

**CESED – CENTRO DE ENSINO SUPERIOR E DESENVOLVIMENTO
UNIFACISA – FACULDADE DE CIÊNCIAS SOCIAIS APLICADAS
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

DIEGO KAZADI KALUNDA

**APLICANDO MODELOS DE APRENDIZAGEM DE MÁQUINA PARA PREVER
HIPERTENSÃO ARTERIAL: UMA ANÁLISE COMPARATIVA**

CAMPINA GRANDE – PB

2020

DIEGO KAZADI KALUNDA

**APLICANDO MODELOS DE APRENDIZAGEM DE MÁQUINA PARA PREVER
HIPERTENSÃO ARTERIAL: UMA ANÁLISE COMPARATIVA**

Trabalho de Conclusão de Curso
apresentado como pré-requisito para a
obtenção do título de Bacharel em
Sistemas de Informação pelo Centro
Universitário de Ciências Sociais
Aplicadas.

Área de Concentração: Aprendizagem
de Máquina.

Orientador: Profa. Msc. Lizianne Souto

CAMPINA GRANDE – PB

2020

Trabalho de Conclusão de Curso,
aplicando modelos de machine learning
para prever hipertensão arterial: uma
análise comparativa, apresentado por
Diego Kazadi Kalunda como parte dos
requisitos para obtenção do título de
Bacharel em Sistemas de Informação
outorgado pela Unifacisa – Centro
Universitário.

APROVADO EM ____/____/____

BANCA EXAMINADORA

Prof^a Msc. Lizianne Souto,
Orientadora

Prof^o

Membro

*Dedico este trabalho à minha família KALUNDA MUTABA e aos meus filhos que
são bênçãos vindas de Deus*

AGRADECIMENTOS

A JESUS CRISTO o criador de todas as coisas visíveis e invisíveis, por ter me dado saúde e força para superar as dificuldades.

À Unifacisa, seu corpo docente, direção e administração que oportunizaram a janela que hoje vislumbro um horizonte superior, pelo ambiente criativo e amigável que proporciona a confiança no mérito e ética aqui presente.

À professora Lizianne Souto pelo paciente trabalho de revisão da redação.

Aos meus pais, pelo amor, incentivo e apoio incondicional apesar da longa distância que nós se para.

E a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

“De tudo o que se tem ouvido, o fim é: Teme a Deus, e guarda os seus mandamentos; porque isto é o dever de todo o homem.”

Rei Salomão (Ecle. 12.13)

RESUMO

O presente trabalho propõe a investigação de modelos de aprendizagem de máquina aplicadas ao problema do diagnóstico de hipertensão arterial, com intuito de prever a taxa probabilística da ocorrência da doença cardíaca em pacientes. O trabalho possui como objetivo investigar quais modelos dentre os selecionados, alcançam melhores taxas de classificação para o problema. A aplicação foi organizada de acordo com as seguintes etapas: coleta e pré-processamento de dados onde o conjunto de dados foi tratado de forma adequada na busca da qualidade de dados, classificação e avaliação dos modelos com o uso da ferramenta Weka. Utilizou-se 4 modelos de aprendizado: *Redes Bayesiana*, *LogitBoost*, *Multilayer-Perceptron* e *J48*. Os hiperparâmetros dos algoritmos foram otimizados por validação cruzada 10-fold, com intuito de selecionar aqueles correspondentes aos melhores modelo. Para cada algoritmo, efetuou-se ajuste na configuração padrão que a ferramenta apresenta para alcançar o melhor desempenho. Todos modelos apresentarem acurácia superior a 70%, e o modelo Redes Bayesiana atingir o melhor resultado entre os quatro à sua taxa de acerto foi superior à 82%, demonstrando a viabilidade de identificar com maior probabilidade de ter hipertensão arterial nos pacientes diagnosticados. Por fim, as métricas de desempenho para os modelos construídos a partir das técnicas de desempenho descritos são comparados e contrastados, juntamente com um resumo dos resultados no geral.

Palavras-chave: Diagnóstico de Hipertensão Arterial. Aprendizagem de Máquina. Classificação. Análise Comparativa.

RESUMÉ

Le présent travail propose d'étudier des modèles d'apprentissage automatique appliqués au problème du diagnostic de l'hypertension artérielle, afin de prédire le taux probabiliste de survenue d'une maladie cardiaque chez les patients. Le travail vise à étudier quels modèles parmi ceux sélectionnés obtiennent les meilleurs taux de classification pour le problème. L'application a été organisée selon les étapes suivantes: collecte et pré-traitement des données où l'ensemble de données a été correctement traité dans la recherche de la qualité des données, la classification et l'évaluation des modèles à l'aide de l'outil Weka. Quatre modèles d'apprentissage ont été utilisés: les réseaux bayésiens, Logit Boost, Multilayer-Perceptron et J48. Les hyperparamètres des algorithmes ont été optimisés par 10 fois la validation croisée, afin de sélectionner ceux correspondant aux meilleurs modèles. Pour chaque algorithme, des ajustements ont été apportés à la configuration standard que l'outil présente pour obtenir les meilleures performances. Tous les modèles avaient une précision supérieure à 70% et le modèle Réseaux Bayésiens a obtenu le meilleur résultat parmi les quatre; son taux de réussite était supérieur à 82%, démontrant la faisabilité d'identifier les plus susceptibles d'avoir une hypertension artérielle chez les patients diagnostiqués. Enfin, les métriques de performance des modèles construits à partir des techniques de performance décrites sont comparées et contrastées, ainsi qu'un résumé des résultats en général.

Mots-clés: Diagnostic de l'hypertension artérielle. Apprentissage automatique. Classification. Analyse Comparative.

LISTA DE ILUSTRAÇÕES

Figura 1 - Equipamento Esfigmomanômetro para Medição.....	21
Figura 2 - Monitorização Ambulatorial da Pressão Arterial.....	24
Figura 3 - Redes Neurais Biológicas.....	27
Figura 4 - Arquitetura da Rede neural.....	29
Figura 5 - Estrutura da Rede Perceptron.....	30
Figura 6 - Exemplo de uma árvore de decisão.....	33
Figura 7 - Weka GUI Chooser.....	37
Figura 8 - Dados carregados no Weka.....	38

LISTA DE TABELAS

Tabela 1 - Matriz de Confusão.....	20
Tabela 2 - Classificação diagnóstica da hipertensão arterial.....	26
Tabela 3 - Cleveland UCI Repository.....	39
Tabela 4 - Definição dos atributos da Cleveland Heart Disease.....	39
Tabela 5 - Acurácia do algoritmo Rede Bayesiana Experimento 1.....	42
Tabela 6 - Acurácia do algoritmo LogitBoost Experimento 2.....	43
Tabela 7 - Acurácia do algoritmo Multilayer Perceptron Experimento 3.....	44
Tabela 8 - Acurácia do algoritmo J48 Experimento 4.....	45
Tabela 9 - Resultados de classificação final dos quatros modelos.....	46
Tabela 10 - Comparação dos resultados trabalhos relacionados.....	47

LISTA DE ABREVIATURAS E SIGLAS

AUC	<i>Area Under The Curve</i> , no inglês, ou Área abaixo da Curva
AD	Árvore de Decisão
AM	Aprendizagem de Máquina
ARFF	<i>Attribute-Relation File Format</i> , ou Arquivo no Formato Atributo Relação.
CSV	<i>Comma Separated Values</i> , ou Arquivo de Valores separados por Vírgula
FP	Falso Positivo
FN	Falso Negativo
HAS	Hipertensão arterial sistêmica
HDL	<i>High Density Lipoprotein</i> , ou Lipoproteína de Alta Densidade
IA	Inteligência Artificial
IBM	<i>International Business Machines</i>
GPL	<i>General Public License</i>
LDL	<i>Low Density Lipoprotein</i> , ou Lipoproteína de Baixa Densidade MAPA Monitorização Ambulatorial da Pressão Arterial
MLP	Multilayer Perceptron
ML	<i>Machine Learning</i> , no inglês, ou Aprendizagem de Máquina
MRPA	Monitorização Residencial de Pressão Arterial
MS	Ministério da Saúde
KDD	<i>Knowledge-Discovery in Data bases</i> , ou Processo de Extração de Informações de Base de Dados
PA	Pressão Arterial
SIM	Sistema de Informações sobre Mortalidade
SGBD	Sistema de Gerenciamento de Banco de Dados
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
WEKA	Waikato Environment for Knowledge Analysis

SUMÁRIO

1 INTRODUÇÃO.....	13
2 FUNDAMENTAÇÃO TEÓRICA.....	15
2.1 INTELIGÊNCIA ARTIFICIAL.....	15
2.1.1 Aprendizagem de Máquina.....	16
2.1.1.1 Aprendizado Supervisionado.....	17
2.1.1.2 Aprendizado Não-Supervisionado.....	17
2.1.2 Aprendizagem de Máquina na Saúde.....	18
2.1.3 Medidas de Desempenho.....	19
2.1.3.1 Matriz de Confusão.....	19
2.1.3.2 Sensibilidade.....	20
2.1.3.3 Especificidade.....	20
2.1.3.4 Acurácia.....	20
2.2 DIAGNÓSTICO E CLASSIFICAÇÃO DA HIPERTENSÃO ARTERIAL.....	20
2.2.1 Medida da Pressão Arterial.....	21
2.2.1.1 Medida Indireta da Pressão Arterial.....	21
2.2.1.2 Medida Domiciliar e Automedida da Pressão Arterial.....	22
2.2.1.3 Medida Ambulatorial da Pressão Arterial (MAPA).....	23
2.2.2 Classificação de Diagnóstico.....	24
2.3 MODELOS DE APRENDIZAGEM DE MÁQUINA.....	26
2.3.1 Redes Neurais Artificiais (RNAs).....	26
2.3.2 Multilayer Perceptron (MLP).....	28
2.3.3 LogitBoost (LB).....	29
2.3.4 Rede Bayesiana (RB).....	31
2.3.5 Árvore de Decisão.....	31
2.3.5.1 Criação da Árvore.....	33
2.3.5.2 Complexidade.....	33
2.3.5.3 Algoritmo J48.....	33
3 METODOLOGIA.....	35
3.1 CLASSIFICAÇÃO DA PESQUISA.....	35
3.2 WEKA.....	35
3.2.1 Exploração de dados Usando Weka.....	36
3.3 COLETA DE DADOS.....	38
3.3.1 Pré-processamento.....	40
3.4 MODELOS DE CLASSIFICAÇÃO.....	40
4 RESULTADOS E DISCUSSÃO.....	41
4.1 EXPERIMENTO 1.....	41
4.2 EXPERIMENTO 2	42
4.3 EXPERIMENTO 3.....	42
4.4 EXPERIMENTO 4.....	43

4.5 DISCUSSÃO.....	44
5 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS.....	47
REFERÊNCIAS BIBLIOGRÁFICAS.....	48

1 INTRODUÇÃO

Segundo o Ministério da Saúde Brasileiro (MS, 2019), a hipertensão arterial ou pressão alta é uma doença crônica caracterizada pelos níveis elevados da pressão sanguínea nas artérias. Esta doença ocorre quando os valores da pressão sanguínea máxima e mínima são iguais ou ultrapassam os 140/90 mmHg - pode ser lido no formato 14/9 (i.e. quatorze por nove).

A pressão alta faz com que o coração tenha que exercer um esforço maior do que o normal para fazer com que o sangue seja distribuído corretamente no corpo. Ela é um dos principais fatores de risco para a ocorrência de acidente vascular cerebral, infarto, aneurisma arterial e insuficiência renal e cardíaca de acordo com dados do MS. Há uma estimativa de que 15 a 20% da população adulta urbana seja acometida por Hipertensão Arterial Sistêmica (HAS), também chamada de pressão alta.

Atualmente não existem causas claras para pressão alta, no entanto, existem vários fatores e condições que podem desempenhar um papel importante em seu desenvolvimento. Nesse contexto pode-se citar: obesidade, falta de atividade física, dieta salgada, estresse, idade, histórico familiar, doença renal crônica entre outros fatores. Uma vez a hipertensão instalada, as consequências e os prejuízos são grandes para a saúde. Por isso, é fundamental o diagnóstico correto bem como o tratamento precoce.

O diagnóstico da hipertensão é complexo pois depende de vários fatores e condições ambientais que podem contribuir para aumentar as leituras de pressão arterial. Na maioria dos centros hospitalares brasileiros não existe uma padronização para o diagnóstico. Sempre que possível, a medida de pressão arterial deverá ser realizada fora do consultório médico para esclarecer o diagnóstico e afastar a possibilidade do efeito do avental branco no processo de verificação. Portanto, a medida da pressão arterial é o elemento-chave para o estabelecimento do diagnóstico da hipertensão arterial (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2019).

Apesar do grande avanço científico e tecnológico no manejo a hipertensão arterial ocorrido nos últimos anos, uma das grandes dificuldades atuais refere-se aos diagnósticos corretos e claros e, em um segundo momento, na prescrição de

tratamento adequado. Mais recentemente diversos trabalhos aplicaram modelos de Aprendizagem de Máquina (AM) com intuito de melhorar o diagnóstico de hipertensão arterial. Todavia, quando a inteligência artificial é empregada em contextos clínicos mais complexos ainda há um caminho mais longo a ser percorrido (AUSTIN et al., 2017).

Assim, o presente trabalho tem como objetivo realizar uma análise comparativa de modelos de Aprendizagem de Máquina aplicados ao diagnóstico da hipertensão arterial. Para fins de experimentação, será utilizado um dataset (i.e. conjunto de dados estruturados) disponível na internet, obtidos numa plataforma de hospedagem para projeto e competições de Data Science. Na seção 3 falaremos em detalhes.

Adicionalmente, no contexto desse trabalho também pretende-se realizar a análise das principais características e fatores que contribuem para um diagnóstico mais preciso da hipertensão arterial. Para isso, serão avaliados dentro de cada modelo os principais hiperparâmetros (i.e. aquelas características que mais contribuem para conclusão do diagnóstico).

Pretende-se que o resultado desse trabalho tenha contribuições significativas na área de computação e medicina. No primeiro caso, iremos analisar em detalhes seis modelos de Aprendizagem de Máquina e avaliar dentro de um contexto específico os aspectos positivos e negativos do emprego de cada um. No segundo caso espera-se que este trabalho forneça um suporte automatizado para profissionais de saúde como mais um aliado para o correto diagnóstico da hipertensão arterial.

De acordo com o contexto, objetivo e questões de pesquisa do presente trabalho, este encontra-se organizado da seguinte forma: no Capítulo 2 são apresentados os principais conceitos que facilitam o entendimento da presente pesquisa; no Capítulo 3 encontra-se a metodologia científica deste estudo. No Capítulo 4 são apresentados os resultados e discussão associados a aplicação dos modelos de Aprendizagem de Máquina no diagnóstico de hipertensão arterial. Capítulo 5 são apresentadas as principais conclusões deste trabalho, bem como os principais desenvolvimentos de pesquisas futuras.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados os principais conceitos para possibilitar o entendimento da presente pesquisa. Serão descritos os conceitos tecnológicos envolvendo área de Inteligência Artificial na seção 2.1 enquanto na seção 2.2 serão mostrados de uma forma resumida, os problemas ligados ao diagnóstico da hipertensão arterial e suas classificações, finalmente, na seção 2.3 serão descritos os modelos de Aprendizagem de Máquina.

2.1 INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial, muitas vezes citada apenas como IA, é, por definição, a área da ciência da computação que surgiu com o intuito de levar para as máquinas a capacidade de aprendizagem do ser humano. Ela faz parte do progresso tecnológico que tem facilitado a sistemas em diferentes setores de atuação detectar problemas e tomarem decisões com base em enorme banco de dados (RUSSEL; NORVIG, 1995).

Encontram-se diversas definições para Inteligência Artificial, algumas filosóficas, outras mais pragmáticas, dependendo muito da área de atuação dos autores das mesmas. Segundo Shannon (1985), a IA atua em dois grandes campos: i. Imitação das capacidades humanas (visão, fala gestos etc.) e ii. Duplicação de resultados estabelecidos pelo homem através de sua capacidade e/ou experiência. Para exemplificar melhor a colocação acima, apresenta-se algumas áreas de pesquisa da IA:

- a. Solucionadores de Problemas: abordagem examinadora que, sem analisar outras as alternativas, indicam um caminho com boa chance de resultado;
- b. Raciocínio Lógico: um processo de estruturação do pensamento de acordo com as normas da lógica que permite chegar a uma determinada conclusão ou resolver um problema;
- c. Processamento de Linguagem Natural: tradução e compreensão de textos;
- d. Robótica e Visão: manipulação de objetos, sequenciamento de tarefas, reconhecimento e padrões;
- e. Programação Automática: gerador automático de código de programação de computadores;

- f. Aprendizagem: o sistema aprende com sua própria experiência, desenvolvendo seu desempenho. Nesta área que se baseia este trabalho;
- g. Sistemas Especialistas: armazenam o conhecimento de uma área específica de atuação, utilizando-o como suporte à tomada de decisão.

2.1.1 Aprendizagem de Máquina

Segundo Muniz (2019), atualmente, pode-se considerar que esta é a “Era da Informação”. Quantidades massivas de dados são produzidas todos os dias, vindos de: transações comerciais, resultados de pesquisas científicas, tráfego na internet, armazenamento de informações como conversas na rede e muitos outros. A quantidade de dados gerados só tende a aumentar, e com o progresso da tecnologia, estes dados podem ser guardados em grandes repositórios de dados estruturados como *Data Warehouse*, Banco de Dados, estruturas de *blockchain*, entre outros meios. Mesmo assim, esta imensa quantidade de dados não possui importância significativa se estes dados não puderem ser analisados e entendidos.

Nesse contexto surge uma área intitulada de Aprendizagem de Máquina (em inglês *Machine Learning - ML*), a qual dispõe de técnicas para programar computadores para aprender com a experiência passada. Para isso, aplica um princípio de inferência denominado indução. Dessa forma, algoritmos de AM aprendem a induzir uma função ou hipótese capaz de resolver um problema a partir de dados que representam observações do problema a ser resolvido (FACELI et al., 2011).

A AM é composta por tarefas que podem ser divididas em duas categorias: preditivas e descritivas. Para as preditivas, a meta é encontrar uma função, a partir dos dados de treinamento, que possa ser utilizada para prever um rótulo ou valor que caracterize um novo exemplo, com base nos valores de seus atributos de entrada. Os algoritmos utilizados nessa tarefa seguem o paradigma de aprendizado supervisionado. Esse termo provém da simulação da presença de um “supervisor externo”, que conhece a saída desejada para cada exemplo. Com isso, o supervisor externo pode avaliar a capacidade da hipótese induzida de prever o valor de saída para novos exemplos (FACELI et al., 2011).

A respeito às tarefas descritivas, os dados são categorizados conforme as suas características, sendo úteis em atividades de mineração de dados.

2.1.1.1 Aprendizado Supervisionado

Algoritmos de Aprendizado Supervisionado são os mais comuns na área de AM, sendo que a maioria dos problemas em que a área de AM, pode ajudar a resolver é supervisionada por natureza (BRINK; RICHARDS, 2014). Simultaneamente com a abordagem de Aprendizado Não-Supervisionado, a grande parte dos algoritmos de AM pertence a estas duas categorias. Encontram-se outras abordagens menos usuais, como Aprendizado Semi-Supervisionado e Classificação de Multi-Classes, que não são o foco desta pesquisa. Portanto, somente as duas primeiras abordagens serão tratadas neste capítulo.

A abordagem de Aprendizado Supervisionado consiste em utilizar uma série de exemplos (chamados de instâncias), já marcados, para induzir um modelo que seja capaz de classificar novas instâncias de forma precisa, com base no aprendizado obtido com o treinamento com os dados de treinamento. É comum que este modelo seja chamado também de classificador. Portanto, é importante que exista um conjunto de dados de treinamento de qualidade, para que o modelo criado possa ser capaz de prever novas instâncias de forma eficiente.

2.1.1.2 Aprendizado Não-Supervisionado

Assim como a abordagem de Aprendizado Supervisionado é considerado sinônimo da Mineração de Dados Preditivo, a abordagem de Aprendizado Não-Supervisionado equivale à técnica de mineração de Dados Descritiva, também considerada como uma abordagem que é utilizada para descoberta de padrões em dados não-categorizados, sendo muito utilizado para atividades de exploração de dados.

O conjunto de dados utilizado na abordagem de Aprendizado Não-Supervisionado não possui classificação, a saída do conjunto de dados de treinamento não possui uma saída esperada para cada uma de suas instâncias. É a abordagem indicada quando o objetivo do sistema não é construir um modelo de predição, e sim um modelo cuja função seja encontrar regularidades nos dados que possam vir a ser úteis.

Independentemente de ser utilizada para atividades mineração de dados (FURNKRANZ; GAMBERGER; LAVRAC, 2012), em conjunto com algoritmos supervisionados para criar conjuntos de dados de teste ou encontrar novos padrões

em determinado conjunto de dados, o Aprendizado Não-Supervisionado também pode ser utilizado para fins preditivos.

Por sua vez, *Clustering* é uma técnica típica de Aprendizado Não-Supervisionado, mas que pode ser utilizada para classificar um conjunto de instâncias em um determinado subgrupo: neste caso, o *cluster* seria uma espécie de classe para a instância.

Como o foco deste trabalho é a respeito dos algoritmos de Aprendizado Supervisionado, mais especificamente os algoritmos de classificação, os algoritmos Não-Supervisionados não serão tratados neste trabalho.

2.1.2 Aprendizagem de Máquina na Saúde

Diversos estudos têm sido desenvolvidos no setor de aprendizagem de máquina aplicados à saúde. A evolução das pesquisas nessa área ocorre devido à necessidade de ferramentas que possam apoiar o diagnóstico médico de forma eficaz e eficiente.

O estudo de Shubhankar (2019), por exemplo, teve como objetivo compreender os determinantes de saúde utilizando Aprendizagem de Máquina para construir modelos preditivos para a ocorrência de doenças (e.g., hipertensão) a partir de uma base de dados disponível na internet. O modelo desenvolvido fez uso dos dados da pesquisa *Cleveland Heart Disease UCI Repository*, na qual foram coletados dados pessoais de saúde, de cerca de 303 indivíduos. Os resultados atingidos indicam que o uso de dados de saúde para a construção de modelos preditivos apresentou o melhor resultado dentre os testes realizados, com uma acurácia de predição de 80.32%. Dentre os dados de saúde, a atividade física e a presença de algumas condições de saúde foram fortes preceptores individuais.

No contexto brasileiro, Oliveira et al. (2017) desenvolveram modelos preditivos de diabetes não diagnosticada a partir de dados de 12.447 adultos entrevistados para o Estudo Longitudinal de Saúde do Adulto (ELSA), utilizando cinco algoritmos de Aprendizagem de Máquina. Os modelos que atingiram os melhores resultados foram Redes Neurais Artificiais 75.24% e Regressão Logística 74.98%. A frequência de diabetes não diagnosticada foi de 11%. Entre os 403 indivíduos do conjunto de dados de teste que tinham diabetes não diagnosticada,

274 foram identificados como casos positivos.

2.1.3 Medidas de Desempenho

Para avaliar o desempenho dos classificadores, se faz necessário indicadores de desempenho. Nesta seção são apresentadas algumas medidas de desempenho mais comumente utilizadas no contexto de doenças como a hipertensão arterial e, especialmente em algoritmos de classificação.

2.1.3.1 Matriz de Confusão

Também conhecida como matriz de erro ou tabela de contingência, a matriz de confusão é uma tabela que permite a visualização do desempenho de um algoritmo de aprendizagem de máquina sob vários aspectos (HOSSIN, 2015). São dispostas às previsões e os valores reais em linhas e colunas e a partir disso é possível extrair alguns valores, tais como:

- Verdadeiro Positivo (VP): no contexto de hipertensão arterial, VP representa a taxa de exemplos na base de dados classificados corretamente como apresentando a doença;
- Falso Positivo (FP): corresponde aos exemplos na base de dados que foram classificados incorretamente como apresentando a doença;
- Verdadeiro Negativo (VN): corresponde aos exemplos no banco de dados que não possuem a doença e foram classificados corretamente como negativo;
- Falso Negativo (FN): corresponde aos exemplos na base de dados que apresentam a doença e foram classificados incorretamente como negativo.

Na Tabela 1 é mostrado como é definida a matriz de confusão pelas classes desejadas e as classes previstas. Lembrando que a classe positiva corresponde à presença da doença e a classe negativo, a ausência dela.

Tabela 1: Matriz de Confusão

---	Positivos previstos	Negativos previstos
Positivos originais	VP	FP
Negativos originais	FN	VN

Fonte: HOSSIN et al., (2015)

2.1.3.2 Sensibilidade

Sensibilidade, também conhecido por *Recall* ou Taxa de Verdadeiros Positivos (VP), corresponde à taxa de acerto na classe positiva (Hossin et al., 2015). A equação aplicada para esta medida é a seguinte:

$$\text{Sensibilidade} = VP / (VP + FN)$$

2.1.3.3 Especificidade

A medida de avaliação Especificidade, ou taxa de verdadeiros Negativos (VN), corresponde à taxa de acerto na classe negativa. Esta medida quantifica a proporção de casos negativos que foram corretamente classificados (Hossin et al., 2015). A equação aplicada para esta medida é a seguinte:

$$\text{Especificidade} = VN / (VN + FP)$$

2.1.3.4 Acurácia (Taxa de acerto)

Esta medida, do inglês *accuracy*, nos problemas de classificação, corresponde à percentagem de instâncias que o algoritmo previu corretamente, ou seja, a taxa de exemplos positivos e negativos corretamente classificados (Hossin et al., 2015). A equação aplicada para esta medida é a seguinte:

$$\text{Acurácia} = (VP + VN) / (VN + FP + VP + FN)$$

2.2 DIAGNÓSTICO E CLASSIFICAÇÃO DA HIPERTENSÃO ARTERIAL

A realização de diagnóstico da Hipertensão Arterial é feita para detectar os níveis elevados e sustentados de pressão arterial (PA) pela medida casual. Os médicos e profissionais de saúde têm o dever de fazer a medição e avaliação de PA dos pacientes em qualquer unidade de saúde ou até domiciliar com as condições apropriadas. Portanto, a medida da pressão arterial é o elemento-chave para o estabelecimento do diagnóstico da hipertensão arterial.

O estudo feito pela Vigitel (2017), mostra que os idosos com mais de 65 anos

são os mais atingidos pela hipertensão arterial. Ao todo, 60,90% dessa população que vive nas capitais brasileiras afirma ter o diagnóstico. Dados preliminares do Sistema de Informações sobre Mortalidade (SIM), do Ministério da Saúde, também mostram que, em 2017, o Brasil registrou 141.878 mortes devido à hipertensão ou a causas atribuíveis a ela. Essa quantidade enumerada demonstra uma realidade preocupante: todos os dias 388,7 pessoas se tornam vítimas fatais da doença, o que significa 16,2 óbitos a cada hora. Grande parte dessas mortes é evitável e 37% dessas mortes são precoces, isso quer dizer que, pessoas com menos de 70 anos de idade (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2017).

De acordo com Vigitel (2018), os novos dados de Sistema de Vigilância de Fatores de Risco de hipertensão coletados em 2018, mostram também que a parcela da sociedade mais atingidas é formada por idosos: 60,9% disseram se hipertensos, assim como 49,5% na faixa etária de 55 a 65 anos. Essa última edição da pesquisa foi realizada por meio telefônico com 52.395 pessoas maiores de 18 anos, entre fevereiro e dezembro do ano anterior. (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2018).

2.2.1 Medida da Pressão Arterial

A medida da pressão arterial, pela sua importância, deve ser estimulada e realizada, em toda avaliação de saúde, por médicos de todas as especialidades e demais profissionais da área de saúde. Vários padrões medição da pressão arterial são estabelecidos na área de saúde que devem ser considerados para fornecer o bom diagnóstico. A seguir diferentes tipos de medição.

2.2.1.1 Medida Indireta da Pressão Arterial

O Esfigmomanômetro é o ideal para medir a pressão arterial (Figura 1). Os aparelhos do tipo aneróide, quando usados, devem ser periodicamente testados e devidamente calibrados. De acordo com Nogueira et al., (1998), dentre os procedimentos dessa medida indireta da pressão arterial, encontram-se:

- a) Explicar o procedimento ao paciente;
- b) Certificar-se de que o paciente:
 - a. Não está com a bexiga cheia;

- b. Não praticou exercícios físicos;
- c. Não ingeriu bebidas alcoólicas, café, alimentos ou fumou até 30 minutos antes da medida;
- c) Deixar o paciente descansar por 10 minutos em ambiente calmo, com temperatura agradável;
- d) Localizar a artéria braquial por palpação;
- e) Colocar o manguito firmemente cerca de dois a três centímetros acima da antecubital;
- f) Fixar os olhos no mesmo nível da coluna de mostrador manômetro;
- g) Tocar o pulso radial, encher o manguito até o desaparecimento do pulso para estimar o nível da pressão sistólica, desinflar rapidamente e aguardar um minuto antes de encher novamente etc.

Figura 1: Equipamento Esfigmomanômetro para medição.



Fonte: Brasil. Dental Acess (2020)

2.2.1.2 Medida Domiciliar e Automedida da Pressão Arterial

De acordo Ministério da Saúde (2012), O departamento de Hipertensão Arterial da Sociedade Brasileira de Cardiologia, afirma que, a hipertensão é uma

doença silenciosa. “Ela é assintomática na maioria dos casos. Estima-se 30% da população brasileira tenha pressão alta”. Por ser incaracterística, o desafio do diagnóstico é ainda maior. A forma de identificar é através da medida da pressão arterial. Por esse motivo, em toda consulta médica a pressão deve ser medida e avaliada. Porém, devido, a grande variabilidade da pressão, em geral, os diagnósticos devem ser confirmados com as medidas fora do consultório .

Nesses casos, os dois métodos de medir a pressão são recomendados. Um deles é Monitorização Ambulatorial da Pressão Arterial (MAPA), em que o paciente fica 24 horas com um equipamento no braço para monitorar a pressão. A cada 15/30 minutos ocorre uma medida. À noite, esse intervalo aumenta para 20/30 minutos. O outro é Monitorização Residencial de Pressão Arterial (MRPA). O paciente mede três vezes pela manhã e três vezes a noite, durante cinco dias. Há uma orientação prévia ao paciente para que o uso do aparelho seja feito de forma certa.

Os dados são registrados automaticamente e, com eles, o médico faz uma média das medidas para determinar o valor da pressão fora do consultório. É um método eficaz em que o paciente participa ativamente do processo ao notar as suas medidas de pressão arterial, o que aumenta seu engajamento no tratamento (V DIRETRIZ, 2006).

Esses procedimentos são úteis para:

- Identificar a hipertensão do avental branco (hipertensão de consultório isolada);
- Avaliar a eficácia da terapêutica anti-hipertensiva;
- Estimular a adesão ao tratamento; e reduzir custos.

Os aparelhos eletrônicos devidamente validados e calibrados são os mais indicados para a medida da pressão arterial domiciliar.

2.2.1.3 Medida Ambulatorial da Pressão Arterial (MAPA)

A Medida Ambulatorial da Pressão Arterial (MAPA) é um método automático de medida indireta e intermitente da pressão arterial durante 24 horas, enquanto o paciente realiza suas atividades rotineiras, inclusive durante o sono. Estudo tem demonstrado que esse método apresenta melhor correlação com risco

cardiovascular do que a medida da pressão arterial de consultório Rev BRAS, (2000).

É importante ressaltar que ainda existe evidência de que esse método deve ser empregado na avaliação rotineira do paciente hipertenso, não substituindo, portanto, a avaliação clínica do paciente e a medida da pressão arterial de consultório (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2018). A Figura 2 mostra o equipamento de Monitorização Ambulatorial da pressão Arterial (MAPA), feita geralmente a cada 20 minutos durante o período de 24 horas.

Figura 2: (MAPA) Monitorização Ambulatorial da Pressão Arterial.



Fonte: Brasil. Instituto de Cardiologia do Lago (2020)

2.2.2 Classificação de Diagnóstico

Os valores da pressão arterial medidos em consultório podem ser maiores, s durante a vigília pela MAPA ou MRPA. Essas diferenças possibilitam a classificação dos pacientes em n diferentes categorias, por exemplo: normotensão, hipertensão, hipertensão do avental branco (hipertensão isolada de consultório) e hipertensão mascarada (normotensão do avental branco). A normotensão se caracteriza por valores normais de pressão arterial no consultório (abaixo de 140/90 mmHg) e na MAPA de 24 horas (igual ou abaixo de 130/80 mmHg) ou na MRPA (igual ou abaixo

de 135/85 mmHg), enquanto que a hipertensão se caracteriza por valores anormais de pressão arterial no consultório (igual ou acima de 140/90 mmHg) e na MAPA de 24 horas (acima de 130/80 mmHg) ou na MRPA (acima de 135/85 mmHg). (CARMO, 2019).

O número de pressão aferido é facultativo e a classificação depende da sua medição. A necessidade de sistematização obriga uma definição operacional para separar indivíduos curados dos doentes. Na realidade, pode-se ter maior ou menor risco cardiovascular tanto acima como abaixo do número limítrofe, quando o paciente é considerado individualmente. Pelo exposto, enfatiza-se a necessidade de extrema cautela antes de rotular alguém como hipertenso, tanto pelo risco de falso-positivo como pela repercussão na própria saúde dos indivíduos adultos (com mais de 18 anos de idade) cifras inferiores a 85 mmHg de pressão diastólica e inferiores a 130 mmHg de pressão sistólica (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2018).

Na Tabela 2 é apresentada a classificação diagnóstica da hipertensão arterial (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2019). A inclusão do grupo com cifras tensionais normal limítrofe de 130 – 139 mmHg/85 – 89 mmHg deve-se ao fato de que esses indivíduos se beneficiarão com as medidas preventivas.

Tabela 2: Classificação diagnóstica da hipertensão arterial.

Classificação diagnóstica da hipertensão arterial (> 18 anos de idade)		
PAD (mmHg)	PAS (mmHg)	CLASSIFICAÇÃO
<85	<130	Normal
85-89	130-139	Normal limítrofe
90-99	140-159	Hipertensão leve (estágio 1)
100-109	160-179	Hipertensão moderada (estágio 2)
≥110	≥180	Hipertensão grave (estágio 3)
<90	≥140	Hipertensão sistólica isolada

Fonte: Ver Bras Hipertens (2010)

Os valores da pressão arterial medidos em consultório podem ser maiores, semelhantes ou menores do que os obtidos durante a vigília pela MAPA ou MRPA. Essas diferenças possibilitam a classificação dos pacientes em categorias diferentes como representados na Tabela 2, normal, normal limítrofe, hipertensão leve (estágio 1), hipertensão moderada (estágio 2), hipertensão grave (estágio 3) e por fim a hipertensão sistólica isolada. Define-se efeito normal com o valor obtido abaixo do percentil 90, normotensão; entre os percentis 90 e 95 normal limítrofe; acima do percentil 95 considerado hipertensão arterial, assim esses pacientes merecem especial atenção pela necessidade de investigação de causas secundárias e tratamento efetivo, já que possuem habitualmente maior risco cardiovascular.

2.3 MODELOS DE APRENDIZAGEM DE MÁQUINA

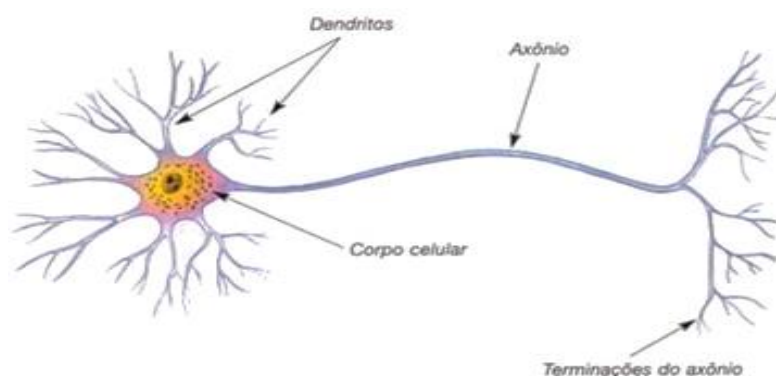
Diversos modelos são encontrados na literatura para aplicação em AM, entre os quais, neste trabalho são destacados: Redes Neurais Artificiais, Multilayer Perceptron, LogitBoost (LB), Redes Bayesianas (BN) e Árvore de Decisão (AD).

2.3.1 Redes Neurais Artificiais (RNAs)

As Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo inspirado na estrutura neural do cérebro humano, e procuram adquirir seu conhecimento através da experiência (HYAKIN, 1998).

Uma RNA é composta por um conjunto de neurônios conectados entre si, sendo que para cada ligação entre neurônios, são associados pesos que representam a influência que aquela informação possui dentro da RNA. Através deste relacionamento, a saída de um neurônio compõe a entrada de outro, como demonstrado na Figura 3, a qual é um exemplo de uma rede neural biológica. Este padrão de repetição resulta no comportamento inteligente para detecção de reconhecimento de padrões e características (CERA, 2005).

Figura 3: Redes Neurais Biológicas.



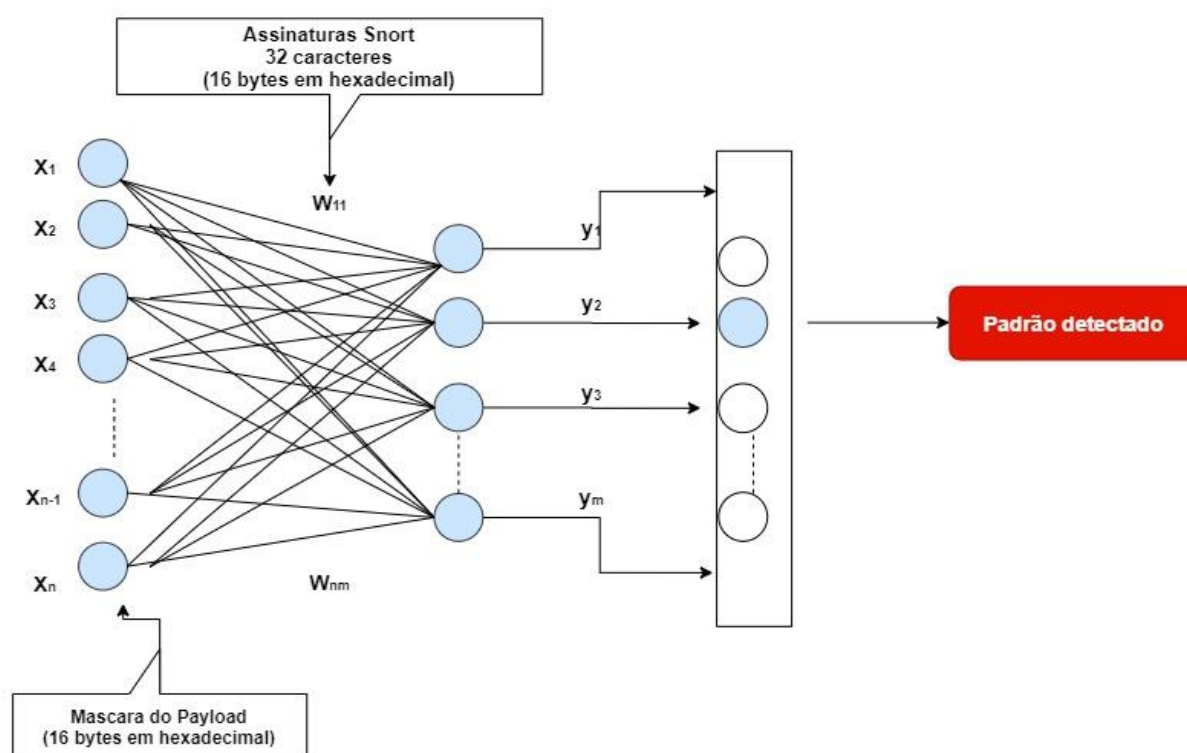
Fonte: Brasil. Malcon Anderson Tafner (1998)

O treinamento de uma RNA através de uma aprendizagem supervisionada consiste em adicionar na camada de entrada às características de um exemplo previamente conhecido, e na camada de saída o resultado esperado da classificação. Esse processo tem início com a escolha dos pesos entre as conexões. Normalmente, os valores originais destes pesos são indicados de maneira aleatória. Para que os neurônios das camadas intermediárias recebam as informações da camada de entrada, é realizada a união de todos os valores das conexões ligadas ao neurônio com seus respectivos pesos, resultando em uma informação conhecida como *Potencial de Ativação*. Este valor representa o conhecimento que aquele neurônio possui até o momento. Porém, para que essa informação seja repassada as próximas camadas, o neurônio no qual recebeu esse conhecimento precisa definir se esse conhecimento é considerado válido ou não. Com isso, esse potencial de ativação é submetido à análise dessa informação, através de uma Função de ativação. (HAYKIN, 2001)

Essa função tem como objetivo limitar o valor para algo de fácil representação, geralmente valores inteiros ou binários. O sinal é então transmitido para as próximas camadas, até atingir a camada de saída da rede. Neste ponto, é verificado se o sinal recebido condiz com a classificação esperada para aquela amostra. Nas situações em que a classificação está incorreta, a diferença entre o valor esperado e o encontrado é retornado as camadas anteriores. Essa diferença será utilizada para que os pesos das conexões possam ser corrigidos. (EYNG, 2008).

Depois o treinamento ser concluído, os valores obtidos para os pesos entre cada neurônio armazenam o conhecimento da rede para identificação das amostras. Através destes pesos, uma amostra que não foi utilizada durante a fase de treinamento poderá ser submetida como entrada à RNA, sendo retornada a classificação desta. (ALVES SOBRINHO et al., 2011). Na Figura 4 é possível observar a arquitetura de uma Rede Neural. A partir da ilustração, verifica-se que a rede recebe as n entradas $\{x_1, x_2, x_3, x_4, x_{(n-1)}, x_{(n)}\}$, as quais representando o comportamento do processo, e prediz então como resposta o respectivo valor esperado para *padrão detectada* fornecido pelo seu neurônio de saída y_2 . Assim, durante o processo de treinamento, a rede tentará ajustar as suas matrizes de pesos visando minimizar o erro y_{nm} produzido pela diferença entre W_{11} e W_{nm} .

Figura 4: Arquitetura da Rede Neural.



Fonte: Bras. Researchgate (2020)

2.3.2 Multilayer Perceptron (MLP)

A rede Perceptron de Multicamadas (do inglês *Multilayer Perceptron* - MLP) é uma generalização do Perceptron, constituída de um conjunto de nós fonte, os quais formam a camada de entrada da rede, uma ou mais camadas ocultas de nós

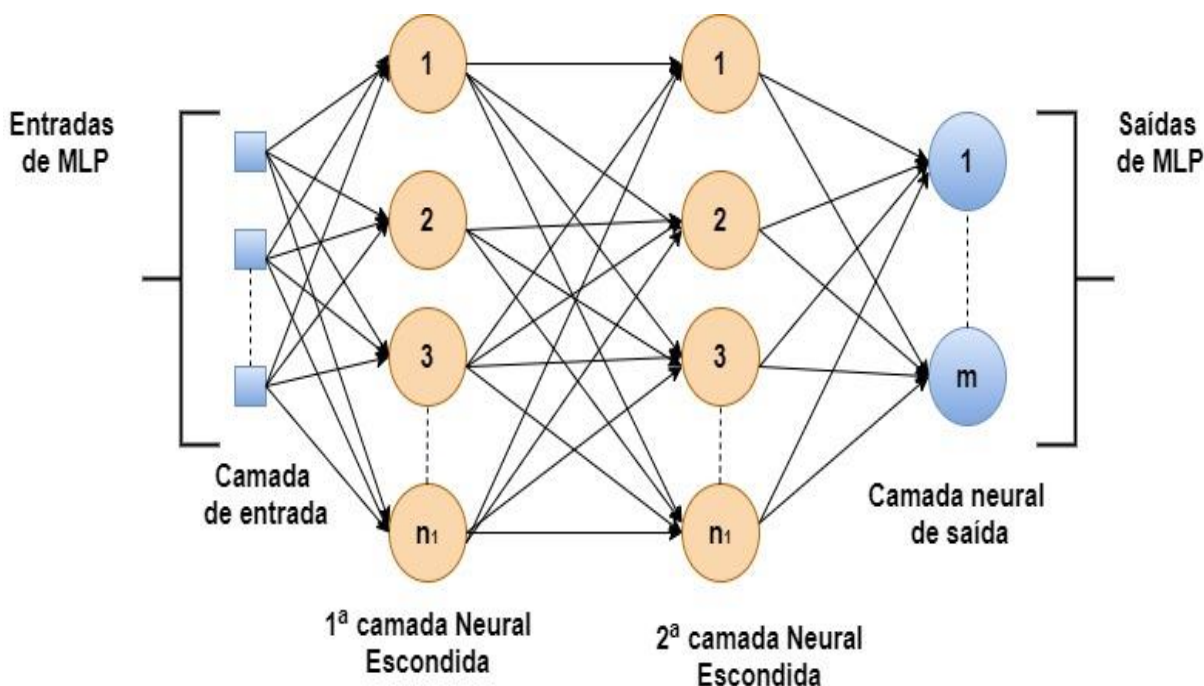
computacionais e uma camada de saída, também de nós computacionais (HAYKIN, 2001). Com exceção da camada de entrada, todas as outras camadas realizam processamento.

As redes MLP são ainda caracterizadas pelas elevadas possibilidades de aplicações em diversos tipos de problemas relacionados com as mais diferentes áreas do conhecimento, sendo também consideradas uma das arquiteturas mais versáteis quanto à aplicabilidade. Entre essas potenciais áreas, têm-se os seguintes destaques:

- Aproximação universal de funções;
- Reconhecimento de padrões;
- Identificação e controle de processos;
- Previsão de séries temporais;
- Otimização de sistemas.

De acordo com Nunes et al., (2016), as redes MLP pertence à arquitetura *feedforward* de camadas múltiplas, cujo treinamento é efetivado de forma supervisionada. Conforme observado na Figura 5, o fluxo de informações na estrutura da rede se inicia na camada de entrada, percorre em seguida as camadas intermediárias, sendo então finalizado na camada neural de saída. Observa-se ainda que na rede MLP convencional inexistente qualquer tipo de realimentação de valores produzidos pela camada neural de saídas ou pelas próprias camadas neurais intermediárias.

Figura 5: Estrutura de Rede Multilayer Perceptron.



Fonte: NUNES (2016)

2.3.3 LogitBoost (LB)

O método conhecido como *boosting* nasceu na área de computação de maneira geral, especificamente na comunidade conhecida pelo nome de Aprendizagem de Máquina. Dentro dessa comunidade, foi proposto um problema teórico chamado de problema de *boosting*, que pode ser informalmente exposto da seguinte maneira: “suponha que existe um método de classificação que é ligeiramente melhor do que uma escolha aleatória, para qualquer distribuição em X . esse método é chamado de *weak learner*, ou classificador fraco. A existência de um classificador fraco implica na existência de um classificador forte (*Strong learner*), com erro pequeno sobre todo o espaço X ”. Em Estatística, isso equivale a perguntar se, dado um método razoável de estimação, é possível obter um método próximo de ótimo.

Mais recentemente, notou-se que, se um algoritmo de *boosting* for executado por um tempo (número de iterações) muito grande, da ordem de dezenas de milhares, isso ocasionará *overfitting*. (Friendman, Hastie e Tibshirani, 2000), dão um exemplo em que isso ocorre. Algumas abordagens para este problema foram tentadas, Jiang (2000) mostrou que, sob certas condições de regularidade, o

algoritmo AdaBoost é consistente em processo (process consistent), no sentido de que, durante o treinamento, ele gera uma sequência de classificadores com erro que converge para erro do classificador (regra) de Bayes. (LUGOSI; VAYATIS, 2004) mostraram um resultado importante de consistência para algoritmos de *boosting* com regularização.

O método *boosting*, possui vários algoritmos, dentre eles encontram-se: Algoritmo *AdaBoost Real*, Algoritmo *AdaBoost Discreto*, e por último Algoritmo *LogitBoost*.

2.3.4 Rede Bayesiana (RB)

As Redes Bayesianas, também conhecidas como Redes Casuais, Rede de crença ou Gráficos de dependência probabilística, surgiram na década de 80 e têm sido aplicadas em uma grande variedade de atividades do mundo real (Bobbio et al, 2001), na saúde (ABICALAFF; AMARAL; DIAS, 2004) (KORD; NICHOLSON, 2003), e desenvolvimento de jogos (FILHO; ALBUQUERQUE, 2007), entre outras.

Segundo Neapolitan (2004), a técnica de Redes Bayesianas surgiu no contexto no qual há um grande número de variáveis e objetivo de verificar qual a influência probabilística não direta de uma variável para as demais. Assim, a teoria de Redes Bayesianas combina com princípios de Teoria de Grafos, teoria de probabilidades, Ciência da Computação e Estatística (BEM-GAL, 2007).

Desta forma, na literatura especializada, uma terminologia específica é utilizada para definir tipos de variáveis, dependências probabilísticas e outras propriedades das Redes Bayesianas.

2.3.5 Árvore de Decisão

A árvore de decisão é montada a partir de um conjunto de dados de treinamento, a princípio tem-se apenas um nó que contém todas as classes. Recursivamente, escolhe-se um atributo que possa dividir esta classe, até que não haja mais divisões e cada nó folha represente uma única classe ou satisfação de um critério (GARCIA, 2000).

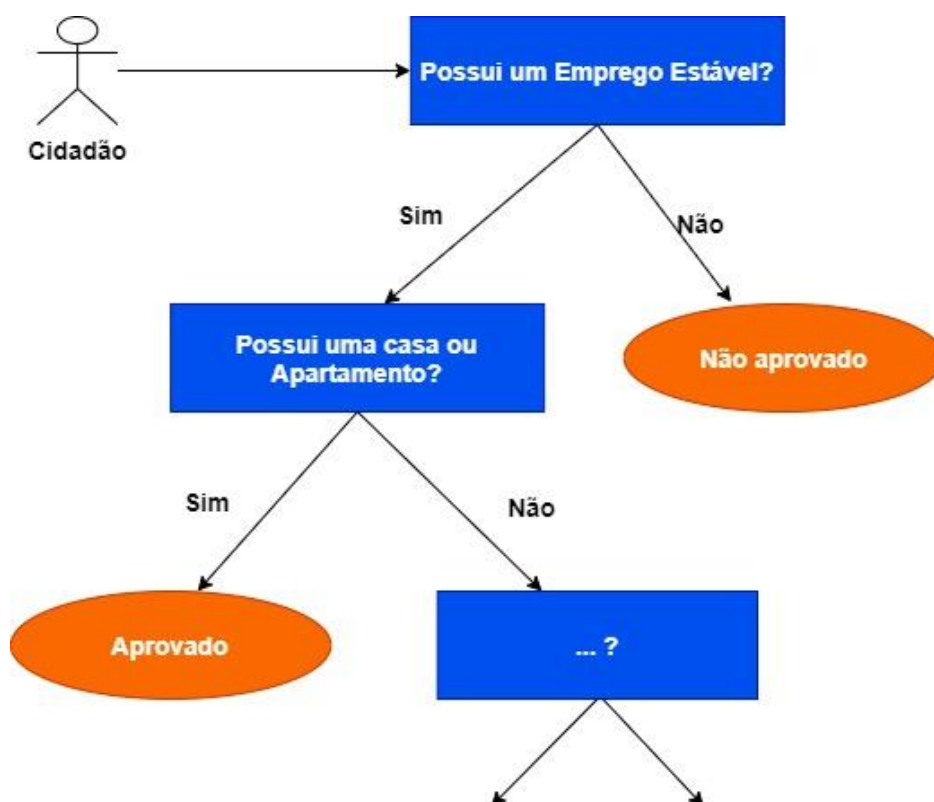
Na árvore de decisão, o primeiro nó é chamado raiz e por conversão fica no topo da árvore. A partir do nó raiz, sucessivas ramificações (também chamadas de links ou ramos) são ligadas a outros nós, chamados de nós internos. As ligações ocorrem até atingir um nó terminal, ou folha.

Para uma boa tomada de decisão utilizando a árvore, o dispositivo é percorrido da raiz para as folhas, testando os valores dos atributos em nós sucessivos e, quando uma folha é alcançada a decisão é tomada (WITTEN; EIBE, 2005).

Na Figura 6 é mostrada uma árvore de decisão cujo objetivo é a validação de uma solicitação de conta bancária, o indivíduo é classificado numa determinada classe. A instância 1 verifica se o candidato possui um emprego estável, caso não

possua, a abertura da conta não será aprovada, no caso afirmativo, ele será submetido a outro teste. A instância 2 consiste em verificar se o candidato possui uma casa ou apartamento, caso seja válido ele recebe um conta, no caso contrário, o candidato passará por outro teste até alcançar o nível mais baixo. Ao verificar o resultado dos dois testes, a árvore rotula a instância de acordo com a classe presente no nó-folha.

Figura 6: Exemplo de uma árvore de decisão.



Fonte: O Autor (2020)

Árvore de Decisão é, provavelmente, o algoritmo de AM mais estudado para aplicações em Mineração de Dados (WITTEN; FRANK, 2000). Os motivos são variados: é um modelo que possui suporte a diversos tipos de características (categóricas e numéricas), sua representação de conhecimento adquirido é facilmente compreendida, e o seu processo de aprendizado e treinamento é relativamente rápido, comparado a outros algoritmos como Redes Neurais Artificiais (ver seção 2.3.1).

2.3.5.1 Criação da Árvore

O motivo crítico na criação de uma Árvore de Decisão é a escolha da característica que irá ser utilizada como nó na árvore. O critério de escolha da característica é definido de uma maneira que cada divisão realizada na árvore seja feita da forma mais “pura” possível, e que cada número máximo possível de instâncias em cada subconjunto pertença a uma única classe (DUDA et al., 2000).

É necessário um critério de seleção para a escolha de uma característica que se tornará um nó na árvore. Geralmente, o critério de seleção de um nó é realizado por meio do cálculo do Ganho de Informação, ou Entropia das características de cada instância presente no subconjunto (RUSSEL; NORVIG, 2002).

Após a construção da árvore, é necessário “podá-la”. Este passo é importante ao se utilizar árvores de decisão, pois este algoritmo é muito suscetível ao problema de *overfitting*. Basicamente, ruído nos dados de treinamento pode levar a construção de ramos e nós que atrapalham a generalização de novas instâncias. Além disso, o modelo construído pode ser muito específico (como, por exemplo, possuir uma folha para cada instância do conjunto de dados de treinamento): este modelo mostra uma precisão enorme nos dados de treinamento, mas irá se mostrar ineficiente ao ser alimentado com instâncias do conjunto de dados de teste (BREIMAN, 1984).

2.3.5.2 Complexidade

A complexidade computacional do algoritmo de Árvore de Decisão é $O(nT\log(T))$ onde n é o número de características de cada instância e T é o número de instâncias que compõem o conjunto de dados de treinamento (HAN; KAMBER, 2006). A árvore de decisão pode crescer no máximo até $nT\log(T)$

2.3.5.3 Algoritmo J48

De acordo com Lokanatha (2010) , o J48 é um algoritmo que pode lidar tanto com atributos contínuos e discretos, quanto com valores categóricos e valores ausentes. O procedimento de atributos contínuos envolve a consideração de todos os valores presentes no conjunto de treinamento, fazendo com que sejam

ordenados de forma crescente considerando todos os valores presentes nos dados de treinamento e, após esta ordenação, seja selecionado o valor que favorecerá a redução da entropia (CAMARGO et al., 2011); (RAMYA et al., 2015).

De acordo com Witten et al. (2005), o algoritmo J48 surgiu da necessidade de recodificar o algoritmo C4.5, que, originalmente, é escrito na linguagem Java. Ele tem como intuito de produzir uma árvore de decisão baseada em uma coleção de dados de treinamento, sendo este modelo usado para classificar as instâncias no conjunto de teste. O algoritmo J48, apresentado por Quinlan (1993), é considerado o que manifesta o melhor resultado na montagem de árvores de decisão, a partir de um conjunto de dados de treinamento.

3. METODOLOGIA

Neste capítulo encontra-se descrita a metodologia utilizada para desenvolver este trabalho. Assim como, a classificação da pesquisa quanto à natureza, objetivos e procedimentos. O público alvo a quem se destina esse trabalho. Uma descrição da ferramenta Weka utilizada para realizar o processo do KDD neste trabalho. Além da descrição da base de dados utilizada.

3.1 CLASSIFICAÇÃO DA PESQUISA

Conforme à natureza, este trabalho classifica-se como uma pesquisa aplicada, uma vez que são aplicados modelos de aprendizagem de máquina a um conjunto de dados para a previsão de diagnóstico médico. Esse conhecimento é aplicado numa realidade bem específica, diagnóstico de doença cardíaca.

Em relação aos objetivos, o presente trabalho visa melhorar o atendimento dos pacientes na realização de diagnóstico de hipertensão arterial sem erro, fazer exames mais precisos e prevenir doença cardíaca e também garantir a acurácia dos valores obtidos.

Quanto aos procedimentos técnicos trata-se de um estudo experimental envolvendo a aplicação dos algoritmos classificadores de aprendizagem de máquina para construção de modelos preditivo de doença cardíaca, hipertensão arterial realizado com os dados (dataset) disponibilizados gratuitamente, utilizando o software *Weka* (*Waikato Environment for Knowledge Analysis*).

3.2 WEKA

Na exploração dos dados não é apenas o grande domínio do conhecimento sobre as variáveis, atributos e outros que são relevantes. Na verdade, é preciso ter um instrumento que possa auxiliar no gerenciamento destes. Existem na literatura ferramentas a realização de mineração de dados, como por exemplo, *IBM Intelligent Miner*, *DBMiner*, *MinerSet* e *Weka*. Optou-se pela utilização da ferramenta *Weka* devido sua simplicidade e interface gráfica que proporciona facilidade de uso (Figura 7).

A Weka é o produto da Universidade de Waikato (Nova Zelândia) e foi implementado pela primeira vez em sua forma moderna em 1997. Ele usa a GNU General Public License (GPL), tem código aberto, e possui vários algoritmos já implementados para realização de mineração de dados (HALL et al., 2009). O principal motivo para a escolha deste software é o fato dele ser implementado em Java, além dele prover boas implementações dos algoritmos de aprendizado.

Neste trabalho, são utilizados os algoritmos de classificação J48, Multilayer Perceptron, Naive Bayes e LogitBoost disponíveis no Weka.

Figura 7: Weka GUI Chooser.



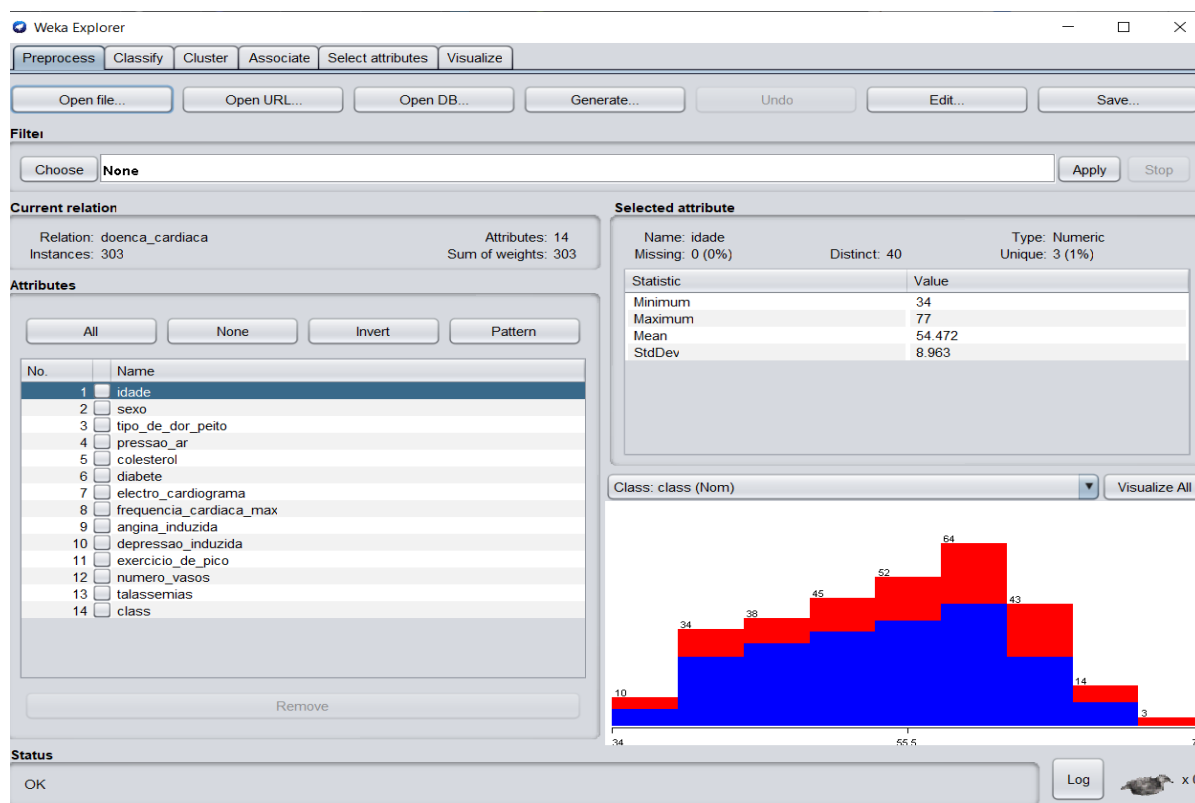
Fonte: Brasil. Devmedia (2020)

3.2.1 Exploração de dados Usando Weka

O software Weka possui duas formas pelas quais acessa os arquivos de dados, a criação de arquivos no formato ARFF e o formato CSV, seguinte, carregar o conjunto de dados em Weka que terá uma conexão direta ao SGBD (Sistema Gerenciamento de Banco de Dados). Na criação de dados, extrair os resultados para um arquivo TXT ou CSV e utilizar estas ações toda vez que um atributo sofre alguma alteração no banco de dados. Já a conexão direta ao SGBD permite ao Weka ter um acesso confiável e direto ao banco de dados, mas para isso é

necessário realizar algumas configurações. Na Figura 8 é mostrado um exemplo de carregamento de dados na ferramenta Weka.

Figura 8: Dados carregados no Weka.



Fonte: O Autor (2020)

3.3 COLETA DE DADOS

Considerando à integridade e consolidação dos dados, optou-se pela coleta de dados no repositório de *Machine Learning* (*UCI Machine Learning Repository*) *Cleveland Clinic* do Estado de Ohio, EUA, o qual encontra-se disponível em: <https://archive.ics.uci.edu/ml/index.php> para a realização de testes e consolidação dos resultados.

O dataset usado neste trabalho é o conjunto de dados da *Cleveland Heart Disease*, UCI Repository. Na Tabela 3 tem-se uma amostra dos seus registros e colunas. No total, a base conta com 15 colunas e 303 registros de pacientes.

Tabela 3: CLEVELAND UCI Repository.

(continua)

idade	sexo	tipo de dor peito	pressão ar	colesterol	diabete	electro cardio	frequência cardíaca máx.
63	1	1	1	233	1	2	150
67	1	4	120	229	0	2	129
67	1	4	160	286	0	108	1
37	1	3	130	250	0	0	187
41	0	2	130	204	0	2	1

Fonte: Cleveland Heart Disease (2019)

Na Tabela 4 os atributos da base de dados *Cleveland Heart* são detalhados.

Tabela 4: Definição dos atributos da Cleveland Heart Disease

Atributo	Descrição	Valores
Idade	Indica a idade do paciente no momento do atendimento.	de 29 à 70 anos
Sexo	Indica o sexo do paciente.	0 - Masculino 1- Feminino
Tipo de dor no peito	Indica o nível de dor que o paciente sente no peito.	1 - Angina típica 2 - Angina atípica 3 - Dor não anginosa

		4 - Assintótico
Pressão de Arterial	Indica a medida da pressão exercida pelo sangue na parede das artérias, expressando a condição do sistema circulatório do paciente.	< 120/80 - Ótima < 130/85 - Normal 130-139/85-89 limítrofe 140-159/90-99 Hipertensão 1 160-179/100-109 Hipertensão 2 >180/>110 Hipertensão 3 >140/<90 Hipertensão sistólica isolada
Colesterol	Indica o nível de colesterol no sangue do paciente na hora da consulta.	> 45 Bom (HDL) > 110 Ruim (LDL)
Diabetes	Indica a presença ou não da condição que resulta em alta o nível de açúcar no sangue do paciente na hora da consulta.	0 - Falso 1 - Verdadeiro
Eletrocardiograma Repouso	Indica a atividade elétrica do coração quando o paciente está em repouso.	0 - Normal 2 - Média ou Alta
Frequência Cardíaca Máximo	Indica o número de batimentos cardíacos por unidade de tempo.	100 e 170 - Normal > 170 - Alta
Angina Induzida	Indica a presença ou não de dor no peito do paciente na hora da consulta.	0 - Negativo 1- Positivo
Depressão induzida	Indica a presença de depressão induzida por medicamentos do paciente.	0 - Negativo > 0 = Positivo
Exercício De Pico	Indica se o paciente faz exercícios com alta intensidade.	1 - Ascendente 2 - Plana 3 - Descendente
Teste Ergométrico	Indica o funcionamento cardiovascular quando submetido a esforço físico gradualmente.	>= 0 - Bom <= 3 - Ruim
Talassemia	Indica a presença ou não da condição de baixa quantidade de hemácias no sangue.	3 - Normal 6 - Defeito fixo 7 - Defeito reversível
Estado angiográfico da doença	Indica o estado das artérias do paciente, medido em percentagem.	0 - Negativo 1 - Positivo

Diagnóstico de doença cardíaca	Indica se o paciente apresenta doença cardíaca	0 - Negativo 1 - Positivo
-----------------------------------	---	------------------------------

Fonte: Cleveland Heart Disease (2019)

3.3.1 Pré-processamento

Após a coleta dos dados, eles foram pré-processados, isto é, submetidos à tarefa de limpeza dos dados, seleção de variáveis e preenchimento de valores ausentes para melhorar a qualidade dos dados para extração de padrões.

Na base de dados original tinha 76 atributos, mas para a realização dos experimentos alguns deles foram excluídos devido à baixa ou à ausência total de utilidade para a descoberta de entendimento. Esses atributos incluem: o código identificador do indivíduo e o ano da realização da consulta, por exemplo. Isto resultou na seleção de um subconjunto de 15 atributos mais relevantes para a realização da classificação. Em relação aos dados faltantes, preencheu-se os com os valores que mais ocorrem (moda) nas respectivas colunas.

3.4 MODELOS DE CLASSIFICAÇÃO

Uma vez que o conjunto de dados foi tratado e construído para a experimentação, definiu-se os classificadores a fim de prever com precisão as classes de novos exemplos. Algoritmos de diferentes paradigmas de aprendizagem foram aplicados. Primeiramente, usou-se a técnica de classificação Rede Bayesiana, logo após a técnica LogitBoost, o Multilayer Perceptron e, por fim, o classificador J48. Os experimentos realizados, simulam a identificação de pacientes hipertensos com base em seus dados coletados na hora de diagnóstico.

É importante mencionar que como todos os métodos são probabilísticos, definiu-se 5 *seeds* que determinam a execução de cada método sob uma semente numérica. Logo, tem-se 5 resultados por método, o qual considera o valor médio dos resultados. Além disso, foi usada a abordagem de validação cruzada (em inglês, *cross validation* - cv). Isto significa que, para cada *seed*, o método realizará 5 execuções, utilizando diferentes porções de base (4/5 e 1/5), onde a primeira porção é passada para o processo interno e (10 fold cv) para o treinamento e

validação, a segunda porção será reservada para testar a melhor solução. No final, o método retornará o valor médio de todo processo.

4. RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados e discutidos os resultados dos experimentos realizados neste trabalho.

4.1 Experimento 1

Para realização do primeiro experimento, utilizou-se o algoritmo de Rede Bayesiana, considerando a configuração padrão do Weka. Como resultado deste experimento tem-se 79.20% das instâncias (positivos e negativos) classificadas corretamente e 20.79% das instâncias (positivos e negativos) classificadas incorretamente, isto é, pacientes hipertensos e pacientes não hipertensos.

Após esse experimento, alterou-se apenas a configuração *useKernalEstimator* para *true* de modo a observar o comportamento do algoritmo e buscar melhor acurácia. Optou-se pela alteração deste parâmetro para *true* pois o mesmo é indicado para atributos numéricos, o que é o caso da base de dados utilizada.

Observou-se o seguinte resultado: 82.17% instâncias foram classificadas corretamente. Um aumento de cerca de três%, resultado superior ao de anterior que foi realizado com a configuração *default* do Weka. O modelo conseguiu o seu desempenho melhor com a taxa de acerto e obteve-se 17.82% a taxa de erro, o algoritmo classificou instância incorretamente. A Tabela 5 apresenta o resultado do melhor cenário deste experimento.

Tabela 5: Acurácia do algoritmo Rede Bayesiana Experimento 1.

Modelo	Métricas		
Rede Bayesiana	Sensibilidade (%)	Especificidade (%)	Acurácia (%)
	79.11	85.51	82.17

Fonte: O Autor (2020)

4.2 EXPERIMENTO 2

Para realização do segundo experimento, usou-se o modelo LogitBoost (LB) com as configurações padrões do Weka, os resultados obtidos foram: 76.23% das instâncias classificadas corretamente e 23.76% das instâncias classificadas incorretamente.

Na busca por melhores resultados alterou-se o parâmetro *useResampling* para *True*. Com esta alteração, foi obtido um resultado inferior comparado ao teste considerando a configuração padrão do algoritmo. Para este cenário, 75.9% das instâncias foram classificadas corretamente, isso gerou uma diferença de 0.33%. Quando o atributo *useResampling* é setado para *True*, ele faz que o algoritmo use o reamostragem que a ponderação na hora do treinamento. E quando o atributo é *False*, usa a ponderação, a cada iteração do processo de treinamento, um peso é atribuído a cada amostra no conjunto de treinamento. Isso implicou o desempenho melhor da configuração padrão do modelo. Assim, o primeiro resultado do modelo com a configuração padrão do Weka que alcançou a melhor taxa. A Tabela 6 apresenta as taxas obtidas neste experimento.

Tabela 6: Acurácia do algoritmo LogitBoost Experimento 2.

Modelo	Métricas		
LogitBoost	Sensibilidade (%)	Especificidade (%)	Acurácia (%)
	77.92	74.49	76.23

Fonte: O Autor (2020)

4.3 EXPERIMENTO 3

Para realização do terceiro experimento, usou-se o modelo Multilayer Perceptron (MLP) com a configuração padrão de Weka. O modelo alcançou uma taxa de 73.92% das instâncias classificadas corretamente e 26.07% das instâncias classificadas incorretamente.

Seguidamente, alterou-se as configurações do modelo com objetivo de melhorar a eficácia do modelo, os parâmetros alterados são, *decay* para o valor *True*, esse atributo fará com que a taxa de aprendizado diminua. Isso dividirá a taxa de aprendizado inicial pelo número da época para determinar qual deve ser a taxa de aprendizado atual, Isso pode ajudar a impedir que a rede se desvie da saída de destino, além de melhorar o desempenho geral; o *hiddenlayers* (quantidade de neurônios) subir para 2, o valor default é “a” (número 0); e o *seed*, que é a semente usada para inicializar o gerador de números aleatórios, são usados para definir os pesos iniciais das conexões entre os nós e também para misturar dados de treinamento trocado para 1.

Após a alteração, o resultado obtido consiste em 77.88% das instâncias classificadas corretamente e 22.11% das instâncias classificadas incorretamente. A Tabela 7 apresenta as medidas de desempenho do algoritmo obtidas neste experimento.

Tabela 7: Acurácia do algoritmo Multilayer Perceptron Experimento 3.

Modelo	Métricas		
Multilayer Perceptron	Sensibilidade (%)	Especificidade (%)	Acurácia(%)
	74.1	82.1	77.88

Fonte: O Autor (2020)

4.4 EXPERIMENTO 4

Para realização do quarto experimento, utilizou-se o algoritmo J48 (C4.5), baseado em árvore de decisão, com as configurações padrão do Weka. O resultado alcançado é de 70.62% instâncias classificadas corretamente e 29.37% instâncias classificadas incorretamente.

Em seguida, decidiu-se alterar o atributo *reducedErrorPrunin* na configuração do modelo com intuito de atingir a melhor performance do modelo, esse atributo foi

trocado para *True*, isto é, ele visa reduzir os erros que a árvore pode gerar durante o treinamento por meio da “poda” da árvore.

O resultado obtido neste experimento é de 72.93% das instâncias classificadas corretamente e 27.06% das instâncias classificadas incorretamente. A Tabela 8 apresenta detalhes da acurácia obtida neste experimento.

Tabela 8: Acurácia do algoritmo J48 Experimento 4.

Modelo	Métricas		
J48	Sensibilidade (%)	Especificidade (%)	Acurácia (%)
	79.74	65.55	72.93

Fonte: O Autor (2020)

4.5 DISCUSSÃO

A Tabela 9 apresenta os resultados da classificação. A principal métrica utilizada para otimização de cada modelo durante o período de aprendizado foi *Acurácia*, sem esquecer da *Sensibilidade e Especificidade*. Finalmente, a avaliação de performance dos modelos selecionados demonstrou desempenho satisfatório, com *Acurácia* superior a 70% para todos modelos. É possível observar que as acurácias geradas dos classificadores na maior parte dos experimentos encontram-se na faixa de 72 a 82%. E os melhores resultados foram encontrados para os modelos Rede Bayesiana (82% de acerto) e Multilayer Perceptron (77% de acerto).

Dentre os métodos utilizados, destaca-se o aprendizado Bayesiano, que utiliza um modelo probabilístico baseado no conhecimento prévio do problema, o qual é combinado com exemplos de treinamento para determinar a probabilidade final de uma hipótese (MITCHELL, 1997). Assim, a Rede Bayesiana é uma abordagem interpretativa e analítica para o raciocínio probabilista e tem sido utilizada recentemente em diversas áreas como, por exemplo, estimação de risco operacional, diagnóstico médico, *credit scoring*, projeto de jogos computacionais, imputação de dados, entre outras.

Também, deve-se levar em consideração outra análise que é em relação às taxas de sensibilidade e especificidade. Um modelo que apresenta uma taxa de sensibilidade que se aproxima de 100%, se caracteriza como sendo um bom modelo para classificar os casos positivos corretamente (com hipertensão arterial). Neste trabalho, o J48 apresentou a maior taxa de sensibilidade comparado aos demais métodos utilizados, cerca de 79,74%. Por outro lado, sua taxa de especificidade foi a mais baixa, 65,55%. Isso significa que a taxa de falso positivo foi elevada, ou seja, o modelo pode classificar um paciente como hipertenso sem de fato ele seja hipertenso, causando desconforto para o paciente. Já a Rede Bayesiana e o MLP obtiveram as maiores taxas de especificidade, ambas acima de 80%.

Tabela 9: Resultados de classificação final dos quatros modelos.

Modelos	Sensibilidade (%)	Especificidade (%)	Acurácia (%)
Rede Bayesiana	79.11	85.51	82.17
LogitBoost	77.92	74.49	76.23
Multilayer Perceptron	74.1	82.1	77.88
J48	79.74	65.55	72.93

Fonte: O Autor (2020)

A Tabela 10 apresenta a comparação dos resultados obtidos neste trabalho e os trabalhos relacionados. Observa-se que o trabalho de Shubhankar (2019) citado na seção 2.1.2, utilizou a mesma base de dados *Cleveland Heart Disease*, porém utilizou-se *python* como linguagem de programação. Ele obteve com o modelo Logistic Regression a taxa de acerto de 80.32%. Já Oliveira et al., (2017) obtiveram os melhores resultados com os modelos Redes Neurais Artificiais e Regressão Logística que atingiram 75.24% e 74.98% de acerto, respectivamente.

Observa-se que o modelo proposto neste trabalho atingiu 82% de acerto, pouco acima da acurácia obtida por Shubhankar (2019). Pode-se supor que a diferença na configuração dos algoritmos, as ferramentas utilizadas para classificação, o pré-processamento dos dados e a seleção dos atributos tenham influenciado nos diferentes resultados obtidos.

Tabela 10: Comparação dos resultados trabalhos relacionados.

Trabalho	Modelos	Base de dados	Atributos	Instâncias	Acurácia
Shubhankar (2019)	Logistic Regression	<i>Cleveland Heart Disease</i>	14	303	80%
Oliveira et al., (2017)	Rede Neurais Artificiais	<i>Elsa (Brasil)</i>	24	403	75%
O Autor (2020)	Rede Bayesiana	<i>Cleveland Heart Disease</i>	15	303	82%

Fonte: O autor (2020)

5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O presente trabalho visou uma exploração do uso de modelos de aprendizado de máquina para utilização no diagnóstico de hipertensão arterial. Com ajuda dos trabalhos realizados, é possível perceber a importância da pesquisa de medidas que possam colaborar no diagnóstico de hipertensão arterial.

Em resumo, a aplicação de modelos de aprendizado de máquina pode provar uma alternativa viável para a construção de modelos preditivos de hipertensão arterial, geralmente com os bons resultados. Considerando os resultados obtidos entre diferentes modelos, de modo geral, mostrou-se que todos eles produziram resultados considerados satisfatórios, demonstrando assim a viabilidade de detectar pacientes diagnosticados como hipertensos através de dados clínicos.

Portanto conclui-se que os resultados obtidos no primeiro experimento com o uso do modelo de Rede Bayesiana com a taxa de acurácia 82.17% demonstraram a eficiência e precisão do modelo de diagnóstico supervisionado com a combinação da ferramenta Weka, que levem benefícios palpáveis no processo de diagnóstico de hipertensão arterial.

Com base no presente trabalho, indica-se algumas recomendações para a continuação da pesquisa nesta área. Alguns pontos merecem aprofundamento em pesquisas ou trabalhos futuros. Destes, os principais são:

- a) Utilização do histórico familiar, se um dos parentes dos pacientes já sofreu hipertensão arterial;
- b) Utilização de outros modelos não aplicados neste trabalho, como por exemplo, Clusterização;
- c) Criação de uma ferramenta de auxílio ao diagnóstico de fácil utilização por parte dos profissionais de saúde, os quais podem fazer uso da ferramenta no momento da consulta, aplicando os dados coletados dos pacientes, apoiando o diagnóstico e diminuindo a incerteza no diagnóstico.

REFERÊNCIAS BIBLIOGRÁFICAS

AUSTIN, **Inteligência artificial: Machine Learning**. 6. ed. São Paulo, 2017. p.119-121.

ABICALAFFE, C.; AMARAL, V. F.; DIAS, J. S. **Aplicação da Rede Bayesiana na Prevenção da Gestão de Alto Risco**. In: Congresso Brasileiro de Informática Médica, Ribeirão Preto. Anais do Congresso Brasileiro de Informática Médica, v. 1. p. 1-1, 2004.

ALVES S. T. et al. **Estimativa da evapotranspiração de referência através de redes neurais artificiais**. Revista Brasileira de Meteorologia, v.26, n.2, 2011. p. 197-203.

BOBBIO, A. et al. **Improving the analysis of dependable systems by mapping fault trees into Bayesian networks**. Reliability Engineering & System Safety. Vol. 71, 2001. p. 249-260.

BREIMAN, L. et al. **Classification and regression trees**. Belmont, Calif.: Wadsworth, 1984.

BRASIL, DENTAL ACESS, **Vendas de produtos**. Disponível em: <<https://www.dentalaccess.com.br/?m=prod&busca=Esfigmoman%C3%B4metro>>. Acesso em: 20 mar. 2020.

BRASIL, INSTITUTO DE CARDIOLOGIA DO LAGO, **Mapa de Pressão Arterial: Para que serve?** Disponível em: <<https://telemedicinamorsch.com.br/blog/o-que-e-mapa-de-pressao-arterial>>. Acesso em: 24 fev. 2020.

BRASIL, MINISTÉRIO DA SAÚDE. **Sociedade Brasileira de Hipertensão**. Hipertensão: Medicamentos. Brasília: Ministério da Saúde; 2012. Disponível em: <http://publicacoes.cardiol.br/2014/diretrizes/2016/05_HIPERTENSAO_ARTERIAL.pdf>. Acesso em: 24 fev. 2020.

BRINK, H.; RICHARDS, J. W. **Real-World Machine Learning**. Disponível em: <http://www.manning.com/brink/RWML_MEAP_CH01.pdf>, 2013. Acesso em: 17 abr. 2020.

CABENA, P. et al. **Intelligent miner for data applications guide**. IBM RedBook SG24-5252-00, 1999.

CABENA, P.; HADJINLIAN, P.; STADLER, R.; VERHEES, J., e ZANASI, A. **Discovering data mining: from concept to implementation**. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1999.

CADERNOS DE SAÚDE PÚBLICA. **Diagnóstico da hipertensão**. Disponível em: <<https://www.scielo.org>> Acesso em: 14 nov. 2019.

CAMARGO, R; MIRANDA Jr.; FERREIRA, R. P. **A hybrid outer-approximation/benders decomposition algorithm for the single allocation hub location problem under congestion**. Operations Research Letters, 39(5), p 329 - 337. 2011.

CARMO. MEDICINANET, **Ataque Isquêmico Transitório e Acidente Vascular Cerebral** Disponível em: <<https://www.medicinanet.com.br/>>. Acesso em: 28 out. 2019.

CERA, M.C. **Uso de Redes Neurais para o Reconhecimento de Padrões**. Disponível em: <<http://www.inf.ufrgs.br/procpar/disc/cmp135/trabs/mccera/t1/padrões.pdf>>. Acesso em : 22 nov. 2019.

DENTAL ACESS, **Produtos Odontológicos**, Equipamento Esfigmomanômetro Disponível em: <<https://www.dentalaccess.com.br/>>. Acesso em: 25 fev. 2020.

DEV MÉDIA, **Mineração de dados no MySQL com a ferramenta Weka**. Disponível em: <<https://www.devmedia.com.br/mineracao-de-dados-no-mysql-com-a-ferramenta-weka/26360>>. Acesso em: 15 abr. 2020.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification** (2nd ed.). New York:Wiley, 2000.

EYING, E. **Controle Feedforward-Feedback aplicado às colunas de absorção de produção de produto de etanol por fermentação**. 2008. Campinas, SP. Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Química, Universidade Estadual de Campinas.

EYING, E., FILETI, A. M.F. Control of absorption columns in the bioethanol process: Influence of measurement uncertainties. **Engineering Applications of Artificial Intelligence**, v. 23, p. 271-282, 2010.

FACELI, K. et al. **Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina**. Rio de Janeiro: LTC, 2011.

FRIEDMAN, J. H.; HASTIE, T.; TIBSHIRANI, R. **Additive Logistic Regression: A statistical view of boosting**. The Annals of Statistics, v. 28, 2000. p. 337 à 407.

FURNKRANZ, J.; GAMBERGER, D.; LAVRAC, N. **Foundations of Rule Learning**.: Springer-Verlag Berlin, 2012.

GARCIA, S.C. **O Uso de árvores de decisão na descoberta de conhecimento na área da saúde**. In: SEMANA ACADÊMICA, 2000. Rio Grande do Sul: Universidade Federal do Rio Grande Do Sul, 2000.

HALL, M. et al. **The Weka data mining software: an update**. ACM SIGKDD Explorations Newsletter, ACM, v. 11, n. 1, 2009. p. 10–18.

HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. 2ª ed. The Morgan Kaufmann Series in Data Management Systems. Elsevier Inc, 2006.

HAYKIN, S. **Redes Neurais-Princípios e Práticas**. BOOKMAN, São Paulo, 2ª ed. 2001. 900p.

SHUBHANKAR R, **Heart Disease Prediction**. Disponível em: <<https://towardsdatascience.com/heart-disease-prediction-73468d630cfc>> Acesso em: set. 2019.

HOSSIN, N.; SULAIMAN MN. **Review on evaluation metrics for data classification evaluations**. Em International Journal of Data Mining & Knowledge Management Process, 2015.

INSTITUTO DE CARDIOLOGIA DE LAGO, **MAPA**, Monitorização Ambulatorial da Pressão Arterial. Disponível em : <<https://cardiolago.com.br/2019/09/06/mapa-significa-monitorizacao-ambulatorial-da-pessoa-arterial-feita-geralmente-a-cada-20/>>. Acesso em: fev. 2020.

KORB, K.; NICHOLSON, A. E. **Bayesian artificial intelligence**. Chapman & Hall/CRC. Florida, 2003.

JIANG, W. **Is Regularization unnecessary for Boosting**: Technical Report 04, Department Of Statistics, Northwestern University, 2000.

LUDENIR, T. B.; CARVALHO, A.C.P.L.F. **Redes Neurais Artificiais: teorias e aplicações**. Livros Técnicos e Científicos, Rio de Janeiro. 2007.

LUGOSI, G.; VAYATIS, N. **On Bayes-risk Consistency of Regularized Boosting Methods**. The Annals of Statistics, v. 32-1, 2004.

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

NEAPOLITAN, R. E. **Learning Bayesian Networks**. Upper Saddle River: Pearson, 2004.

NOGUEIRA, P. C.; CALIRI M. H. L. **Fatores de Risco para Úlcera de Pressão em pacientes com trauma medular**. Análise da literatura científica nacional. 6º Simpósio de Iniciação Científica da USP, Ribeirão Preto, nov. 1998.

NUNES, H. et al. **Redes Neurais Artificiais**. 2 ed. Para engenharia e ciências aplicadas. Universidade de São Paulo. 2016.

OLIVEIRA, et al. **Comparação de algoritmos de aprendizagem de máquina para construir um modelo preditivo para detecção de diabetes não diagnosticada: Elsa-Brasil: Estudo de acurácia**. São Paulo, jun. 2017.

QUINLAN, J.C. **C4.5: Programs for machine learning**. San Mateo: Morgan Kaufmann, 1993.

RAEDER, T. **Model Monitor User's Guide version 1.0**. Department of Computer Science and Engineering, University of Notre Dame. 2008.

RAMYA, M et al. **A predictive model construction for mulberry crop productivity**, *Procedia Computer Science* **45**: p 156–165, 2015.

RUSSEL, S; NORVIG, P. **Artificial Intelligence: A modern approach**. Series in Artificial Intelligence. Prentice Hall. 1995.

SANTOS, R. **Conceitos de mineração de dados na web**. XV Simpósio Brasileiro de Sistemas Multimídia e Web, p. 81–124, 2009.

SHANNON, R.; MAYER, R.; ADELSBERGER, H. **Expert systems and simulation**, Simulation, 1985.

SGAMBATTI, M. S.; PIERIN, A.; MION Jr. D. **A Medida da Pressão Arterial no idoso**. Rev Bras Hipertens, 2000.

SOCIEDADE BRASILEIRA DE CARDIOLOGIA. VII **Diretrizes de Monitorização Ambulatorial da Pressão Arterial (MAPA) e Monitorização Residencial de Pressão Arterial (MRPA)**. Arquivos Brasileiros de Cardiologia, São Paulo, v.107, p. 1-32, 3 set. 2018.

TAFNER, M A. **Redes Neurais Artificiais**: Aprendizado e Plasticidade. Revista “Cérebro & Mente”, 2-5 mar/maio, 1998.

VENKATADRI, M; LOKANATHA, C. **A Comparative study on decision tree Classification Algorithms in data mining**, International Journal of Computer Applications Engineering, Technology and Sciences (Ij-Ca-Ets) Issn: 0974-3596(2): 24, 2010.

VIEIRA FILHO, V.; ALBUQUERQUE, M. T. C. F. **Abordagem Bayesiana para Simulação de Jogos Complexos**. In: SBGames, São Paulo. Proceedings of games, 2007.

VIGITEL BRASIL. MINISTÉRIO DE SAÚDE BRASILEIRO. **Diagnóstico de Hipertensão**: Fatores de Risco e Proteção para Doenças Crônicas por inquérito telefônico. Disponível em: <<http://www.saude.gov.br/noticias/agencia-saude/45446-no-brasil-388-pessoas-morrem-por-dia-por-hipertensao>>. Acesso em: abr. 2020.

WANG, J et al. DBMiner: **A system for mining knowledge in large relational data bases**. In: Proc. Intl. Conf. on **Data Mining and Knowledge Discovery (KDD'96)**., 1996. p. 250–255.

WANG, Y. J.; XIN, Q.; COENEN, F. **Mining efficiently significant classification association rules**. In: STUDIES IN COMPUTATIONAL INTELLIGENCE, 2008 U467 Springer-Verlag p. 443 .

WITTEN, I.H.; FRANK, E. **Data mining:** Practical machine learning tools and techniques with Java implementations. 2nd. ed. San Francisco, California: Morgan Kaufmann, 2000 ISBN 0-12-088407-0. p. 18,21,34 – 39, 42 – 48,51 e 59.

WITTEN, I. H.; FRANK, E. **Data mining:** Practical machine learning tools and techniques. 2 ed. San Francisco: Morgan Kaufmann Publishers, 2005.