

UNIVERSIDADE FEEVALE

MAICON BASILIO TEIXEIRA UNZER

MACHINE LEARNING PARA O APOIO AO DIAGNÓSTICO DE
DEPRESSÃO

Novo Hamburgo
2020

MAICON BASILIO TEIXEIRA UNZER

MACHINE LEARNING PARA O APOIO AO DIAGNÓSTICO DE
DEPRESSÃO

Trabalho de Conclusão de Curso
apresentado como requisito parcial
à obtenção do grau de Bacharel em
Ciência da Computação pela
Universidade Feevale

Orientador: Prof. Dr. Juliano Varella de Carvalho

Novo Hamburgo
2020

AGRADECIMENTOS

Gostaria de agradecer a todos os que, de alguma maneira, contribuíram para a realização desse trabalho de conclusão, em especial:

Aos meus pais, Paulo e Liamar, por compreenderem e me apoiarem no desenvolvimento do meu trabalho.

Ao meu filho, Enzo Gabriel, por compreender as ausências e por ser meu porto seguro, dando motivação nos momentos mais difíceis.

A empresa *BeeTouch*, por fornecer os dados e o suporte necessário para a construção deste trabalho.

Ao professor Dr. Juliano, pela disposição, pelas sugestões, parceria e paciência investida neste período de construção do trabalho de conclusão, para que pudéssemos desenvolver algo de qualidade.

A todos que, de alguma forma, contribuíram para a conclusão deste trabalho.

Muito obrigado!

RESUMO

A depressão é considerada o mal do século e é a responsável por mais anos perdidos de uma pessoa que qualquer outra doença. Muitas pessoas em quadro depressivo acabam cometendo suicídio por não saberem como lidar com essa situação. A depressão quando diagnosticada no estágio inicial, torna o tratamento mais efetivo. Existem diversas aplicações que com base em questionários são capazes de apoiar a detecção de um indivíduo com depressão. A plataforma AVAXpsi foi desenvolvida com o intuito de auxiliar os psicólogos em suas tarefas cotidianas, porém, além disso, desenvolveu um recurso que consegue detectar pacientes em depressão através das respostas inseridas na avaliação psicológica. Esse recurso de detecção, porém, não é feito com técnicas de predição de dados. Este trabalho utilizou os dados coletados das avaliações psicológicas, contidas na plataforma AVAXpsi e construiu modelos preditivos utilizando as técnicas de *Machine Learning Support: Vector Machine* e *Decision Tree*, com a finalidade de identificar pacientes com depressão. Os modelos desenvolvidos neste trabalho obtiveram uma excelente acurácia, tanto com dados balanceados quanto desbalanceados.

Palavras-chave: *Machine Learning*. Depressão. Análise preditiva. Psicologia. Tecnologia.

ABSTRACT

The depression disorder is the greatest blight of the century and is responsible for more lost years than any other disease. Many people with depression end up committing suicide because they do not know how to deal with this situation. Depression when diagnosed at an early stage, makes treatment more effective. There are several applications that based on questionnaires can support the detection of an individual with depression. The AVAXpsi platform was developed with the intention of helping psychologists in their daily tasks, however, in addition, it has developed a resource that can detect patients in depression through the responses inserted in the psychological assessment. This detection feature, however, is not done with data prediction techniques. This work used the data collected from psychological evaluations, contained in the AVAXpsi platform and built predictive models using the Machine Learning Support Vector Machine and Decision Tree techniques, in order to identify patients with depression. The models developed in this work obtained excellent accuracy, with both balanced and unbalanced data.

Keywords: Machine Learning. Depression. Predictive Analysis. Psychology. Technology.

LISTA DE FIGURAS

FIGURA 1 - PAÍSES E A PREVALÊNCIA DA DEPRESSÃO	18
FIGURA 2 - FASES DO TRATAMENTO DA DEPRESSÃO E SUAS INTERCORRÊNCIAS.	22
FIGURA 3 - TELA DE CADASTRO DA PLATAFORMA AVAXPSI	25
FIGURA 4 - TELA DO FORMULÁRIO DE ENTREVISTA DA PLATAFORMA AVAXPSI.....	26
FIGURA 5 - QUESTÕES EXISTENTES NA PLATAFORMA AVAXPSI QUE APONTAM O RISCO DE DEPRESSÃO	27
FIGURA 6 - TOTAL DE ARTIGOS RETORNADOS PELAS BASES DE DADOS PESQUISADAS.....	33
FIGURA 7 - RESULTADO OBTIDO APÓS A EXECUÇÃO DA SEGUNDA E TERCEIRA ETAPA DA SELEÇÃO DOS ESTUDOS	33
FIGURA 8 - RESULTADO DA SELEÇÃO DOS ARTIGOS PARA LEITURA INTEGRAL.....	34
FIGURA 9 - DESEMPENHO OBTIDO NO ESTUDO DE (KIM ET AL., 2018) UTILIZANDO A TÉCNICA <i>DECISION TREE</i>	38
FIGURA 10 - DESEMPENHO DAS DEMAIS TÉCNICAS ML QUE FORAM UTILIZADAS NO ESTUDO DE (KIM ET AL., 2018)	39
FIGURA 11 - RESULTADO DO DESEMPENHO DE CADA CLASSIFICADOR UTILIZADO NO ESTUDO DE (BYUN ET AL., 2019).....	40
FIGURA 12 - DESEMPENHO OBTIDO EM CADA TÉCNICA ML UTILIZADA NO ESTUDO DE (SAU; BHAKTA, 2017)	40
FIGURA 13 - PROCESSO DE EXTRAÇÃO DE CONHECIMENTO.....	45
FIGURA 14 - EXEMPLO DE <i>BINNING</i>	48
FIGURA 15 - EXEMPLO DE CLASSES NÃO SEPARÁVEIS LINEARMENTE	54
FIGURA 16 - CLASSES EM UMA DIMENSÃO MAIS ALTA COM USO DO KERNEL	54
FIGURA 17 - HIPERPLANO E MARGENS EM SVM.....	55
FIGURA 18 - ARVORE DE DECISÃO PARA COMPRA DE UM CARRO	56
FIGURA 19 - MATRIZ DE CONFUSÃO	58
FIGURA 20 - EQUAÇÃO DA ACURÁCIA.....	59
FIGURA 21 - EQUAÇÃO DA SENSIBILIDADE	59
FIGURA 22 - EQUAÇÃO DA ESPECIFICIDADE.....	60
FIGURA 23 - ÁREA SOB A CURVA ROC	60
FIGURA 24 - <i>DATASET</i> DE PERCENTUAL DE VALORES NA POR ATRIBUTO	65
FIGURA 25 - BOXPLOT DE PERCENTUAL DE VALORES FALTANTES POR ATRIBUTO.....	66

FIGURA 26 - FLUXO DE ANÁLISE PARA SELEÇÃO DOS ATRIBUTOS E GERAÇÃO DA CORRELAÇÃO	66
FIGURA 27 - EQUAÇÕES PARA DETERMINAR FAIXA INICIAL E FINAL DO PERCENTUAL DE VALORES NA ACEITOS PARA ANÁLISE.....	67
FIGURA 28 - RESULTADO DA CORRELAÇÃO UTILIZANDO A ESTRATÉGIA UM DA SELEÇÃO DOS ATRIBUTOS E A ESTRATÉGIA UM DA MANIPULAÇÃO DOS VALORES AUSENTES	67
FIGURA 29 - RESULTADO DA CORRELAÇÃO UTILIZANDO A ESTRATÉGIA UM DA SELEÇÃO DOS ATRIBUTOS E A ESTRATÉGIA DOIS DA MANIPULAÇÃO DOS VALORES AUSENTES	68
FIGURA 30 - RESULTADO DA CORRELAÇÃO UTILIZANDO A ESTRATÉGIA DOIS DA SELEÇÃO DOS ATRIBUTOS E A ESTRATÉGIA UM DA MANIPULAÇÃO DOS VALORES AUSENTES	69
FIGURA 31 - RESULTADO DA CORRELAÇÃO UTILIZANDO A ESTRATÉGIA DOIS DA SELEÇÃO DOS ATRIBUTOS E A ESTRATÉGIA DOIS DA MANIPULAÇÃO DOS VALORES AUSENTES	69

LISTA DE QUADROS

QUADRO 1 - ARTIGOS SELECIONADOS PARA LEITURA INTEGRAL	34
QUADRO 2 - ALGORITMOS DE ML UTILIZADOS POR CADA ARTIGO	36
QUADRO 3 - LINGUAGENS E FERRAMENTAS UTILIZADAS PELOS ARTIGOS SELECIONADOS	36
QUADRO 4 - MÉTRICAS E FORMAS DE VALIDAÇÃO DO DESEMPENHO DO MODELO ML UTILIZADO EM CADA ARTIGO SELECIONADO.	37
QUADRO 5 - TÉCNICAS ML UTILIZADAS PELOS ARTIGOS SELECIONADOS	38
QUADRO 6 - QUESTIONÁRIO REALIZADO AO CTO DA <i>BEETOUCH</i> E SUAS RESPOSTAS	75

LISTA DE TABELAS

TABELA 1 – TOTAL DE OBSERVAÇÕES E ATRIBUTOS POR <i>DATASET</i>	63
TABELA 2 - MÉTRICAS DE PERFORMANCE DE CADA MODELO DE PREDIÇÃO <i>SVM</i>	71
TABELA 3 - MÉTRICAS DE PERFORMANCE DE CADA MODELO DE PREDIÇÃO <i>DECISION TREE</i>	72
TABELA 4 - RESULTADO OBTIDO POR TÉCNICA <i>ML</i> UTILIZANDO A TÉCNICA DE VALIDAÇÃO QUE OBTIVE MELHOR RESULTADO DENTRE AS DEMAIS	73
TABELA 5 - MÉTRICAS DE PERFORMANCE DE CADA MODELO DE PREDIÇÃO <i>SVM</i> COM <i>DATASET</i> BALANCEADO	74
TABELA 6 - MÉTRICAS DE PERFORMANCE DE CADA MODELO DE PREDIÇÃO <i>DECISION TREE</i> COM <i>DATASET</i> BALANCEADO	74

LISTA DE ABREVIATURAS E SIGLAS

ML	<i>Machine Learning</i>
SVM	<i>Support Vector Machine</i>
TDM	Transtorno Depressivo Maior
TCC	Terapia Cognitivo-Comportamental
TIP	Terapia Interpessoal
AD	Antidepressivo
ISRS	Inibidores Seletivos da Recaptação de Serotonina
IMAO	Inibidores da Monoaminoxidase
TEC	Terapia Eletroconvulsiva
ADDS	Apoio ao Diagnóstico de Depressão e Avaliação do Risco de Suicídio
UFRGS	Universidade Federal do Rio Grande do Sul
CRP	Conselho Regional de Psicologia
DSM	<i>Diagnostic and Statistical Manual</i>
PICOC	População, Intervenção, Comparação, Resultados e Contexto
RS	Revisão Sistemática
EEG	Eletroencefalograma
AUC	<i>Area Under Curve</i>
VPP	Valor Preditivo Positivo
VPN	Valor Preditivo Negativo
DT	<i>Decision Tree</i>
RT	<i>Random Forest</i>
rs-fMRI	<i>Resting State fMRI</i>
OMS	Organização Mundial da Saúde

SUMÁRIO

1 INTRODUÇÃO	13
2 DEPRESSÃO	17
2.1 TIPOS DE DEPRESSÃO	18
2.1.1 Transtorno depressivo maior.....	19
2.1.2 Transtorno depressivo persistente (Distímia)	19
2.1.3 Depressão perinatal	19
2.1.4 Transtorno afetivo sazonal.....	20
2.1.5 Depressão psicótica	20
2.2 TRATAMENTOS.....	20
2.3 TECNOLOGIAS DE APOIO A DEPRESSÃO	24
2.4 PLATAFORMA “AVAXPSI” NA DEPRESSÃO.....	25
3 REVISÃO SISTEMÁTICA DE TÉCNICAS DE MACHINE LEARNING APLICADAS PARA A PREDIÇÃO DE PACIENTES COM DEPRESSÃO	28
3.1 PROTOCOLO DA REVISÃO SISTEMÁTICA.....	28
3.1.1 O protocolo	28
3.2 DESENVOLVIMENTO DA REVISÃO SISTEMÁTICA	32
3.3 AVALIAÇÃO DOS CRITÉRIOS DE QUALIDADE	35
3.4 AVALIAÇÃO DOS RESULTADOS DOS ARTIGOS SELECIONADOS	41
4 MACHINE LEARNING	44
4.1 SELEÇÃO DOS DADOS.....	45
4.2 PRÉ-PROCESSAMENTO	46
4.2.1 Limpeza dos dados.....	47
4.2.2 Transformação	49
4.3 DATA MINING	49
4.3.1 Associação.....	50
4.3.2 Regressão	50
4.3.3 Clustering.....	51
4.3.4 Classificação	52
4.4 PÓS-PROCESSAMENTO	57
4.4.1 Knowledge filtering.....	57
4.4.2 Avaliação dos modelos.....	58

5 MODELOS DE PREDIÇÃO DE PACIENTES COM DEPRESSÃO A PARTIR DOS DADOS DA PLATAFORMA AVAXPSI.....	62
5.1 QUALIDADE DE DADOS.....	62
5.1.1 Seleção dos atributos.....	63
5.2 GERAÇÃO DOS MODELOS DE PREDIÇÃO	70
5.2.1 SVM.....	71
5.2.2 <i>Decision tree</i>	72
5.3 LIMITAÇÃO DOS MODELOS.....	73
5.4 ENTREVISTA COM O CTO DA BEETOUCH.....	75
6 CONCLUSÃO.....	76
REFERÊNCIAS BIBLIOGRÁFICAS	78

1 INTRODUÇÃO

Depressão é considerada pela Organização Mundial da Saúde (OMS) como o "Mal do Século" (MINISTÉRIO DA SAÚDE, 2018). Trata-se de uma desordem mental que faz com que o indivíduo entre em estado de tristeza profunda. Ela pode parecer devastadora e ser uma experiência muito solitária, fazendo com que a pessoa não sinta vontade de fazer as mesmas coisas (ARDEN, 2015). É globalmente a responsável por mais "anos perdidos" por incapacidade do que qualquer outra condição (KERRI SMITH, 2014).

A causa de depressão está ligada a alterações químicas no cérebro do indivíduo deprimido, principalmente em relação aos neurotransmissores: serotonina, noradrenalina e, em menor proporção, dopamina. Ao contrário do que normalmente se pensa, os fatores psicológicos e sociais muitas vezes são consequência e não a causa da depressão (MINISTÉRIO DA SAÚDE, 2018).

A depressão é caracterizada por humor baixo, falta de energia, tristeza, insônia e uma desabilidade de aproveitar a vida (CUI, 2015). O indivíduo com quadro depressivo também pode sofrer mudanças no apetite, dormir mais ou menos, redução da concentração, indecisão, inquietação, sentimentos de inutilidade, culpa ou desesperança e pensamentos de autoagressão ou suicídio (WORLD HEALTH ORGANIZATION, 2017).

Segundo a OMS, 300 milhões de pessoas sofrem do transtorno no planeta. No Brasil, quase 6 % da população, um total de 11,5 milhões de pessoas sofrem com a doença, tornando o Brasil o país campeão em casos de depressão na América Latina (TUCHLINSKI, 2018). Infelizmente ainda há um estigma sobre saúde mental, algumas pessoas ainda veem essas desordens, tal como depressão, como uma fraqueza (MORIN, 2019).

Já existem aplicações que auxiliam na prevenção, diagnóstico, apoio e acompanhamento às pessoas que sofrem de depressão, ansiedade e transtorno bipolar. O uso da tecnologia na área da saúde mental não tem como objetivo substituir os profissionais, e sim complementar o tratamento, com o propósito de atingir os pacientes em pontos que a abordagem tradicional não tem surtido efeito (OLIVETO, 2018). Existe, por exemplo, um aplicativo chamado ADDS (Apoio ao Diagnóstico de Depressão e na definição de risco de Suicídio), construído pelo projeto de pesquisa TelesaúdeRS, projeto esse desenvolvido pela Universidade Federal do Rio Grande do Sul, que visa, através de perguntas, auxiliar o profissional da saúde a chegar em um diagnóstico (SOCIAL GOOD BRASIL, 2017) .

Outro aplicativo desenvolvido no apoio a pacientes com quadro depressivo é o "Sanvello - Stress & Anxiety Help" que foi desenvolvido pela empresa Pacífica. Esse aplicativo não se

propõe a apoiar os profissionais no diagnóstico de seus pacientes, e sim, apoiar e monitorar os pacientes. Esse aplicativo oferece exercícios de relaxamento, registro de humor, pensamentos e metas diárias (SOCIAL GOOD BRASIL, 2017).

Existem também tecnologias que fazem o acompanhamento de pacientes com depressão a partir do monitoramento cardíaco. Através de um dispositivo vestível, o qual é colocado no peito do paciente (ROH; HONG; YOO, 2014). Esse sistema é baseado em uma placa de circuito muito fina, adequada para ser colocada direto na roupa (ROH; HONG; YOO, 2014).

Outra tecnologia bastante utilizada atualmente na área da saúde é *Machine Learning*. "*Machine Learning* (ML) foi introduzido no campo da medicina com o objetivo de prover ferramentas capazes de melhorar a acurácia e precisão dos diagnósticos, minimizando tarefas que requerem intervenção humana [...]" (VICTOR, 2019, tradução nossa).

Existem diversas tecnologias que utilizam técnicas de *Machine Learning* para apoiar os pacientes e os profissionais da saúde no tratamento e diagnóstico de pacientes com depressão. Há muitas técnicas de análise de dados que podem ser usadas em *Machine Learning*, tais como classificadores Bayesianos, SVM (*Support Vector Machine*), Redes Neurais e métodos baseados em amostra (MARTINEZ, 2019).

É possível criar modelos em *Machine Learning* para detectar se o indivíduo está com depressão analisando os posts do *Twitter*. Um modelo criado por Martinez (2019) para um experimento, utilizou o teorema de *Bayes* e foi desenvolvido utilizando a linguagem de programação *Python*. Para a construção do modelo foi utilizada uma amostra de 8.000 tweets extraídos a partir do *Dataset Sentiment140*¹, uma base de dados que contém mais de 1.600.000 *tweets*, classificados através do atributo (MARTINEZ, 2019) como, 0 sendo um *tweet* negativo, 2 sendo um neutro e 4 um *tweet* positivo.

Outro *Dataset* utilizado para construção do modelo, foi o construído a partir da extração de *posts* do *Twitter*, os quais continham *hashtags* com termos como, *Depressed*, *Feeling Down*, *Depression*, *Anxiety*, entre outros. Com o auxílio da ferramenta de *Web Scraping* chamada *Twint*², foi possível extrair 2.314 *Tweets* para auxiliar na construção do modelo de previsão de usuários com depressão (MARTINEZ, 2019).

Outro trabalho apresenta modelos de ML que foram aplicados na classificação de indivíduos com depressão, analisando sinais extraídos da eletroencefalografia (HOSSEINIFARD; MORADI; ROSTAMI, 2013). Eletroencefalografia (EEG) é um método de monitoramento de

¹ <https://www.kaggle.com/kazanova/sentiment140>

² <https://github.com/twintproject/twint>

atividade elétrica do cérebro, onde, colocando eletrodos no escalpo do paciente com algum componente condutivo como gel ou pasta é possível extrair as ondas elétricas produzidas pelo cérebro (WIKIPEDIA, [s.d.]).

O estudo para a criação do modelo de classificação, a partir de dados coletados da EGG, contou com 45 pacientes normais e 45 pacientes com depressão não medicados, e utilizou técnicas de classificação como, *K-nearest Neighbor*, *Linear Discriminant Analysis* e *Logistic Regression* (HOSSEINIFARD; MORADI; ROSTAMI, 2013).

Existem também plataformas de entrevistas e avaliações psicológicas, que visam apoiar os profissionais a serem mais precisos em seus diagnósticos. Um exemplo disso, já previamente citado, é o ADDS, plataforma desenvolvida pelo projeto TelessaúdeRS da UFRGS, criada para auxiliar os profissionais das Unidades Básicas de Saúde (UBS) no diagnóstico de pacientes com depressão ou tendência suicida.

Outra plataforma criada para auxiliar os profissionais da área da saúde mental, é a plataforma AVAXpsi. Essa plataforma oferece um pacote completo aos psicólogos que vai desde auxiliar o profissional em suas atividades administrativas, como emissão de atestados, laudos e relatórios até as atividades mais técnicas do profissional, como o diagnóstico de um paciente.

A plataforma AVAXpsi foi desenvolvida pela empresa *Bee Touch*³. Atualmente a AVAXpsi conta com mais de mil avaliações psicológicas, coletadas de diversos psicólogos e pacientes. A avaliação psicológica da plataforma conta com perguntas de diferentes tópicos, como, o contexto de trabalho do paciente, se o paciente é exposto a riscos, quais as atividades, *hobbies* que ele gosta de praticar, dentre outras. A partir dessa base de dados, é possível a criação de modelos de predição que podem apoiar o diagnóstico e com isto, oferecer mais um recurso tecnológico na plataforma.

A AVAXpsi também possui perguntas que tem o intuito de identificar se o paciente possui ou não depressão. Perguntas como, “Sentiu-se triste quase todos os dias?”, e “Sentiu-se sem interesse e prazer pelas coisas que lhe agradam habitualmente, quase todos os dias?”, são alguns exemplos do conteúdo existente na plataforma. Apesar de ela utilizar algumas regras para “diagnosticar” os pacientes com depressão, essa análise não é feita a partir de técnicas de *Machine Learning*, ou seja, caso as características dos pacientes mudem, não é possível fazer com que o algoritmo se adapte ou que aprenda com os novos dados inseridos no sistema.

³ <http://www.beetouch.com.br/>

O objetivo deste trabalho foi desenvolver modelos preditivos utilizando técnicas de ML para prever pacientes que possam estar em quadro depressivo, utilizando os dados disponíveis na plataforma AVAXpsi. O intuito disso foi trazer uma nova perspectiva em relação aos dados quando se trata da análise feita para identificar pacientes com depressão.

No capítulo dois foi falado sobre o que é depressão, quais os principais sintomas dessa doença, as causas, os tipos de depressão, os tratamentos existentes, as aplicações voltadas ao tratamento e diagnóstico de depressão a fim de entender o cenário para o qual os modelos de predição foram construídos. No terceiro capítulo foi feita uma revisão sistemática, a fim de encontrar trabalhos que utilizaram técnicas de *Machine Learning* na classificação de pessoas com depressão.

No capítulo quatro foi detalhado como funciona o processo de extração de conhecimento, tratamento de valor ausentes, seleção de atributos e outras estratégias garantindo que os dados selecionados para análise fossem de qualidade, e que os modelos de predição atingissem uma boa performance. Neste capítulo também foram abordadas técnicas de *Machine Learning* utilizadas para o desenvolvimento dos modelos, técnicas essas que foram selecionadas a partir dos artigos escolhidos na revisão sistemática.

No capítulo cinco foi descrito o processo de desenvolvimento dos modelos de predição, quais atributos foram selecionados para classificarem os pacientes com depressão, como foi feito o tratamento dos valores ausentes encontrados, quais as métricas utilizadas para medir a performance dos modelos e os resultados obtidos por cada um nos diferentes cenários apresentados. E por fim, um relato do CTO (*Chief Technology Officer*) da *Beetouch*, empresa que disponibilizou os dados para análise. O último capítulo faz uma conclusão a respeito do trabalho.

2 DEPRESSÃO

Depressão é uma doença médica crônica que pode afetar pensamentos, humor e a saúde física do indivíduo. Ela é caracterizada por falta de energia, tristeza, insônia e humor baixo (CUI, 2015). Ela atinge as pessoas com golpe duplo, pois há a agonia dos próprios sintomas, como a paralisia da depressão, a frustração da psicose e o pânico da ansiedade. Depois há a maneira como os sintomas interferem na maneira como a pessoa gostaria de viver sua vida (NATURE INTERNATIONAL JOURNAL OF SCIENCE, 2014).

O MINISTÉRIO DA SAÚDE (2018) diz que,

De acordo com a OMS, a depressão situa-se em 4º lugar entre as principais causas de ônus, respondendo por 4,4% dos ônus acarretados por todas as doenças durante a vida. Ocupa 1º lugar quando considerado o tempo vivido com incapacitação ao longo da vida (11,9%).

A depressão pode fazer com que o indivíduo tenha uma experiência muito solitária. Pacientes em quadro depressivo normalmente não tem vontade de fazer atividades que podem servir para aliviar os sintomas da depressão, pior ainda, eles não estão convencidos de que as abordagens orientadas pelos terapeutas fazem sentido. É esperado que em 2030 a depressão seja o distúrbio com o custo mais alto entre todos os outros distúrbios psicológicos nos países de alta renda (ARDEN, 2015).

Os fatores de risco para depressão incluem pessoas do gênero feminino, meia idade, solteiros, baixa renda e incapacitados, bem como o histórico familiar de depressão, experiências adversas na infância, outros transtornos psiquiátricos e doenças crônicas ou incapacitantes (SEMENKOVICH et al., 2015). “A depressão não é apenas muito sofrimento; mas sofrimento demais pode virar depressão” (SOLOMON, 2014, p. 16).

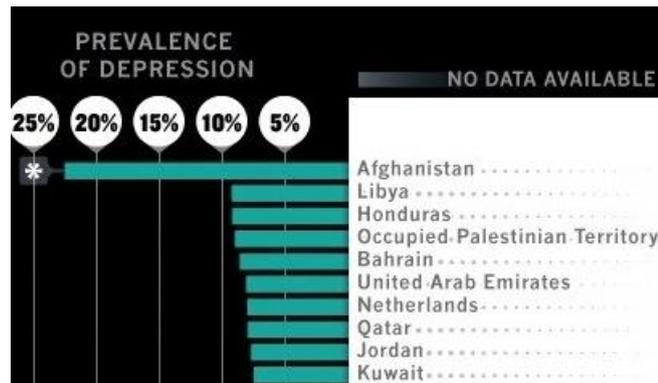
Segundo a OMS, aproximadamente 340 milhões de pessoas no mundo sofrem de depressão (CHIRIȚĂ et al., 2015). Ela afeta em torno de 12% dos homens, enquanto nas mulheres, a prevalência da doença é duas vezes maior, podendo chegar a 25% (CRAMER et al., 2017). Apesar dos dados demonstrarem um número significativo na quantidade de adultos que sofrem de depressão, é importante salientar que, aproximadamente 25% do total de adultos que sofrem com a doença, relatam que o primeiro episódio depressivo ocorreu antes mesmo dos dezoito anos de idade (BAHLS, 2002).

Até pouco tempo não se pensava que crianças e adolescentes poderiam sofrer de depressão, e que ela era um mal exclusivo dos adultos. Porém, isso começou a mudar, e a partir do século XX a depressão infanto-juvenil começou a ser reconhecida, e que crianças e adolescentes

também podem sofrer de depressão (STERZ; CARDOSO DA SILVA, 2017). Aproximadamente 20% dos adolescentes sofrem de transtorno depressivo nos dias atuais (CAMPOS; PRETTE; PRETTE, 2018).

Outro fato importante de destacar, é como a depressão atinge os países e o quanto está presente em cada lugar do globo terrestre.

Figura 1 - Países e a prevalência da depressão



Fonte: Imagem adaptada de (KERRI SMITH, 2014)

Como ilustrado pela Figura 1 o Afeganistão é o país que mais possui casos de depressão, chegando aproximadamente a um total de 22,5% da população que sofre com a doença. O Afeganistão ocupa o primeiro lugar, pelo fato de ser um país que passa por um período de guerra. Com a tensão, pessoas vivendo na pobreza, morrendo de fome ou por falta de atendimento médico. São fatores que contribuem com que as pessoas desenvolvam problemas psicológicos, pois acabam vivendo experiências traumáticas. Na Holanda, conhecida por Países Baixos, por mais que seja um país bem desenvolvido e com uma forte economia, o que pode ser uma causa dos moradores sofrerem de depressão, é o seu clima, no qual pode acabar fazendo com que as pessoas desenvolvam a depressão sazonal. A depressão sazonal caracteriza-se por pessoas que tem episódios de depressão maior principalmente no inverno e outono. Este tipo de depressão será esclarecido logo mais neste artigo.

O Brasil está em 64º lugar, com aproximadamente 5% da população, algo em torno de 11,5 milhões de brasileiros que sofrem de transtorno depressivo.

2.1 TIPOS DE DEPRESSÃO

Os tipos mais comuns de depressão são: transtorno depressivo maior e transtorno depressivo persistente, também conhecido como distímia. Existem também quadros depressivos que são desenvolvidos através de circunstâncias específicas, como depressão perinatal, transtorno

afetivo sazonal e depressão psicótica, entre outros diferentes tipos (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

2.1.1 TRANSTORNO DEPRESSIVO MAIOR

O transtorno depressivo maior (TDM) é altamente prevalente na população em geral e está associado a graves consequências, incluindo mortalidade excessiva, incapacidade, morbidade secundária e altos custos socioeconômicos (LOPEZ; KOS; TURECKI, 2018).

Neste tipo de transtorno o indivíduo apresenta quadro depressivo na maior parte do dia quase todos os dias por 2 semanas, no qual acaba interferindo em sua habilidade de trabalhar, estudar, comer e de aproveitar a vida (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

2.1.2 TRANSTORNO DEPRESSIVO PERSISTENTE (DISTIMIA)

A prevalência da distimia ao longo da vida de um indivíduo é estimada entre 3% a 6% (LIVNE et al., 2018). Com sintomas de depressão que duram pelo menos 2 anos, uma pessoa diagnosticada com esta forma de depressão pode ter episódios de depressão maior, juntamente com períodos de sintomas menos graves (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

É um distúrbio depressivo leve e prolongado, normalmente associado a uma resposta heterogênea a psicoterapia e tratamentos psicofarmacológicos (FRÍAS IBÁÑEZ et al., 2016). Para ser comprovado que o paciente está sofrendo de distimia, no período de dois anos, os sintomas e episódios de depressão não podem estar ausentes por mais de dois meses a cada vez (ABDO, 2010, apud, APA, 2002).

2.1.3 DEPRESSÃO PERINATAL

A depressão perinatal pode ser definida como um episódio de depressão maior ou menor com início dos sintomas durante a gravidez ou nos primeiros 12 meses após o parto. A depressão perinatal é prevalente, subdiagnosticada e pode ter efeitos sérios a longo prazo no bem-estar das mulheres, seus parceiros e bebês. Na ausência de estratégias de identificação ativa, a maioria das mulheres com depressão perinatal não procura nem receber ajuda (MILGROM; GEMMILL, 2014).

2.1.4 TRANSTORNO AFETIVO SAZONAL

O transtorno afetivo sazonal é uma combinação de distúrbios biológicos e de humor com um padrão sazonal, ocorrendo normalmente no outono e inverno com remissão na primavera ou no verão (KURLANSIK; IBAY, 2012).

Este transtorno também é conhecido como depressão ou tristeza de inverno, é um distúrbio de humor no qual pessoas com saúde mental saudável durante a maior parte do ano, apresentam sintomas depressivos no inverno ou, menos comumente, no verão (KURLANSIK; IBAY, 2012). É predominante em mulheres e os sintomas são atípicos, tais como hiperfagia, aumento de peso, letargia e hipersonia (ABDO, 2010).

2.1.5 DEPRESSÃO PSICÓTICA

Esse tipo de depressão ocorre quando uma pessoa tem depressão grave e mais alguma forma de psicose, como por exemplo, crenças falsas e perturbadoras (delírios) e ouvir ou ver coisas que outras pessoas não podem (alucinações) (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

A depressão psicótica é amplamente aceita como um subtipo específico de depressão maior unipolar (O'CONNOR; AGIUS, 2015). Embora sintomas depressivos e psicóticos possam existir em outros diagnósticos, como esquizofrenia, transtorno esquizoafetivo e transtorno bipolar I, a marca registrada da depressão psicótica é o surgimento da psicose durante um episódio depressivo maior e a recessão da psicose como o principal episódio depressivo remite (NEUFELD et al., 2018).

2.2 TRATAMENTOS

Apesar da depressão ser uma doença devastadora, ela na maioria das vezes responde bem aos tratamentos. A chave é realizar uma avaliação específica e um plano de tratamento. Um plano de tratamento pode incluir exercícios, medicações, psicoterapias, terapia da luz e abordagens alternativas, como por exemplo, meditação, acupuntura e nutrição (NATIONAL ALLIANCE OF MENTAL ILLNESS, 2017). Se esses tratamentos não reduzirem os sintomas, a Terapia Eletroconvulsiva (TEC) pode ser outra opção a ser explorada (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

O primeiro passo para obter o plano de tratamento correto é visitar um profissional da saúde ou um profissional de saúde mental, como um psiquiatra ou psicólogo. O médico poderá

solicitar exames, realizar entrevistas e fazer exames laboratoriais para descartar outras condições de saúde que podem apresentar os mesmos sintomas da depressão. Uma vez diagnosticada, a depressão pode ser tratada (NATIONAL INSTITUTE OF MENTAL HEALTH, 2016).

Existem diversos fatores que devem ser considerados no momento da escolha de um tratamento antidepressivo. Alguns desses fatores são: a gravidade do quadro, o perfil e estilo de vida do paciente, o curso da doença (evolução natural), a satisfação com o tratamento (tolerância ao medicamento), entre outros (ABDO, 2010, apud APA, 1993).

A Psicoterapia é recomendada para pacientes com depressão de gravidade leve a moderada (KOK; REYNOLDS, 2017). Os 2 dois tipos de psicoterapia que têm os melhores índices de sucesso no tratamento da depressão, são a terapia cognitivo-comportamental (TCC) e a terapia interpessoal (TIP).

SCHESTATSKY e FLECK (1999) comentam que terapia interpessoal (TIP) é

[...] Um tratamento de tempo limitado, planejado especificamente para tratar a fase aguda das depressões maiores, unipolares e não-psicóticas, e que foi desenvolvido e testado através de testes clínicos randomizados por Gerald Klerman e colaboradores, que publicaram seu manual em 1984.

A TIP não considera os fatores interpessoais como causa da depressão e sim a conceitua como sendo um fenômeno complexo e multideterminado que deve ser objeto de investigação de outras áreas. A terapia interpessoal procura concentrar como foco do tratamento uma área problema, como por exemplo, luto, disputas interpessoais (com filhos, parentes etc.), mudança de papéis (novo emprego, saída de casa, etc.) e déficits interpessoais, como solidão e isolamento social.

Já a terapia cognitivo-comportamental (TCC) procura ajudar os pacientes a identificar e mudar os pensamentos negativos e padrões de pensamentos autodestrutivos (ANTHES, 2014). A TCC é uma forma de terapia psicodinâmica, a qual se baseia nas reações emocionais e mentais do paciente a eventos externos do presente e da infância, e é dirigida para a obtenção de metas específicas. O sistema foi desenvolvido por Aaron Beck, da Universidade da Pensilvânia (SOLOMON, 2014).

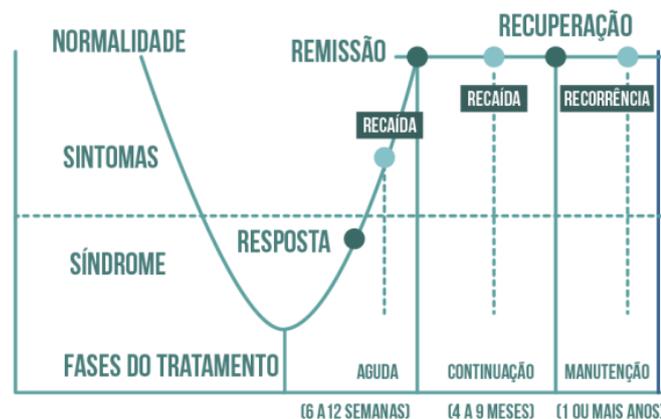
Outro tipo de terapia muito utilizada no tratamento da depressão é a Psicofarmacoterapia. A Psicofarmacoterapia é um tratamento direcionado à base de medicamentos chamados antidepressivos (ADs). Os ADs modificam os níveis de certas substâncias do cérebro alterando a forma de como um paciente sente, age e conseqüentemente baixam os níveis de ansiedade (SOLOMON, 2014). Há quatro classes de medicações disponíveis, as mais populares são os

ISRSs (Inibidores Seletivos da Recaptação de Serotonina), que acarretam níveis mais altos de serotonina no cérebro, como é o exemplo dos medicamentos Prozac, Luvox, Paxil, Zoloft e Celexa (SOLOMON, 2014). Os medicamentos de categoria ISRS têm eficácia semelhante aos tricíclicos e aos inibidores da monoaminoxidase (IMAO), porém, com a vantagem de menos efeitos colaterais aos pacientes (ABDO, 2010).

Existem dois tipos de antidepressivos mais antigos, os tricíclicos e os inibidores de monoaminoxidase (IMAOs). Essas duas classes de ADs possuem mais efeitos colaterais que os ISRSs. Os tricíclicos são denominados assim por sua estrutura química, que afetam a serotonina e a dopamina. Elavil, Anafanil, Norpramin, Tofranil e Pamelor são todos da classe dos tricíclicos. Os inibidores de monoaminoxidase (IMAOs) inibem o colapso da serotonina, dopamina e norepinefrina, alguns dos remédios de categoria IMAOs são Nardil e Parnate.

O tratamento farmacológico (Psicofarmacoterapia) divide-se em três fases, sendo elas: fase aguda, fase de continuação e a fase de manutenção. A fase aguda se caracteriza pela remissão dos sintomas e reinserção psicossocial do paciente, essa fase tem duração média de 6 a 12 semanas. Após a fase aguda, existe então a fase de continuação, que se caracteriza pela prevenção de recaídas e recuperação do paciente, com retorno ao funcionamento psicossocial anterior a depressão. Este tratamento possui duração de aproximadamente 4 a 9 meses. E por fim a fase de manutenção, que tem o principal objetivo trabalhar a prevenção de recorrências de novos episódios de depressão, assim evitando que a doença volte a perturbar o paciente. Essa fase tem duração de 1 ano ou mais, dependendo do quadro evolutivo do paciente (ABDO, 2010). A Figura 2 ilustra como funcionam essas fases do tratamento.

Figura 2 - Fases do tratamento da depressão e suas intercorrências.



Fonte: Núcleo de Telessaúde da Universidade Federal do Rio Grande Do Sul (2017)

Para pacientes com um grave quadro de depressão, onde eles não respondem aos medicamentos antidepressivos e possuem forte tendência suicida, a terapia eletroconvulsiva (TEC) se torna uma forte opção. Enquanto o tratamento com antidepressivos apresentam uma eficácia de cerca de 50%, a TEC apresenta um impacto significativo em 75% a 90% dos casos. O eletrochoque funciona rapidamente, muitos pacientes sentem-se melhor em poucos dias de tratamento. Para realizar esta terapia o paciente deve se submeter a alguns exames, como por exemplo, um exame de sangue, um eletrocardiograma, radiografia do tórax e exames relacionados a anestesia. O paciente comprovando aptidão para ser submetido ao procedimento, realiza sessões de eletrochoque, aplicados ou de forma unilateral, do lado não dominante do cérebro, geralmente do lado direito ou de forma bilateral (SOLOMON, 2014).

Para realizar a terapia eletroconvulsiva, o paciente precisa, desde a noite anterior ao procedimento, permanecer em jejum. No dia da terapia, o paciente é anestesiado e eletrodos são aplicados em suas têmporas. O paciente então recebe um choque de um segundo de aproximadamente duzentos joules, equivalente a uma lâmpada de cem watts, que causa uma convulsão na têmpora e no vértice do cérebro, e dura aproximadamente 30 segundos, tempo suficiente para mudar a química do cérebro. O médico que administra o tratamento eletroconvulsivo, pode escolher entre o estímulo de ondas seno, que provoca um estímulo mais continuado ou entre o estímulo de impulsos breves no qual induz convulsões com efeitos colaterais menores (SOLOMON, 2014).

Acima de tratamentos, a depressão deve ser prevenida. Existem evidências de que exercícios físicos podem ser benéficos no alívio dos sintomas da depressão e auxiliar no tratamento, sendo usado como uma ferramenta de apoio no combate a depressão. Embora sejam necessárias pesquisas para investigar e detectar os mecanismos responsáveis por essa condição, permanece o fato de que atividade física é boa para a saúde física e mental. Apesar dos estudos sugerirem que exercícios podem aliviar os sintomas da depressão, ainda não é claro o quanto, com que frequência e intensidade devem ser as práticas (BALCHIN et al., 2016).

Programas escolares que desenvolvem pensamentos positivos nas crianças e adolescentes também são ótimas ferramentas para o combate a depressão (PAVAN, 2017). Ações como estas junto com o apoio da família, amigos e redes de auxílios, são muito importantes e têm papel fundamental na recuperação e na prevenção de novos episódios depressivos.

2.3 TECNOLOGIAS DE APOIO A DEPRESSÃO

Atualmente a tecnologia tem exercido papel importante e fundamental na área da saúde em geral. Na área da saúde mental a tecnologia vem facilitando o acompanhamento e o tratamento de pacientes com depressão e outros transtornos mentais.

Aplicativos como o ADDS (Apoio ao Diagnóstico de Depressão e Avaliação do Risco de Suicídio), *Sanvello - Stress & Anxiety Help*, *7cups – Anxiety & Stress Chat*, *Mind Shift* e até mesmo a plataforma que será objeto de estudo neste trabalho, a plataforma AVAXpsi, oferecem diferentes recursos, cada um com suas características e usabilidades diferentes, mas que tem um objetivo em comum, apoiar os pacientes e os profissionais da saúde mental quando o assunto se trata de depressão e suicídio.

O aplicativo *Mind Shift* visa auxiliar principalmente os pré-adolescentes e adolescentes e se propõe a preencher as lacunas existentes no acesso ao tratamento. O aplicativo basicamente ajuda os usuários a aprender a relaxar e estar atentos, a desenvolver maneiras mais eficazes de pensar e a usar medidas ativas para controlar sua ansiedade com explicações e conselhos sobre o que é ansiedade. Este aplicativo foi desenvolvido pela *Anxiety Canada Association*⁴ com o apoio da *British Columbia Children's Hospital*⁵ (SOCIAL GOOD BRASIL, 2017).

Já o aplicativo *7cups – Anxiety & Stress Chat* foi desenvolvido pela *7 Cups of Tea*⁶. Segundo PAVAN (2017) o *app* “oferece terapia online por meio de mensagens privadas ou conversas em grupo com ouvintes treinados ou terapeutas, vinte e quatro horas por 69 dias”. Os ouvintes são voluntários que procuram o serviço para oferecer sua ajuda a indivíduos que querem desabafar e conversar sobre um determinado problema ou situação. Os ouvintes do projeto são treinados pelo fundador do mesmo, o psicólogo Glen Moriarty (PAVAN, 2017). “São oferecidos exercícios para controlar a ansiedade, vídeos que tratam sobre o tema e disponibilizadas salas de conversa e fóruns de discussão para os usuários” (SOCIAL GOOD BRASIL, 2017).

A plataforma *Sanvello - Stress & Anxiety Help* foi desenvolvida pela companhia *Pacifica Labs Inc.* Esta plataforma, segundo PAVAN (2017), “apropria-se das técnicas de terapia cognitivo-comportamental e da atenção plena para criar ferramentas e atividades que podem auxiliar o usuário a gerenciar o estresse, a ansiedade e a depressão”. Esse aplicativo oferece

⁴ <https://anxietycanada.com>

⁵ <http://www.bcchildrens.ca/>

⁶ <https://www.7cups.com/>

exercícios de relaxamento, registro de humor, pensamentos e metas diárias (SOCIAL GOOD BRASIL, 2017).

Diferente dos exemplos citados até o momento, no qual os aplicativos se caracterizam por prestar suporte a indivíduos que estejam sofrendo de tristeza profunda, depressão e que possuem grande risco de cometer suicídio, o aplicativo ADDS (Apoio ao Diagnóstico de Depressão e Avaliação do Risco de Suicídio) tem como objetivo principal auxiliar os profissionais da área da saúde mental, principalmente os profissionais do sistema público de saúde, o SUS (Sistema Único de Saúde), no diagnóstico da depressão e na definição do risco de suicídio (PAVAN, 2017). Essa plataforma foi construída pelo projeto de pesquisa TelessaúdeRS⁷, projeto esse vinculado a Universidade Federal do Rio Grande do Sul (UFRGS). O aplicativo visa, a partir das respostas obtidas através de um questionário, auxiliar o profissional da saúde a chegar em um diagnóstico (SOCIAL GOOD BRASIL, 2017).

2.4 PLATAFORMA “AVAXPSI” NA DEPRESSÃO

A plataforma AVAXpsi⁸ será objeto de estudo e aplicação deste trabalho. Ela se propõe a apoiar os psicólogos na execução de suas tarefas, oferecendo-lhe recursos e ferramentas com o intuito de auxiliar esses profissionais na identificação de sintomas e diagnóstico dos pacientes através das informações obtidas das avaliações psicológicas.

Figura 3 - Tela de cadastro da plataforma AVAXpsi

AVAX^{PSI}

Bem-vindo à AVAXpsi. Você já possui uma conta? Se sim, clique aqui: [Entrar](#)

Para criar sua conta, preencha os dados abaixo. Você poderá testar gratuitamente o plano por 10 dias.

Se você é um [cliente corporativo](#), entre em contato conosco através do e-mail: contato@avaxpsi.com.br

Nome

E-mail

Telefone

CPF

RG

Número de Registro no CRP

(Procurar Cidade)

Gênero

Usudário na VOL (Plataforma Vetor Online)

Cupom de desconto ou código do afiliado

Escolha o plano* *teste gratuitamente por 10 dias.

Senha

Confirma sua senha

Você poderá testar o plano selecionado gratuitamente por 10 dias após a ativação. Não se preocupe, você não precisará informar dados de pagamento ao testar o plano.

[Mais informações sobre o plano](#)

Fonte: https://aplicacao.avaxpsi.com.br/users/sign_up

⁷ <https://www.ufrgs.br/telessauders/>

⁸ <http://www.avaxpsi.com.br/>

Para utilizar a plataforma, é necessário ser um psicólogo e possuir um registro profissional no Conselho de Psicologia, apenas com o CRP o profissional poderá realizar seu cadastro no site. A tela de cadastro da plataforma é ilustrada na Figura 3.

Depois do profissional estar devidamente cadastrado e com seu plano ativo, ele estará apto a utilizar os recursos da plataforma, como os formulários para avaliação psicológica existentes, onde os profissionais podem definir qual o formulário mais adequado para sua demanda e percepção.

Figura 4 - Tela do formulário de entrevista da plataforma AVAXpsi

Fonte: Imagem adaptada de AVAXpsi

A Figura 4, mostra como é o formulário padrão, utilizado como um *screening*, na plataforma AVAXpsi. O formulário possui vários blocos, divididos por diferentes tópicos. As perguntas pertinentes a serem validadas no momento de apontar o risco de o paciente estar com depressão estão contidas no tópico “Condições e Sintomas”, como é mostrado na Figura 5.

A partir das respostas obtidas dos questionários dos pacientes, é possível detectar se o paciente se encontra em quadro depressivo ou não. Para que a plataforma AVAXpsi identifique o risco de um paciente estar com depressão, é necessário que no questionário respondido, as questões como “Sentiu-se triste quase todos os dias?” e “Sentiu-se sem interesse e prazer pelas coisas que lhe agradam habitualmente, quase todos os dias?” possuam valores positivos, seguindo as diretrizes do DSM (*Diagnostic and Statistical Manual*). Qualquer uma das questões mencionadas acima, sempre que positiva, abre um bloco com mais alguns sintomas que auxiliam no levantamento do risco do paciente estar com depressão. A regra funciona da seguinte forma, cada sintoma selecionado corresponde a um ponto. Se o avaliado somar um *score* maior ou igual a 3 (selecionou pelo menos 3 sintomas), o sistema então “liga” o risco de depressão.

Figura 5 - Questões existentes na plataforma AVAXpsi que apontam o risco de depressão

DEPRESSÃO

Nas duas últimas semanas:

Sentiu-se triste quase todos os dias?

Não

Sentiu-se sem interesse e prazer pelas coisas que lhe agradam habitualmente, quase todos os dias?

Sim

Em relação aos seguintes sintomas, quais você apresentou nas duas últimas semanas, quase todos os dias?

<input type="checkbox"/> Tristeza	<input type="checkbox"/> Sentiu-se sem valor
<input type="checkbox"/> Hipoatividade	<input type="checkbox"/> Sentiu-se inútil
<input type="checkbox"/> Perda de interesse / prazer	<input type="checkbox"/> Sentiu-se abandonado / sozinho
<input type="checkbox"/> Irritabilidade	<input type="checkbox"/> Sentiu-se culpado

Fonte: Imagem adaptada de AVAXpsi

A AVAXpsi foi desenvolvida pela empresa *Bee Touch*⁹, especializada no desenvolvimento de ferramentas tecnológicas que tem o objetivo de apoiar os profissionais da área da saúde e bem-estar. A plataforma possui mais de mil avaliações psicológicas em sua base de dados, sendo esses extraídos de diferentes cenários, pacientes e profissionais.

Usando as mais de mil avaliações disponíveis na plataforma AVAXpsi, o objetivo deste trabalho é então desenvolver uma ferramenta que fará o uso destes dados usando anonimização e privacidade dos dados, para então submetê-los a técnicas de *Machine Learning* com a finalidade de extrair conhecimento da base, detectar padrões para então classificar pacientes que possam estar com depressão.

⁹ <http://www.beetouch.com.br/>

3 REVISÃO SISTEMÁTICA DE TÉCNICAS DE *MACHINE LEARNING* APLICADAS PARA A PREDIÇÃO DE PACIENTES COM DEPRESSÃO

Segundo (KITCHENHAM, 2007), revisão sistemática é

[...] Uma forma de estudo secundário que utiliza uma metodologia bem definida para identificar, analisar e interpretar todas as evidências disponíveis relacionadas a uma questão de pesquisa específica de uma maneira que seja imparcial e (até certo ponto) repetível.

O objetivo desta revisão sistemática é identificar as técnicas de *Machine Learning* mais utilizadas atualmente, a fim de identificar pacientes com depressão. Com base nos resultados obtidos aqui, será possível escolher as técnicas mais promissoras para a construção da ferramenta desejada. Além disso, as linguagens, pacotes, *frameworks* e validações utilizadas nos trabalhos descobertos, auxiliarão os *insights*, a concepção e o desenvolvimento da ferramenta.

3.1 PROTOCOLO DA REVISÃO SISTEMÁTICA

Esta revisão sistemática foi desenvolvida com base no protocolo proposto e aplicado por (MEDEIROS, 2016).

3.1.1 O PROTOCOLO

O planejamento desta revisão seguiu a seguinte estrutura:

a) Título

Revisão sistemática sobre técnicas de *Machine Learning* aplicadas na detecção de pessoas com depressão

b) Resumo

Depressão é uma doença médica crônica que pode afetar pensamentos, humor e a saúde física do indivíduo. Ela é caracterizada por falta de energia, tristeza, insônia e humor baixo (CUI, 2015).

Machine Learning, segundo (NILSSON, 1998),

[...] Refere-se às mudanças nos sistemas que executam tarefas associadas à inteligência artificial (IA). Essas tarefas envolvem reconhecimento, diagnóstico, planejamento, controle de robôs, previsão etc. As alterações podem ser

aprimoramentos para sistemas já em execução ou síntese ab initio de novos sistemas.

[...]

c) Objetivo

O principal objetivo da proposta e execução dessa revisão sistemática é buscar na literatura especializada, técnicas de *Machine Learning* que foram aplicadas para problemas semelhantes a detecção de pacientes com depressão. Nesta revisão, optou-se por utilizar o acrônimo PICOC (População, Intervenção, Comparação, Resultados e Contexto) proposto por Kitchenham (2007) para auxiliar na criação da *string* de busca.

3.1.1.1 Formulação da pesquisa

a) Foco da questão

Encontrar técnicas, pacotes, *frameworks*, processos, metodologias de *Machine Learning* que são mais utilizadas e possuem resultados promissores para auxiliar na identificação de pessoas com depressão.

b) Questões de interesse

- Quais as técnicas de ML aplicadas sobre o contexto de depressão?
- Quais as métricas de validação dos modelos utilizados?
- Quais as linguagens e pacotes utilizados?
- Quais atributos relevantes para treinamento do modelo?

c) Palavras-chave

Machine Learning, Data Mining, Knowledge Discovery, Classifier, Prediction, Mental Illness, Depression diagnosis.

As palavras-chave citadas acima foram escolhidas com base no conteúdo existente nas referências bibliográficas utilizadas no primeiro capítulo deste trabalho.

d) Intervenção

Identificar técnicas de *Machine Learning* que são utilizadas na área da saúde mental com o objetivo de detectar pacientes com transtorno depressivo.

e) Controle

Inexistente

f) Efeito

Identificar as oportunidades de pesquisa na área de aprendizagem de máquina com o foco no apoio ao diagnóstico e detecção de pessoas com depressão.

g) Medida de resultado

Produzir embasamento e referencial teórico para o trabalho de conclusão de curso e para a elaboração de artigos científicos.

h) População de interesse

Pesquisadores, professores, desenvolvedores e profissionais da área da saúde e da computação.

i) Aplicação

O conhecimento extraído desta revisão sistemática será utilizado para a construção da ferramenta que detectará pacientes com depressão utilizando técnicas de *Machine Learning*.

j) Desenho do experimento

Não será desenvolvido

k) Financiamento

Não há financiamento

3.1.1.2 Seleção das bases de dados

a) Definição dos critérios de seleção das fontes de dados

Para a seleção das fontes de dados foi observado se a fonte possui publicações nas áreas da computação e na área da saúde. Para buscar trabalhos relacionados à área da computação, serão utilizadas as bases de dados *Web of Science*¹⁰ e *IEEEExplore*¹¹. Nesta primeira base de dados somente pesquisadores CAPES possuem acesso integral aos trabalhos ou pessoas vinculadas a instituições de ensino. A base de dados *IEEEExplore* fornece acesso para mais de três milhões de documentos de algumas das publicações mais citadas no mundo em Engenharia Elétrica, Ciência da Computação e Eletrônica (MEBDEIROS, 2016). Em relação à área da saúde, será utilizada a base de dados Pubmed¹², que contém periódicos publicados da área (MEDEIROS, 2016). Essa base de dados possui periódicos do Medline (*Medical Literature Analysis and Retrieval System Online*), possivelmente o banco de dados bibliográfico mais importante do mundo em ciências da vida e informações biomédicas (EDHLUND; MCDOUGALL, 2014).

b) Idiomas das fontes de dados

Serão considerados somente os materiais bibliográficos no idioma em inglês.

c) String de busca

¹⁰ <http://www.periodicos.capes.gov.br/>

¹¹ <https://ieeexplore.ieee.org/Xplore/home.jsp>

¹² <https://www.ncbi.nlm.nih.gov/pubmed/>

Com base nas palavras-chave definidas anteriormente gerou-se a *string* de busca abaixo. Esta *string* é aplicada nos motores de buscas das bibliotecas definidas anteriormente.

(“*machine learning*” or “*data mining*” or “*knowledge discovery*”) and (“*predictive analysis*” or “*predictive modelling*” or “*classifier*”) and (“*mental illness*” or “*depression diagnosis*” or “*depressive disorder*” or “*depressive disorder diagnosis*” or “*bipolar disorder*” or “*depression prediction*”)

d) Artigos de controle

Optou-se por não utilizar nenhum artigo de controle para esta revisão sistemática.

3.1.1.3 Seleção dos estudos

i. Critérios para inclusão/exclusão dos resultados

- a) Ser um artigo publicado em revistas ou congresso;
- b) O ano de publicação do artigo deve estar entre os anos de 2014 a 2019;
- c) Ser um artigo escrito em inglês;
- e) O artigo deve possuir relação entre assuntos obrigatórios da pesquisa.

ii. Procedimentos para seleção dos estudos

Primeiramente, a *string* de busca será executada nos motores de busca das bases mencionadas anteriormente, após esse procedimento, os trabalhos encontrados serão exportados das bases de dados e importados na ferramenta StArt¹³. Após isso, serão executadas as fases definidas a seguir.

iii. Fases de seleção de artigos

- Fase 1 – Validar os critérios de inclusão/exclusão;
- Fase 2 – Identificar os artigos duplicados;
- Fase 3 – Leitura do título, palavras-chave e resumo;
- Fase 4 – Leitura da introdução e conclusão;
- Fase 5 – Leitura integral dos artigos e elaboração das respostas para os critérios de qualidade.

iv. Critérios de qualidade

- Quais foram os algoritmos mais utilizados?
- Quais foram as técnicas de *Machine Learning* mais abordadas?
- Os artigos foram aplicados à área da saúde mental?
- Os artigos foram aplicados à depressão?

¹³ http://lapes.dc.ufscar.br/tools/start_tool

- Quais foram as formas de validação do modelo?
- Quais foram as métricas utilizadas para avaliar o desempenho do modelo?
- Quais foram as linguagens de programação mais utilizadas?
- A proposta teve resultados positivos?
- Quais foram as ferramentas mais utilizadas?

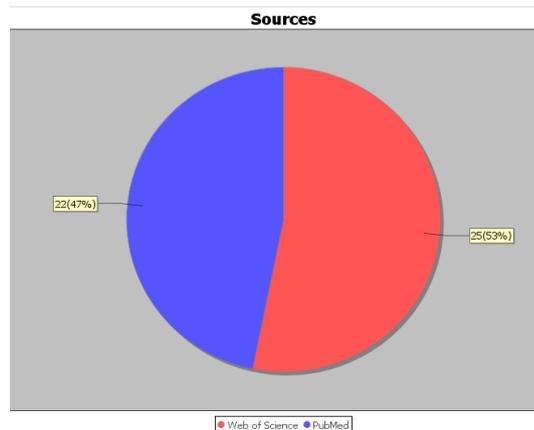
3.2 DESENVOLVIMENTO DA REVISÃO SISTEMÁTICA

Após a elaboração do protocolo da revisão sistemática (RS), iniciou-se a fase de execução de validação dos artigos com base nos critérios definidos no mesmo. Para auxiliar no processo de execução e na condução da RS, contou-se com o apoio da ferramenta StArt. A ferramenta StArt, segundo (HECKLER, 2018), “foi desenvolvida para auxiliar os pesquisadores na aplicação da técnica de revisão sistemática, desde o cadastro do protocolo até a fase final da revisão. Também são gerados gráficos com informações sobre o andamento do trabalho”.

A primeira etapa da RS então foi cadastrar o protocolo no StArt, e executar a *string* de busca nos motores de pesquisa das bases de dados mencionadas no protocolo. As consultas as bases de dados foram executadas no dia 6 de outubro de 2019. Depois da execução da *string* de busca, os resultados obtidos das bases foram exportados e importados no StArt.

Na base IEEE Explore, nenhum artigo foi encontrado, ele não é exibido na Figura 6 por conta disso. Já na base de dados *Web of Science*, após a execução da *string* de busca, foram retornados 25 artigos, estes artigos já estão sendo filtrados pelo período de publicação entre os anos de 2014 a 2019, e pelo critério de que os conteúdos dos artigos retornados devem estar disponíveis na íntegra, caso o usuário possua vínculo a alguma instituição de ensino. Na base de dados PubMed, foram retornados 22 artigos, também, artigos que foram publicados entre os anos de 2014 a 2019 e têm seu conteúdo disponibilizado na íntegra de forma gratuita sem necessidade de vínculo a qualquer instituição de ensino. Foi obtido, ao final da primeira etapa, um total de 47 artigos, os quais já respeitam os critérios de inclusão/exclusão impostos pelo protocolo, a Figura 6 ilustra os resultados.

Figura 6 - Total de artigos retornados pelas bases de dados pesquisadas.

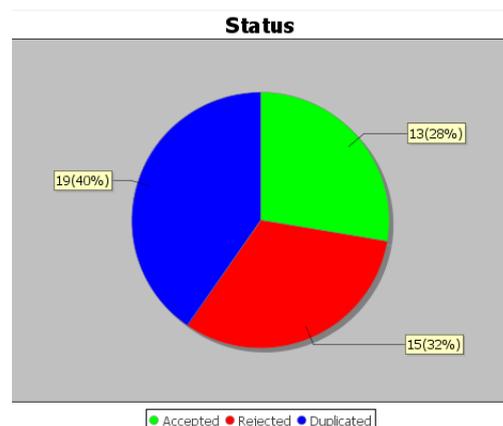


Fonte: Elaborado pelo autor

A segunda etapa foi eliminar os artigos duplicados. Nessa fase foram eliminados 19 artigos, restando um total de 28 artigos. A terceira etapa então foi constituída pela leitura do título, palavras-chave e resumo de cada artigo, buscando com essa leitura selecionar os artigos que possuem relação com os assuntos obrigatórios da pesquisa. Nessa fase 15 artigos foram rejeitados e 13 foram aceitos. Os artigos que foram rejeitados nessa fase são os que usavam *Deep Learning*, classificavam tipos de doença tais como Alzheimer, autismo, esquizofrenia e transtorno bipolar ou que buscavam prever o resultado de um tratamento para a depressão e não de fato a identificação de indivíduos com depressão.

Os treze artigos selecionados são os que possuem mais proximidade da área de estudo deste trabalho, que é a aplicação de técnicas de ML em dados para detectar pacientes que sofrem de transtorno depressivo maior, seja através de Eletroencefalograma (EEG), dados oriundos de mídias sociais ou entrevistas. O resultado obtido ao final da segunda e terceira etapa são ilustrados na Figura 7.

Figura 7 - Resultado obtido após a execução da segunda e terceira etapa da seleção dos estudos

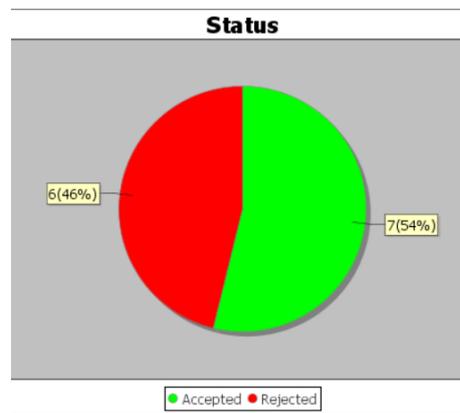


Fonte: Elaborado pelo autor

A quarta etapa de seleção de artigos foi constituída pela leitura da introdução e conclusão dos artigos aceitos na terceira etapa. O que foi priorizado nesta fase, foi a seleção de artigos que buscavam detectar os pacientes com depressão utilizando técnicas de *Machine Learning* e algoritmos de classificação.

Nesta etapa foram rejeitados 6 artigos, dentre eles artigos onde o foco era melhorar performance e precisão de algoritmos de ML existentes, prever o resultado de um tratamento para depressão. Também foram deixados de fora os artigos que buscavam comparar um público restrito, como por exemplo, identificar pacientes com depressão de pacientes bipolares. Ao final desta etapa foram selecionados para a leitura integral 7 artigos. A Figura 8 ilustra o resultado ao final da quarta fase.

Figura 8 - Resultado da seleção dos artigos para leitura integral.



Fonte: Elaborado pelo autor

No Quadro 1 são exibidos o título e os autor(es) dos artigos selecionados para leitura integral e uso como referência para as estratégias que serão aplicadas neste trabalho.

Quadro 1 - Artigos selecionados para leitura integral

Título artigo	Autor(es)(as)
Automatic detection of major depressive disorder using electrodermal activity	Kim, Ah Young Jang, Eun Hye Kim, Seunghwan Choi, Kwan Woo Jeon, Hong Jin Yu, Han Young Byun, Sangwon
Exploring the utility of community-generated social media content for detecting depression: An analytical study on instagram	Ricard, Benjamin J. Marsch, Lisa A. Crosier, Benjamin Hassanpour, Saeed

Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity	Bhaumik, Runa Jenkins, Lisanne M. Gowins, Jennifer R. Jacobs, Rachel H. Barba, Alyssa Bhaumik, Dulal K. Langenecker, Scott A.
Entropy analysis of heart rate variability and its application to recognize major depressive disorder: A pilot study	Byun, Sangwon Kim, Ah Young Jang, Eun Hye Kim, Seunghwan Choi, Kwan Woo Yu, Han Young Jeon, Hong Jin
Cortical thickness predicts the first onset of major depression in adolescence	Foland-Ross, Lara C. Sacchet, Matthew D. Prasad, Gautam Gilbert, Brooke Thompson, Paul M. Gotlib, Ian H.
Seeking insights about cycling mood disorders via anonymized search logs	Yom-Tov, Elad White, Ryen W. Horvitz, Eric
Predicting anxiety and depression in elderly patients using machine learning technology	Sau, Arkaprabha Bhakta, Ishitaa

Fonte: Elaborado pelo autor

3.3 AVALIAÇÃO DOS CRITÉRIOS DE QUALIDADE

Após a leitura integral dos artigos selecionados, foi realizada a validação dos critérios de qualidade estabelecidos no protocolo. As respostas obtidas para cada uma das perguntas serão apresentadas neste capítulo.

Referente as perguntas “Os artigos foram aplicados à área da saúde mental?” e “Os artigos foram aplicados à depressão?” contidas no protocolo, após a leitura dos artigos, foi possível confirmar que, todos os artigos selecionados para leitura integral aplicaram seus estudos voltados a área da saúde mental e depressão.

No Quadro 2 são exibidos os dados que respondem à pergunta do protocolo “Quais foram os algoritmos mais utilizados?”. Nota-se que dos 7 artigos, 2 mencionaram o algoritmo que foi utilizado para a construção do modelo de predição. Sendo que (YOM-TOV; WHITE; HORVITZ, 2014) não especifica qual o algoritmo usado, porém, menciona que faz uso de um algoritmo de propriedade da Microsoft. Os outros 5 artigos não mencionam qual o algoritmo aplicado para a construção do modelo. Este critério de qualidade acaba por não ter relevância na decisão das estratégias e algoritmos que serão utilizadas neste trabalho.

Quadro 2 - Algoritmos de ML utilizados por cada artigo

Artigos	Algoritmos ML utilizados
(KIM et al., 2018)	Não informado
(RICARD et al., 2018)	Não informado
(BHAUMIK et al., 2017)	Não informado
(BYUN et al., 2019)	Não informado
(FOLAND-ROSS et al., 2015)	Não informado
(YOM-TOV; WHITE; HORVITZ, 2014)	Algoritmo proprietário da Microsoft
(SAU; BHAKTA, 2017)	J48, sequential minimal optimization (SMO)

Fonte: Elaborado pelo autor

No Quadro 3 temos as informações que respondem as questões "Quais foram as linguagens de programação mais utilizadas?" e "Quais foram as ferramentas mais utilizadas?" respectivamente. Dentro dos artigos que informaram a linguagem utilizada, a linguagem R foi a que mais se destacou, sendo utilizada por (KIM et al., 2018), (RICARD et al., 2018) e (BYUN et al., 2019), os demais artigos não informaram qual foi a linguagem utilizada ou não fizeram uso de nenhuma linguagem de programação para a construção do modelo, onde optaram por fazer uso de ferramentas prontas para aplicação de algoritmos de ML e cálculos estatísticos, como o WEKA e MATLAB. Como mostrado no Quadro 3, o *software* MATLAB foi a ferramenta mais utilizada dentre os estudos selecionados.

Quadro 3 - Linguagens e ferramentas utilizadas pelos artigos selecionados

Artigos	Linguagem de Programação Utilizada	Ferramentas utilizadas
(KIM et al., 2018)	R	MATLAB
(RICARD et al., 2018)	R	glmnet R <i>package</i> , ROCR, pROC
(BHAUMIK et al., 2017)	Não informado	MATLAB
(BYUN et al., 2019)	R	MATLAB, SPSS 25, Library for SVM (LIBSVM), nparLD.
(FOLAND-ROSS et al., 2015)	Não informado	Não informado
(YOM-TOV; WHITE; HORVITZ, 2014)	Não informado	Não informado
(SAU; BHAKTA, 2017)	Não informado	WEKA

Fonte: Elaborado pelo autor

Para responder as questões “Quais foram as formas de validação do modelo?” e “Quais foram as métricas utilizadas para avaliar o desempenho do modelo?”, foi construído o Quadro 4 que mostra que, estratégias de *Cross-Validation* foram as formas de validação mais utilizadas, sendo mencionadas em 6 dos 7 artigos selecionados. (BHAUMIK et al., 2017) foi o único que utilizou mais de uma forma de validação, utilizando *leave-one-out*, *10-fold cross-validation* e *hold-out*.

Quadro 4 - Métricas e formas de validação do desempenho do modelo ML utilizado em cada artigo selecionado.

Artigos	Métricas utilizadas para avaliar o desempenho do modelo	Forma de validação do modelo
(KIM et al., 2018)	Acurácia; Sensibilidade; Especificidade; Valor Preditivo Positivo (VPP); Valor Preditivo Negativo (VPN); Área Sob a Curva (AUC).	<i>5-fold cross-validation</i>
(RICARD et al., 2018)	AUC	<i>20-fold cross validation</i>
(BHAUMIK et al., 2017)	Acurácia; Sensibilidade; Especificidade.	<i>leave-one-out, 10-fold cross-validation, hold-out</i>
(BYUN et al., 2019)	Acurácia; Sensibilidade; Especificidade; Valor Preditivo Positivo (VPP); Valor Preditivo Negativo (VPN).	<i>leave-one-out</i>
(FOLAND-ROSS et al., 2015)	Acurácia; Sensibilidade; Especificidade.	<i>10-fold cross-validation</i>
(YOM-TOV; WHITE; HORVITZ, 2014)	AUC	<i>5-fold cross-validation</i>
(SAU; BHAKTA, 2017)	Acurácia; Sensibilidade; Taxa Falso Positivo; Valor Preditivo Positivo (VPP); <i>F-measures</i> ; AUC.	<i>10-fold cross-validation</i>

Fonte: Elaborado pelo autor

O Quadro 5 mostra quais foram as técnicas de *Machine Learning* utilizadas por cada artigo, com isso conseguimos responder à questão “Quais foram as técnicas de Machine Learning mais abordadas?” que foi mencionada no protocolo. É possível perceber que SVM foi a técnica mais utilizada na criação de um modelo de predição, 4 artigos entre os 7 selecionados para a leitura integral aplicaram essa técnica. A segunda técnica mais utilizada foi *Naive Bayes*, que foi utilizada em 3 artigos. As técnicas, *Decision Tree* e K-NN ficaram empatadas, ambas as técnicas foram utilizadas em 2 artigos. As outras técnicas contidas no Quadro 5 foram utilizadas apenas em 1 artigo cada uma.

Quadro 5 - Técnicas ML utilizadas pelos artigos selecionados

Artigos	Técnicas ML utilizadas
(KIM et al., 2018)	SVM, <i>Decision Tree</i> , K-NN, <i>Naive Bayes</i>
(RICARD et al., 2018)	<i>Linear regression</i>
(BHAUMIK et al., 2017)	SVM
(BYUN et al., 2019)	SVM, Linear Discriminant Analysis (LDA), K-NN, <i>Naive Bayes</i>
(FOLAND-ROSS et al., 2015)	SVM
(YOM-TOV; WHITE; HORVITZ, 2014)	<i>Decision Tree</i>
(SAU; BHAKTA, 2017)	<i>Bayesian Network, Logistic regression, multiple layer perceptron, Naive Bayes, Random Forest, Random Tree, Random Sub-Space, K Star.</i>

Fonte: Elaborado pelo autor

Outro fator que pode ser destacado com as informações contidas no Quadro 5 é que, dos 7 estudos selecionados, 3 implementaram mais de uma técnica de *Machine Learning*, sendo eles os artigos de (KIM et al., 2018), (BYUN et al., 2019) e (SAU; BHAKTA, 2017).

Figura 9 - Desempenho obtido no estudo de (KIM et al., 2018) utilizando a técnica *Decision Tree*.

Classifier	Number of features	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Decision tree	1	64.03	66.03	54.53	65.32	72.17
	5	72.57	72.67	69.47	75.60	75.23
	11	73.71	73.74	71.15	76.97	75.90
	15	73.30	73.04	70.70	76.64	75.53
	20	72.00	71.94	68.87	74.99	74.69
	25	71.1	71.46	67.47	73.91	74.21
	30	70.38	70.85	66.00	72.97	74.13
	35	69.30	70.48	64.55	71.90	74.06

Fonte: Imagem adaptada de (KIM et al., 2018)

Dentre as 4 técnicas utilizadas em (KIM et al., 2018), foi obtido um desempenho melhor utilizando a técnica de ML *Decision Tree* (DT), o resultado obtido com DT é ilustrado na Figura 9 e os resultados obtidos com o uso das outras técnicas utilizadas no estudo são ilustrados na Figura 10.

Na Figura 10 é evidente que a técnica *Decision Tree* superou as demais independentemente do número de características utilizadas para a análise. DT se destacou quando avaliou 11 características do *Dataset*, obtendo 73.71% de acurácia, 73.74% de sensibilidade, 71.15% de especificidade, 76.97% de valor preditivo positivo (PPV) e 75.90% de valor preditivo negativo (NPV),

Figura 10 - Desempenho das demais técnicas ML que foram utilizadas no estudo de (KIM et al., 2018)

Classifier	Number of features	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
SVM	1	57.35	54.05	61.65	65.56	53.67
	5	63.62	63.69	54.65	65.56	71.30
	11	62.49	61.53	55.07	64.94	68.85
	15	61.69	60.90	53.80	64.04	68.46
	20	61.24	60.42	52.88	63.54	68.40
	25	60.95	60.05	52.58	63.28	68.13
	30	60.81	59.87	52.37	63.15	68.04
	35	60.79	59.86	52.30	63.12	68.07
k-NN	1	46.64	46.34	88.02	54.20	46.00
	5	58.37	56.70	49.82	60.91	65.70
	11	62.44	62.24	49.98	63.56	61.30
	15	62.35	63.73	47.83	63.19	61.32
	20	61.33	62.12	46.53	62.17	60.27
	25	59.71	59.89	45.03	60.82	58.66
	30	56.71	54.92	42.33	58.50	55.68
	35	53.72	49.63	40.62	56.37	52.79
Naïve Bayes	1	57.35	48.86	76.98	63.04	31.39
	5	63.62	49.36	84.20	63.23	25.49
	11	62.49	49.34	83.95	64.36	25.83
	15	61.69	50.36	84.72	68.12	28.09
	20	61.24	50.69	84.93	69.42	28.77
	25	60.95	50.81	84.92	69.62	29.06
	30	60.81	50.93	85.55	69.63	28.76
	35	60.79	50.93	85.57	69.55	28.80

Fonte: Imagem adaptada de (KIM et al., 2018)

No estudo de (BYUN et al., 2019), que também optou por utilizar 4 técnicas de *Machine Learning*, foi possível obter um desempenho melhor utilizando SVM na construção do modelo, resultando em uma acurácia de 70%, sensibilidade de 64%, especificidade de 76%, valor preditivo positivo de 72% e valor preditivo negativo de 68%, superando as outras técnicas aplicadas no mesmo estudo. Os resultados obtidos por (BYUN et al., 2019) são ilustrados na Figura 11.

Legenda da Figura 11:

ACC Acurácia

SEN Sensibilidade

SPE Especificidade

PPV Valor preditivo positivo

NPV Valor preditivo negativo

NF Number of features

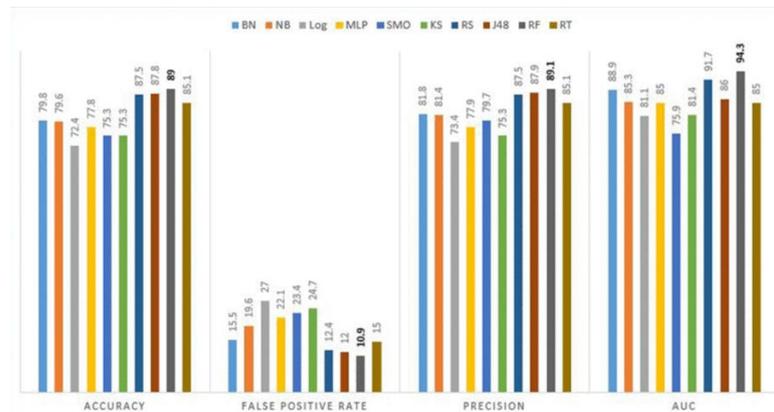
Figura 11 - Resultado do desempenho de cada classificador utilizado no estudo de (BYUN et al., 2019)

Classifier	NF	ACC (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)
SVM	3	70	64	76	72	68
LDA	3	68	64	73	70	67
k-NN	1	67	67	67	67	67
NB	9	67	58	76	70	64

Fonte: Imagem adaptada de (BYUN et al., 2019)

Em (SAU; BHAKTA, 2017), a técnica que obteve um melhor desempenho para o *Data Set* apresentado foi a técnica *Random Forest* (RT), com essa técnica foi possível obter uma acurácia de 89%, uma taxa de falso positivo de 89%, valor preditivo positivo de 89.1% e AUC (*Area Under Curve*) de 94.3%. O desempenho de cada classificador utilizado em (SAU; BHAKTA, 2017) é ilustrado na Figura 12.

Figura 12 - Desempenho obtido em cada técnica ML utilizada no estudo de (SAU; BHAKTA, 2017)



Fonte: Imagem adaptada de (SAU; BHAKTA, 2017)

Todos os artigos que foram selecionados para leitura integral responderam com êxito os critérios de qualidade. O único critério que acabou por não ter relevância foi referente aos algoritmos de ML utilizados nos estudos, visto que a maioria dos estudos selecionados não informaram tal dado. Sobre o critério que avalia o resultado positivo dos artigos, será esclarecido em uma seção separada, contendo a justificativa do autor do porquê seu estudo ter obtido êxito na prática e as métricas obtidas com a estratégia abordada.

3.4 AVALIAÇÃO DOS RESULTADOS DOS ARTIGOS SELECIONADOS

Todos os artigos selecionados na revisão sistemática conseguiram alcançar o objetivo proposto de seu estudo. Os artigos conseguiram de alguma forma, ampliar as possibilidades do uso de técnicas de *Machine Learning* na área da saúde mental, principalmente no apoio a identificação de pacientes com depressão. Cada estudo aplicou suas técnicas de ML a diferentes *DataSet*. Dados oriundos de eletroencefalograma, resposta galvânica da pele e informações e atividades do paciente em redes sociais, são algumas das fontes de dados utilizadas pelos artigos selecionados. Esses estudos trazem o benefício de detectar pacientes sofrendo de depressão, possibilitando o início do tratamento o quanto antes, evitando assim, maiores danos na vida do paciente que vive com essa doença.

(KIM et al., 2018) realizaram um estudo no qual o foco era identificar pacientes com depressão a partir de biomarcadores extraídos da resposta galvânica da pele. (KIM et al., 2018, tradução nossa) comentam que obtiveram sucesso porque,

Demonstramos aqui, através de experimentos de prova de princípio, que os recursos da EDA podem ser usados como um biomarcador para o MDD. Pacientes com TDM e participantes de controle saudáveis foram classificados com acurácia de 74% usando um algoritmo de árvore de decisão.

(KIM et al., 2018) obtiveram 73.71% de acurácia, 73.74% de sensibilidade e 71.15% de especificidade no modelo de predição criado usando a técnica de *Machine Learning* (ML) *Decision Tree*, avaliando 11 das mais relevantes características do *Dataset*.

Em (RICARD et al., 2018), com seu estudo que tem a proposta de identificar usuários com depressão, a partir de atividades registradas em seu *Instagram* e a reação de seus seguidores em suas postagens, sendo analisado comentários e curtidas, conseguiram obter com seu modelo de predição *Linear Regression*, uma *Area Under Curve* (AUC) de 0,71 quando analisado apenas o *Data Set* que continha as informações geradas pela comunidade. Já no *Data Set* que combina os dados gerados pelo usuário com os dados gerados pela comunidade, conseguiram obter uma AUC de 0,72. Por fim, quando avaliado o *Data Set* que continha apenas os dados gerados pelo usuário foi obtida uma AUC de 0,63 que acabou não alcançando os resultados estatisticamente significativos.

(RICARD et al., 2018, tradução nossa) comentam que, “Os resultados apresentados neste estudo indicam que a alavancagem de dados gerados pela comunidade a partir de mídias sociais, além de dados gerados por usuários, pode ser informativa para prever a depressão entre os usuários de mídias sociais.”. (RICARD et al., 2018, tradução nossa) também comentam que:

Os resultados apresentados neste artigo indicam que os dados gerados por pessoas que interagem com postagens feitas por outros usuários de mídia social contêm informações sobre a saúde mental desses usuários, especificamente sobre o status de depressão.

Em (BHAUMIK et al., 2017), buscaram identificar pacientes com depressão de controle normal utilizando métodos de ML para a extração das características mais relevantes a partir dos dados extraídos do *Resting State fMRI (rs-fMRI)*. (BHAUMIK et al., 2017) obtiveram uma acurácia de 76.1%, sensibilidade de 81.5% e especificidade de 68.9% em um modelo de predição que utilizou a técnica ML SVM com o método de extração de características *Elastic Net*. (BHAUMIK et al., 2017, tradução nossa) comentam que, “O presente estudo é o primeiro a demonstrar que novas técnicas de SVM podem ser usadas com dados de rs-fMRI para distinguir indivíduos com histórico de MDD e controles com uma acurácia razoavelmente boa.”

(BYUN et al., 2019) propuseram um estudo que busca identificar pacientes com depressão a partir da entropia da variabilidade da frequência cardíaca. (BYUN et al., 2019) construíram seu modelo de predição utilizando SVM e dizem que,

Demonstramos a diferenciação de pacientes com TDM de controles saudáveis com acurácia de 70%, sensibilidade de 64% e especificidade de 76%, usando os três recursos de entropia mais relevantes. Até onde sabemos, o presente estudo é o primeiro a classificar pacientes com TDM e controles saudáveis usando apenas recursos de entropia da VFC

(FOLAND-ROSS et al., 2015) realizaram um estudo onde buscaram identificar e prever em até 5 anos o aparecimento da depressão em adolescentes a partir da espessura cortical. (FOLAND-ROSS et al., 2015) conseguiram obter com seu modelo preditivo utilizando SVM uma acurácia de 69.7%, sensibilidade de 69.3% e especificidade de 70%. (FOLAND-ROSS et al., 2015, tradução nossa) dizem que:

[...] usando métodos de aprendizado de máquina, demonstramos que os padrões de estrutura cerebral na adolescência podem ajudar a prever o início subsequente da depressão. Nossas descobertas representam um primeiro passo importante para a compreensão dos processos neurais patológicos subjacentes ao surgimento desse distúrbio e destacam vários caminhos importantes para pesquisas futuras.

(YOM-TOV; WHITE; HORVITZ, 2014) desenvolveram um estudo para identificar os pacientes com distúrbios de humor usando registros anônimos de pesquisa na *Web*. Segundo (YOM-TOV; WHITE; HORVITZ, 2014, tradução nossa),

[...] Encontramos evidências de mudanças significativas no comportamento e nos interesses da pesquisa, coincidindo com os dias em que são feitas consultas no MSD (Droga Estabilizadora de Humor). Isso inclui grandes aumentos (> 100%) no acesso a informações nutricionais, comerciais e materiais para adultos. Uma pesquisa com pacientes diagnosticados com transtornos de humor forneceu evidências de que consultas repetidas sobre MSD podem vir com exacerbações do transtorno de humor.

(YOM-TOV; WHITE; HORVITZ, 2014, tradução nossa) também mencionam que:

Identificamos uma população de usuários que demonstram um forte interesse em MSD e mostram mudanças significativas em seu comportamento de pesquisa online na época em que manifestam interesse pelo medicamento. Mostramos que podemos criar modelos preditivos que podem ser usados para prever a aparência futura de consultas de pesquisa centradas em MSD, que podem estar associadas ao início de um episódio bipolar.

(YOM-TOV; WHITE; HORVITZ, 2014) conseguiram com seu modelo obter uma AUC de 0.78 utilizando a técnica *Decision Tree* para a construção dele.

(SAU; BHAKTA, 2017) desenvolveram um modelo preditivo adequado para diagnosticar ansiedade e depressão em pacientes idosos por fatores sociodemográficos e relacionados a saúde usando *Machine Learning*. As informações contidas no *Dataset* foram extraídas de um questionário de 14 perguntas onde cada paciente teve que responder. Cada resposta do questionário teve uma pontuação de 0 a 3 a pontuação final ajudou a classificar os pacientes com depressão dos controles saudável. (SAU; BHAKTA, 2017) utilizaram 10 técnicas/algoritmos de ML para o desenvolvimento do modelo, porém, o melhor resultado foi apresentado usando a técnica *Random Forest*, e com isso conseguiram obter uma acurácia de 89%, uma taxa de falso positivo de 10.9%, precisão de 89.1% e AUC de 94.3%.

Nesta seção apresentamos a justificativa e os resultados obtidos de cada artigo selecionado para leitura integral de forma a evidenciar se eles obtiveram sucesso nas práticas e estratégias realizadas. Como podemos ver, todos conseguiram bons resultados na predição utilizando técnicas diferentes. A técnica que mais obteve resultados positivos foi SVM, essa técnica foi utilizada nos estudos de (BHAUMIK et al., 2017), (BYUN et al., 2019) e (FOLAND-ROSS et al., 2015) e os mesmos obtiveram uma taxa de acurácia de 76.1%, 70% e 69.7% respectivamente.

A partir desta revisão sistemática e avaliações dos artigos é possível afirmar que as estratégias demonstradas pelos estudos podem ser reaproveitadas e aplicadas no contexto deste trabalho.

4 MACHINE LEARNING

(NILSSON, 1998, tradução nossa) diz que, “*Machine Learning* geralmente se refere a alterações em sistemas que executam alguma determinada tarefa associada com inteligência artificial (IA). Tais tarefas podem envolver, reconhecimento, diagnóstico, planejamento, controle de robô, predição, entre outros.”. *Machine Learning* permite que os sistemas se adaptem e melhorem sua performance através da experiência acumulada dos dados que são observados (LEDESMA et al., 2018).

Machine Learning mescla ideias da neurociência, biologia, estatística, matemática e física, para fazer os computadores aprenderem (MARSLAND, 2014). (LEDESMA et al., 2018) dizem que:

Frequentemente, o aprendizado de máquina é usado para criar modelos preditivos, extraindo padrões de grandes conjuntos de dados. Esses modelos são usados em aplicativos preditivos de análise de dados, como: previsão de preços, avaliação de riscos, previsão do comportamento do cliente, design de mecanismos, vigilância de tráfego e classificação de documentos.

ML se divide em 3 principais categorias, sendo elas, aprendizado supervisionado, aprendizado não-supervisionado e aprendizado por reforço (LEDESMA et al., 2018). O aprendizado supervisionado basicamente aprende por exemplo, onde um conjunto de exemplos de treinamento com as respostas corretas é fornecido e, com base nesses dados de treinamento, o algoritmo generaliza para responder corretamente a todas as entradas possíveis (MARSLAND, 2014).

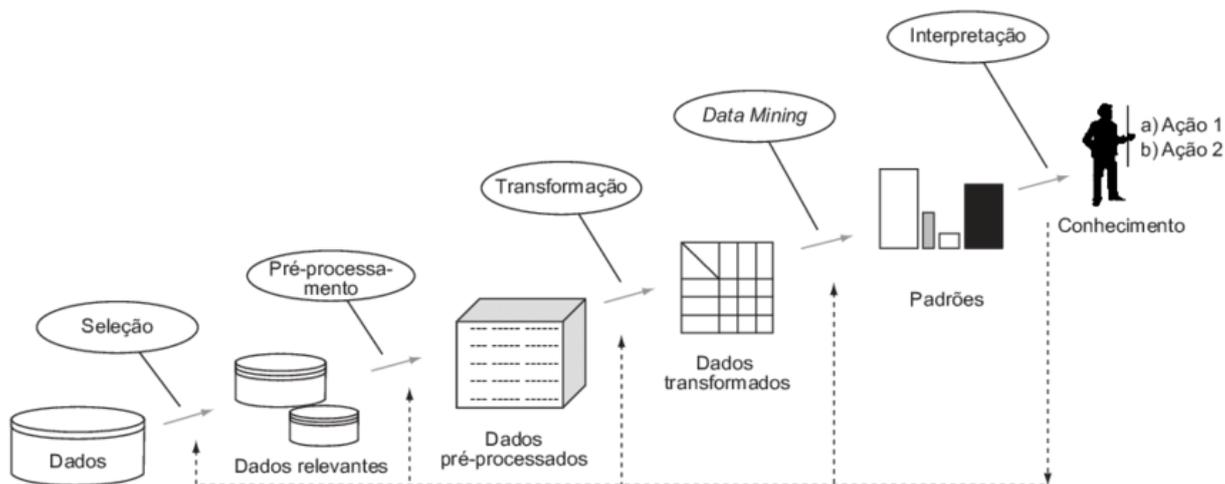
O aprendizado não-supervisionado não possui um conjunto de dados para comparação, porém, o algoritmo tenta identificar semelhanças entre os dados de entrada no qual os dados de entrada semelhantes são categorizados juntos (MARSLAND, 2014). De fato, o aprendizado não-supervisionado é frequentemente utilizado como parte de uma análise exploratória (LEDESMA et al., 2018).

O aprendizado por reforço está entre o aprendizado supervisionado e o não supervisionado. O algoritmo é informado quando a resposta está errada, porém, não é informado como corrigir e, portanto, o algoritmo experimenta diferentes possibilidades até encontrar a melhor resposta para o cenário analisado (MARSLAND, 2014).

Os algoritmos de *Machine Learning* precisam de uma quantidade significativa de dados e de preferência sem ruídos (MARSLAND, 2014). Os dados podem-se originar de diferentes

fontes, como por exemplo, *Data Warehouses*, bancos relacionais e arquivos simples (HAN; KAMBER, 2006).

Figura 13 - Processo de extração de conhecimento



Fonte: (STEINER et al., 2006)

Os dados do mundo real tendem a ser incompletos, barulhentos e inconsistentes (HAN; KAMBER, 2006). Etapas como seleção dos dados, pré-processamento, transformação dos dados (considerado um processo da etapa de pré-processamento), *Data Mining*, interpretação e avaliação dos resultados são aplicados aos dados com o objetivo de eliminar os ruídos, dados inconsistentes e relacionar dados de diferentes fontes (STEINER et al., 2006). Ao final das etapas de extração do conhecimento, como ilustrado pela Figura 13, é possível identificar padrões e extrair conhecimento da base de dados (HAN; KAMBER, 2006).

4.1 SELEÇÃO DOS DADOS

Esta etapa se refere a seleção de um conjunto de dados ou um subconjunto de variáveis ou amostras de dados no qual se deseja extrair conhecimento (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Nada mais é onde os dados relevantes para a tarefa de análise são recuperados do banco de dados ou de qualquer outra fonte (HAN; KAMBER, 2006).

(CIOS et al., 2007, tradução nossa) dizem que, “Aqui, o minerador de dados seleciona um subconjunto de variáveis (atributos) e pontos de dados (exemplos) que serão usados para executar tarefas de descoberta. Essa etapa geralmente inclui consultar os dados existentes para selecionar o subconjunto desejado”. Essa etapa é uma forma de pré-processamento (HAN; KAMBER, 2006).

4.2 PRÉ-PROCESSAMENTO

Em muitos campos da ciência da computação, como reconhecimento de padrões, recuperação de informações, *Machine Learning*, *Data Mining* e inteligência na Web, é necessário preparar dados de qualidade realizando operações de limpeza e pré-processamento nos dados brutos. Na prática, verificou-se geralmente que a limpeza e preparação de dados representa aproximadamente 80% do esforço total na engenharia de dados (ZHANG; ZHANG; YANG, 2003).

(FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996, tradução nossa) dizem que esta etapa:

inclui operações básicas, como remover ruídos ou *outliers*, se apropriado, coletar as informações necessárias para modelar ou contabilizar ruídos, decidir estratégias para lidar com campos de dados ausentes e contabilizar informações de sequência de tempo e alterações conhecidas, além de decidir problemas de DBMS (*Database Management System*), como a tipagem de dados, *schema* e mapeamento de valores ausentes e desconhecidos.

Segundo (ZHANG; ZHANG; YANG, 2003), o pré-processamento de dados compreende técnicas relacionadas à análise de dados brutos, a fim de gerar dados de qualidade, no qual inclui principalmente, limpeza, integração, transformação, redução e discretização dos dados.

A limpeza de dados pode ser aplicada para remover ruídos e corrigir erros. A integração de dados mescla dados de várias fontes em um armazenamento de dados coerente, como um *Data Warehouse* (este processo não será abordado neste trabalho). Transformações de dados, como normalização, pode ser aplicada. Por exemplo, a normalização, onde os atributos são redimensionados para se enquadrarem em um pequeno intervalo especificado, pode melhorar a precisão e a eficiência dos algoritmos de mineração que envolvem medições de distância (HAN; KAMBER, 2006).

A redução de dados pode eliminar recursos redundantes ou os agrupando, reduzindo assim a massa total de dados a ser analisada. Essas técnicas não são mutuamente exclusivas, elas podem trabalhar juntas. Por exemplo, a limpeza de dados pode envolver transformações para corrigir dados incorretos, como transformar todas as entradas de um campo de data em um formato comum (HAN; KAMBER, 2006).

Pré-processamento gera dados de qualidade, o que leva a padrões de qualidade (ZHANG; ZHANG; YANG, 2003). As técnicas de pré-processamento, quando aplicadas antes do *Data Mining*, podem melhorar substancialmente a qualidade geral dos padrões extraídos e/ou o tempo necessário para a mineração real (HAN; KAMBER, 2006).

4.2.1 LIMPEZA DOS DADOS

Esse processo é executado com o objetivo de melhorar a qualidade dos dados que serão analisados (ZHANG; ZHANG; YANG, 2003). Dados do mundo real são incompletos, cheio de ruídos e inconsistências. Técnicas básicas para a limpeza dos dados como, manipulação de valores faltantes e suavização de ruídos são importantes nesse processo (HAN; KAMBER, 2006).

4.2.1.1 Manipulação de valores faltantes

Para a manipulação de valores faltantes existem métodos como, “ignore a tupla”, que nada mais é que desconsiderar aquele registro na hora da análise. Esse método normalmente é usado quando falta a classe *label* (classe de rótulo), porém, não é muito efetivo, ao menos que a tupla contenha muitos atributos sem valor. Também para tratar de valores faltantes pode se optar por preencher os atributos que não possuem informação de forma manual, porém, esse método consome muito tempo (HAN; KAMBER, 2006).

Outra abordagem para se tratar de valores faltantes é o preenchimento desses atributos com o valor obtido de uma constante, como se fosse um rótulo, como por exemplo, “*undefined*”. Esse método não é infalível, porém, é um método simples para identificar os atributos sem valores. Dependendo do tipo da informação, outro tipo de abordagem que pode ser utilizada para tratar dos valores faltantes é, preenchê-los com a média da soma do atributo da tupla. Para este tipo de abordagem deve se usar a média, moda ou mediana do atributo para todas as amostras que pertencem a mesma classe que a tupla (HAN; KAMBER, 2006).

Outra técnica para tratar dos valores faltantes é, preencher esses atributos com o valor mais provável. Isso pode ser determinado com regressão, indução por árvore de decisão e por inferência bayesiana. Essa é a estratégia mais popular (HAN; KAMBER, 2006).

4.2.1.2 Suavização de ruídos

As técnicas para suavização de ruídos, servem para remover erros ou valores externos que desviam do esperado. Dados incorretos também podem resultar de inconsistências nas convenções de nomenclatura ou códigos de dados usados ou formatos inconsistentes para campos de entrada como data (BASKAR; AROCKIAM; CHARLES, 2013). Algumas técnicas são utilizadas para suavizar os ruídos dos dados, são elas, *Binning*, *Regressão*, *Clustering*.

A suavização por regressão envolve encontrar a melhor linha para ajustar dois atributos ou variáveis, para que um possa ajudar a prever o outro. A suavização por *clustering* serve para detectar *outliers* e agrupar os dados com características similares (HAN; KAMBER, 2006).

Na suavização por *Binning* os métodos suavizam um valor de dados classificados consultando sua “vizinhança”, isto é, os valores à sua volta. Os valores classificados são distribuídos em um número de “*buckets*” ou *bins*, como mostra a Figura 14. Métodos de *Binning*, como consultam apenas os valores mais próximos, acabam por realizar uma suavização local (HAN; KAMBER, 2006).

Figura 14 - Exemplo de *Binning*

Sorted data for <i>price</i> (in dollars): 4, 8, 15, 21, 21, 24, 25, 28, 34
Partition into (equal-frequency) bins:
Bin 1: 4, 8, 15
Bin 2: 21, 21, 24
Bin 3: 25, 28, 34
Smoothing by bin means:
Bin 1: 9, 9, 9
Bin 2: 22, 22, 22
Bin 3: 29, 29, 29
Smoothing by bin boundaries:
Bin 1: 4, 4, 15
Bin 2: 21, 21, 24
Bin 3: 25, 25, 34

Fonte: Imagem adaptada de (HAN; KAMBER, 2006)

Os métodos de *Binning* realizam uma divisão por frequência igual, onde cada *bin* contém o mesmo número de elementos. A suavização pode ser feita pela média dos *bins*, onde cada valor original dentro do *bin* é substituído pela média da soma dos valores contidos no mesmo. A mesma estratégia vale para a suavização por mediana, no qual os valores originais contidos no *bin* são substituídos pela mediana. Outra suavização utilizando métodos de *binning* é a suavização pelos limites (*boundaries*), onde, os valores são substituídos pelo valor mínimo ou máximo contido no *bin*, caso o valor seja mais próximo do valor menor contido no *bin*, esse valor receberá o valor mínimo, o mesmo acontece para a substituição pelo valor máximo (HAN; KAMBER, 2006).

4.2.2 TRANSFORMAÇÃO

No processo de transformação, os dados são consolidados em formas apropriadas para a análise. Esse processo pode envolver etapas como, suavização, agregação, generalização, normalização e construção de atributo (HAN; KAMBER, 2006).

A etapa de suavização serve para remover os ruídos e pode envolver técnicas como *clustering*, regressão e *binning* (detalhado anteriormente). A etapa de agregação é onde é aplicado métricas estatísticas como, a média e variância que são necessárias para sumarizar os dados (BHAYA; ALASADI, 2017). Esta etapa é normalmente usada na construção de um cubo de dados para realizar a análise em várias granularidades, ou seja, em diferentes categorias (HAN; KAMBER, 2006). A etapa de generalização dos dados é onde os dados “brutos” de nível inferior são substituídos por conceitos mais detalhados de nível superior através do uso de hierarquia de conceito, por exemplo, atributos categóricos como rua, podem ser generalizados para nível superior como, cidade ou país (HAN; KAMBER, 2006).

A etapa de normalização serve para ajustar a escala dos atributos em uma faixa de valores menores. Este método é útil em técnicas de *Machine Learning* como, classificação, redes neurais e técnicas de *clustering*. *Minimum-maximum*, *z-score* e escala decimal são formas populares de normalização (BHAYA; ALASADI, 2017). Na etapa de construção de atributo ou construção de característica é criado um atributo a partir de informações contidas em outros do *dataset* e que podem ajudar no processo de extração do conhecimento (HAN; KAMBER, 2006).

4.3 DATA MINING

A mineração de dados consiste em uma enumeração de modelos sobre os dados (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Segundo (ZHANG; WANG; ZHANG, 2019),

Data mining pode ser considerado como um tipo particular de processo de descoberta de conhecimento. Pode ser definido como a análise de um conjunto de dados para descobrir relações potencialmente relevantes e resumir essas relações de uma maneira nova e compreensível

Data mining envolve o ajuste de modelos ou a determinação de padrões a partir dos dados observados. Os modelos ajustados desempenham o papel do conhecimento alcançado (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Essa etapa é um processo onde que métodos inteligentes e técnicas são aplicados para extrair conhecimento dos dados (HAN; KAMBER, 2006)

As técnicas clássicas de *data mining* são associação, regressão, *clustering* e classificação. Essas técnicas geralmente são usadas apenas para analisar uma etapa específica do processo geral (VAN DER AALST, 2012). Os algoritmos e técnicas de classificação supervisionada serão abordados com maior detalhe já que essa técnica será o foco dessa pesquisa que visa a construção de um modelo de predição de pacientes com depressão a partir de algoritmos utilizados para classificação. As outras técnicas serão abordadas de maneira mais superficial.

4.3.1 ASSOCIAÇÃO

Essa técnica consiste em relacionar os atributos/campos das classes ou tabelas da base de dados, buscando obter correlações de vários campos satisfazendo os limites de suporte e confiança (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Existem muitos algoritmos que vem sendo usados para gerar regras de associação, sendo eles, *Apriori*, *Tertius* e *Predictive Apriori* (versão atualizada do Apriori) (ALI et al., 2019).

Para validar a utilidade e a certeza das regras de associação geradas pelos algoritmos é utilizado duas medidas, sendo elas, suporte e confiança. Normalmente, a regra de associação é considerada interessante ou “forte” se satisfizer um limite mínimo de suporte e um limite mínimo de confiança (HAN; KAMBER, 2006).

O suporte de uma regra de associação $X \Rightarrow Y$ é definido como a porcentagem de registros que contêm X e Y para o número total de transações (N) no banco de dados. O suporte da regra de associação é representado pela equação $Support(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{N}$. A confiança de uma regra de associação é definida como a porcentagem do número de transações que contêm X e Y ao número total de transações que contêm X, essa operação é representada pela equação $Confidence(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}$. O símbolo σ na equação representa a operação de seleção.

Os limites de suporte e confiança das regras de associação podem ser definidos por usuários ou especialistas no domínio. Outras análises podem ser realizadas para descobrir correlações estatísticas interessantes entre itens associados (HAN; KAMBER, 2006).

4.3.2 REGRESSÃO

As técnicas de regressão, segundo (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996), “mapeiam um item de dados para uma variável de previsão de valor real”. Alguns dos famosos métodos de regressão em *Machine Learning* incluem regressão linear e redes neurais (ANOOPKUMAR; MD ZUBAIR RAHMAN, 2016).

Regressão é uma metodologia estatística desenvolvida por Sir Frances Galton. É a abordagem mais utilizada para a previsão numérica. (HAN; KAMBER, 2006) dizem que, “A previsão numérica é a tarefa de prever valores contínuos (ou ordenados) para uma determinada entrada.”, por exemplo, prever o potencial de venda de um produto a partir do seu preço.

A técnica de regressão pode ser usada para modelar o relacionamento entre uma ou mais variáveis preditoras e uma variável de resposta. As variáveis preditoras são os atributos de interesse que descrevem a tupla (ou seja, compõem o vetor de atributo). A variável de resposta é o que queremos prever. Dada uma tupla descrita por variáveis preditoras, queremos prever o valor associado da variável de resposta (HAN; KAMBER, 2006).

Algumas das técnicas de regressão são, *Multiple Linear Regression, Linear Regression, Elastic Net Regression, Ada Boosting Regression, Gradient Boosting Regression* (MADHURI; ANURADHA; PUJITHA, 2019).

4.3.3 CLUSTERING

Essa técnica faz parte da categoria de algoritmos não supervisionados em *Machine Learning*, pois analisa os dados sem se basear por uma classe *label*, ou seja, sem amostra/exemplo de dados para avaliar se a resposta é certa ou errada. Tal técnica consiste em agrupar os dados conforme a proximidade de suas características. Os dados que possuem grande similaridade são considerados parte de um *cluster*. Membros de um cluster não possuem nenhuma proximidade com os dados de outro. Essa técnica pode servir para construir uma classe *label* (rótulo) e transformar os dados permitindo assim que sejam utilizados em técnicas que se baseiam em rotulagem, como por exemplo, os algoritmos de classificação e predição (HAN; KAMBER, 2006).

Segundo (YAN et al., 2019, tradução nossa),

O armazenamento em *cluster* é uma das técnicas mais amplamente usadas para particionar espaços de dados com base na similaridade dos mesmos e descobrir padrões dos fluxos de dados. É uma ferramenta eficaz para capturar a natureza dos dados, separando-os com base em similaridades, o que pode ajudar muito os usuários a entender o conceito por trás das informações.

Na notação estatística, o *clustering* é o mais importante algoritmo de *Machine Learning* não supervisionado (DUTT; ISMAIL; HERAWAN, 2017). A clusterização é uma técnica de mineração de dados multivariada que é amplamente usada em várias áreas. Como dito por

(PUGGINA BIANCHESI et al., 2019, tradução nossa), as áreas onde a técnica de clusterização pode ser aplicada são:

marketing, identificando participação de mercado; medicina, identificando pacientes com uma causa comum de doença; educação, medindo características psicológicas para identificar grupos de alunos que precisam de atenção especial; biologia, construindo uma taxonomia de grupos e subgrupos de plantas semelhantes; climatologia, fornecendo novas idéias sobre tendências climatológicas e ambientais.

O objetivo desta técnica é agrupar automaticamente os n elementos de um banco de dados em k clusters, usando apenas as informações das variáveis de cada caso (PUGGINA BIANCHESI et al., 2019). *K-means*, *Fuzzy C-means*, *Single Link (SLINK)* e *Nearest Neighbor* são alguns algoritmos utilizados para clusterização (DUTT; ISMAIL; HERAWAN, 2017).

4.3.4 CLASSIFICAÇÃO

A técnica de classificação pode ser dividida em 2 categorias, supervisionada e não supervisionada (DUTT; ISMAIL; HERAWAN, 2017). O foco deste capítulo será em cima das técnicas de classificação supervisionada, pois serão as abordagens aplicadas neste trabalho em função do que foi encontrado nos trabalhos relacionados.

A classificação supervisionada é uma técnica de *Data Mining* muito simples sendo também a mais utilizada. Na classificação supervisionada é dado uma coleção de padrões e dados rotulados que irão servir como base para determinar o rótulo dos novos dados não classificados (DUTT; ISMAIL; HERAWAN, 2017).

O processo de classificação se divide em duas etapas, no qual a primeira etapa é onde o modelo classificador é construído descrevendo um conjunto predeterminado de classes ou conceitos a partir dos dados. Esta é a etapa de treinamento do modelo classificador, em que um algoritmo de classificação cria o classificador analisando ou “aprendendo com” um conjunto de treinamento composto de tuplas de banco de dados e seu rótulo associado (HAN; KAMBER, 2006). O conhecimento dos dados de treinamento é obrigatório para a compreensão da classificação (JALOTA; AGRAWAL, 2019).

A segunda etapa do processo de classificação é avaliar o modelo utilizando dados de testes (JALOTA; AGRAWAL, 2019). Primeiro, a acurácia preditiva do modelo é apurada. Se usarmos o conjunto de treinamento para medir a acurácia do classificador, essa estimativa provavelmente será otimista, porque o classificador tende a superestimar os dados (ou seja, durante o aprendizado, ele pode incorporar algumas anomalias particulares dos dados de

treinamento que não estão presentes no conjunto geral de dados). Portanto, é usado um conjunto de testes composto de tuplas de teste e sua classe *label* associada. Essas tuplas são selecionadas aleatoriamente no conjunto de dados geral. Eles são independentes das tuplas de treinamento, o que significa que não são usados para construir o classificador (HAN; KAMBER, 2006).

SVM (*Support Vector Machine*) e árvore de decisão são algumas das técnicas utilizadas na construção de modelos classificadores supervisionados (ANOOPKUMAR; MD ZUBAIR RAHMAN, 2016).

4.3.4.1 *Support Vector Machine*

Essa técnica pode ser utilizada para análise de classificação e regressão. Mas na maioria das vezes, é utilizado para desafios de classificação. Essa técnica pode ser aplicada em um grande volume de dados para analisar e encontrar padrões ocultos (JALOTA; AGRAWAL, 2019).

O algoritmo SVM é um algoritmo de *Machine Learning* desenvolvido a partir da teoria estatística do aprendizado. Baseado no kernel, é um método projetado inicialmente para resolver os problemas de classificação binária e foi desenvolvido por Vapnik no início dos anos 90 (SONKAMBLE; DOYE, 2008).

Support Vector Machine (SVM) é um método de *Machine Learning* que vem sendo amplamente utilizado na classificação de textos e imagens. Os algoritmos SVM podem processar dados de amostra pequenos não lineares, de alta dimensão e desequilibrados e ainda sim resultar em boa generalização (DING, 2019).

(HAN; KAMBER, 2006, tradução nossa) dizem que:

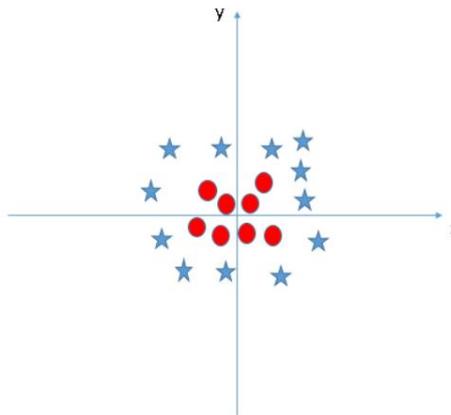
SVM é um algoritmo que usa um mapeamento não linear para transformar os dados de treinamento originais em uma dimensão mais alta (z). Dentro dessa nova dimensão, ele procura o hiperplano de separação ideal linear (ou seja, um "limite de decisão" que separa as tuplas de uma classe da outra). Com um mapeamento não linear apropriado para uma dimensão suficientemente alta, os dados de duas classes sempre podem ser separados por um hiperplano. O SVM encontra esse hiperplano usando vetores de suporte (tuplas de treinamento "essenciais") e margens (definidas pelos vetores de suporte).

O mapeamento não linear é uma técnica SVM conhecida como truque do *kernel*. Ele converte um problema não separável em um problema separável. São funções que ocupam um espaço de entrada dimensional baixo e o transformam em um espaço dimensional mais alto, essas funções são chamadas de núcleos (*kernels*). Simplificando, ele faz algumas

transformações de dados extremamente complexas e, em seguida, descobre o processo para separar os dados com base nos rótulos ou saídas definidas (SOUZA, 2019a).

A Figura 15 mostra um exemplo de classes onde que não é possível classificá-las a partir de um hiperplano linear. Quando isso acontece, SVM então os transforma para uma dimensão mais alta para que sejam classificados (SOUZA, 2019a).

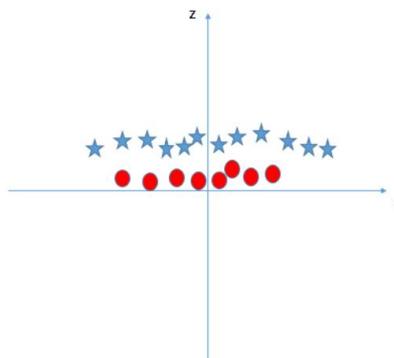
Figura 15 - Exemplo de classes não separáveis linearmente



Fonte: (SOUZA, 2019a)

Para transformar os dados e colocá-los em uma dimensão mais alta, o recurso adicional $z = x^2 + y^2$ é aplicado aos dados. Depois disso o eixo y é substituído pelo eixo z, onde que todos os valores de z seriam positivos pois é a soma quadrática de x e y. A Figura 16 mostra o resultado após a aplicação do recurso adicional, onde que, os círculos vermelhos que estavam próximos da origem dos eixos x e y, levaram a um valor menor de z e as estrelas que estavam relativamente longe do resultado da origem dos eixos obtiveram valor maior de z (SOUZA, 2019a). Depois da transformação dos dados é possível encontrar o melhor hiperplano de separação das classes.

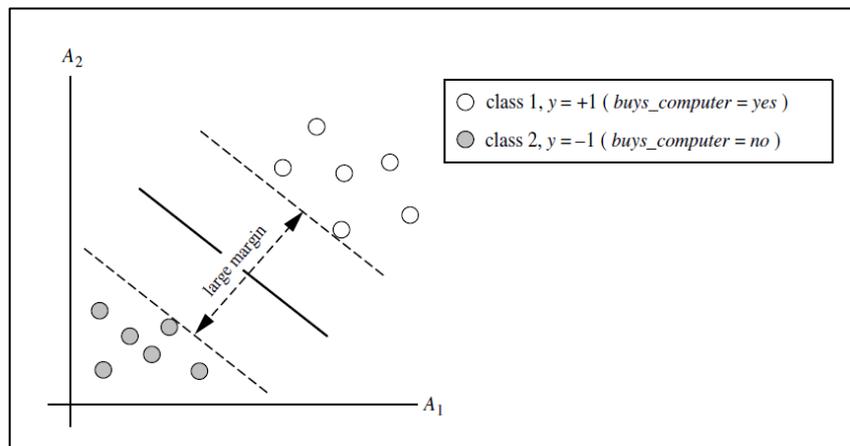
Figura 16 - Classes em uma dimensão mais alta com uso do kernel



Fonte: (SOUZA, 2019a)

O hiperplano pode ser representado pela seguinte equação, $\mathbf{W} \cdot \mathbf{X} + b = 0$, onde que \mathbf{W} é um vetor de peso, sendo representado por $\mathbf{W} = (W_1, W_2, \dots, W_n)$. Cada elemento dentro do vetor \mathbf{W} corresponde a um peso. O parâmetro b da equação é um escalar, geralmente chamado de viés, pode se pensar que b é um peso adicional. E o \mathbf{X} representa os dados de treinamento, onde que X_1 e X_2 são os valores para os atributos W_1 e W_2 respectivamente (HAN; KAMBER, 2006).

Figura 17 - Hiperplano e margens em SVM



Fonte: Imagem adaptada de (HAN; KAMBER, 2006)

A margem fornece uma maior separação entre classes. A menor distância entre um hiperplano e um lado de sua margem é igual à menor distância entre o hiperplano e o outro lado de sua margem, onde os “lados” da margem são paralelos ao hiperplano (HAN; KAMBER, 2006).

A Figura 17, mostra as margens (linhas tracejadas) e o hiperplano (traço em negrito), onde que os elementos que se encontram abaixo do hiperplano (H) pertencem a classe -1 e os elementos que se encontram acima do hiperplano pertencem a classe 1. Os elementos classificados para qualquer hiperplano e que ficam próximos a extremidade da margem são chamados de vetores de suporte (HAN; KAMBER, 2006).

4.3.4.2 Árvore de decisão

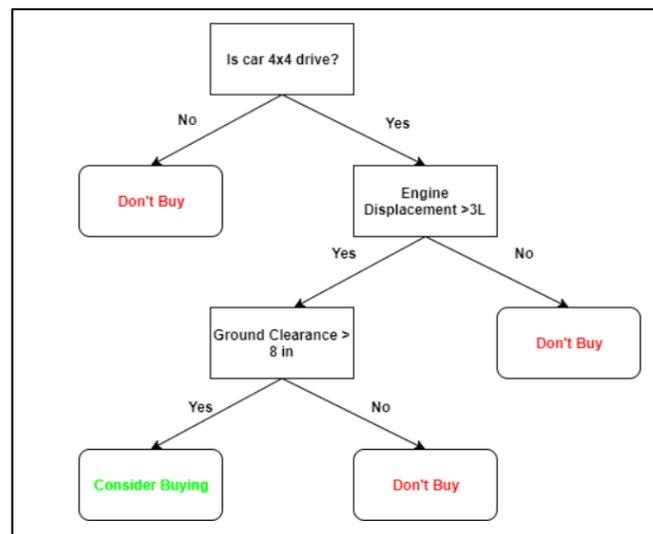
As árvores de decisão são ferramentas populares e eficazes para mineração, classificação e previsão de dados. Uma árvore de decisão é um modelo preditivo que pode ser usado para representar o modelo dos classificadores e os modelos de regressão. As árvores de

decisão são uma forma simples, mas poderosa, de análise de múltiplas variáveis (LAKSHMI; DEVI, 2018).

A árvore de decisão é uma estrutura de árvore do tipo fluxograma. O aprendizado dos algoritmos de árvore de decisão se dá a partir de tuplas de treinamento rotuladas (HAN; KAMBER, 2006). Um algoritmo de aprendizado em árvore de decisão precisa um conceito de destino usando uma representação em árvore, onde cada nó interno corresponde a um recurso e cada nó terminal corresponde a uma classe (LAKSHMI; DEVI, 2018).

O nó mais alto de uma árvore é o nó raiz e se trata de um tipo diferente de nó (HAN; KAMBER, 2006). Existem outros dois tipos de nós em uma árvore de decisão, sendo eles, nós internos e nós terminais. Os nós internos também são conhecidos como, nós não-folha e representam o teste em um determinado atributo dos dados (HAN; KAMBER, 2006). Cada ramificação gerada a partir do nó interno é o resultado dos testes realizados dando origem ao nó terminal (LAKSHMI; DEVI, 2018).

Figura 18 - Arvore de decisão para compra de um carro



Fonte: (SOUZA, 2019b)

O nó terminal, conhecido como nó folha, tem um rótulo de classe atribuído a ele (LAKSHMI; DEVI, 2018). O nó terminal denota as diferentes classes nos quais os dados podem ser separados e onde cada observação é mapeada para sua classe correspondente (WANG et al., 2015).

A Figura 18 mostra o exemplo de um modelo de predição utilizando árvores de decisão para avaliar se a compra do carro é viável considerando os atributos, tração, altura e capacidade do motor. O primeiro retângulo, é o nó raiz, é de onde a árvore começa a ser construída, de cima para baixo. Cada retângulo abaixo do nó raiz simboliza um nodo interno que se divide em

branches. Cada retângulo com as bordas arredondadas representa um nó folha, ou seja, o nó terminal, que é a classe/rótulo resultante do teste realizado na *branch* e que é atribuído a tupla avaliada.

4.4 PÓS-PROCESSAMENTO

A etapa de pós processamento serve para verificar a qualidade do conhecimento (padrões) descoberto, procurando identificar se ele auxilia na resolução do problema original que motivou a realização do processo de KDD (*Knowledge Discovery in Databases*) (APARECIDO DE ALMEIDA CALIL et al., 2008). É nesta etapa onde técnicas de visualização e representação dos dados são usadas para apresentar o conhecimento extraído para o usuário (HAN; KAMBER, 2006).

Segundo (APARECIDO DE ALMEIDA CALIL et al., 2008),

O pós-processamento contempla a depuração e/ou síntese dos padrões descobertos. Na grande maioria dos processos de KDD, a etapa de pós-processamento se justifica, pois o volume de conhecimento descoberto é tão grande/extenso que dificulta a sua análise e fundamentalmente inviabiliza o seu uso no apoio a tomada de decisão. Isso se deve a vários fatores, como por exemplo: padrões redundantes, relações irrelevantes, entre outros.

A etapa de pós processamento inclui técnicas como *knowledge filtering*, avaliação dos modelos e visualização dos dados (a visualização de dados não será abordada neste trabalho) (TOMAR; AGARWAL, 2014).

4.4.1 KNOWLEDGE FILTERING

Knowledge filtering tem como principal objetivo extrair resultados significativos para executar determinadas etapas da tomada de decisão. Os dois principais métodos de *knowledge filtering* são truncamento de regras e pós poda, normalmente usados em modelos de classificação de árvore de decisão (TOMAR; AGARWAL, 2014).

Por exemplo, uma árvore de decisão pode gerar várias regras com certo grau de redundância, dificultando a tradução em termos de ações úteis para a tomada de decisão. O truncamento de regras é um método de pós-processamento aplicado nas regras de decisão para melhorar o desempenho do classificador (TOMAR; AGARWAL, 2014).

Outro problema nas árvores de decisão é que muitos ramos podem acabar refletindo ruídos ou irregularidades nos dados de treinamento. A “poda” da árvore de decisão tenta

identificar e remover as arestas irregulares, com o objetivo de melhorar a precisão da classificação. A poda se dá pela remoção das *branches* e substituição pelo rótulo mais frequente da sub-árvore que está sendo substituída (HAN; KAMBER, 2006).

4.4.2 AVALIAÇÃO DOS MODELOS

É essencial avaliar o desempenho do modelo (TOMAR; AGARWAL, 2014). A avaliação do modelo é o processo de escolher entre diferentes tipos de modelos, parâmetros de ajuste e recursos. Melhores processos de avaliação levam a modelos melhores e mais precisos (SOUZA, 2019c).

A matriz de confusão é uma ferramenta útil para analisar quão bem o seu classificador pode reconhecer tuplas de diferentes classes. Uma matriz de confusão para duas classes é mostrada na Figura 19. Dadas "m" classes, uma matriz de confusão é uma tabela de pelo menos tamanho "m" por "m". Uma entrada, $CM(i, j)$ nas primeiras "m" linhas e "m" colunas indica o número de tuplas da classe "i" que foram rotuladas pelo classificador como classe "j". Para que um classificador tenha boa acurácia, idealmente, a maioria das tuplas seria representada ao longo da diagonal da matriz de confusão, da entrada $CM(1, 1)$ à entrada $CM(m, m)$ com o restante das entradas próximo de zero. A tabela pode ter linhas ou colunas adicionais para fornecer totais ou taxas de reconhecimento por classe (HAN; KAMBER, 2006).

Figura 19 - Matriz de confusão

		Predicted class	
		C_1	C_2
Actual class	C_1	true positives	false negatives
	C_2	false positives	true negatives

Fonte: (HAN; KAMBER, 2006)

Uma matriz de confusão pode ser definida livremente como uma tabela que descreve o desempenho de um modelo de classificação em um conjunto de dados de teste para os quais os valores verdadeiros são conhecidos. Uma matriz de confusão é altamente interpretativa e pode ser usada para estimar várias outras métricas (SOUZA, 2019c).

As métricas que podem ser estimadas pela matriz de confusão são, acurácia, sensibilidade e especificidade. A partir da sensibilidade e a especificidade é possível obter outra métrica de desempenho, a AUC (*Area Under Curve*). Para obter o valor das métricas, técnicas de validação são aplicadas aos modelos. Técnicas como *k-fold* e *leave-one-out cross validation*

foram algumas das técnicas mais utilizadas nos trabalhos encontrados na revisão sistemática e servirão como base para a validação dos modelos construídos neste trabalho.

4.4.2.1 Métricas de validação

A acurácia é a taxa do número de previsões corretas em relação ao número total de amostras de entrada (ADITYA MISHRA, 2018). A acurácia pode ser representada pela equação mostrada na Figura 20. Essa métrica foi utilizada em 5 artigos dos 7 selecionados na revisão sistemática.

Figura 20 - Equação da acurácia

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

Fonte: (ADITYA MISHRA, 2018)

A sensibilidade conhecida também como taxa de verdadeiro positivo, corresponde à proporção de pontos de dados positivos que são corretamente considerados positivos, com relação a todos os pontos de dados positivos (ADITYA MISHRA, 2018). A Figura 21 mostra a equação que representa a sensibilidade.

Figura 21 - Equação da sensibilidade

$$TruePositiveRate = \frac{TruePositive}{FalseNegative + TruePositive}$$

Fonte: (ADITYA MISHRA, 2018)

Já a especificidade, conhecida como a taxa de falso positivo, corresponde à proporção de pontos de dados negativos que são erroneamente considerados positivos, com relação a todos os pontos de dados negativos (ADITYA MISHRA, 2018). A Figura 22 mostra a equação para se chegar ao valor da especificidade do modelo.

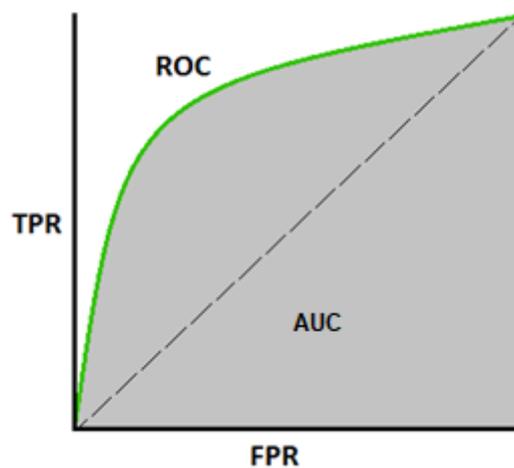
Figura 22 - Equação da especificidade

$$FalsePositiveRate = \frac{FalsePositive}{FalsePositive + TrueNegative}$$

Fonte: (ADITYA MISHRA, 2018)

Quando precisamos verificar ou visualizar o desempenho do problema de classificação multi-classe, usamos a área sob a curva (AUC) ROC (*Receiver Operating Characteristics*). Essa é uma das métricas de avaliação mais importantes para verificar o desempenho de qualquer modelo de classificação (NARKHEDE, 2018).

Figura 23 - Área sob a curva ROC



Fonte: Imagem adaptada de (NARKHEDE, 2018)

ROC é uma curva de probabilidade e AUC representa o grau ou medida de separabilidade, ou seja, diz quanto o modelo é capaz de distinguir as classes. A curva ROC é plotada com a sensibilidade contra a especificidade em uma faixa de valores de 0 a 1. Quanto maior a AUC, melhor o modelo em prever 0 como 0 e 1 como 1 (NARKHEDE, 2018). Como é ilustrado pela Figura 23, a sensibilidade (*True Positive Rate*) fica no eixo y e a especificidade (*False Positive Rate*) é representada no eixo x.

4.4.2.2 Técnicas de validação

A técnica *k-fold cross validation* faz com que o conjunto de dados seja dividido aleatoriamente em k subconjuntos. Cada vez que um dos subconjuntos k é usado como conjunto de teste, os outros subconjuntos k-1 são usados para formar um conjunto de treinamento.

Portanto, obtemos a propensão de falha para todas as “n” classes existentes (SINGH; KAUR; MALHOTRA, 2010). Dada uma base de dados hipotética em que conste 100 registros, e definindo o $k=10$ a base de dados será dividida em 10 subconjuntos onde cada subconjunto terá 10 registros cada. Após a divisão em subconjuntos, será utilizado um subconjunto, para ser utilizado na validação do modelo e os conjuntos restantes são utilizados como treinamento. O processo de validação cruzada é então repetido k (10) vezes, de modo que cada um dos k subconjuntos sejam utilizados exatamente uma vez como teste para validação do modelo (NELSON CORLETA SCHREIBER et al., 2017).

A validação cruzada *Leave-One-Out* ocorre da mesma maneira que o método *k-fold* com a principal diferença, o treinamento é realizado com $n-1$ dados e o teste com 1 dos registros somente. O método *Leave-One-Out* define o número de subconjuntos igual ao número de registros da base de dados. Então, se a base de dados tiver 100 registros dentro dela, serão definidos 100 subconjuntos cada um com 1 registro. Após a divisão dos subconjuntos o mesmo processo do *k-fold* é realizado, utiliza-se o subconjunto 1 para teste e o restante para treinamento, no caso do exemplo seriam 99 subconjuntos para treinamento, e assim sucessivamente (NELSON CORLETA SCHREIBER et al., 2017).

A validação cruzada é uma técnica muito útil para avaliar a eficácia do modelo, principalmente nos casos em que se precisa reduzir o excesso de ajuste (*overfitting*) (GUPTA, 2017). As técnicas de validação mais utilizadas nos artigos estudados na revisão sistemática servirão como base para avaliação da performance dos modelos que serão construídos neste trabalho.

5 MODELOS DE PREDIÇÃO DE PACIENTES COM DEPRESSÃO A PARTIR DOS DADOS DA PLATAFORMA AVAXPSI

Nesta etapa será colocado em prática o conhecimento adquirido no decorrer do desenvolvimento deste trabalho, tais como: técnicas e estratégias de pré-processamento; seleção dos atributos a serem utilizados para a construção do modelo preditivo; a construção do modelo utilizando as técnicas *Support Vector Machine* e *Decision Tree*; e a avaliação dos resultados dos modelos. Os modelos foram construídos com o objetivo de identificar o risco de um paciente estar com depressão, baseado nas respostas obtidas da entrevista *online* realizada na plataforma AVAXpsi.

As duas técnicas de ML utilizadas neste trabalho foram selecionadas com base nos artigos estudados na etapa de revisão sistemática e foram as mais aplicadas dentre os artigos. Após a construção dos modelos foi realizada a avaliação da performance deles com o intuito de avaliar a classificação dos pacientes com depressão. As técnicas de validação e métricas utilizadas para identificar a performance do modelo são as que foram estudadas no capítulo anterior, sendo elas, acurácia, sensibilidade, especificidade e AUC. A linguagem utilizada para o desenvolvimento deste trabalho foi R, linguagem bastante utilizada nos artigos lidos da revisão sistemática.

5.1 QUALIDADE DE DADOS

Nesta etapa foram realizados os procedimentos necessários para que os atributos a serem avaliados sejam selecionados e tratados, eliminando assim àqueles que não agregam valor na predição de pacientes com depressão e aplicando as transformações necessárias para análise. Para realizar a análise, foram recebidos 3 *datasets* em formato csv da empresa *BeeTouch*, desenvolvedora da plataforma AVAXpsi. Um *dataset* é o que contém as informações pessoais do paciente; o segundo *dataset* foi o de entrevista, que contém as respostas para cada uma das perguntas contidas na plataforma da AVAXpsi. Este *dataset* possui as características que desejamos avaliar para então apontar o risco de o paciente estar com depressão.

O último *dataset* que compõe a análise é o de avaliação. Este *dataset* possui o resultado para as respostas contidas no *dataset* de entrevista, ou seja, conforme a resposta do usuário na entrevista, o *dataset* de avaliação pode apontar se o paciente está sofrendo com problemas no trabalho, transtorno de ansiedade e o risco de o paciente estar com depressão, sendo este último item o atributo alvo deste trabalho. A Tabela 1 ilustra o total de informações contidas em cada

um dos *datasets* recebidos, porém, no *dataset* de avaliação, o atributo que será considerado para construção do modelo será apenas o atributo *depression_flag*.

Tabela 1 – Total de observações e atributos por *dataset*

Dataset	Total de Observações	Total de Atributos
Paciente	1713	6
Entrevista	1330	259
Avaliação	2040	88

Fonte: Elaborado pelo autor

O atributo *depression_flag* foi utilizado para construir o modelo de predição oferecendo suporte para identificar os pacientes com depressão. Este atributo trata de uma variável do tipo booleana que indica se o paciente pode estar sofrendo de transtorno depressivo ou não, conforme as respostas obtidas da entrevista, seguindo as diretrizes descritas na Seção 2.4.

Para iniciar a análise, primeiramente foi realizada uma operação de união entre o *dataset* de entrevista com o de avaliação. Do *dataset* de avaliação foram recuperados apenas os atributos *depression_flag* e id, e do *dataset* de entrevista, neste primeiro momento, foram recuperados todos os atributos. Nesta união, todas as entrevistas que possuíam avaliação foram consideradas para análise, sendo as demais descartadas. Após a união foi obtido um *dataset* com um total de 1220 observações e 257 atributos.

5.1.1 SELEÇÃO DOS ATRIBUTOS

Nesta etapa foram realizados os procedimentos necessários para identificar os atributos do *dataset* de entrevista a serem selecionados para classificar pacientes com depressão. A seleção dos atributos é uma etapa importante no processo de ML pois ajuda a construir um modelo simples, de qualidade e de fácil implementação em um ambiente de produção (AGARWAL, 2019).

Neste trabalho foi decidido utilizar a correlação Pearson para avaliar a correlação dos demais atributos em relação a propriedade *depression_flag*. A correlação mostra o quanto uma variável pode interferir no valor de outra, ou seja, o quanto estão correlacionadas. Os coeficientes de correlação variam de -1 até 1, sendo zero indicando de que não há nenhuma correlação entre as variáveis. Quando o valor é acima de zero quer dizer que possuem uma correlação positiva, ou seja, quando uma variável aumenta a outra também aumentará. Quando o coeficiente de correlação fica abaixo de zero quer dizer que as variáveis possuem uma correlação negativa, ou seja, quando uma variável aumenta a outra diminui (MAGIYA, 2019).

Antes de avaliar a correlação dos atributos percebeu-se que alguns atributos do *dataset*, resultantes da união do *dataset* de entrevista e a avaliação não possuíam valores em nenhuma tupla, ou possuíam muitas tuplas com valores ausentes, o que pode interferir na qualidade da correlação e futuramente no modelo. Para contornar estes problemas algumas estratégias foram utilizadas para selecionar os atributos e preencher os valores ausentes.

5.1.1.1 Preenchimento de valores ausentes

Para evitar que as tuplas do *dataset* permanecessem com valores ausentes, foram utilizadas duas estratégias de preenchimento para substituí-los por outros valores. A primeira estratégia preenche os valores NA (*Not Available*) de forma diferente para cada tipo de dado, onde que, para atributos do tipo booleano, em que em suas tuplas os únicos valores possíveis são 0 e 1 foi utilizado uma estratégia diferente da utilizada nos atributos numéricos com valores variados.

Para os atributos booleanos foi coletado o percentual de tuplas com o valor 1 e o percentual de tuplas com valor 0, desconsiderando as tuplas com valor ausente. Caso o percentual de um ou de zero fosse maior que 80%, os valores NA daquele atributo são substituídos pela moda, caso contrário, é atribuído o valor 2. O valor 2 foi escolhido para substituir os valores ausentes do tipo booleano por conta de que ele não será considerado como um dado válido no momento da predição. Para os atributos de outros tipos de dados, as tuplas com valor NA são substituídas com o valor da mediana do atributo, desconsiderando as tuplas de valores ausentes.

A segunda estratégia simplesmente preenche todos os valores faltantes independentemente do tipo de dado do atributo com o valor da mediana. O objetivo de utilizar duas estratégias diferentes nesta etapa é para comparar a qualidade da correlação para cada uma, fazendo com que a melhor seja aplicada. As estratégias desta etapa serão utilizadas em conjunto com as que foram aplicadas para selecionar os atributos para correlação e terão seus resultados comparados na próxima seção.

5.1.1.2 Seleção dos atributos a partir dos resultados da correlação

Nesta etapa foi executado o processo de geração da matriz de correlação com o intuito de detectar os atributos que tem grande influência em classificar um indivíduo que sofre de transtorno depressivo e utilizá-los na construção dos modelos de predição.

Antes de gerar e avaliar a correlação dos atributos, o primeiro passo foi eliminar os atributos que poderiam acabar interferindo na qualidade da correlação ou do modelo. Os

primeiros atributos a serem eliminados foram os que não possuíam nenhum valor em suas tuplas, restando assim 184 atributos para análise dos 257 iniciais. Mesmo depois deste procedimento muito dos atributos ainda possuíam um número significativo de tuplas com valores NA (*Not Available*), onde, mesmo que técnicas para substituir estes valores fossem utilizadas, os valores inseridos acabariam resultando em dados não relevantes para análise, já que se tratariam apenas de tuplas com muitos rótulos e dados não oriundos da entrevista realizada com o paciente.

Para diminuir o número de rótulos nas tuplas e melhorar a qualidade dos dados, a estratégia aplicada foi de selecionar os atributos que possuíam o menor percentual de valores faltantes. Para isso foi descoberto o percentual de NA para cada atributo e colocado esses dados em outro *dataset*. A Figura 24 mostra alguns dos dados contidos no *dataset*, onde o nome dos atributos está contido na coluna “colunas_percentual_valores_na” e o percentual de valores NA na coluna “percentuais_valores_na”.

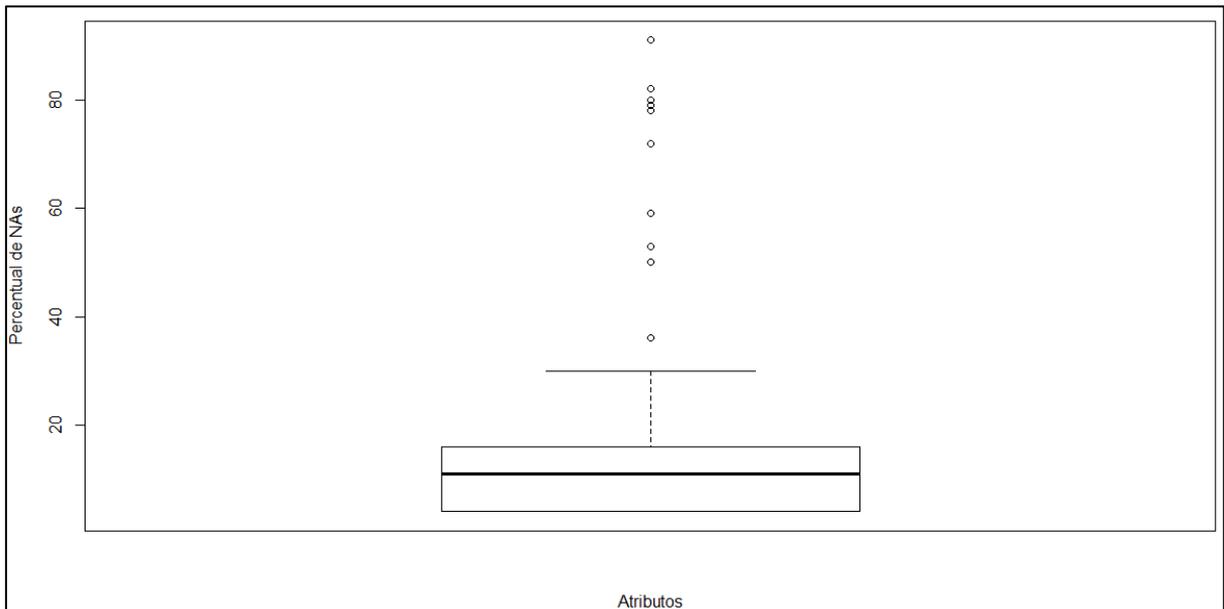
Figura 24 - Dataset de percentual de valores NA por atributo

	colunas_percentual_valores_na	percentuais_valores_na
experience_time	experience_time	7
experience_time_detail	experience_time_detail	29
demand_increase	demand_increase	4
promotion	promotion	4
relationship_difficulty	relationship_difficulty	4
boss_difficulty	boss_difficulty	4
transfer	transfer	4
job_loss	job_loss	4
new_job	new_job	4
risks	risks	15
risk_height	risk_height	11
risk_pressure	risk_pressure	11
risk_high_temperature	risk_high_temperature	11
risk_low_temperature	risk_low_temperature	11
risk_electricity	risk_electricity	11
risk_noise	risk_noise	11
risk_contamination	risk_contamination	11
risk_confinement	risk_confinement	11
risk_explosion	risk_explosion	11
risk_collapse	risk_collapse	11
risk_flood	risk_flood	11

Fonte: Elaborado pelo autor

Posteriormente, foi gerado um boxplot em cima deste *dataset* para revelar os *outliers*, ou seja, aqueles atributos com percentual NA que destoam dos quartis gerados pelo *boxplot*. Como é ilustrado pela Figura 25, os atributos considerados *outliers* são os que possuem um percentual de valores faltantes maior que 30%, sendo então eliminados do *dataset*.

Figura 25 - Boxplot de percentual de valores faltantes por atributo

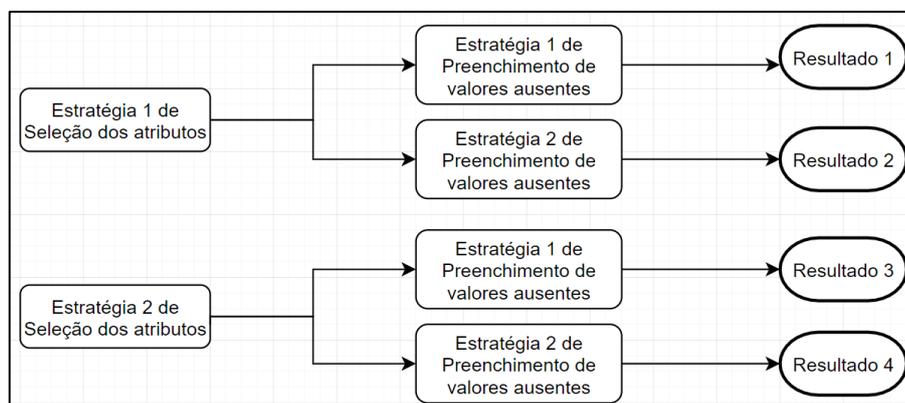


Fonte: Elaborado pelo autor

Depois de eliminar os atributos *outliers*, restaram 168 atributos. Com os atributos restantes foi decidido utilizar 2 estratégias diferentes para selecionar os atributos a serem considerados para correlação. O uso dessas estratégias visa garantir uma matriz de correlação de qualidade e um modelo onde apenas os atributos relevantes para a classificação sejam avaliados, garantindo uma melhor performance do modelo.

Cada estratégia utilizada para seleção dos atributos foi aplicada em conjunto com as estratégias de preenchimento dos valores faltantes, estratégias essas que foram documentadas na Seção 5.1.1.1. O esquema apresentado pela Figura 26 ilustra como foi feito o uso dessas combinações e mostra como os atributos utilizados nos modelos de predição foram selecionados.

Figura 26 - Fluxo de análise para seleção dos atributos e geração da correlação



Fonte: Elaborado pelo autor.

O uso dessas combinações tem como objetivo encontrar quais as técnicas que obtiveram uma qualidade superior na correlação. Para comparar a qualidade da correlação gerada para cada combinação, foram considerados os atributos que obtiveram valores de correlação ≥ 0.30 ou ≤ -0.30 , desconsiderando a correlação da variável *depression_flag* para ela mesma.

A primeira estratégia de seleção dos atributos faz uso do desvio padrão e a média dos valores NA para definir uma faixa de valores NA permitido por atributo. A equação realizada para definir a faixa inicial e final de valores ausentes nos atributos é ilustrada pela Figura 27. O percentual de valores ausentes aceitos nesta estratégia foram os que ficavam dentro da faixa de -1% até 23% resultando assim em 156 atributos.

Figura 27 - Equações para determinar faixa inicial e final do percentual de valores NA aceitos para análise

$$\begin{aligned} \text{FaixaInicial} &= \text{media} - 2 * \text{desvioPadrao} \\ \text{FaixaFinal} &= \text{media} + 2 * \text{desvioPadrao} \end{aligned}$$

Fonte: Elaborado pelo autor

Essa estratégia combinada com a primeira estratégia de preenchimento dos valores ausentes resultou em uma correlação onde 6 atributos acabaram obtendo valores NA e 150 atributos obtiveram valores de correlação válidos. A Figura 28 ilustra o resultado obtido para essa combinação de estratégias no qual os atributos considerados para avaliar são destacados na imagem.

Figura 28 - Resultado da correlação utilizando a estratégia um da seleção dos atributos e a estratégia um da manipulação dos valores ausentes

coluna_origem	coluna_destino	valor
129 anxiety_dizziness	depression_flag	0.30
130 stress_procrastination	depression_flag	0.30
131 anxiety_accelerated_heart	depression_flag	0.33
132 stress_lack_control	depression_flag	0.33
133 stress_accumulation_problems	depression_flag	0.33
134 stress_irritation	depression_flag	0.33
135 stress_feeling_low_capacity	depression_flag	0.33
136 anxiety_headache	depression_flag	0.36
137 anxiety_nervousness	depression_flag	0.37
138 stress_impairment_decision_making	depression_flag	0.37
139 anxiety_bad_thoughts	depression_flag	0.38
140 depression_symptoms_hypoactivity	depression_flag	0.46
141 depression_symptoms_useless	depression_flag	0.56
142 depression_symptoms_worthless	depression_flag	0.58
143 depression_no_interest	depression_flag	0.59
144 depression_symptoms_alone	depression_flag	0.64
145 depression_symptoms_guilty	depression_flag	0.66
146 depression_sad_every_day	depression_flag	0.70
147 depression_symptoms_irritability	depression_flag	0.70
148 depression_symptoms_sadness	depression_flag	0.73
149 depression_symptoms_loss_of_interest	depression_flag	0.76
150 depression_flag	depression_flag	1.00

Fonte: Elaborado pelo autor

No uso dessa técnica de seleção de atributos com a segunda estratégia descrita na seção de preenchimento dos valores ausentes, acabou resultando em uma correlação onde 6 atributos resultaram em valor de correlação NA e 150 atributos valores válidos. O resultado para esta combinação é ilustrado na Figura 29. É possível perceber que entre essas duas combinações não houve nenhuma diferença nos resultados obtidos.

Figura 29 - Resultado da correlação utilizando a estratégia um da seleção dos atributos e a estratégia dois da manipulação dos valores ausentes

	coluna_origem	coluna_destino	valor
129	anxiety_dizziness	depression_flag	0.30
130	stress_procrastination	depression_flag	0.30
131	anxiety_accelerated_heart	depression_flag	0.33
132	stress_lack_control	depression_flag	0.33
133	stress_accumulation_problems	depression_flag	0.33
134	stress_irritation	depression_flag	0.33
135	stress_feeling_low_capacity	depression_flag	0.33
136	anxiety_headache	depression_flag	0.36
137	anxiety_nervousness	depression_flag	0.37
138	stress_impairment_decision_making	depression_flag	0.37
139	anxiety_bad_thoughts	depression_flag	0.38
140	depression_symptoms_hypoactivity	depression_flag	0.46
141	depression_symptoms_useless	depression_flag	0.56
142	depression_symptoms_worthless	depression_flag	0.58
143	depression_no_interest	depression_flag	0.59
144	depression_symptoms_alone	depression_flag	0.64
145	depression_symptoms_guilty	depression_flag	0.66
146	depression_sad_every_day	depression_flag	0.70
147	depression_symptoms_irritability	depression_flag	0.70
148	depression_symptoms_sadness	depression_flag	0.73
149	depression_symptoms_loss_of_interest	depression_flag	0.76
150	depression_flag	depression_flag	1.00

Fonte: Elaborado pelo autor

Já a segunda estratégia de seleção de atributos selecionou aqueles que obtiveram um percentual de valores ausentes menor ou igual ao valor da mediana obtida do *boxplot*, ou seja, que possuíam valores ausentes menor ou igual a 11%, o que acabou resultando em 102 atributos.

Essa estratégia combinada com a primeira estratégia de preenchimento dos valores ausentes acabou resultando em uma correlação onde que também 6 atributos resultaram em valor de correlação NA, porém, 96 atributos acabaram gerando valores de correlação válidos. A Figura 30 ilustra os resultados obtidos e os atributos que são levados em consideração para análise nessa combinação.

Figura 30 - Resultado da correlação utilizando a estratégia dois da seleção dos atributos e a estratégia um da manipulação dos valores ausentes

	coluna_origem	coluna_destino	valor
75	anxiety_dizziness	depression_flag	0.30
76	stress_procrastination	depression_flag	0.30
77	anxiety_accelerated_heart	depression_flag	0.33
78	stress_lack_control	depression_flag	0.33
79	stress_accumulation_problems	depression_flag	0.33
80	stress_irritation	depression_flag	0.33
81	stress_feeling_low_capacity	depression_flag	0.33
82	anxiety_headache	depression_flag	0.36
83	anxiety_nervousness	depression_flag	0.37
84	stress_impairment_decision_making	depression_flag	0.37
85	anxiety_bad_thoughts	depression_flag	0.38
86	depression_symptoms_hypoactivity	depression_flag	0.46
87	depression_symptoms_useless	depression_flag	0.56
88	depression_symptoms_worthless	depression_flag	0.58
89	depression_no_interest	depression_flag	0.59
90	depression_symptoms_alone	depression_flag	0.64
91	depression_symptoms_guilty	depression_flag	0.66
92	depression_sad_every_day	depression_flag	0.70
93	depression_symptoms_irritability	depression_flag	0.70
94	depression_symptoms_sadness	depression_flag	0.73
95	depression_symptoms_loss_of_interest	depression_flag	0.76
96	depression_flag	depression_flag	1.00

Fonte: Elaborado pelo autor

Utilizando essa segunda estratégia de seleção dos atributos com a segunda estratégia para preenchimento dos valores ausentes acabou resultando em uma correlação que também obteve 6 atributos com valor de correlação NA e 96 atributos com valores válidos. A Figura 31 ilustra os atributos que obtiveram valor de correlação ≥ 0.30 nesta combinação.

Figura 31 - Resultado da correlação utilizando a estratégia dois da seleção dos atributos e a estratégia dois da manipulação dos valores ausentes

	coluna_origem	coluna_destino	valor
75	anxiety_dizziness	depression_flag	0.30
76	stress_procrastination	depression_flag	0.30
77	anxiety_accelerated_heart	depression_flag	0.33
78	stress_lack_control	depression_flag	0.33
79	stress_accumulation_problems	depression_flag	0.33
80	stress_irritation	depression_flag	0.33
81	stress_feeling_low_capacity	depression_flag	0.33
82	anxiety_headache	depression_flag	0.36
83	anxiety_nervousness	depression_flag	0.37
84	stress_impairment_decision_making	depression_flag	0.37
85	anxiety_bad_thoughts	depression_flag	0.38
86	depression_symptoms_hypoactivity	depression_flag	0.46
87	depression_symptoms_useless	depression_flag	0.56
88	depression_symptoms_worthless	depression_flag	0.58
89	depression_no_interest	depression_flag	0.59
90	depression_symptoms_alone	depression_flag	0.64
91	depression_symptoms_guilty	depression_flag	0.66
92	depression_sad_every_day	depression_flag	0.70
93	depression_symptoms_irritability	depression_flag	0.70
94	depression_symptoms_sadness	depression_flag	0.73
95	depression_symptoms_loss_of_interest	depression_flag	0.76
96	depression_flag	depression_flag	1.00

Fonte: Elaborado pelo autor

Com base nos resultados obtidos nas combinações ilustradas nesta seção foi possível perceber que não há diferença no resultado da correlação nas diferentes combinações e estratégias escolhidas, pois todas as combinações obtiveram os mesmos atributos onde o valor da correlação positiva é ≥ 0.30 . Outro ponto avaliado nos resultados é que nenhuma combinação gerou atributos com correlação negativa, ou seja, não obtiveram valor de correlação ≤ -0.30 e, portanto, não foram selecionados para análise e nem ilustrados nos resultados.

Como não há diferença nos resultados apresentados, a combinação escolhida como sendo a melhor para este estudo foi aquela que possuía menos atributos a serem avaliados e a que aplicou o preenchimento dos valores ausentes com mais coerência com o tipo de dado manipulado, trazendo mais sentido a informação inserida na tupla. Seguindo essas premissas, os atributos selecionados correspondem ao resultado 3 ilustrados pela Figura 26 e são os escolhidos para a construção dos modelos de predição.

5.2 GERAÇÃO DOS MODELOS DE PREDIÇÃO

Nesta etapa iniciou-se o processo de criação dos modelos de predição com os atributos selecionados na seção 5.1. Antes de iniciar a criação dos modelos de predição, o *dataset* utilizado na criação dos mesmos foi dividido em 70% para treinamento e 30% para teste. Essa divisão ocorreu de forma aleatória, porém, este procedimento teve que ser feito de forma onde houvesse um balanceamento dos dados, ou seja, a distribuição de uma proporção dos registros de pacientes com depressão e sem depressão.

O *dataset* possuía um total de 1220 observações, sendo 1161 observações de pacientes sem depressão e 59 observações de pacientes com depressão. Diante deste cenário, poderia ocorrer do *dataset* de treinamento não possuir nenhum registro de paciente com depressão impactando diretamente na performance do modelo e no resultado da classificação. Para contornar esse problema, foi necessário, primeiramente, separar todos os pacientes sem depressão dos pacientes com depressão.

Depois disso, foi selecionado 70% dos pacientes com depressão e 70% dos pacientes sem depressão e unido esses dados em um único *dataset*, formando assim, o *dataset* de treinamento. O restante dos registros, ou seja, os 30% dos pacientes com depressão e sem depressão, também foram unidos em um único *dataset* com o objetivo de criar o *dataset* de teste.

Depois da criação dos *datasets* de treinamento e de teste, dois modelos de predição foram gerados, um utilizando a técnica de *Machine Learning Support Vector Machine* (SVM) e outro

utilizando a técnica *Decision Tree*, técnicas essas que foram as mais aplicadas nos artigos selecionados na revisão sistemática, e escolhidas para serem implementadas neste trabalho. Após a criação e treinamento dos modelos, os dados de testes foram aplicados a eles, e com isso, a matriz de confusão e as métricas de performance para cada modelo foram obtidas. Os dados positivos da matriz de confusão foram representados pelos pacientes sem depressão e os dados negativos da matriz foram representados pelos pacientes com depressão. Todos os modelos desenvolvidos neste trabalho foram construídos utilizando a linguagem R no R Studio.

5.2.1 SVM

Para a construção do modelo SVM foram utilizadas as bibliotecas *e1071*, *caret* e *klar* que se encontram no repositório CRAN¹⁴. As etapas para a construção dos modelos SVM foram divididas em treinamento e predição.

5.2.1.1 Treinamento

Na etapa de treinamento, foram criados 3 modelos, cada modelo aplicou uma estratégia de validação diferente. As estratégias de validação utilizadas foram, *10-fold cross validation*; *leave-one-out* e *holdout*. Os resultados obtidos de cada modelo são ilustrados na seção 5.2.1.2. Para o treinamento dos modelos SVM, foi utilizado o método Linear para classificação.

5.2.1.2 Predição

Após o treinamento, cada modelo treinado foi aplicado ao conjunto de dados de teste para realizar a predição dos pacientes com depressão, utilizando-se a variável *depression_flag* como atributo alvo. Para cada método de validação foi gerada a matriz de confusão, e as métricas como, acurácia, especificidade e sensibilidade, foram extraídas para análise e comparação. As métricas de performance de classificação de cada modelo construído são ilustradas na Tabela 2.

Tabela 2 - Métricas de performance de cada modelo de predição SVM

Método de Validação	Acurácia	Sensibilidade	Especificidade	AUC
Holdout	0,992	0,997	0,889	0,943
10-fold	0,984	0,997	0,889	0,943
Leave-one-out	0,995	0,994	1	0,997

Fonte: Elaborado pelo autor

¹⁴ <https://cran.r-project.org/>

É possível perceber que o modelo que utilizou o método de validação cruzada *leave-one-out* para treinamento obteve uma performance melhor na classificação de pacientes com e/ou sem depressão e levou aproximadamente 25 segundos para realizar o treinamento.

5.2.2 DECISION TREE

Para a construção do modelo *Decision Tree* foram utilizadas as bibliotecas *rpart*, *caret* e *klar* que se encontram no repositório CRAN¹⁵. O algoritmo *Decision Tree* utilizado neste trabalho foi *Adaboost*, sendo selecionado por ser o que se encontra em primeiro lugar na lista de algoritmos disponíveis na biblioteca *caret*. As etapas para a construção dos modelos *Decision Tree* também foram divididas em treinamento e predição.

5.2.2.1 Treinamento

Na etapa de treinamento utilizando a técnica *Decision Tree*, 3 modelos foram criados, onde cada modelo aplicou uma estratégia de validação diferente. As estratégias de validação utilizadas foram as mesmas da construção dos modelos SVM. Com isso é possível avaliar qual das técnicas obteve um resultado melhor. Os resultados obtidos de cada modelo utilizando a técnica *Decision Tree* são ilustrados na seção 5.2.2.2. Para o treinamento dos modelos *Decision Tree*, foi utilizado o método de classificação *adaboost*.

5.2.2.2 Predição

Após o treinamento, cada modelo treinado foi aplicado ao conjunto de dados de teste para realizar a predição dos pacientes com depressão utilizando a variável *depression_flag*. Para cada método de validação foi gerado a matriz de confusão, e as métricas como, acurácia, especificidade e sensibilidade, foram extraídas para análise e comparação. As métricas de performance de classificação de cada modelo construído são ilustradas na Tabela 3.

Tabela 3 - Métricas de performance de cada modelo de predição *Decision Tree*.

Método de Validação	Acurácia	Sensibilidade	Especificidade	AUC
Holdout	0,981	0,997	0,677	0,832
10-fold	0,981	0,989	0,833	0,911
Leave-one-out	0,986	0,989	0,944	0,966

Fonte: Elaborado pelo autor

¹⁵ <https://cran.r-project.org/>

É possível perceber que o modelo *Decision Tree* que utilizou o método de validação cruzada *leave-one-out* para treinamento obteve uma performance melhor na classificação de pacientes com e/ou sem depressão, ou seja, o mesmo cenário ocorrido na técnica SVM, porém, o tempo para treinamento deste modelo levou aproximadamente 2 horas e 40 minutos.

Ao final da etapa de criação dos modelos é possível realizar um comparativo das técnicas de ML utilizadas e identificar qual foi a melhor delas na tarefa de classificação. Comparando o melhor resultado de cada uma delas é possível perceber que as duas técnicas obtiveram resultados melhores quando utilizadas com a técnica de validação cruzada *leave-one-out*.

Tabela 4 - Resultado obtido por técnica ML utilizando a técnica de validação que obteve melhor resultado dentre as demais

Técnica ML	Método de Validação	Acurácia	Sensibilidade	Especificidade	AUC
SVM	Leave-one-out	0,995	0,994	1	0,997
<i>Decision Tree</i>	Leave-one-out	0,986	0,989	0,944	0,966

Fonte: Elaborado pelo autor

Conforme é ilustrado pela Tabela 4, com o cenário apresentado, as informações coletadas e submetidas para análise de classificação, a técnica que obteve uma performance melhor foi a técnica de *Machine Learning SVM* em conjunto com a técnica de validação cruzada *leave-one-out*. O que chama a atenção neste modelo é que ele obteve o valor de especificidade máximo, isso significa que o modelo SVM, utilizando validação cruzada *leave-one-out*, foi excelente classificando pacientes com depressão, ou seja, dentre 10 pacientes marcados com depressão na base de teste, os 10 foram classificados com depressão.

5.3 LIMITAÇÃO DOS MODELOS

No decorrer do desenvolvimento dos modelos de predição notou-se que o *dataset* utilizado para a análise de classificação de pacientes com depressão e sem depressão possuía muitos registros de pacientes sem depressão comparado aos com depressão. Mesmo com esse cenário, os modelos, tanto de SVM quanto *Decision Tree*, acabaram resultando em uma boa performance na hora da classificação.

Desta forma, o trabalho procurou avaliar como seria a performance dos modelos com os dados balanceados, a fim de observar se eles obteriam boas métricas ou se a classificação não seria eficiente. Para isso, foi obtido aleatoriamente 59 registros de pacientes sem depressão e unidos esses dados com os 59 registros de pacientes com depressão, formando um *dataset*

balanceado para análise. Após a criação do *dataset* balanceado, ele foi dividido em 70% para treinamento e 30% para teste.

A Tabela 5 ilustra os resultados obtidos para os modelos de predição que utilizaram a técnica ML SVM com os dados balanceados. Os modelos que utilizaram validação *holdout* e *10-fold* obtiveram uma AUC e uma especificidade maior do que quando aplicado no *dataset* desbalanceado, porém, a acurácia e a sensibilidade do modelo se mostraram inferiores. É possível perceber que esses modelos não classificaram muito bem os pacientes sem depressão e acabaram classificando-os como se tivessem depressão, interferindo assim na acurácia deles.

Tabela 5 - Métricas de performance de cada modelo de predição SVM com dataset balanceado

Método de Validação	Acurácia	Sensibilidade	Especificidade	AUC
Holdout	0,943	0,888	1	0,944
10-fold	0,943	0,888	1	0,944
Leave-one-out	0,943	0,888	1	0,944

Fonte: Elaborado pelo autor

Tabela 6 - Métricas de performance de cada modelo de predição *Decision Tree* com dataset balanceado

Método de Validação	Acurácia	Sensibilidade	Especificidade	AUC
Holdout	1	1	1	1
10-fold	0,971	0,944	1	0,972
Leave-one-out	0,971	0,944	1	0,972

Fonte: Elaborado pelo autor

Na validação *leave-one-out* o modelo acabou obtendo acurácia, sensibilidade e AUC menor em relação ao modelo treinado com os dados desbalanceados, porém, a especificidade se mostrou a mesma, concluindo que o modelo conseguiu classificar muito bem os pacientes com depressão nos dois cenários.

Já nos modelos que utilizaram a técnica de ML *Decision Tree* o resultado foi diferente. A Tabela 6 ilustra os resultados obtidos nos modelos de predição *Decision Tree* quando treinados com dados balanceados. É possível perceber que a performance destes modelos foi melhor do que a dos modelos que foram treinados com os dados desbalanceados. Neste cenário, o modelo utilizando a validação *holdout* superou os modelos que utilizaram as estratégias de validação *leave-one-out* e *10-fold*.

Nesta seção podemos concluir que com os dados balanceados, o modelo *Decision Tree* utilizando a técnica de validação *holdout* se saiu melhor que SVM com os dados desbalanceados, e melhor que DT com *leave-one-out* quando treinado com o *dataset* desbalanceado.

5.4 ENTREVISTA COM O CTO DA BEETOUCH

Esta seção aborda a entrevista que foi realizada com o CTO da empresa *BeeTouch*. Essa entrevista teve como foco principal apresentar os resultados obtidos no trabalho, mostrar o processo de seleção dos atributos utilizados para classificação, explicar as métricas utilizadas para medir a performance dos modelos e as técnicas ML que foram aplicadas para a criação deles. Depois disso, um questionário foi encaminhado ao CTO para que fosse possível saber a opinião dele em relação a este trabalho. O Quadro 6 mostra as perguntas que foram feitas e as respostas obtidas para cada uma delas.

Quadro 6 - Questionário realizado ao CTO da *Beetouch* e suas respostas

Perguntas	Respostas
Como o TCC agrega valor para a plataforma AVAXpsi?	O trabalho agrega uma nova visão para a massa de dados armazenada. Até então, a plataforma olhava apenas para os dados realizando análises conforme a literatura. O trabalho trouxe uma nova visão para esta informação, mostrando correlações que até então não eram exploradas
O que achou dos resultados apresentados?	Como citado acima, os resultados trouxeram uma nova visão para uma série de dados, que até então eram analisados separadamente. Gostei muito do desempenho apresentado pelos algoritmos de predição. Como nossa base de dados não era uma base grande, fiquei preocupado com uma possível taxa baixa de sucesso
Pretendes utilizar os resultados do trabalho no AVAXpsi? Como?	Certamente. Neste momento, estamos em fase de definição da uma grande mudança na plataforma. Pretendo utilizar os resultados como base para um módulo preditivo da plataforma
O que achaste dos atributos escolhidos para compor os modelos?	Alguns dos atributos escolhidos já estavam presentes em nosso algoritmo, porém, a análise trouxe outros atributos, que são utilizados apenas em outros fatores, como relevantes. É neste ponto que está a maior riqueza do trabalho

Fonte: Elaborado pelo autor

Como pode ser observado, de uma maneira geral, a partir das respostas obtidas do questionário feito ao CTO, este trabalho agregou valor a plataforma, trouxe uma nova visão em relação aos dados contidos na base e trouxe visibilidade para atributos que antes eram utilizados em um contexto diferente, não relacionado a depressão, mas que, possuem forte correlação e importância no momento da classificação de pacientes com depressão. O trabalho também apresentou ótimos resultados de performance dos modelos e trouxe a possibilidade da criação de um módulo preditivo dentro da plataforma AVAXpsi.

6 CONCLUSÃO

Este trabalho se propôs a construir modelos de predição utilizando técnicas de *Machine Learning* nos dados obtidos da plataforma AVAXpsi a fim de identificar pacientes com depressão. Antes de iniciar o processo de construção dos modelos preditivos, foi realizado um levantamento bibliográfico com o objetivo de aprofundar os conhecimentos sobre depressão. Para isso, foi realizada uma pesquisa para descobrir os diferentes tipos de depressão (neste trabalho citamos os mais comuns), os tratamentos disponíveis para as pessoas que sofrem com essa doença e os lugares com maiores casos de depressão através de dados estatísticos. Também, no capítulo sobre depressão, foi discutido sobre as aplicações existentes que servem para apoiar os pacientes e os profissionais da área da saúde mental quando o assunto é depressão, com o objetivo de saber se os modelos aqui construídos em conjunto com a plataforma AVAXpsi agregam valor para a área da saúde e tecnologia.

Este trabalho também realizou uma revisão sistemática que serviu para identificar as técnicas e algoritmos mais utilizados na construção de modelos preditivos e que serviram como base para o desenvolvimento dos modelos de predição deste trabalho. Os trabalhos avaliados na revisão sistemática foram aqueles que aplicaram *Machine Learning* em dados da área da saúde, principalmente aqueles que buscaram classificar pacientes com depressão. Também procurou-se identificar quais foram os métodos de validação utilizados nos artigos para avaliar o desempenho dos modelos preditivos.

Com o estudo feito na revisão sistemática, as técnicas SVM e *Decision Tree* foram selecionadas para a construção dos modelos preditivos. Todas as técnicas foram aplicadas a 3 métodos de validação diferentes, sendo eles, *holdout*, *leave-one-out* e *10-fold cross validation*. Essa estratégia de combinar diferentes métodos de validação foi utilizada para identificar qual foi o modelo que obteve a melhor performance.

Para garantir a qualidade dos dados a serem analisados, estratégias de preenchimento de valores ausentes e técnicas de seleção de atributos foram utilizadas. Durante a criação dos *datasets* de treinamento e teste percebeu-se que os dados eram muito desbalanceados, existiam muito mais registros de pessoas sem depressão do que com depressão, sendo 1161 contra 59, respectivamente. Mesmo com este cenário, os modelos com dados desbalanceados obtiveram ótimos resultados, sendo o modelo SVM utilizando *leave-one-out* o melhor em comparação aos outros.

Sabendo do limite do modelo em relação a quantidade de dados e em relação ao desbalanceamento deles, foi realizado um outro estudo com o objetivo de avaliar como os modelos se sairiam em uma base onde os dados estivessem balanceados. Para isso, foram obtidos 59 registros de forma aleatória dentre os registros de pacientes sem depressão, esses dados foram combinados em um outro *dataset* com os de pacientes com depressão. Utilizando este novo *dataset*, as etapas de treinamento e predição foram as mesmas antes aplicadas no *dataset* desbalanceado. Com os dados balanceados, a técnica *Decision Tree* utilizando validação *houldout* obteve o melhor resultado dentre todos os outros.

Para entender a importância e relevância deste estudo, um questionário foi feito para o CTO da *BeeTouch* buscando entender se este trabalho pode evoluir e se tornar algo aplicável dentro da plataforma e saber se os resultados e os dados apresentados trouxeram um conhecimento de qualidade e uma nova perspectiva diante as informações contidas na base. Diante das respostas foi possível perceber que este trabalho supriu as expectativas e trouxe uma nova visão aos dados que antes não eram explorados. Atributos que antes não eram considerados importantes para identificar pacientes com depressão agora podem ser analisados e serem utilizados como peça fundamental no apoio ao diagnóstico de pacientes com depressão.

Este trabalho abre a oportunidade da criação de um módulo preditivo dentro da AVAXpsi enriquecendo mais ainda a plataforma e aprimorando a qualidade das análises feitas pelo sistema. Conforme mais dados forem sendo inseridos na plataforma mais os modelos irão aprender e se aperfeiçoar. Novas técnicas de classificação também podem ser adicionadas e novos atributos explorados. Outro aspecto importante deste trabalho é a contribuição para a área de predição de casos de depressão a partir de dados originados de avaliações psicológicas, auxiliando outras pesquisas que tem um objetivo similar ao deste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDO, Carmita. **Da depressão à disfunção sexual (e vice-versa)**. 3. ed. São Paulo.
- ADITYA MISHRA. **Metrics to Evaluate your Machine Learning Algorithm**. 2018. Disponível em: <<https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234>>. Acesso em: 1 mar. 2020.
- AGARWAL, Rahul. **The 5 Feature Selection Algorithms every Data Scientist should know**. 2019. Disponível em: <<https://towardsdatascience.com/the-5-feature-selection-algorithms-every-data-scientist-need-to-know-3a6b566efd2>>. Acesso em: 29 mar. 2020.
- ALI, Yasir et al. Detection of schistosomiasis factors using association rule mining. **IEEE Access**, [s. l.], v. 7, p. 186108–186114, 2019.
- ANOOPKUMAR, M.; MD ZUBAIR RAHMAN, A. M. J. A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration. In: PROCEEDINGS OF 2016 INTERNATIONAL CONFERENCE ON DATA MINING AND ADVANCED COMPUTING, SAPIENCE 2016 2016, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2016.
- ANTHES, Emily. Depression: A change of mind. **Nature**, [s. l.], v. 515, n. 7526, p. 185–187, 2014.
- APARECIDO DE ALMEIDA CALIL, Leonardo et al. **MINERAÇÃO DE DADOS E PÓS-PROCESSAMENTO EM PADRÕES DESCOBERTOS DATA MINING AND POS-PROCESSING IN PATTERNS DISCOVERABLE** *Exatas Terra, Ci. Agr. Eng.* [s.l: s.n.]. Disponível em: <<http://ri.uepg.br/riuepg/handle/123456789/142>>. Acesso em: 29 fev. 2020.
- ARDEN, John B. **Brain2Brain: Enacting Client Change Through the Persuasive Power of Neuroscience**. [s.l: s.n.]. Disponível em: <<https://doi.org/10.1002/9781119171683.ch9>>
- BAHLS, Saint-Clair. **Aspectos Clínicos da Depressão em Crianças e Adolescentes** *Jornal de Pediatria*. [s.l: s.n.]. Disponível em: <<http://www.scielo.br/pdf/0D/jped/v78n5/7805359.pdf>>.
- BALCHIN, Ross et al. Sweating away depression? the impact of intensive exercise on depression. **Journal of Affective Disorders**, [s. l.], v. 200, p. 218–221, 2016.
- BASKAR, S. S.; AROCKIAM, L.; CHARLES, S. **A Systematic Approach on Data Pre-processing In Data Mining** *An international journal of advanced computer technology*. [s.l: s.n.].
- BHAUMIK, Runa et al. Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity. **NeuroImage: Clinical**, [s. l.], v. 16, p. 390–398, 2017. Disponível em: <<https://www.sciencedirect.com/periodics/capes.gov.br/science/article/pii/S2213158216300390?via%3DiHub>>. Acesso em: 3 nov. 2019.
- BHAYA, Wesam S.; ALASADI, Suad A. Review of Data Preprocessing Techniques in Data Mining Some of the authors of this publication are also working on these related projects: Influential Nodes Based Alleviation of User Cold-Start Problem in Recommendation System View project TCP and UDP to . [s. l.], 2017. Disponível em: <<https://www.researchgate.net/publication/319990923>>. Acesso em: 18 fev. 2020.
- BYUN, Sangwon et al. Entropy analysis of heart rate variability and its application to recognize major depressive disorder: A pilot study. In: TECHNOLOGY AND HEALTH CARE 2019,

- Anais... : IOS Press, 2019. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6597986/>>. Acesso em: 3 nov. 2019.
- CAMPOS, Josiane Rosa; PRETTE, Zilda Aparecida Pereira Del; PRETTE, Almir Del. Relações entre depressão, habilidades sociais, sexo e nível socioeconômico em grandes amostras de adolescentes. **Psicologia: Teoria e Pesquisa**, [s. l.], v. 34, n. 0, 2018.
- CHIRIȚĂ, Anca Livia et al. Current understanding of the neurobiology of major depressive disorder. **Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie**, [s. l.], v. 56, n. 2 Suppl, p. 651–8, 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26429155>>. Acesso em: 21 set. 2019.
- CIOS, Krzysztof J. et al. The Knowledge Discovery Process. In: **Data Mining**. Boston, MA: Springer US, 2007. p. 9–24.
- CRAMER, Holger et al. **A systematic review of yoga for major depressive disorder** *Journal of Affective Disorders* Elsevier B.V., , 2017.
- CUI, Ranji. Editorial: A Systematic Review of Depression. **Current Neuropharmacology**, [s. l.], v. 13, n. 4, 2015. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4790400/>>
- DING, Weiping. SVM-Based feature selection for differential space fusion and its application to diabetic fundus image classification. **IEEE Access**, [s. l.], v. 7, p. 149493–149502, 2019.
- DUTT, Ashish; ISMAIL, Maizatul Akmar; HERAWAN, Tutut. **A Systematic Review on Educational Data Mining** *IEEE Access* Institute of Electrical and Electronics Engineers Inc., , 2017.
- EDHLUND, Bengt M.; MCDUGALL, Allan G. **PubMed essentials : mastering the world's health research database**. [s.l.] : Form & Kunskap AB, 2014.
- FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. The KDD Process for Extracting Useful Knowledge from Volumes of Data. **Communications of the ACM**, [s. l.], v. 39, n. 11, p. 27–34, 1996. Disponível em: <<http://portal.acm.org/citation.cfm?doid=240455.240464>>. Acesso em: 12 fev. 2020.
- FOLAND-ROSS, Lara C. et al. Cortical thickness predicts the first onset of major depression in adolescence. **International Journal of Developmental Neuroscience**, [s. l.], v. 46, p. 125–131, 2015. Disponível em: <<https://www-sciencedirect.ez310.periodicos.capes.gov.br/science/article/pii/S0736574815300368?via%3DiHub>>. Acesso em: 3 nov. 2019.
- FRÍAS IBÁÑEZ, Álvaro et al. Motivation for change as a predictor of treatment response for dysthymia. **Psicothema**, [s. l.], v. 28, n. 2, p. 156–60, 2016. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/27112812>>. Acesso em: 15 set. 2019.
- GUPTA, Prashant. **Cross-Validation in Machine Learning**. 2017. Disponível em: <<https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f>>. Acesso em: 21 abr. 2020.
- HAN, Jiawei.; KAMBER, Micheline. **Data mining : concepts and techniques**. [s.l.] : Elsevier, 2006.
- HECKLER, WESLLEI FELIPE. **ANÁLISE PREDITIVA SOBRE PACIENTES DO “PROJETO DE EXTENSÃO REABILITAÇÃO PULMONAR” DA UNIVERSIDADE FEEVALE**. 2018. Universidade Feevale, [s. l.], 2018. Disponível em: <https://tconline.feevale.br/NOVO/tc/files/0001_4637.pdf#page=30&zoom=100,0,665>
- HOSSEINIFARD, Behshad; MORADI, Mohammad Hassan; ROSTAMI, Reza. Classifying

depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. **Computer Methods and Programs in Biomedicine**, [s. l.], v. 109, n. 3, p. 339–345, 2013. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169260712002507>>. Acesso em: 25 ago. 2019.

JALOTA, Chitra; AGRAWAL, Rashmi. Analysis of Educational Data Mining using Classification. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON MACHINE LEARNING, BIG DATA, CLOUD AND PARALLEL COMPUTING: TRENDS, PRESPECTIVES AND PROSPECTS, COMITCON 2019 2019, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2019.

KERRI SMITH. **Mental health: a world of depression**. 2014. Disponível em: <<https://www.nature.com/news/mental-health-a-world-of-depression-1.16318>>.

KIM, Ah Young et al. Automatic detection of major depressive disorder using electrodermal activity. **Scientific Reports**, [s. l.], v. 8, n. 1, 2018. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6242826/>>. Acesso em: 3 nov. 2019.

KITCHENHAM, Elisabete. **Guidelines for performing Systematic Literature Reviews in Software Engineering version 2**. Durham: Department of Computer Science University of Durham Durham, UK, 2007.

KOK, Rob M.; REYNOLDS, Charles F. **Management of depression in older adults: A review****JAMA - Journal of the American Medical Association**American Medical Association, , 2017.

KURLANSIK, Stuart L.; IBAY, Annamarie D. Seasonal affective disorder. **American family physician**, [s. l.], v. 86, n. 11, p. 1037–41, 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23198671>>. Acesso em: 15 set. 2019.

LAKSHMI, G. Seetha; DEVI, P. Isakki. Prediction of anti-retro viral for HIV and STD patients using data mining technique. In: PROCEEDINGS OF THE 2017 IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNIQUES IN CONTROL, OPTIMIZATION AND SIGNAL PROCESSING, INCOS 2017 2018, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2018.

LEDESMA, Sergio et al. Analysis of data sets with learning conflicts for machine learning. **IEEE Access**, [s. l.], v. 6, p. 45062–45070, 2018.

LIVNE, Ofir et al. The association between lifetime cannabis use and dysthymia across six birth decades. **Journal of affective disorders**, [s. l.], v. 234, p. 327–334, 2018. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/29605793>>. Acesso em: 15 set. 2019.

LOPEZ, Juan Pablo; KOS, Aron; TURECKI, Gustavo. **Major depression and its treatment: MicroRNAs as peripheral biomarkers of diagnosis and treatment response****Current Opinion in Psychiatry**Lippincott Williams and Wilkins, , 2018.

MADHURI, C. H. Rag.; ANURADHA, G.; PUJITHA, M. Vani. House Price Prediction Using Regression Techniques: A Comparative Study. In: 6TH IEEE INTERNATIONAL CONFERENCE ON &AMP;AMP;AMP;AMP;QUOT;SMART STRUCTURES AND SYSTEMS&AMP;AMP;AMP;AMP;QUOT;, ICSSS 2019 2019, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2019.

MAGIYA, Joseph. **Pearson Coefficient of Correlation Explained. - Towards Data Science**. 2019. Disponível em: <<https://towardsdatascience.com/pearson-coefficient-of-correlation-explained-369991d93404>>. Acesso em: 29 mar. 2020.

MARSLAND, S. Machine learning: an algorithmic perspective. [s. l.], 2014. Disponível em:

- <<https://content.taylorfrancis.com/books/download?dac=C2012-0-12750-0&isbn=9780429102509&format=googlePreviewPdf>>. Acesso em: 8 fev. 2020.
- MARTINEZ, Viridiana Romero. **A machine learning approach for the detection of depression and mental illness in Twitter**. 2019. Disponível em: <<https://medium.com/datadriveninvestor/a-machine-learning-approach-for-detection-of-depression-and-mental-illness-in-twitter-3f3a32a4df60>>. Acesso em: 20 ago. 2019.
- MEDEIROS, EDNA RAMOS DE. **REVISÃO SISTEMÁTICA SOBRE OS DISPOSITIVOS VESTÍVEIS NA ÁREA DA SAÚDE**. 2016. Universidade Feevale, [s. l.], 2016. Disponível em: <https://tconline.feevale.br/NOVO/tc/files/0002_4240.pdf#page=42&zoom=100,0,136>
- MILGROM, Jeannette; GEMMILL, Alan W. Screening for perinatal depression. **Best practice & research. Clinical obstetrics & gynaecology**, [s. l.], v. 28, n. 1, p. 13–23, 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24095728>>. Acesso em: 15 set. 2019.
- MINISTÉRIO DA SAÚDE. **Depressão: causas, sintomas, tratamentos, diagnóstico e prevenção**. 2018. Disponível em: <<http://www.saude.gov.br/saude-de-a-z/saude-mental/depressao>>. Acesso em: 3 ago. 2019.
- MORIN, Amy. **Depression Statistics Everyone Should Know**verywell mind, , 2019. Disponível em: <<https://www.verywellmind.com/depression-statistics-everyone-should-know-4159056>>
- NARKHEDE, Sarang. **Understanding AUC - ROC Curve - Towards Data Science**. 2018. Disponível em: <<https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>>. Acesso em: 1 mar. 2020.
- NATIONAL ALLIANCE OF MENTAL ILLNESS. **Depression**. 2017. Disponível em: <<https://www.nami.org/Learn-More/Mental-Health-Conditions/Depression>>. Acesso em: 16 set. 2019.
- NATIONAL INSTITUTE OF MENTAL HEALTH. **Depression Basics**. 2016. Disponível em: <<https://www.nimh.nih.gov/health/publications/depression/index.shtml#pub2>>. Acesso em: 15 set. 2019.
- NATURE INTERNATIONAL JOURNAL OF SCIENCE. The burden of depression. **Nature**, [s. l.], v. 515, n. 7526, p. 163–163, 2014. Disponível em: <<http://www.nature.com/articles/515163a>>. Acesso em: 15 set. 2019.
- NELSON CORLETA SCHREIBER, Jacques et al. **TÉCNICAS DE VALIDAÇÃO DE DADOS PARA SISTEMAS INTELIGENTES: UMA ABORDAGEM DO SOFTWARE SDBAYES**. [s.l.: s.n.]. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/181199/102_00071.pdf?sequence=1&isAllowed=y>. Acesso em: 2 mar. 2020.
- NEUFELD, Nicholas H. et al. Resting state functional connectivity in patients with remitted psychotic depression: A multi-centre STOP-PD study. **EBioMedicine**, [s. l.], v. 36, p. 446–453, 2018.
- NILSSON, NJ. **Introduction to machine learning: An early draft of a proposed textbook**. Stanford: Department of Computer Science Stanford University Stanford, CA, 1998. Disponível em: <<http://ai.stanford.edu/~nilsson/MLBOOK.pdf>>. Acesso em: 6 out. 2019.
- O'CONNOR, Stuart; AGIUS, Mark. A systematic review of structural and functional MRI differences between psychotic and nonpsychotic depression. **Psychiatria Danubina**, [s. l.], v. 27 Suppl 1, p. S235-9, 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26417770>>. Acesso em: 15 set. 2019.

OLIVETO, Paloma. **Aplicativos de celular são as novas armas no combate à depressão** Correio Braziliense, , 2018. Disponível em: <https://www.correiobraziliense.com.br/app/noticia/ciencia-e-saude/2018/07/29/interna_ciencia_saude,698103/aplicativos-de-celular-contra-a-depressao.shtml>. Acesso em: 18 ago. 2019.

PAVAN, Marina. **DESIGN E TECNOLOGIA: CRIAÇÃO DE FERRAMENTA DE CONSCIENTIZAÇÃO SOBRE DEPRESSÃO**. 2017. UNIVATES - Universidade do Vale do Taquari, [s. l.], 2017. Disponível em: <<https://www.univates.br/bdu/bitstream/10737/1881/1/2017MarinaPavan.pdf>>

PUGGINA BIANCHESI, Natalia Maria et al. A design of experiments comparative study on clustering methods. **IEEE Access**, [s. l.], v. 7, p. 167726–167738, 2019.

RICARD, Benjamin J. et al. Exploring the utility of community-generated social media content for detecting depression: An analytical study on instagram. **Journal of Medical Internet Research**, [s. l.], v. 20, n. 12, 2018. Disponível em: <<https://www.jmir.org/2018/12/e11817/>>. Acesso em: 3 nov. 2019.

ROH, Taehwan; HONG, Sunjoo; YOO, Hoi Jun. Wearable depression monitoring system with heart-rate variability. In: 2014 36TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, EMBC 2014 2014, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2014.

SAU, Arkaprabha; BHAKTA, Ishitaa. Predicting anxiety and depression in elderly patients using machine learning technology. **Healthcare Technology Letters**, [s. l.], v. 4, n. 6, p. 238–243, 2017. Disponível em: <<https://digital-library.theiet.org/content/journals/10.1049/htl.2016.0096>>. Acesso em: 3 nov. 2019.

SCHESTATSKY, Sidnei; FLECK, Marcelo. Psicoterapia das depressões. **Revista Brasileira de Psiquiatria**, [s. l.], v. 21, n. suppl 1, p. 41–47, 1999.

SEMENKOVICH, Katherine et al. Depression in type 2 diabetes mellitus: Prevalence, impact, and treatment. **Drugs**, [s. l.], v. 75, n. 6, p. 577–587, 2015.

SINGH, Yogesh; KAUR, Arvinder; MALHOTRA, Ruchika. Empirical validation of object-oriented metrics for predicting fault proneness models. **Software Quality Journal**, [s. l.], v. 18, n. 1, p. 3–35, 2010.

SOCIAL GOOD BRASIL. **13 iniciativas que ajudam no combate à depressão**. 2017. Disponível em: <<https://socialgoodbrasil.org.br/laboratorio/13-iniciativas-que-ajudam-no-combate-a-depressao/>>. Acesso em: 18 ago. 2019.

SOLOMON, Andrew. **O Demônio do meio-dia: uma anatomia da depressão**. 2. ed. São Paulo.

SONKAMBLE, Balwant A.; DOYE, D. D. An overview of speech recognition system based on the support vector machines. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTER AND COMMUNICATION ENGINEERING 2008, ICCCE08: GLOBAL LINKS FOR HUMAN DEVELOPMENT 2008, **Anais...** [s.l: s.n.]

SOUZA, Alex. **Algoritmo SVM (Máquina de Vetores de Suporte) a partir de exemplos e código (Python e R) | ALEX SOUZA**. 2019a. Disponível em: <<https://pessoalex.wordpress.com/2019/04/10/algoritmo-svm-maquina-de-vetores-de-suporte-a-partir-de-exemplos-e-codigo-python-e-r/>>. Acesso em: 29 fev. 2020.

SOUZA, Alex. **Decision Trees e Random Forests para Classificação e Regressão | ALEX SOUZA**. 2019b. Disponível em: <<https://pessoalex.wordpress.com/2019/09/30/decision-trees-e-random-forests-para-classificacao-e-regressao/>>. Acesso em: 29 fev. 2020.

SOUZA, Alex. **Avaliação do Modelo de Classificação** | ALEX SOUZA. 2019c. Disponível em: <<https://pessoalex.wordpress.com/2019/04/02/avaliacao-do-modelo-de-classificacao/>>. Acesso em: 1 mar. 2020.

STEINER, Maria Teresinha Arns et al. Abordagem de um problema Médico por meio do processo de KDD com ênfase Á análise exploratória dos dados. **Gestao e Producao**, [s. l.], v. 13, n. 2, p. 325–337, 2006.

STERZ, Gabriela Anita; CARDOSO DA SILVA, Jerto. **DEPRESSÃO NA INFÂNCIA E NA ADOLESCÊNCIA** *Boletim EntreSIS*. [s.l.: s.n.]. Disponível em: <<https://online.unisc.br/acadnet/anais/index.php/boletimsis/article/view/16902>>.

TOMAR, Divya; AGARWAL, Sonali. A Survey on Pre-processing and Post-processing Techniques in Data Mining. **International Journal of Database Theory and Application**, [s. l.], v. 7, n. 4, p. 99–128, 2014. Disponível em: <<http://dx.doi.org/10.14257/ijdta.2014.7.4.09>>. Acesso em: 29 fev. 2020.

TUHLINSKI, Camila. **Depressão será a doença mental mais incapacitante do mundo até 2020** ESTADO DE S.PAULO, , 2018. Disponível em: <<https://emails.estadao.com.br/noticias/bem-estar,depressao-sera-a-doenca-mental-mais-incapacitantes-do-mundo-ate-2020,70002542030>>. Acesso em: 17 ago. 2019.

VAN DER AALST, Wil. **Process mining: Overview and opportunities** *ACM Transactions on Management Information Systems*, 2012. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2229156.2229157>>. Acesso em: 22 fev. 2020.

VICTOR, Ezekiel et al. Detecting Depression Using a Framework Combining Deep Multimodal Neural Networks With a Purpose-Built Automated Evaluation. **Psychological Assessment**, [s. l.], n. 8, p. 1019–1027, 2019.

WANG, Tong et al. Decision tree based online stability assessment scheme for power systems with renewable generations. **CSEE Journal of Power and Energy Systems**, [s. l.], v. 1, n. 2, p. 53–61, 2015.

WIKIPEDIA. **Electroencephalography**. [s.d.]. Disponível em: <<https://en.wikipedia.org/wiki/Electroencephalography>>. Acesso em: 25 ago. 2019.

WORLD HEALTH ORGANIZATION. **Depression: let's talk**. 2017. Disponível em: <https://www.who.int/mental_health/management/depression/en/>. Acesso em: 17 ago. 2019.

YAN, Xuyang et al. A novel streaming data clustering algorithm based on fitness proportionate sharing. **IEEE Access**, [s. l.], v. 7, p. 184985–185000, 2019.

YOM-TOV, Elad; WHITE, Ryan W.; HORVITZ, Eric. Seeking insights about cycling mood disorders via anonymized search logs. **Journal of Medical Internet Research**, [s. l.], v. 16, n. 2, 2014. Disponível em: <<https://www.jmir.org/2014/2/e65/>>. Acesso em: 3 nov. 2019.

ZHANG, Lili; WANG, Wenjie; ZHANG, Yuqing. Privacy Preserving Association Rule Mining: Taxonomy, Techniques, and Metrics. **IEEE Access**, [s. l.], v. 7, p. 45032–45047, 2019.

ZHANG, Shichao; ZHANG, Chengqi; YANG, Qiang. Data preparation for data mining. **Applied Artificial Intelligence**, [s. l.], v. 17, n. 5–6, p. 375–381, 2003. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/713827180>>. Acesso em: 15 fev. 2020.