de características das palavras, e na identificação da polaridade que as mesmas apresentam, assim avaliando o grau que cada texto tem em diferentes categorias.

Dao et al. (2015) utilizaram os conjuntos de humor definidos nas postagens do *LiveJournal* para catalogar a base, e a dividiu em três grupos: depressão com 38,400 postagens, autismo que contém 10,000 postagens e geral que possui 220,000 postagens.

No estudo de Zuorba, Olan e Cantara (2017) foi realizada uma busca nos últimos 100 *tweets* de 2,151 usuários, e as últimas 100 postagens de 515 usuários públicos do *Facebook*, após isso a base foi rotulada utilizando vocabulários léxicos de duas línguas diferentes Emolex(Inglês) e IsaWika Lexicon(Tagalo). Esse vocabulários léxicos foram utilizados para determinar a tristeza de uma pessoa.

Vedula e Parthasarathy (2017) identificaram usuários do *Twitter* com depressão, com auxílio médico. Após essa identificação a base de treinamento foi montada com os *tweets* do período de julho de 2016 a janeiro de 2017 dos usuários supracitados, bem como dos seus seguidores.

Islam et al. (2018) coletaram os dados do *Facebook* com o auxílio do NCapture. Após a coleta, foi utilizado o vocabulário léxico LIWC para calcular o valor por meio do padrão escalas do LIWC2015.

A base de dados do estudo de Wu et al. (2018) foram constituídas com a co-relação de postagens coletadas do *Facebook* de universitários de Tawian que responderam o questionário *online* CES-D. Os resultados de triagem dos testes do CES-D. E ainda com dados de ambiente disponíveis ao público tais como: histórico de acidentes de transito,

acidentes em geral, densidade da população, entre outros. Na base de dados deste estudo cada usuário é rotulado como deprimido ou não de acordo com a pontuação total atingida no CES-D, onde atingir mais de 60 pontos significa uma maior frequência de experiências depressivas.

Hasan, Rundensteiner e Agu (2019) utilizaram palavras simples para realizar a busca de informações chamadas também de uni-gramas. Após a coleta das informações, para selecionar e manter no conjunto de dados apenas palavras com sinais de emoções, foram utilizados diferentes dicionários léxicos tais como: LIWC, ANEW, AFINN. Este último foi criado para incluir uma nova lista de palavras, focado especialmente para microblogs.

O estudo Nguyen et al. (2017) utilizou a "busca de comunidades por interesse" presente no *LiveJournal*. Foram selecionadas apenas as comunidades que possuíam no mínimo 200 postagens e pelo menos uma atualização no mês anterior a busca. Ao final foram selecionadas 24 comunidades com um total de 38.401 postagens, as quais foram agrupadas nos seguintes subgrupos: depressão, transtorno bipolar, auto-mutilação, luto e suicídio. Para que criar a base de dados de forma balanceada foram escolhidos aleatoriamente 1.000 postagens de cada grupo.

Os autores Birjali, Beni-Hssane e Erritali (2017) optaram por buscar postagens do *Twitter* e as rotular de forma automática utilizando para isso o método de similaridade semântica *WordNet* tendo como base um "vocabulário da depressão" definido previamente.

Para criar a base de dados do estudo de Ma, Wang e Zhang (2017) foram monitorados postagens do *Twitter* contendo a palavra "depressão". Postagens providas de contas de profissionais de saúde mental. E ainda *blogs* de depressão. Com isso totalizando a base de dados com mais de 54 milhões de postagens referentes a depressão.

Observa-se ainda que dentre todas as bases que foram criadas pelos estudos selecionados, apenas quatro foram disponibilizadas, o estudo de Hassan et al. (2017) utilizou o *dataset 20newsgroups* que é popular para experimentos de técnicas de aprendizado de máquina, como classificação e agrupamento de textos. O estudo de Larsen et al. (2015) disponibiliza sua base de dados por meio da API¹. A base de dados utilizada pelos estudos Leiva e Freire (2017), Briand, Almeida e Meurs (2018) está disponível para fins de pesquisa sob contratos de adequação de usuário².

¹<http://wefeel.csiro.au/#/api>

²<http://tec.citius.usc.es/ir/code/dc.html>

3.3.3 ANÁLISE E DISCUSSÃO

Durante a execução deste MS foram encontrados dois estudos com o mesmo propósito. O primeiro é um *survey* (JOHANNSEN; BIEMANN, 2018) publicado no *Springer* no ano de 2019. O segundo, consiste em uma Revisão Sistemática (ALONSO et al., 2018) publicada em 2018 no *Scopus*.

Em (JOHANNSEN; BIEMANN, 2018), o autor fez uma revisão de estudos relacionados a técnicas e algoritmos de *Data Mining* em saúde mental até Março de 2018. O enfoque deste estudo era mais amplo, pois abrangia além da depressão, doenças como demência, Alzheimer, Esquizofrenia, e ainda saúde mental em geral. Além disso, o estudo não filtrou somente estudos que se propuseram a análise textual das doenças, condicionando na inclusão de estudos que realizaram a análise de exames cerebrais de imagens. Entre os resultados referentes a depressão, os autores identificaram que SVM e NB foram os algoritmos mais utilizados para compor as abordagens dos estudos encontrados.

Alonso et al. (2018) fez a coleta de estudos também de forma ampla, buscando estudos que verificaram se a existência de alterações cognitivas influenciavam na forma de utilizar a linguagem. Os autores criaram uma seção para descrever as ferramentas utilizadas pelos estudos, onde foram identificados dicionários léxicos como LIWC e CELEX, algoritmos de redução de dimensionalidade como LDA. *Frameworks* como *Natural Language Toolkit* (NLTK) e *Scikit-Learn*. Questionários como CES-D e *The Beck Depression Inventory* (BDI). E ainda abordagens de AM supervisionadas como SVM, *Hidden Markov Models* (HMM), *Recurrent Neural Networks* (RNN) e *Long short-term memory Neural Network* (LSTM).

Visando uma melhor contextualização das respostas de cada QP, a seguir são apresentados os objetivos de alguns dos estudos primários selecionados. Essa análise visa compreender quais técnicas têm sido empregadas, as bases de dados que são utilizadas e como os estudos alcançaram seus respectivos resultados na utilização de AS. Por fim, é importante ressaltar que algumas dessas informações não estavam disponíveis nos estudos, assim as discussões apresentadas a seguir serão de acordo com a interpretação dos condutores do MS.

Diferentes estudos (Hassan et al., 2017; Deshpande; Rao, 2017; Aldarwish; Ahmad, 2017; Larsen et al., 2015; ZUORBA; OLAN; CANTARA, 2017; VEDULA; PARTHASARATHY, 2017; HASAN; RUNDENSTEINER; AGU, 2019; TAO et al., 2016; BIRJALI; BENI-HSSANE; ERRITALI, 2017; MA; WANG; ZHANG, 2017) utilizaram

o *Twitter* como fonte de dados para criação das bases de dados. Dentre tais estudos, é importante destacar o estudo de Vedula e Parthasarathy (2017), o qual teve como objetivo identificar se existe uma diferenciação no comportamento *online* entre usuários potencialmente deprimidos, com relação a usuários normais. Os autores observaram que existem desvios significativos no comportamento dos usuários deprimidos, alguns exemplos são: atividade *online* reduzida e noturna, uso de emoções negativas nos textos e uso de pronomes auto-focados. Para detectar os sinais sintomáticos da depressão, foram examinados sinais como: estilos linguísticos, onde pode ser destacado aumento do uso de pronomes pessoais associados e derivados, e redução de pronomes de terceira pessoa ou que possuem conotação coletiva; sinais emocionais; interação do usuário e localização geográfica.

Os estudos (Islam et al., 2018; Aldarwish; Ahmad, 2017; Rosa et al., 2019; Noureen; Qamar; Ali, 2017; ZUORBA; OLAN; CANTARA, 2017; WU et al., 2018) utilizaram postagens do *Facebook* para compor suas respectivas base de dados de treinamento e testes. Dentre eles destaca-se o estudo de Rosa et al. (2019) por ser o único estudo dentre todos os 25 que utilizou uma fonte de dados em português. Os autores realizaram o desenvolvimento de uma aplicação *Web* que tem como objetivo servir como um sistema de recomendação para aliviar sintomas da depressão, para isso foi criada uma base de textos a partir de ontologias. O sistema inclui uma fase de monitoramento que visa detectar depressão e estresse nos usuários. A operação proposta é de coletar sentenças e filtrar por meio de AM para detectar sinais de depressão e estresse. As sentenças filtradas então são classificadas de acordo com a sua intensidade, podendo ser: extremo, intermediário e inferior. O sistema de recomendação se baseia nessa intensidade para definir quais itens irá sugerir para o usuário.

Somente quatro estudos (Aldarwish; Ahmad, 2017; Nguyen et al., 2014; Dao et al., 2015; NGUYEN et al., 2017) utilizaram o *LiveJournal*. O estudo que merece destaque é o de Nguyen et al. (2014), que visa demonstrar o potencial que as redes sociais possuem para identificar e monitorar a depressão em usuários. Para isso, foram investigadas comunidades de depressão online, visando encontrar os fatores diferenciais com relação a outras comunidades online. Foi identificado que aspectos como afeto e processos psicolinguísticos afetam o conteúdo escrito e o estilo da escrita em usuários depressivos.

O estudo de Leiva e Freire (2017) é o que mais se destaca dentre os seus similares (Seah; Jin Shim, 2018; BRIAND; ALMEIDA; MEURS, 2018) que utilizaram o

Reddit como fonte de dados. Neste estudo o objetivo é melhorar o estado da arte referente a detecção antecipada de depressão em redes sociais. Os experimentos foram executados em um conjunto de dados criado com postagens do *Reddit*. Os textos foram extraídos de maneira mais precisa se comparado com as abordagens que utilizam apenas *Term Frequency–Inverse Document Frequency* (TF-IDF). Foram explorados o comportamento de técnicas clássicas e modernas de AM para prever de uma melhor forma os casos positivos. Ainda foi analisado a influencia de algoritmos genéticos nos resultados.

O estudo de Wang, Zhang e Sun (2013) utilizou o Sina micro blog, já que o objetivo deste estudo era analisar dados textuais escritos em chinês, e esta fonte de dados é uma especie de twitter existente na china , onde a quantidade majoritária das informações postadas são em chinês.

Notas em árabe foram utilizados por Boukil et al. (2019) para compor sua base de dados, a justificativa da escolha desta fonte de dados foi a dificuldade de encontrar informações referentes a depressão e ao suicídio escritas em árabe, já que em países árabes o tabu referente a estes problemas é maior que no restante do mundo. O objetivo dos autores é criar um sistema automático para extrair e prever a intenção de suicídio de um indivíduo, com base em mensagens e notas textuais do mesmo. Para isso, foi utilizado o método TF-IDF para reformular as notas textuais na forma que o sistema possa reconhecer e entender. Posteriormente foi criado um modelo utilizando *Convolutional Neural Network* (CNN) para reconhecer e classificar notas que apresentam potencial suicida e notas que não apresentam este potencial.

Em Hassan et al. (2017) utilizaram a base de dados *20newsgroups* para testar a assertividade de seus resultados, já que os mesmos tinham como principal objetivo realizar a comparação entre as técnicas: SVM, NB e *Maximum Entropy* (ME) para detectar depressão em diferentes bases de dados. Para combinar os múltiplos classificadores, este estudo utilizou o modelo de voto majoritário, que consiste em escolher a classe que obteve o maior número de classificações entre as técnicas utilizadas no processamento dos dados.

O fórum de discussões PTT foi utilizado por TTung e Lu (2016), onde os autores utilizam técnicas de PLN, recuperação de informações e mineração de textos, com o objetivo de transferir conhecimento especializado e, conseqüentemente, aprimorar a análise e dicção da tendência de depressão em postagens na *internet*. Este estudo ainda utilizada léxicos para realizar a análise, e uma de suas contribuições é a expansão de um léxico de eventos negativos de forma automática.

Em Ma, Wang e Zhang (2017) foram utilizados *blogs* de depressão em geral, onde

o objetivo dos autores era de extrair e resumir fatores incomuns, mas que sejam relevantes para a identificação de sintomas depressivos, com isso aumentando a possibilidade da ferramenta criada pelos mesmos de conseguir identificar de forma correta perfis considerados depressivos.

Vocabulários *online* foram utilizados por Tao et al. (2016) para aumentar a quantidade de termos que estão relacionados a depressão. Onde estes termos auxiliaram na análise das informações publicadas pelos usuários do *Twitter*. Foram utilizadas técnicas de PLN, mais especificamente análise semântica. A abordagem proposta pelos autores possui ainda o objetivo de oferecer a assistentes sociais a capacidade de acessar potenciais pessoas depressivas que precisam de ajuda no estágio inicial da depressão.

O estudo de Wu et al. (2018) foi o único que utilizou dados do CES-D para auxiliar na identificação dos perfis depressivos dos participantes do estudo, onde foram relacionadas as respostas dos fornecidas pelos mesmos no questionário com as informações coletadas das redes sociais.

Os estudos de Hassan et al. (2017), Aldarwish e Ahmad (2017), Zuorba, Olan e Cantara (2017), Tao et al. (2016), Ma, Wang e Zhang (2017) utilizaram mais de uma fonte de dados para criar suas respectivas bases de dados. O motivo disso pode estar ligado ao fato de que apenas uma fonte de dados pode ser insuficiente para conseguir generalizar um perfil depressivo. Ou ainda com o objetivo de validar a abordagem criada em uma fonte de dados diferente da de treinamento, como no caso de Hassan et al. (2017), no qual o *20newsgroups* foi utilizado para testar o modelo.

Fazer o uso de tecnologia para ajudar pessoas que possuem depressão é de grande relevância, pois nos últimos anos a quantidade de portadores desta doença aumentou em números alarmantes (ORGANIZATION et al., 2017). Os estudos encontrados durante a execução deste MS provaram ser possível identificar sintomas depressivos em dados textuais, portanto é possível alertar indivíduos que não tem o conhecimento de possuir a patologia e conseqüentemente não procuram auxílio profissional.

Após a análise dos resultados, observou-se que as duas técnicas mais recorrentes entre os estudos foram NB e SVM, os quais foram utilizadas por 12 estudos (Hassan et al., 2017; Deshpande; Rao, 2017; Aldarwish; Ahmad, 2017; Rosa et al., 2019; Noureen; Qamar; Ali, 2017; ISLAM et al., 2018; HASAN; RUNDENSTEINER; AGU, 2019; TAO et al., 2016; LEIVA; FREIRE, 2017; BRIAND; ALMEIDA; MEURS, 2018, 2018; Keumhee Kang; Chanhee Yoon; Eun Yi Kim, 2016). A maioria dos estudos relataram resultados satisfatórios utilizando as técnicas citadas na Tabela 2, pois em geral as mesmas atingiram

uma boa acurácia durante as classificações realizadas.

Zuorba, Olan e Cantara (2017) destaca-se como o estudo que obteve o melhor resultado entre os demais. Este estudo utilizou um modelo de Regressão Logística (*Logistic Regression*) que foi treinado com os vetores de características gerados a partir dos léxicos. O LDA foi utilizado para melhorar a capacidade de discriminação entre as classes, reduzindo a dimensionalidade por meio da geração de tópicos.

Vale ressaltar que em Noureen, Qamar e Ali (2017) é realizada uma análise comparativa das técnicas mencionadas, e por isso possui um elevado número de técnicas relacionadas. Com a análise dos resultados, foi possível observar que Seah e Jin Shim (2018), Zuorba, Olan e Cantara (2017) e Keumhee Kang, Chanhee Yoon e Eun Yi Kim (2016) (LDA) e o Larsen et al. (2015) (PCA) utilizaram técnicas de redução de dimensionalidade das bases de dados. Essas técnicas são utilizadas para melhorar a capacidade de discriminação entre as classes existentes na base de dados.

3.4 AMEAÇAS À VALIDADE

A principal ameaça à condução de um MS está relacionada ao fator humano que pode introduzir viés no processo de seleção dos estudos primários, além de provocar imprecisão na extração dos dados. Neste MS foram identificadas algumas ameaças à validade, bem como possíveis melhorias para minimizar ou extinguir essas ameaças.

A primeira ameaça a ser discutida é o número de pesquisadores. Somente um pesquisador realizou este MS, aumentando o risco de vieses no trabalho. Com o objetivo de minimizar este risco, os resultados encontrados, bem como todas as fases do MS foram acompanhadas e validadas pelo orientador deste trabalho.

A segunda ameaça está relacionada com a criação e adaptação da *string* de busca. Como as palavras e as expressões que compõem são derivadas a partir das QPs, a correta construção desta é vital para a efetividade da pesquisa. Da mesma forma, a adaptação da *string* para cada base pode mudar o foco da pesquisa caso não seja feita de forma precisa. Para mitigar esta ameaça, foi solicitado a um especialista que avaliasse a *string* para validá-la e melhorar sua efetividade.

Por fim, a última ameaça é referente à extração dos dados. Neste caso, devido a omissão de informações importantes em vários estudos, houve a necessidade do pesquisador buscar outras publicações relacionadas a fim de auxiliar na análise dessas informações.

3.5 CONSIDERAÇÕES FINAIS

Utilizar AS para identificar traços depressivos em usuários de redes sociais mostra-se uma atividade promissora, pois os estudos selecionados neste MS demonstraram que pessoas que possuem depressão, tristeza excessiva, ou sinais depressivos, tendem a apresentar um comportamento *online* diferente se comparado com o comportamento de usuários que não possuem tais características. Alguns aspectos desse comportamento são: forma se manifestar, estilos linguísticos tais como, aumento da utilização de pronomes auto-focados e expressões negativas, mudança no horário de utilização das redes sociais, onde o mesmo torna-se em grande maioria noturno, entre outros.

Os resultados alcançados após a execução deste MS, demonstram ainda, que apesar da depressão ser um tema atualmente muito discutido no âmbito global, e o uso de AS vir se provando ao longo dos anos como uma excelente maneira de analisar e entender opiniões dos usuários das redes sociais, poucos estudos tem sido realizados envolvendo essas duas áreas de pesquisa nos últimos anos, principalmente quando se trata da análise dos dados em português.

4 PROPOSTA

Este capítulo apresenta a proposta de AM e AS para detecção de perfis depressivos em redes sociais. Na seção 4.1 é descrita a proposta de pesquisa desse projeto. A descrição da estratégia criada juntamente com o detalhamento de cada parte que a compõe são apresentados na Seção 4.2. Ainda nesta Seção são apresentadas as tecnologias e ferramentas utilizadas para o desenvolvimento da proposta. Por fim, são apresentadas as considerações finais deste capítulo.

4.1 PROPOSTA DE PESQUISA

Assim como qualquer doença, se diagnosticado no início, um quadro de depressão pode ser tratado e revertido com menos dificuldade do que se o transtorno for descoberto em um estágio mais avançado. Esse diagnóstico é complicado de se realizar, principalmente pelo fato de que os sintomas da depressão se analisados separadamente são apontados como problemas cotidianos, que na maioria das vezes passam despercebidas pelas pessoas devido a falta de informação. Alguns desses sintomas são: alteração de apetite, peso e sono, dificuldade de concentração e de tomar decisões, perda de energia repentina, perda de interesse generalizado e retraimento social (American Psychiatric Association , 2013).

Com a popularização das redes sociais, cada vez mais tem se falado abertamente sobre depressão. Essa atitude reduz o problema da desinformação, ajuda a quebrar o tabu que está atrelado à doença, e também gera uma quantidade enorme de dados públicos referente ao assunto. Tais dados são obtidos por meio de publicações dos usuários em mídias sociais. Posteriormente, estes dados podem ser processados e analisados no intuito de se identificar e extrair emoções e informações pertinentes dessas publicações.

O processo de extração de emoções em textos de redes sociais é realizado por meio da AS, que utiliza técnicas de AM e PLN para realizar o processamento e categorização dos dados. Normalmente os modelos são treinados com o auxílio de psicólogos e linguistas que ajudam a identificar padrões comportamentais e gramaticais presentes nos dados.

Dessa forma, é possível treinar algoritmos de AM para identificar sinais ou sintomas de depressão de um usuário.

O principal problema enfrentado pela literatura no contexto deste trabalho esta na criação da base de dados para treinamento. Isso ocorre uma vez que a identificação de sentimentos, positivos e negativos presentes nos textos, não é suficiente para que o algoritmo de AM aprenda de um texto o que podem ser sinais depressivos. Para que o algoritmo obtenha um bom nível de aprendizado, e conseqüentemente, um elevado índice de predições corretas, é necessário também, que os demais sintomas da depressão sejam identificados nos dados textuais.

Nesse contexto, para esse trabalho foi aplicado um método não identificado em trabalhos anteriores, cujo objetivo é realizar o treinamento e classificação com base nos sintomas da depressão. Pois segundo o DSM-V, sintomas psíquicos e comportamentais tem maior relevância para sinalizar um quadro depressivo do que sintomas fisiológicos. Realizando a análise desta forma é possível controlar a relevância de cada tipo de sintoma presente em um perfil de usuário. O método ainda inclui uma análise de forma temporal dos perfis, com o objetivo de verificar se os sintomas identificados no primeiro passo ocorrem de uma forma contínua e relevante o suficiente para sinalizar que o perfil analisado possui indícios de depressão.

4.2 ESTRATÉGIA *DP-SYMPTOM-IDENTIFIER*

A estratégia proposta neste trabalho foi denominada de *DP-Symptom-Identifier*. A mesma foi criada com o objetivo de identificar sintomas depressivos em postagens do *Twitter*, utilizando técnicas de PLN e algoritmos de AM. A visão geral da estratégia é apresentada na Figura 13, a mesma é dividida em duas etapas: **A - Geração do Modelo** e **B - Predição dos Dados**.

A etapa **A**, composta pelos itens: 1, 2, 3 e 4, consiste na realização do treinamento do algoritmo, ou seja, a geração do modelo de AM propriamente dito. Nesta etapa, o item 1 representa a base de dados de treinamento utilizada para treinar o algoritmo de AM, a qual é composta por postagens previamente rotuladas com as categorias sintomáticas da depressão. O item 2 representa a etapa de pré-processamento dos dados de treinamento. Já no item 3 é realizado o processo de extração de características. Finalmente no item 4 é realizado o treinamento do algoritmo gerando o modelo de AM.

Já os itens 5, 6 e 7 compõem a etapa **B**, na qual é realizada a predição dos dados.

Nesta etapa, o item 5 representa o processo de coleta dos dados de entrada, extraídos do *Twitter*. No item 6, é realizado o pré-processamento dos dados. Já no item 7 são extraídas as características dos dados. Por fim, estes são processados pelo modelo gerado anteriormente pela etapa **A**, cuja a saída deste processamento são as predições realizadas. Estas predições obtidas assumem uma entre quatro classes possíveis: **0 - Nenhum**, **1 - Fisiológico**, **2 - Comportamental** e **3 - Psíquico**.

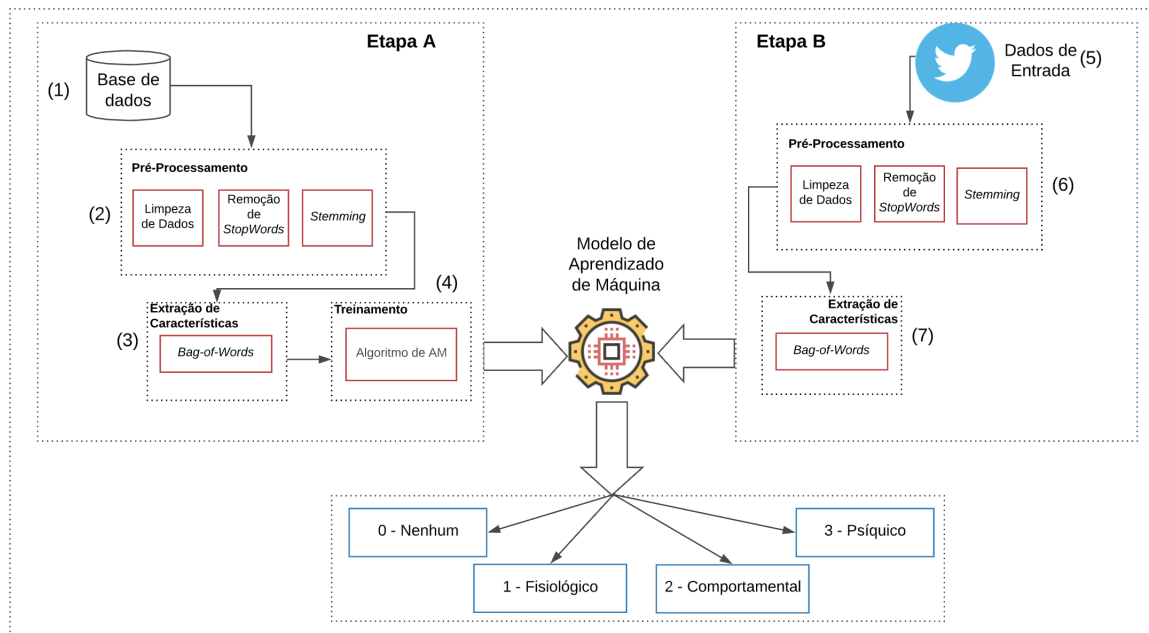


Figura 13: *DP-System-Identifier*: uma estratégia para identificação de sintomas depressivos.

Fonte: Autoria própria.

A estratégia *DP-Symptom-Identifier* foi desenvolvida com base em algumas etapas do processo de aprendizagem, e são descritas nas próximas subseções.

4.2.1 COLETA DE DADOS

Com base nas revisão bibliográfica optou-se por utilizar o *Twitter* como fonte de dados para a execução deste trabalho, visto que é a maior fonte de informação de dados abertos para o problema. Para realizar a coleta dos dados foi criada uma aplicação utilizando a linguagem de programação Java, denominada Aplicação de Coleta de Dados (ACD).

As funcionalidades implementadas na ACD possibilitaram utilizar palavras e sentenças como parâmetros de filtragem durante a seleção dos *tweets*. Com isso, só eram

retornadas postagens que continham os parâmetros desejados. Dentre os *tweets* retornados, foram excluídos os que representavam um *retweet*¹. As sentenças passadas como parâmetro para a busca das postagens na ACD foram criadas pela psicóloga que acompanhou o desenvolvimento deste trabalho. Todas as sentenças foram previamente rotuladas de acordo com o sintoma que representavam. Alguns exemplos de sentenças e seus respectivos rótulos são apresentadas na Tabela 5.

Tabela 5: Sintomas presentes em sentenças rotuladas com ajuda de profissional da área de psicologia.

Sentença	Categoria Sintomática
Eu quero morrer	Comportamental
Não desejo sair de casa	Comportamental
Tenho dificuldade para dormir	Fisiológico
Estou sempre cansado	Fisiológico
Me sinto inútil	Psíquico
Sou infeliz	Psíquico

Com o intuito de facilitar a visualização dos dados, foi desenvolvida uma Aplicação Web (AW). A ferramenta ajuda também na análise e rotulação das postagens coletadas, juntamente da psicóloga. A seguir são apresentadas algumas das telas desta aplicação, em conjunto a alguns comentários pontuais sobre a informação que sintetizam.

A Figura 14 demonstra a tela de listagem dos dados. Esta tela objetiva facilitar a visualização das informações de cada um dos *tweets*. Dentre elas: quais sintomas estão presentes no texto (colunas Psíquico, Fisiológico e Comportamental); se o dado foi analisado por um especialista; e por fim um botão de ação, caso houve-se a necessidade de excluir a postagem da base coletada.

Já a Figura 15 apresenta a tela criada para analisar as predições realizadas. Nesta tela, as postagens apresentadas são limitadas a um mesmo usuário, o qual pode ser filtrado pelo campo de texto disponível. Visando facilitar a análise, os registros apresentados são ordenados por data da publicação de forma decrescente, simulando uma *timeline* da atividade do usuário ao longo do tempo.

Já por sua vez, a Figura 16 exhibe a tela criada para facilitar o processo de rotulação dos dados coletados. Por meio desta tela é possível rotular cinco postagens por vez, marcando ou desmarcando a opção para cada um dos sintomas. Após a rotulação de todas as postagens, o botão presente no canto inferior direito é acionado para gravar as

¹ *Retweet* é uma republicação de um *Tweet*.

Projeto TCC

Dataset Categorizar Timelines

Texto	Psíquico	Fisiológico	Comportamental	Validado Especialista	Ação
só eu que já fiquei com dúvida se só tô muito triste ou é depressão?	✓	✗	✗	✓	🗑️
Crises de ansiedade Depressão Tô me afundando e n sei como me recuperar	✓	✗	✗	✓	🗑️
hj vai passar uma materia sobre depressao no globo reporter, ja to ate vendo o tanto q vou chorar	✗	✗	✗	✓	🗑️
Hoje o Profissão Repórter fala sobre Depressão. ASSISTAM	✗	✗	✗	✓	🗑️
Essa depressão ainda vai me matar	✗	✗	✓	✓	🗑️

Figura 14: Captura da Tela de Listagem dos Dados.

Fonte: Autoria própria.

informações alteradas.

Após o processo de rotulação, cada postagem pode apresentar os seguintes valores para rótulo:

- **Nenhum:** quando a postagem não possui nenhum dos sintomas;
- **Psíquico:** quando a postagem possui apenas indícios de sintomas psíquicos;
- **Comportamental:** nos casos onde a postagem possui indicação de evidências comportamentais;
- **Fisiológico:** quando o sintoma identificado na postagem é de origem fisiológica.

Para a avaliação apropriada deste trabalho foram criadas duas bases de dados distintas:

- **Base de dados de treinamento:** criada utilizando as postagens de vários usuários coletadas e rotuladas conforme o processo descrito acima. Esta base possui um total de 2008 linhas e 7 colunas. A Figura 17 demonstra uma fração da mesma. Além do dado textual extraído da postagem “tweet”, as demais colunas descritivas são: “idTweet” coluna utilizada para identificação dos *tweets*, “fisiologico” coluna a qual remete se o *tweet* pertence a categoria de sintomas fisiológicos. A coluna “comportamental” é utilizada para controlar o fato do *tweet* ser ou não um sintoma da categoria comportamental. Já a coluna “psiquico” demonstra a presença ou não

Projeto TCC

Dataset

Categorizar

Timelines

Usuário

Buscar Timeline

Texto	Classificação	Validado Especialista	Data
Porque eu tenho tanta insegurança nessa 'relação'?	Nenhum	✓	2019-10-28
Vou fazer um brigadeiro pra comer sozinha porque eu sou a minha melhor companhia	Psíquico	✓	2019-10-09
Eu tava disposta a fazer diferente dessa vez, caralho eu falei de nós até pra minha mãe..	Nenhum	✓	2019-08-11
Sobre os relatórios https://t.co/kZ7MShCNju	Psíquico	✓	2019-11-01
O comentário que recebi hoje: "Sorte de que tá 'comendo'. Eu: verdade, um total de 0 pessoa Mereço mesmo kkkkkkkkkkkkkk	Nenhum	✓	2019-10-26

Figura 15: Captura da Tela de Análise dos Dados.

Fonte: Autoria própria

de sintomas psíquicos. A coluna “label” contém informações referente a rotulação do dado. A coluna “dhTweet” apresenta e data da postagem. Por fim a coluna “validadoEspec” informa se a postagens foi validade pela especialista que acompanha este trabalho.

- **Base de dados de testes:** criada utilizando a funcionalidade de busca de *timelines* por ID de usuário, existente na ACD. Esta base possui as mesmas colunas que a base de dados de treinamento. Os dados armazenados nesta base foram utilizados para testar os modelos de AM induzidos, verificando a existência ou não de indícios de depressão nos perfis (*timelines*) coletados.

Aqui cabe salientar que a geração de duas bases de dados distintas, uma para treinamento/validação e outra para testes, tira qualquer tipo de enviesamento na avaliação dos modelos induzidos. Isso é garantido, pois os dados usados para teste, não foram nunca vistos pelos algoritmos durante as etapas de treinamento e validação.

Por fim, a Figura 18 demonstra a distribuição das classes existente na base de dados de treinamento. A classe que representa os sintomas psíquicos é a que possui a maior

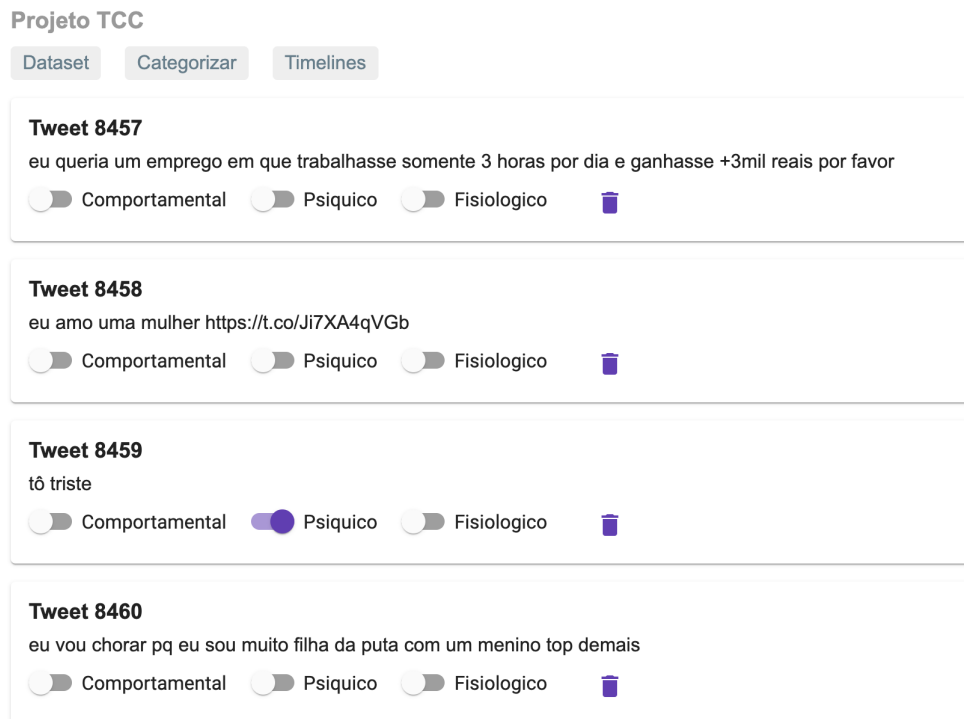


Figura 16: Captura da Tela de Rotulação dos Dados.

Fonte: Autoria própria

quantidade de exemplos dentre as demais. Isso pode ser justificado pelo fato de que esta categoria sintomática é a que possui uma maior quantidade de exemplos no (American Psychiatric Association , 2013), e ainda, alguns sintomas da categoria comportamental são relacionados com os da categoria psíquico.

4.2.2 PRÉ-PROCESSAMENTO

Visando reduzir o tamanho da base de dados, e também remover dados redundantes e irrelevantes, foi realizada uma etapa de pré-processamento dos dados. No geral, foram feitas duas tarefas, descritas a seguir.

A primeira tarefa de pré-processamento consistiu na limpeza dos dados: todos os caracteres das postagens foram convertidos para caracteres minúsculos. Em seguida, caracteres numéricos foram removidos, assim como pontuações, *links* e espaços desnecessários.

tweet	idTweet	fisiologico	comportamental	psiquico	label	dhTweet	validadoEspec
1 só eu que já fiquei com dúvida se só tô muito triste o...	11	FALSE	FALSE	TRUE	3	2019-09-12	TRUE
2 Crises de ansiedade Depressão Tô me afundando e n ...	20	FALSE	FALSE	FALSE	0	2019-09-11	TRUE
3 hj vai passar uma materia sobre depressao no globo r...	22	FALSE	FALSE	FALSE	0	2019-09-11	TRUE
4 Hoje o Profissão Repórter fala sobre Depressão. ASSI...	23	FALSE	FALSE	FALSE	0	2019-09-11	TRUE

Figura 17: Exemplificação da Base de Dados de Treinamento.

Fonte: Autoria própria

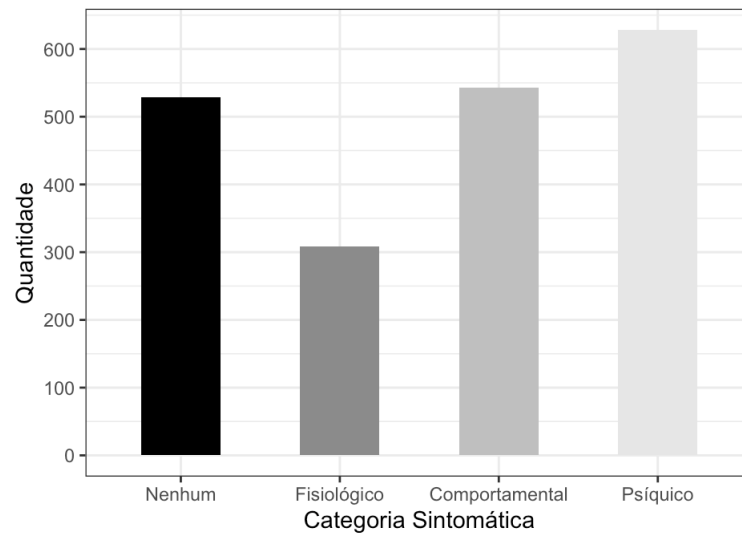


Figura 18: Distribuição das classes presentes no conjunto de treinamento.

Fonte: Autoria própria

Além disso, citações de outros usuários e *retweets* também foram removidos. Um último passo removeu todos os caracteres especiais, tais como *emojis*, etc.

A segunda tarefa foi de fato responsável pelo pré-processamento dos dados. Para isso foram utilizadas algumas técnicas de PLN citadas na literatura, como remoção de *stopwords* e *stemming*.

4.2.3 EXTRAÇÃO DE CARACTERÍSTICAS E TREINAMENTO

Com o objetivo de encontrar a melhor forma de extrair as características dos dados textuais da base, foram realizados testes com duas técnicas citadas na literatura: TF-IDF e BoW.

A primeira técnica utilizada foi o TF-IDF, a qual possui como objetivo expressar a relevância de uma palavra em um dado *corpus*. A mesma foi utilizada com o auxílio da

função `TfidfVectorizer` do pacote `superml`² do R. Ao executar a vetorização dos 2008 exemplos presentes na base de dados, o processo retornava uma matriz com mais de 10 mil atributos descritivos (colunas). Para tentar diminuir esta quantidade, foram identificados e removidos todos os atributos correlacionados, em uma porcentagem igual ou superior a 95%. Mesmo após realizar esta operação de redução, ainda ocorria a geração de um *dataset* longitudinal, onde a quantidade de atributos (colunas) é maior que a quantidade de exemplos (linhas) disponível na base de dados. Nos primeiros testes realizados, tais bases longitudinais geravam modelos com *overfitting*: eles prediziam sempre a mesma classe (majoritária) e erravam a predição de todas as outras classes.

Por conta desses resultados com o TF-IDF, foram realizados testes com a técnica BoW, que consiste em realizar a contagem da quantidade de vezes que cada termo se repete dentro da base de dados. Utilizando esta técnica a quantidade de atributos gerados foi de 4650. Após o processamento principal da técnica, foram removidos os atributos com uma frequência menor que 10, reduzindo a quantidade total de atributos descritivos para 340. Tal mudança foi muito significativa, pois foi possível induzir modelos mais precisos, obtendo resultados superiores em até 50% quando comparado com os resultados obtidos com TF-IDF. Sendo assim, o BoW foi a técnica escolhida para realizar a extração de características dos dados textuais.

4.2.4 ESTRATÉGIAS DE CLASSIFICAÇÃO EXPLORADAS

Visando explorar diferentes formas de identificar sintomas depressivos em dados textuais, foram criadas duas estratégias de AM. Tais estratégias se diferenciam na forma como é feita a modelagem do problema:

- **Múltiplos modelos de classificação binária:** cada categoria de sintoma foi tratado de forma separada. Sendo assim, foram criados três modelos binários individuais, cuja predição é a presença ou ausência do sintoma em questão. Dessa forma, cada modelo induzido é responsável pela predição de um sintoma da depressão.
- **Classificação Multi-classe:** aqui as categorias sintomáticas foram tratadas em conjunto, criando-se apenas um único modelo. O rótulo dos exemplos pode apresentar quaisquer uma das categorias sintomáticas: Fisiológico, Comportamental, Psíquico, ou ausência completa de sintomas (Nenhum).

²<https://www.rdocumentation.org/packages/superml/versions/0.4.0>

Uma terceira estratégia foi também investigada, que partia da hipótese de existência de correlação entre a ocorrência dos sintomas, isto é, se um ocorre o outro também tem alta probabilidade de ocorrer. Foram conduzidos alguns experimentos usando **Stacking**, que consiste em “empilhar” classificadores, adicionando a saída de um modelo como atributo descritivo do próximo. Entretanto, os resultados experimentais foram muito ruins, abaixo dos *baselines* aleatórios, o que refutou a hipótese de correlação dos sintomas. Por isso, tais resultados e métodos não foram adicionados e discutidos no trabalho.

4.2.5 AVALIAÇÃO

Para a avaliar os modelos induzidos foi escolhida a metodologia de validação cruzada (*Cross-Validation* (CV)). Tal método consiste em dividir a base de dados em k partições de dados de mesmo tamanho. Assim, o algoritmo é treinado utilizando $k - 1$ subconjuntos, e a parte restante é utilizada para validar o modelo induzido. O processo é repetido k vezes, de forma que em cada iteração o subconjunto de validação é alterado. A quantidade de subconjuntos (k) escolhida para validar os modelos criados para este trabalho é de $k = 10$ subconjuntos, valor frequentemente utilizado pela literatura específica no contexto deste trabalho.

Após treinamento e validação, os melhores modelos induzidos foram aplicados no conjunto de testes, visando verificar se é possível identificar um perfil depressivo analisando informações contidas em *timelines* de usuários do *Twitter*. Os dados de teste foram pré-processados e vetorizados da mesma forma que os dados de treinamento. As previsões obtidas são então manipulados por uma etapa de pós-processamento, que separa os valores de acordo com os identificadores de usuários, e verifica a ocorrência dos sintomas ao longo do tempo.

4.2.6 TECNOLOGIAS E FERRAMENTAS

Nesta Subseção são apresentadas as tecnologias utilizadas no desenvolvimento deste trabalho. As mesmas são divididas entre as aplicações citadas anteriormente (ACD e AW) e ainda a Estratégia de Classificação (EC) que tem implementadas as funções necessárias para o pré-processamento, extração de características, treinamento, validação e testes das abordagens. Na Tabela 6 são descritas as linguagens, *frameworks* e pacotes/bibliotecas usados na implementação das aplicações que compõem a proposta do trabalho. A primeira coluna apresenta cada aplicação, seguida das ferramentas utilizadas e suas funcionalidades práticas.

Tabela 6: Tecnologias e ferramentas utilizadas para o desenvolvimento do trabalho.

Aplicação	Ferramenta	Funcionalidade
ACD	JAVA ³	Linguagem de programação <i>back-end</i>
	Spring ⁴	<i>Framework</i> para desenvolvimento JAVA
	Twitter4J ⁵	Biblioteca JAVA para operações de API do Twitter
	PostgreSQL ⁶	Banco de dados relacional
AW	Angular 6 ⁷	<i>Framework front-end</i> para desenvolvimento <i>Web</i>
EC	R ⁸	Linguagem para manipulação, análise e visualização gráfica dos dados
	superml ⁹	Pacote em R para criar modelos de AM
	ggplot2 ¹⁰	Pacote em R para visualização de dados
	httr ¹¹	Pacote em R para operações com URLs e HTTP
	jsonlite ¹²	Pacote em R para manipulação de JSON
	qdapRegex ¹³	Pacote em R para manipulação de <i>strings</i>
	tm ¹⁴	Pacote em R para mineração de textos
	mlr ¹⁵	<i>Framework</i> de aprendizado de máquina em R

4.3 CONSIDERAÇÕES FINAIS

Neste Capítulo foi apresentada a estratégia *DP-Symptom-Identifier*, criada para identificar sintomas depressivos em dados textuais. A mesma foi criada utilizando algoritmos e técnicas de PLN e AM referenciados na literatura encontrada no contexto deste trabalho.

³https://www.java.com/pt_BR/

⁴<https://spring.io/projects/spring-boot>

⁵<http://twitter4j.org/en/>

⁶<https://www.postgresql.org/>

⁷<https://angular.io/>

⁸<https://www.r-project.org/>

⁹<https://www.rdocumentation.org/packages/superml/versions/0.4.0>

¹⁰<https://www.rdocumentation.org/packages/ggplot2/versions/3.2.1>

¹¹<https://www.rdocumentation.org/packages/httr/versions/1.4.1>

¹²<https://www.rdocumentation.org/packages/jsonlite/versions/1.6>

¹³<https://www.rdocumentation.org/packages/qdapRegex/versions/0.7.2>

¹⁴<https://www.rdocumentation.org/packages/tm/versions/0.7-6>

¹⁵<https://www.rdocumentation.org/packages/mlr/versions/2.15.0>

5 AVALIAÇÃO EXPERIMENTAL

Neste Capítulo é apresentada a avaliação experimental do trabalho, cujo objetivo é a eficiência das estratégias de aprendizado discutidas no capítulo anterior. Na Seção 5.1 é detalhado o planejamento dos experimentos, as questões de pesquisa (QPs) e suas respectivas hipóteses. Na Seção 5.3 são respondidas as QPs, e os resultados obtidos apresentados e discutidos. Já na Seção 5.4 são descritas as ameaças à validade das avaliações empíricas conduzidas. Por fim, na Seção 5.5 são apresentadas as considerações finais desta Capítulo.

5.1 DEFINIÇÃO DO EXPERIMENTO

Para a realização do experimento, seguiu-se o modelo *Goal-Question-Metric* (GQM) (BASILI; WEISS, 1984). Este modelo de experimentação é dividido em cinco partes:

- **Objeto de estudo:** técnicas e algoritmos de AM;
- **Propósito:** avaliar a eficiência das técnicas/algoritmos de AM;
- **Perspectiva:** o experimento é executado do ponto de vista de pesquisadores;
- **Foco qualitativo:** o principal efeito sob investigação é a eficiência das técnicas/algoritmos de AM em termos da Área sob a curva (*Receiver Operating Characteristic* - ROC) (AUC); e
- **Contexto:** o contexto deste experimento é a análise de sintomas depressivos.

A AUC é uma medida de desempenho utilizada em problemas de classificação, que representa a medida de separabilidade, ou seja, mostra até que ponto o modelo avaliado é capaz de separar corretamente as classes (FLACH; HERNANDEZ-ORALLO; FERRI, 2011). O valor da AUC pode variar de 0.5 a 1.0. Quanto mais próximo de 1.0 o valor da AUC, melhor o modelo consegue separar as classes. Já um modelo com previsões ruins

possui uma AUC que se aproxima de 0.5. O valor da AUC pode ser representado em forma de gráfico como na Figura 19, neste caso, quanto maior a curva melhor é o algoritmo.

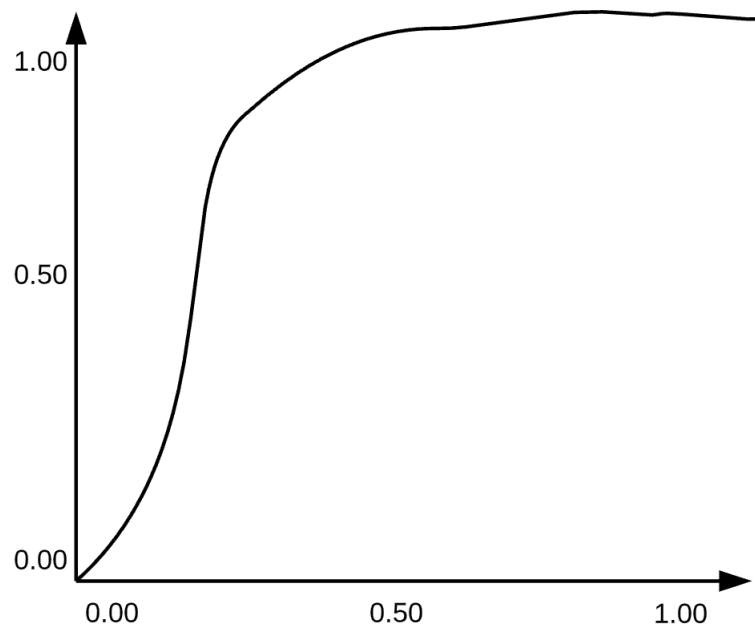


Figura 19: Exemplo de curva AUC.

Fonte: Autoria própria

A fim de avaliar os resultados das estratégias elaboradas para identificar sintomas depressivos, as seguintes questões de pesquisa foram definidas:

- QP_1 - Qual algoritmo de AM é mais preciso para identificar sintomas depressivos em textos?

Para responder a QP_1 , foram selecionados quatro algoritmos de AM com base na revisão bibliográfica: *Support Vector Machines* (SVMs), *Náive Bayes* (NB), *Multilayer Perceptron* (MLP) e *Random Forest* (RF). Os mesmos foram avaliados com relação a precisão, representada pela medida de desempenho AUC.

Em um segundo momento, o melhor algoritmo avaliado é selecionado para compor as estratégias preditivas descritas no trabalho, e que serão utilizadas para resolução das demais questões de pesquisa propostas a seguir.

Para esta questão de pesquisa as seguintes hipóteses foram definidas:

$H1_0$: Não existe diferença de desempenho (μE) entre os modelos induzidos pelos

algoritmos.

$$H1_0 : \mu E_{SVM} = \mu E_{NB} = \mu E_{MLP} = \mu E_{RF} \quad (1)$$

$H1_1$: Os modelos induzidos com SVMs são superiores aos demais induzidos pelos outros algoritmos.

$$H1_1 : \mu E_{SVM} > \mu E_{NB} > \mu E_{MLP} > \mu E_{RF} \quad (2)$$

$H1_2$: Os modelos induzidos com NB são superiores aos demais induzidos pelos outros algoritmos.

$$H1_2 : \mu E_{NB} > \mu E_{SVM} > \mu E_{MLP} > \mu E_{RF} \quad (3)$$

$H1_3$: Os modelos induzidos com MLP são superiores aos demais induzidos pelos outros algoritmos.

$$H1_3 : \mu E_{MLP} > \mu E_{SVM} > \mu E_{NB} > \mu E_{RF} \quad (4)$$

$H1_4$: Os modelos induzidos com RF são superiores aos demais induzidos pelos outros algoritmos.

$$H1_4 : \mu E_{RF} > \mu E_{SVM} > \mu E_{NB} > \mu E_{MLP} \quad (5)$$

- QP_2 - Com qual frequência cada categoria de sintomas é identificada?

Para responder a QP_2 foram analisados os resultados das predições realizadas pelas duas estratégias propostas para este trabalho (descritas na subseção 4.2.4).

$H2_0$: Todas as categorias de sintomas são identificadas com a mesma frequência (μF).

$$H2_0: \mu F\text{-Psíquico} = \mu F\text{-Comportamental} = \mu F\text{-Fisiológico}$$

$H2_1$: A frequência com que cada categoria de sintomas é identificada difere entre as demais.

$$H2_1: \mu F\text{-Psíquico} \neq \mu F\text{-Comportamental} \neq \mu F\text{-Fisiológico}$$

- QP_3 - É possível identificar um perfil com tendências depressivas analisando as postagens de forma temporal?

Para responder a QP_3 foi utilizada a estratégia com melhor eficiência identificada na QP_1 , para tal, a mesma foi submetida a classificar os exemplos presentes na base de dados de teste.

5.2 DESIGN DO EXPERIMENTO

Foram realizados dois experimentos diferentes: ($E = e_1; e_2$). O primeiro experimento (e_1) tem por objetivo responder a questão de pesquisa QP_1 . O segundo experimento (e_2) objetivou responder as questões de pesquisa QP_2 e QP_3 . Para os experimentos relatados nesse estudo, foi manipulada apenas uma variável independente: identificação de sintomas depressivos; e quatro variáveis dependentes: as distribuições de desempenho dos algoritmos SVM, NB, RF e MLP;

Para execução dos experimentos foi utilizado um MacBook Pro 13 com um processador Intel Core i5 2,3 GHz, 8 GB de memória RAM, no sistema macOS Mojave versão 10.14.5.

O procedimento realizado durante a execução do experimento consistiu nas seguintes etapas:

1. **coleta dos dados:** nesta etapa foram coletados os dados da rede social *Twitter*, resultando em 2008 postagens coletadas.
2. **pré-Processamento dos dados:** os dados coletados foram pré-processados utilizando as técnicas de PLN evidenciadas na literatura.
3. **extração de Características:** após os dados serem pré-processados, foi utilizado a técnica de PLN chamada BoW, esta realiza a contagem de cada termo presente na base de dados. Os termos que possuíam uma frequência inferior a 10 aparições foram descartados.
4. **treinamento:** com o vetor de características gerado, as estratégias foram treinadas de acordo com sua especificação, detalhes desta etapa estão descritos na Seção 4.2.3 deste trabalho.
5. **validação das Estratégias:** utilizando a validação cruzada de 10-*folds* foi extraído o valor da medida AUC, a qual foi utilizada para responder as QPs definidas anteriormente.

6. **criação da base de testes:** novamente foram coletadas postagens do *Twitter*, com a diferença que nesta etapa foram coletadas postagens de *timelines* completas, visando verificar se o dono do perfil pode ou não conter sinais de depressão.
7. **identificação de Perfis Depressivos:** as estratégias propostas foram treinadas de acordo com sua especificação, e as mesmas foram apresentadas a base de testes para identificar a existência ou não de sinais depressivos nos dados apresentados.

5.3 RESULTADOS E DISCUSSÕES

Os resultados obtidos durante os experimentos são apresentados e discutidos a seguir. O algoritmo utilizado pelas estratégias de aprendizado com múltiplos modelos binários ou classificação multi-classe foi definido com base nos resultados obtidos no experimento e_1 , apresentado na Subseção 5.3.1.

5.3.1 DESEMPENHO DOS ALGORITMOS DE CLASSIFICAÇÃO (QP_1)

Para responder esta questão de pesquisa, foi realizado um experimento comparativo (*benchmark*). Explorou-se a validação cruzada com *10-folds* para cada um dos algoritmos selecionados: SVM, NB, MLP e RF. Todos os algoritmos utilizados foram selecionados com base no levantamento bibliográfico do trabalho.

Os resultados obtidos da execução dos mesmos são apresentados na Figura 20. No eixo Y é apresentada a AUC alcançada por cada uma das estratégias e algoritmos, localizados no eixo X.

Observando os resultados apresentados na Figura 20, percebe-se que existem diferenças entre as medianas de AUC atingidas por cada um dos algoritmos em cada uma das estratégias avaliadas. Vale a pena destacar que o algoritmo RF atingiu os maiores valores em ambas as estratégias. Já o algoritmo NB, um dos mais frequentemente utilizados na literatura (9 de 26 estudos), não obteve bons resultados, alcançado em suas melhores performances, aproximadamente 80% de assertividade nas predições. O SVM por sua vez, obteve resultados próximos aos alcançados pelo RF, com cerca de 90% de AUC em ambas as estratégias, justificando sua utilização em 11 dos 26 estudos encontrados. Por fim, o algoritmo MLP mostrou uma eficiência próxima a alcançada pelo SVM e RF.

Para avaliar a confiabilidade dos dados coletados nesta QP, foi aplicado o teste não paramétrico de *Nemenyi* (NEMENYI, 1963), seguido de um *Post Hoc Test*, que identifica

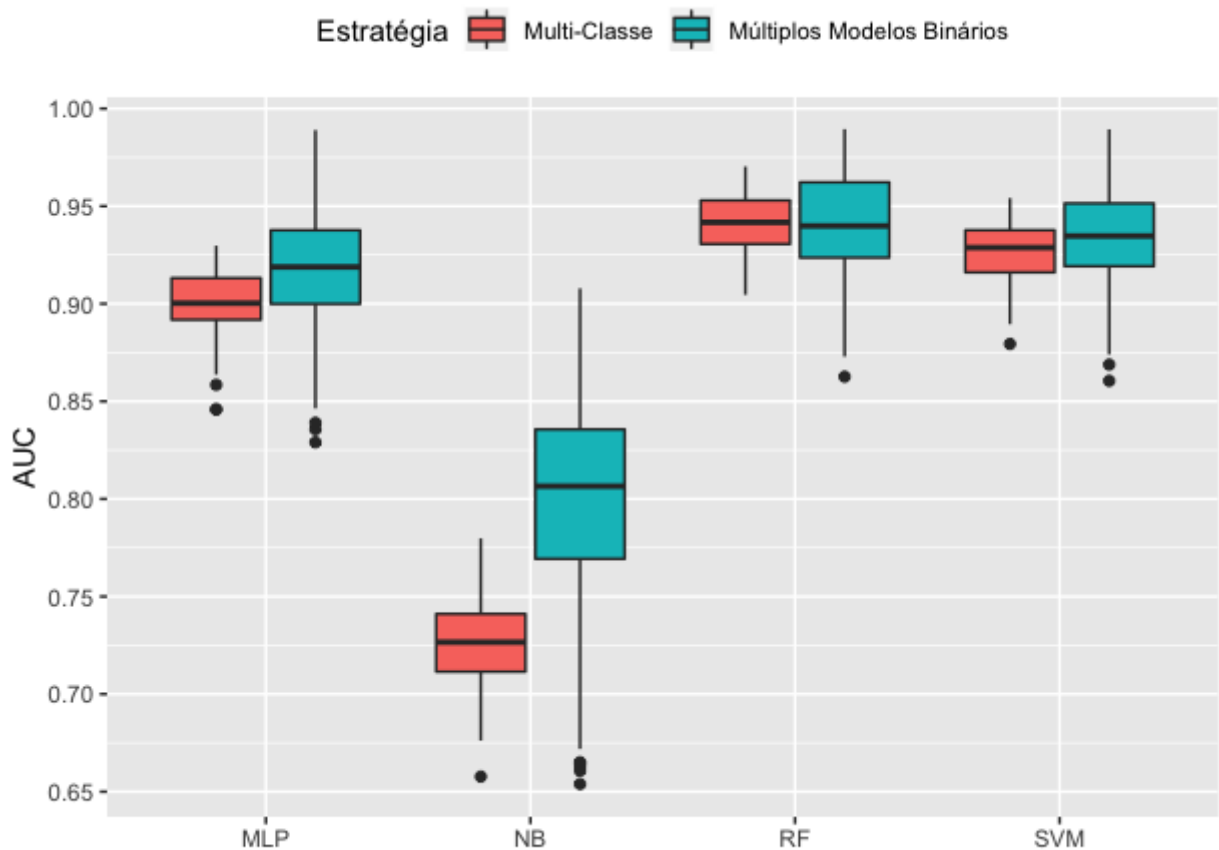


Figura 20: AUC obtida por cada estratégia de aprendizado com cada um dos algoritmos de AM.

Fonte: Autoria própria

onde ocorrem as diferenças entre os resultados. O resultado *Nemenyi test* com 95% de confiança aponta que apenas existe diferença estatística entre os algoritmos RF e NB, para ambas as estratégias, os valores de p-value para estas comparações são demonstrados na Tabela 7. Ou seja, SVM, RF e MLP são estatisticamente equivalentes em termos de eficiência. Como o objetivo desta QP é identificar e selecionar um algoritmo para compor a estratégia proposta, foi necessário analisar outra variável. Sendo assim, os algoritmos também foram analisados em termos de tempo de processamento.

Tabela 7: Resultados do *Nemenyi test*

Comparação dos algoritmos e estratégias	<i>P-Value</i>
RF ~ NB	0,0023
RF ~ NB Multi-Classe	0,0012
RF Multi-Classe ~ NB	0,0023
RF Multi-Classe ~ NB Multi-Classe	0,0012

Para identificar o algoritmo mais viável em termos de tempo de processamento, foram analisados os tempos gerais de processamento de cada um deles, apresentados na Figura 21. Para melhorar a representação do gráfico foi utilizada escala logarítmica nos valores do eixo Y. Já no eixo X estão identificados os algoritmos com cada estratégia.

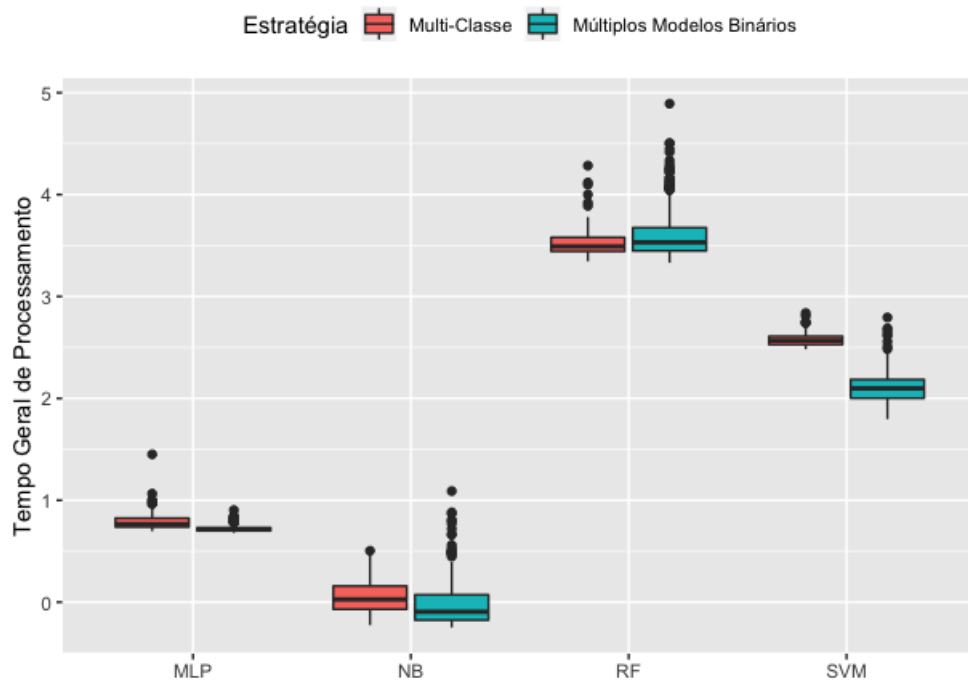


Figura 21: Tempo de Processamento de cada Estratégia com cada um dos Algoritmos.

Fonte: Autoria própria.

Os resultados apresentados na Figura 21, indicam que o algoritmo RF, apesar de obter melhores resultados de AUC, foi o que atingiu o pior tempo de processamento entre todos os algoritmos, com uma mediana de 40 segundos, sem muitas variações, mas com vários *outliers* superiores a este tempo. Isso ocorre pelo fato do mesmo gerar um comitê de árvores de decisão, induzindo muitos modelos, e a seleção da classe é feita com base na

maioria das predições individuais desses modelos. Os algoritmos MLP e SVM, os quais obtiveram resultados estatisticamente semelhantes ao RF, não atingiram grandes tempos de processamento. O tempo de processamento obtido pelo MLP aplicado na estratégia Multi-Classe foi inferior a 5 segundos e não teve grandes variações. Já para o SVM na mesma estratégia, o tempo de processamento foi próximo a 13 segundos. Por fim, o NB, o qual obteve a pior mediana AUC dentre os algoritmos, atingiu tempos de processamento inferiores a 5 segundos, também com poucas variações.

Por fim, para identificar de fato o melhor algoritmo para o contexto deste trabalho, os resultados da AUC e tempo de processamento foram comparados em conjunto para cada um dos algoritmos e estratégias. Esse processo de análise resultou na seleção do algoritmo MLP, o qual obteve um desempenho estatisticamente semelhante ao do RF com relação a AUC, e foi cerca de 98% mais rápido ao realizar o processo de identificação de perfis depressivos como um todo, o que confirmou a hipótese alternativa $H1_3$ (Os modelos induzidos com MLP são superiores aos demais induzidos pelos outros algoritmos).

5.3.2 FREQUÊNCIA DE IDENTIFICAÇÃO DAS CATEGORIAS SINTOMÁTICAS (QP_2)

Com o objetivo de responder a QP_2 , o mesmo *benchmark* anterior, com CV de 10-*folds*, disponibilizou dados das predições realizadas. Após a execução dos algoritmos, analisou-se a quantidade de predições corretas para cada um dos sintomas modelados.

A Figura 22 apresenta a frequência de predições corretas para cada um dos sintomas. No eixo Y são apresentadas as frequências de identificação de cada uma das categorias de sintomas. Já no eixo X é apresentado a disposição das categorias e estratégias. Como pode ser observado nesta Figura, para a estratégia de Múltiplos Modelos Binários, os sintomas fisiológicos são identificados com uma frequência superior aos demais. Já para a estratégia Multi-Classe, os sintomas que são identificados com mais frequência são os da categoria psíquica. Com base na análise realizada, confirma-se a hipótese alternativa $H2_1$ - A frequência com que cada categoria de sintomas é identificada difere entre as demais.

5.3.3 IDENTIFICAÇÃO DE PERFIL COM TENDÊNCIAS DEPRESSIVAS ($QP3$)

Com o objetivo de responder esta QP foram coletadas as últimas quinhentas postagens de cinco usuários diferentes, criando cinco *timelines* para serem avaliadas pela estratégia *DP-Symptom-Identifier*.

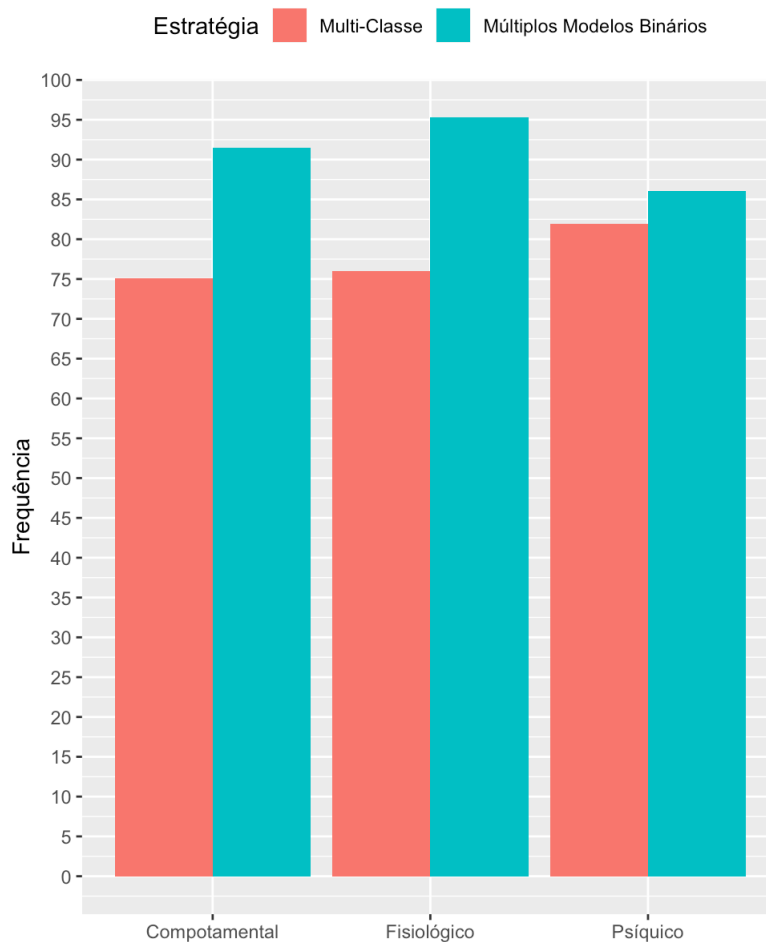


Figura 22: Frequência de Predições corretas em cada categoria sintomática.

Fonte: Autoria própria

Ressalva-se que as discussões apresentadas a seguir foram realizadas em conjunto com a psicóloga que acompanha este trabalho. Primeiramente as *timelines* foram analisadas manualmente para identificar indícios depressivos. Em seguida, as predições do modelo foram analisadas. Finalmente ambas as análises foram comparadas com o objetivo de mensurar a assertividade do modelo. Visando preservar a identidade de cada um dos usuários, os mesmos foram identificados com números de 1 a 5.

Na Figura 23 estão sintetizadas as predições realizadas pelo modelo induzido com relação aos dados de cada um dos usuários. Em seguida são apresentadas as discussões pertinentes para cada um dos usuários.

- **Usuário 1:** após realizar a análise da *timeline* deste usuário verificou-se que o mesmo possui algumas tendências depressivas em determinados momentos. A análise pode ser complementada olhando-se a Tabela 8, cuja primeira coluna identifica o

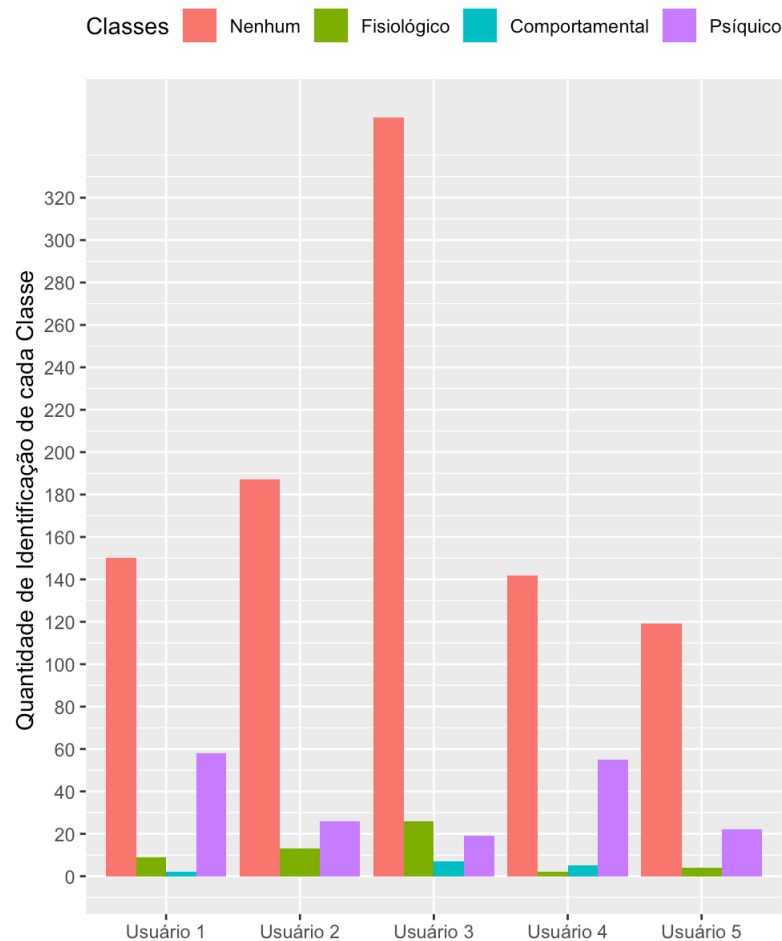


Figura 23: Síntese das predições realizadas pelo modelo para cada um dos usuários.

Fonte: Autoria própria.

tweet, e as demais colunas mostram a rotulação inferida pelo *DP-Symptom-Identifier*, e a data da publicação do *tweet*. Alguns sintomas depressivos são identificados e estão em uma sequência temporal similar a descrita pelo (American Psychiatric Association, 2013). Comprovando a efetividade que o modelo proposto tem em identificar sintomas depressivos, e que quando estes estão organizados de uma forma temporal podem caracterizar tendências depressivas no usuário em questão.

- **Usuário 2:** para este usuário, poucas postagens foram classificadas com algum sintoma, sendo que dentre estes, nenhum remete aos sintomas comportamentais, apenas alguns fisiológicos e psíquicos. Após realizar uma análise aprofundada da *timeline* coletada, foram verificadas e confirmadas pela profissional as predições realizadas pelo modelo. Indicando de forma correta que não existem sequências anormais que possam vir representar sintomas depressivos, apenas foram encontrados relatos do dia a dia.

Tabela 8: Postagens do usuário de id 1

<i>Tweet</i>	Classificação	Data
Eu quero desistir, mas não posso fazer chorar as poucas pessoas que ainda se importam..	Psíquico	07/07/2019
Que Deus me elimine essa noite, amém.	Psíquico	15/07/2019
Nada te destrói mais que seus próprios pensamentos	Psíquico	22/07/2019
Depois de vários remédios, espero que consiga dormir	Fisiológico	22/07/2019

- **Usuário 3:** este usuário possui relatos recorrentes sobre sono, que foram em sua maioria classificados pelo modelo como um sintoma fisiológico, mas podem ter vários outros fatores envolvidos. De uma forma geral, a *timeline* deste usuário apresenta poucos *tweets* que precisam de atenção. Não foram encontradas sequências anormais entre as publicações que possam caracterizar um possível quadro depressivo.
- **Usuário 4:** em geral a *timeline* deste usuário é composta por postagens aleatórias, sem identificação de situações diárias reais. Apesar de alguns *tweets* apresentarem classificações de sintomas, após a análise realizada, não foram identificadas evidências que possibilitassem demonstrar sinais depressivos presentes nas postagens.

Usuário 5: para este usuário, a maioria dos sintomas encontrados foram do tipo psíquico, mas não foi encontrada nenhuma sequência anormal que possa caracterizar indícios depressivos presentes na *timelines* deste usuário. Este usuário também apresenta uma grande aleatoriedade entre as publicações.

Como pode ser observado durante as discussões apresentadas previamente, apenas o Usuário 1 foi identificado com tendências depressivas. Isso ocorre pois em dados momentos, algumas das postagens deste usuário que contém sintomas da depressão, estão dentro do intervalo de duas semanas definido pelo (American Psychiatric Association , 2013). Já para os demais usuários, não foram identificadas tendências depressivas, pois em nenhum momento foram identificadas sequências de postagens contendo algum dos sintomas da depressão dentro do tempo estipulado pelo DSM-V.

Para esta QP, as análises das *timelines* foram realizadas de forma manual para facilitar a validação do processo como um todo. Estas análises tiveram como objetivo verificar se a estratégia seria capaz de identificar os sintomas depressivos e também para validar se realizando a análise de forma temporal é possível identificar traços depressivos, já que não foram encontradas evidências na literatura que se assemelham com o processo

realizado pela estratégia proposta.

Com isso, foi possível validar a efetividade da estratégia proposta que verifica a existência dos sintomas da depressão nas postagens e posteriormente analisa a *timeline* de forma temporal a fim de identificar se os sintomas encontrados no primeiro momento ocorrem de forma contínua e relevante o suficiente durante um período de duas semanas para sinalizar que a *timeline* analisada possui traços depressivos.

Para automatizar o processo de identificar os traços depressivos, faz-se necessário criar uma função de pós-processamento, que análice os *tweets* verificando a existência de possíveis sintomas de depressão em períodos temporais. Para que com base nisso a estratégia apresente uma saída informando se a *timeline* analisada possui traços depressivos ou não. Analisar de forma temporal as postagens é de suma importância pois somente a existência de sintomas da depressão nos *tweets* não caracterizam necessariamente um usuário como detentor de traços depressivos, já que é necessário a presença dos sintomas por pelo menos duas semanas para que isso ocorra.

5.4 AMEAÇAS À VALIDADE

Nesta seção são detalhadas as possíveis ameaças à validade que podem afetar os valores e as conclusões dos estudos empíricos conduzidos.

As ameaças à validade externa estão relacionadas à generalização dos resultados. Analisar dados textuais é uma tarefa complexa e difícil de generalizar, principalmente tratando-se de dados coletados de redes sociais. Com o objetivo de minimizar este problema, foram coletadas postagens de diversos tamanhos, formas de escrita e usuários diferentes. Contudo, não é possível afirmar que os resultados obtidos podem ser generalizados para todas as postagens do *Twitter*. Nesse contexto, replicações futuras são necessárias para corroborar os resultados obtidos.

Já as ameaças à validade por construção possuem um enfoque na relação entre a teoria e o que é observado. Durante a criação da estratégia proposta e na análise das *timelines*, possíveis equívocos podem ter sido cometidos. A fim de mitigar este risco, foi contado com o auxílio de uma psicóloga para avaliar de forma apropriada cada uma das *timelines*, bem como rotular a base de dados utilizada para treinar os algoritmos de AM.

Por fim as ameaças à validade interna caracterizam o grau de confiabilidade entre os resultados esperados e os resultados obtidos. Para mitigar esse risco, o experimento foi executado de forma independente cem vezes para cada estratégia e algoritmo. Além disso,

foram executados testes estatísticos para assegurar que a confiabilidade dos resultados apresentados por meio de tabelas e gráficos sejam realmente coerentes e interpretados da forma apropriada.

5.5 CONSIDERAÇÕES FINAIS

Neste Capítulo foi apresentado um estudo experimental visando investigar a eficiência da estratégia *DP-Symptom-Identifier*, para identificar sintomas depressivos em postagens do *Twitter*. O primeiro experimento realizado teve como objetivo identificar qual o melhor classificador e a melhor estratégia para identificar sintomas depressivos em dados textuais. Os algoritmos utilizados no experimento foram avaliados com relação a medida de desempenho AUC. Os resultados deste primeiro experimento demonstraram que o algoritmo MLP combinado com a estratégia Multi-Classe possuem uma melhor eficiência ao classificar os sintomas depressivos, atingindo uma AUC de 90%.

Em um segundo experimento realizado, objetivou-se verificar se as categorias sintomáticas eram identificadas com a mesma frequência. Como resultado, percebeu-se que os sintomas psíquicos são mais frequentemente identificados pela estratégia multi-classe, e os fisiológicos aparecem com uma maior frequência na estratégia de múltiplos modelos binários.

Por fim, foi executado um experimento visando verificar se a estratégia selecionada no primeiro experimento seria capaz de distinguir de forma correta perfis com tendências depressivas, e perfis sem tais tendências. O resultado deste experimento demonstrou que mesmo com uma quantidade relativamente baixa de exemplos na base de dados de treinamento, de uma forma geral, a estratégia criada conseguiu prever de forma correta os sintomas depressivos, e quando estes são analisados de forma cronológica, torna-se possível verificar a existência ou não de tendências depressivas.

6 CONSIDERAÇÕES FINAIS

A Análise de Sentimentos (AS) vem se provando uma área de pesquisa pertinente para a identificar opiniões ou sentimentos de pessoas em geral para os mais variados assuntos. Entretanto, o uso de AS para o contexto deste trabalho ainda é um tema emergente, principalmente quando se trata do cenário brasileiro, com poucos estudos publicados. Essa defasagem de trabalhos relacionados ao contexto, é também uma grande motivação para a execução deste trabalho.

Neste trabalho foi criada uma estratégia intitulada de *DP-Symptom-Identifier*, capaz de auxiliar no processo de identificação de perfis com sintomas depressivos presentes no *Twitter*. A mesma consiste em analisar postagens do *Twitter* a fim de identificar quais destas possuem algum dos três tipos de sintomas da depressão, que são: Psíquicos, Fisiológicos e Evidências Comportamentais. Para tal fim, são utilizadas técnicas de PLN e algoritmos de AM. Além disso, a estratégia seguida é composta por uma aplicação de coleta de dados, que facilita o processo de coleta das postagens do *Twitter* e uma aplicação *web* que auxilia a análise dos dados gerados como resultado das predições realizadas pelo modelo de AM.

Como parte do trabalho, foi conduzido um estudo experimental com o objetivo de selecionar o algoritmo de AM mais adequado para a tarefa de identificação dos sintomas depressivos. Os resultados mostraram que o MLP é o algoritmo com melhor custo benefício entre desempenho preditivo dos sintomas e tempo de execução.

O estudo experimental ainda objetivou avaliar a efetividade da estratégia apresentada, por meio da análise de *timelines* de usuários do *Twitter*. Os resultados desta parte do estudo demonstram que o *DP-Symptom-Identifier* é capaz de reconhecer sintomas de depressão nas postagens, auxiliando no processo de identificação de perfis depressivos. Pois, após os dados serem classificados pelo modelo de AM, os mesmos podem ser consultados na aplicação *Web* que apresenta as postagens em ordem cronológica, facilitando assim a identificação de possíveis sinais depressivos contidos nas *timelines*.

Realizando uma comparação com os trabalhos existentes na literatura, foi identificado que a estratégia de analisar as postagens com o objetivo de identificar os sintomas depressivos, e ao fim, realizar a análise da *timeline* como um todo, buscando indícios de depressão, não foram utilizadas pelos trabalhos encontrados. Em grande maioria os mesmos analisavam os dados com relação a polaridade, podendo ser negativo, neutro ou positivo, ou ainda analisavam com base nas informações semânticas dos dados.

As contribuições deste trabalho são apresentadas na Seção 6.1. Já as limitações da estratégia e as dificuldades encontradas no desenvolvimento deste trabalho são apresentadas na Seção 6.2. Por fim na Seção 6.3 são descritos os possíveis trabalhos futuros a serem abordados.

6.1 CONTRIBUIÇÕES

Além da criação de uma estratégia que se provou capaz de identificar sintomas depressivos em postagens do *Twitter*, e que demonstrou a possibilidade de analisar perfis depressivos por meio dos sintomas isolados desta doença, este trabalho ainda contribui com:

- criação de uma base de dados em português, rotulada por uma psicóloga, a qual pode ser utilizada por modelos de AM no contexto de problemas psicológicos;
- uma aplicação para a coleta de dados do *Twitter*;
- uma aplicação *web* para auxiliar na análise e rotulação de postagens;
- um MS da literatura no contexto deste trabalho; e
- realização de estudos experimentais empíricos que endossam a literatura da área.

6.2 LIMITAÇÕES E DIFICULDADES ENCONTRADAS

A principal limitação existente na estratégia *Dp-Symptom-Identifier* é em relação a necessidade da avaliação temporal de forma manual, já que a mesma classifica os sintomas nas postagens, e as ordena em ordem cronológica, para que um profissional avalie se existem tendências depressivas entre as postagens. A solução para resolver esta limitação esta descrita na Subseção de trabalhos futuros (Subseção 6.3).

A principal dificuldade encontrada na realização deste trabalho foi a criação da base de dados de treinamento, já que a mesma possui uma grande influência no resultado das predições dos modelos de AM. O processo de identificação de palavras e sentenças relevantes, a coleta e rotulação dos dados de forma correta foi o mais demorado e complexo dentre todos os executados neste trabalho.

Também foram encontradas dificuldades para definir o melhor método para extrair as características dos dados, foram testadas várias configurações do TF-IDF e do BoW a fim de melhorar os resultados obtidos.

6.3 TRABALHOS FUTUROS

Com base no MS conduzido e na solução proposta no Capítulo 4.1, ainda existem alguns trabalhos futuros que podem ser realizados no contexto deste trabalho, os quais são descritos a seguir:

- aumentar a quantidade de exemplos na base de dados de treinamento;
- realizar experimentos adicionais utilizando outros métodos de extração de características;
- realizar experimentos adicionais utilizando n-gramas;
- validar a solução proposta em mais usuários;
- desenvolver funcionalidades na aplicação *web* visando permitir o ajuste da classificação pelo especialista, assim as predições realizadas podem ser incorporadas na base de dados;
- ajustar os hiperparâmetros dos algoritmos de AM visando atingir melhores taxas de assertividade nas predições;
- integrar o modelo de AM em uma ferramenta para que seja possível realizar as predições e análises dos perfis de forma automatizada;
- desenvolver uma aplicação, que integrada ao modelo de AM possa realizar as predições de determinado usuário, e com base no resultado destas, sugerir conteúdos motivacionais a fim de aliviar os sintomas depressivos; e
- expandir a coleta de dados para outras redes sociais.

REFERÊNCIAS

- AL-MOSAIWI, M.; JOHNSTONE, T. In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. **Clinical Psychological Science**, v. 6, n. 4, p. 529–542, 2018.
- Aldarwish, M. M.; Ahmad, H. F. Predicting depression levels using social media posts. In: **2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)**. [S.l.: s.n.], 2017. p. 277–280.
- ALONSO, S. G. et al. Data mining algorithms and techniques in mental health: A systematic review. **Journal of Medical Systems**, v. 42, n. 9, p. 161, Jul 2018. ISSN 1573-689X. Disponível em: <<https://doi.org/10.1007/s10916-018-1018-2>>.
- ALPAYDIN, E. **Introduction to machine learning**. [S.l.]: MIT press, 2009.
- American Psychiatric Association . Book. **Diagnostic and statistical manual of mental disorders : DSM-5**. 5th ed.. ed. [S.l.]: American Psychiatric Association Arlington, VA, 2013. xlv, 947 p. ; p. ISBN 089042554 0890425558 9780890425541 9780890425558.
- BASIL, V.; WEISS, D. A methodology for collecting valid software engineering data. **IEEE Trans. Software Eng.**, v. 10, p. 728–738, 11 1984.
- BIRJALI, M.; BENI-HSSANE, A.; ERRITALI, M. A method proposed for estimating depressed feeling tendencies of social media users utilizing their data. In: . Cham: Springer International Publishing, 2017. p. 413–420. ISBN 978-3-319-52941-7.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: **Proceedings of the Fifth Annual Workshop on Computational Learning Theory**. [S.l.]: ACM, 1992. (COLT '92), p. 144–152.
- BOUKIL, S. et al. Deep learning algorithm for suicide sentiment prediction. In: . Cham: Springer International Publishing, 2019. p. 261–272.
- BREIMAN, L. Random forests. **Machine Learning**, Kluwer Academic Publishers, Hingham, MA, USA, v. 45, n. 1, p. 5–32, out. 2001.
- BRERETON, P. et al. Lessons from applying the systematic literature review process within the software engineering domain. **Systems and Software**, New York, NY, USA, v. 80, n. 4, p. 571–583, 2007.
- BRIAND, A.; ALMEIDA, H.; MEURS, M.-J. Analysis of social media posts for early detection of mental health conditions. In: . Cham: Springer International Publishing, 2018. p. 133–143. ISBN 978-3-319-89656-4.
- BUDGEN, D. et al. Using mapping studies in software engineering. In: **PPIG**. [S.l.: s.n.], 2008. v. 8, p. 195–204.

- Dao, B. et al. Nonparametric discovery of online mental health-related communities. In: **2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)**. [S.l.: s.n.], 2015. p. 1–10.
- Deshpande, M.; Rao, V. Depression detection using emotion artificial intelligence. In: **2017 International Conference on Intelligent Sustainable Systems (ICISS)**. [S.l.: s.n.], 2017. p. 858–862.
- FLACH, P.; HERNANDEZ-ORALLO, J.; FERRI, C. A coherent interpretation of auc as a measure of aggregated classification performance. In: . [S.l.: s.n.], 2011. p. 657–664.
- HASAN, M.; RUNDENSTEINER, E.; AGU, E. Automatic emotion detection in text streams by analyzing twitter data. **International Journal of Data Science and Analytics**, v. 7, n. 1, p. 35–51, Feb 2019. ISSN 2364-4168.
- Hassan, A. U. et al. Sentiment analysis of social networking sites (sns) data using machine learning approach for the measurement of depression. In: **2017 International Conference on Information and Communication Technology Convergence (ICTC)**. [S.l.: s.n.], 2017. p. 138–140.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. [S.l.: s.n.], 1999.
- HAYKIN, S. S. **Neural networks and learning machines**. Third. Upper Saddle River, NJ: Pearson Education, 2009.
- INDURKHYA, N.; DAMERAU, F. J. (Ed.). **Handbook of Natural Language Processing, Second Edition**. [S.l.]: Chapman and Hall/CRC, 2010.
- ISLAM, M. R. et al. Depression detection from social network data using machine learning techniques. **Health Information Science and Systems**, v. 6, n. 1, p. 8, Aug 2018.
- Islam, M. R. et al. Detecting depression using k-nearest neighbors (knn) classification technique. In: **2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)**. [S.l.: s.n.], 2018. p. 1–4.
- JOHANNSEN, D.; BIEMANN, C. Between the lines: Machine learning for prediction of psychological traits - a survey. In: HOLZINGER, A. et al. (Ed.). **Machine Learning and Knowledge Extraction**. Cham: Springer International Publishing, 2018. p. 192–211. ISBN 978-3-319-99740-7.
- Keumhee Kang; Chanhee Yoon; Eun Yi Kim. Identifying depressive users in twitter using multimodal analysis. In: **2016 International Conference on Big Data and Smart Computing (BigComp)**. [S.l.: s.n.], 2016. p. 231–238. ISSN 2375-9356.
- KITCHENHAM, B. et al. Systematic literature reviews in software engineering—a tertiary study. **Information and software technology**, Elsevier, v. 52, n. 8, p. 792–805, 2010.
- Larsen, M. E. et al. We feel: Mapping emotion on twitter. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 4, p. 1246–1252, July 2015. ISSN 2168-2194.

- LEIVA, V.; FREIRE, A. Towards suicide prevention: Early detection of depression on social media. In: **Internet Science**. Cham: Springer International Publishing, 2017. p. 428–436. ISBN 978-3-319-70284-1.
- LIDDY, E. D. Natural language processing. **Encyclopedia of Library and Information Science**, Inc, New York, NY, USA, n. 2, 2001.
- LIU, B. Sentiment analysis and opinion mining. **Synthesis lectures on human language technologies**, Morgan & Claypool Publishers, v. 5, n. 1, p. 1–167, 2012.
- LOSADA, D. E.; CRESTANI, F. A test collection for research on depression and language use. In: FUHR, N. et al. (Ed.). **Experimental IR Meets Multilinguality, Multimodality, and Interaction**. Cham: Springer International Publishing, 2016. p. 28–39. ISBN 978-3-319-44564-9.
- LOSADA, D. E.; CRESTANI, F.; PARAPAR, J. erisk 2017: Clef lab on early risk prediction on the internet: Experimental foundations. In: JONES, G. J. et al. (Ed.). **Experimental IR Meets Multilinguality, Multimodality, and Interaction**. Cham: Springer International Publishing, 2017. p. 346–360. ISBN 978-3-319-65813-1.
- MA, L.; WANG, Z.; ZHANG, Y. Extracting depression symptoms from social networks and web blogs via text mining. In: . Cham: Springer International Publishing, 2017. p. 325–330. ISBN 978-3-319-59575-7.
- MITCHELL, T. M. **Machine Learning**. 1. ed. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN 0070428077, 9780070428072.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. In: **Sistemas Inteligentes Fundamentos e Aplicações**. 1. ed. Barueri-SP: Manole Ltda, 2003. p. 89–114. ISBN 85-204-168.
- NEMENYI, P. **Distribution-free Multiple Comparisons**. [S.l.]: Princeton University, 1963.
- NGUYEN, T. et al. Using linguistic and topic analysis to classify sub-groups of online depression communities. **Multimedia Tools and Applications**, v. 76, n. 8, p. 10653–10676, Apr 2017. ISSN 1573-7721.
- Nguyen, T. et al. Affective and content analysis of online depression communities. **IEEE Transactions on Affective Computing**, v. 5, n. 3, p. 217–226, July 2014. ISSN 1949-3045.
- Noureen, A.; Qamar, U.; Ali, M. Semantic analysis of social media and associated psychotic behavior. In: **2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)**. [S.l.: s.n.], 2017. p. 1621–1630.
- ORGANIZATION, W. H. et al. **Depression and other common mental disorders: global health estimates**. [S.l.], 2017.
- PROVOST, F.; KOHAVI, R. Glossary of terms. **Journal of Machine Learning**, v. 30, n. 2-3, p. 271–274, 1998.

Rosa, R. L. et al. A knowledge-based recommendation system that includes sentiment analysis and deep learning. **IEEE Transactions on Industrial Informatics**, v. 15, n. 4, p. 2124–2135, April 2019.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.

Seah, J. H. K.; Jin Shim, K. Data mining approach to the detection of suicide in social media: A case study of singapore. In: **2018 IEEE International Conference on Big Data (Big Data)**. [S.l.: s.n.], 2018. p. 5442–5444.

TAO, X. et al. Sentiment analysis for depression detection on social networks. In: . [S.l.: s.n.], 2016. p. 807–810. ISBN 978-3-319-49585-9.

TUNG, C.; LU, W. Analyzing depression tendency of web posts using an event-driven depression tendency warning model. **Artif. Intell. Med.**, Elsevier Science Publishers Ltd., Essex, UK, v. 66, n. C, p. 53–62, jan. 2016. ISSN 0933-3657.

VEDULA, N.; PARTHASARATHY, S. Emotional and linguistic cues of depression from social media. In: **DH**. [S.l.: s.n.], 2017.

WANG, X. et al. A depression detection model based on sentiment analysis in micro-blog social network. In: **PAKDD Workshops**. [S.l.: s.n.], 2013.

Wang, X.; Zhang, C.; Sun, L. An improved model for depression detection in micro-blog social network. In: **2013 IEEE 13th International Conference on Data Mining Workshops**. [S.l.: s.n.], 2013. p. 80–87. ISSN 2375-9232.

WU, M. Y. et al. A deep architecture for depression detection using posting, behavior, and living environment data. **Journal of Intelligent Information Systems**, Oct 2018. ISSN 1573-7675.

ZUORBA, H. D.; OLAN, C. L. O.; CANTARA, A. D. A framework for identifying excessive sadness in students through twitter and facebook in the philippines. In: **ICBRA 2017**. [S.l.: s.n.], 2017.