



UNIVERSIDADE FEDERAL DO MARANHÃO - UFMA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA - CCET
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PPGE

BRENDA IRLA CARDOSO FEITOSA SOARES

**MÉTODO DE DIAGNÓSTICO DA SÍNDROME DA APNEIA OBSTRUTIVA DO
SONO POR APRENDIZADO DE MÁQUINA**

SÃO LUÍS

2022

BRENDA IRLA CARDOSO FEITOSA SOARES

**MÉTODO DE DIAGNÓSTICO DA SÍNDROME DA APNEIA OBSTRUTIVA DO
SONO POR APRENDIZADO DE MÁQUINA**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica do Centro de Ciências Exatas e Tecnologias, da Universidade Federal do Maranhão, como requisito para obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Dr. Allan Kardec Duailibe Barros Filho

São Luís

2022

Ficha gerada por meio do SIGA/Biblioteca com dados fornecidos pelo (a) autor (a) Núcleo Integrado de Bibliotecas/ UFMA

Cardoso Feitosa Soares, Brenda Irla.

MÉTODO DE DIAGNÓSTICO DA SÍNDROME DA APNEIA OBSTRUTIVA DO SONO POR APRENDIZADO DE MÁQUINA / Brenda Irla Cardoso Feitosa Soares. - 2022. 78 f.

Orientador(a): Allan Kardec Duailibe Barros Filho.

Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, 2022.

1. Apneia. 2. Diagnóstico. 3. Eletrocardiograma. 4. Polissonografia. I. Kardec Duailibe Barros Filho, Allan. II. Título.

CCET – Centro de Ciências Exatas e Tecnologias
Universidade Federal do Maranhão

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica intitulada ***Método de Diagnóstico da Síndrome da Apneia Obstrutiva do Sono por Aprendizado de Máquina*** de autoria de Brenda Irla Cardoso Feitosa Soares, aprovada pela banca examinadora constituída pelos seguintes professores:

Prof. Dr. Allan Kardec Duailibe Barros Filho – Orientador

Prof. Dr. Carlos Alberto Bezerra Tomaz – Membro da Banca

Prof. Dr. Ewaldo Eder Carvalho Santana – Membro da Banca

São Luís, 22 de março de 2022

Dedico essa dissertação aos meus pais que foram responsáveis por tudo que conquistei até aqui, por me incentivarem a estudar e sempre persistir em todos os desafios, não importando a distância ou a dificuldade para que eu tivesse sucesso na vida. Dedico também ao meu amado primo e grande amigo, Ivo Aguiar (*in memoriam*), a saudade aqui está grande, você foi e sempre será uma base muito importante na minha vida.

AGRADECIMENTOS

Citando aqui Nanal, um autor ainda desconhecido: “A vida tem seus momentos difíceis, mas tem suas compensações. As vezes com um abraço ou uma mensagem de um amigo distante, coisas que parecem insignificantes, mas que tem uma importância tão grande, fazem toda a diferença”. Assim, venho agradecer a todos por cada etapa desta conquista.

Primeiramente, quero agradecer a Deus pelo seu amor e compaixão, pelas bençãos de toda ordem que me foram concedidas e por ter me abençoado com capacidades físicas e habilidades intelectuais que me ajudaram a superar todas as adversidades que encontrei ao longo da minha jornada.

Quero agradecer imensamente a toda minha família, especialmente, aos meus pais Marlene Cardoso e Feitosa Soares, por todos os sacrifícios que fizeram, por todas as dificuldades que enfrentaram e por toda a perseverança que tiveram para que eu chegasse onde estou. Agradeço também aos meus irmãos, Bárbara Soares e Feitosa Soares Júnior, por sempre estarem junto comigo em todos os momentos, inclusive nas discussões típicas de irmãos.

Ao meu melhor amigo e namorado Bruno Nascimento, não tenho palavras para expressar o quanto sou grata por tudo que ele fez por mim, através do seu amor e cuidados sei o que realmente quer dizer as palavras de companheirismo, paciência e resiliência, pois ele nunca me desamparou emocionalmente durante a construção de estudo. Obrigada imensamente por sempre estar ao lado e ser o meu refúgio quando sempre precisei.

Jamais poderia deixar de agradecer a todos os meus amigos que de alguma forma me ajudaram ao longo da minha vida. No entanto, quero expressar minha gratidão a alguns, em especial, que ajudaram durante o processo desta dissertação: Jéssica Almeida, Luís Eduardo, Bruno Cardoso, Cenira Marinho, Wanderson Cintra, Crislaine Oliveira, Marcelo Martins, José Lucas, Edhymally e Edhymilkson Aquino (*Pessoal, desculpe por alguns momentos de estressantes que eu fiz vocês me aguentarem. Mas, afinal, amigo serve para isso também, né, aguentar uma amiga em pequenos “surtos” temporários?! Prometo que no meu próximo estudo não ficarei assim... Eu espero! Rs...*)

Ao meu orientador e coorientador, Allan Kardec e Jonathan Queiroz, quero agradecer todas as orientações e ensinamentos que me foram dados, pois sem eles o presente estudo não se seria possível. Obrigada! Não posso deixar de agradecer também ao meu amigo e colega de mestrado Luis Fillype da Silva.

E por fim, mas não mesmo importante, sou extremamente grata a CNPQ por me ter auxiliado durante do todo esse processo de pós-graduação e operacionalização do estudo.

A todos, o meu muito, muito, Obrigada!

“Não fui eu que ordenei a você? Seja forte e corajoso! Não se apavore nem desanime, pois o Senhor, o seu Deus, estará com você por onde você andar”

Josué 1:9

RESUMO

A Síndrome da Apneia Obstrutiva do Sono (SAOS) é caracterizada pela fragmentação e hipoxia repetitiva durante o sono, quando não diagnosticada e tratada da forma adequada se torna causa de complicações graves associadas a doenças cardiovasculares. O diagnóstico dessa síndrome requer um estudo clínico detalhado, denominado polissonografia, composto por uma série de exames, dificultando o processo de diagnóstico da SAOS. O exame de polissonografia é considerado padrão para diagnóstico e avaliação dos distúrbios respiratórios do sono, como o próprio nome sugere, este exame reúne diversos testes que realizam análise da atividade cerebral (EEG), cardíaca (ECG), ocular (EOG) e muscular (EMG). Devido à complexidade na realização da polissonografia, o presente trabalho objetiva classificar e diagnosticar dois grupos de indivíduos, saudáveis e com SAOS em grau leve, a partir dos sinais de ECG aplicados em rede neural MLP juntamente à análise de componentes principais (PCA). Ao utilizar a metodologia de extração de características, adaptada para diagnóstico da apneia obstrutiva do sono, os resultados foram amostrados em duas e três dimensões com classificação de 95% de acurácia.

PALAVRAS-CHAVE: Apneia; eletrocardiograma; polissonografia; diagnóstico.

ABSTRACT

Obstructive sleep apnea syndrome (OSAS) is characterized by fragmentation and repetitive hypoxia during sleep, if this syndrome is not properly diagnosed and treated, it becomes the cause of serious complications such as cardiovascular problems. The diagnosis of this syndrome requires a detailed clinical examination called polysomnography, which consists of several tests that perform an analysis of brain (EEG), heart (ECG), muscle (EMG) and eye (EOG) activity. Due to the complexity of performing polysomnography, the present study aims to classify and diagnose two groups of subjects, healthy and with normal apnea, based on the use of ECG signals applied in a supervised machine learning algorithm along with Principal Component Analysis (PCA). Using the feature extraction methodology adapted for the diagnosis of obstructive sleep apnea, the results were sampled in two and three dimensions with 95% accuracy.

KEYWORDS: Apnea; electrocardiogram; polysomnography; diagnosis.

LISTA DE FIGURAS

Figura 1- Exame polissonográfico com episódios de apneia obstrutiva	24
Figura 2- Sistema de condução do coração	27
Figura 3- Ilustração de eletrocardiograma normal, representação das ondas P, T e completo QRS.	27
Figura 4- Medidas de tendência central (média, mediana e moda) coincidindo no mesmo ponto dentro da curva	30
Figura 5- Curtose tipo mesocúrtica – Representa curva com distribuição normal.....	31
Figura 6- Curtose tipo leptocúrtica – Representa curva com distribuição anormal pontiaguda	31
Figura 7- Curtose tipo platicúrtica – Representa curva com distribuição anormal achatada	31
Figura 8- Esquema de modelo de aprendizado supervisionado.	33
Figura 9- Esquema de aprendizado de máquina.....	33
Figura 10- Problema de classificação	34
Figura 11- Mapeamento das entradas a partir de função matemática	35
Figura 12- Matriz de confusão	36
Figura 13- Diferentes valores da variável K para classificação de amostra através do algoritmo KNN.....	37
Figura 14- Classificação de dado por algoritmo KNN	38
Figura 15- Demonstração dos vetores de suporte e margem utilizadas para algoritmo SVM.	39
Figura 16- Tamanhos diferentes de margem ao considerar o mesmo problema.	40
Figura 17- Classificação não linear do algoritmo SVM.....	41
Figura 18- Neurônio biológico em (a) e Neurônio artificial em (b).....	42
Figura 19- Estrutura de uma rede neural artificial.....	43
Figura 20- Rede neural de camada única.....	44
Figura 21- Rede neural MLP	45
Figura 22- Rede neural recorrente	46
Figura 23- Batimentos cardíacos de indivíduo saudável	52
Figura 24- Batimentos cardíacos de indivíduo com SAOS.....	53
Figura 25- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau leve	54

Figura 26- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve	55
Figura 27- Representação em 2D da dispersão de assimetria e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve	56
Figura 28- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau leve apresentados em espaço 3D a partir da mediana, assimetria e curtose.....	57
Figura 29- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau moderado.....	58
Figura 30- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado.....	59
Figura 31- Representação em 2D da dispersão de assimetria e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado.....	60
Figura 32- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau moderado apresentados em espaço 3D a partir da mediana, assimetria e curtose.....	61
Figura 33- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau grave	62
Figura 34- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau grave.....	63
Figura 35- Representação em 2D da dispersão de assimetria e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau grave	64
Figura 36- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau grave apresentados em espaço 3D a partir da mediana, assimetria e curtose.....	65
Figura 37- Valores de classificação em porcentagem para SAOS em grau leve para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação	66
Figura 38- Valores de classificação em porcentagem para SAOS em grau moderado para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação.	67
Figura 39- Valores de classificação em porcentagem para SAOS em grau grave para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação.	68
Figura 40- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve sem aplicação de PCA.	69
Figura 41- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado sem aplicação de PCA.....	70
Figura 42- Representação em 2D da dispersão de curtose e mediana para os sinais dos	

indivíduos saudáveis e com SAOS em grau grave sem aplicação de PCA..... 71

LISTA DE TABELAS

Tabela 1- Número de gravações para cada grupo portador da SAOS	47
Tabela 2- Configurações da rede neural artificial MLP	50

LISTA DE ABREVIATURAS

Aprendizado de Máquina	-	AM
Classificação Internacional dos Distúrbios do Sono	-	ICSD
Eletrocardiograma	-	ECG
Eletroencefalograma	-	EEG
Eletromiografia	-	EMG
Eletro-oculograma	-	EOG
Falso Negativo	-	FN
Falso Positivo	-	FP
Índice de Apneia-Hipopneia	-	IAH
K-Vizinhos mais Próximos	-	KNN
Máquina de Vetores de Suporte	-	SVM
<i>Multilayer Perceptron</i>	-	MLP
Polissonografia	-	PSG
<i>Principal Component Analysis</i>	-	PCA
Rede Neural de Múltiplas Camadas	-	MLP
Redes Neurais Artificiais	-	ANNs
Redes Neurais Simuladas	-	SNNs
Síndrome da Apneia Obstrutiva do Sono	-	SAOS
Sonolência Diurna Excessiva	-	SDE
<i>Support Vector Machine</i>	-	SVM
Vias Aéreas Superiores	-	VAS
Verdadeiro Negativo	-	TN
Verdadeiro Positivo	-	TP

SUMÁRIO

1 INTRODUÇÃO	18
1.1 OBJETIVO GERAL.....	19
1.2 OBJETIVOS ESPECÍFICOS	19
2 DISTÚRBIOS DO SONO	20
2.1 CLASSIFICAÇÃO INTERNACIONAL DOS DISTÚRBIOS DO SONO	20
2.1.1 Dissonias	21
2.1.2 Parassonias	21
2.1.3 Distúrbios do sono relacionados a alterações médico-psiquiátricas	21
2.1.4 Distúrbios do sono propostos	21
2.2 SÍNDROME DA APNEIA OBSTRUTIVA DO SONO (SAOS)	22
2.2.1 Consequências da SAOS	22
2.2.2 Epidemiologia	23
2.3 ESTUDO POLISSONOGRÁFICO	24
3 ELETROCARDIOGRAMA	26
3.1 ATIVAÇÃO ELÉTRICA DO CORAÇÃO.....	26
3.2 ARRITMIAS CARDÍACAS	28
4 MOMENTOS ESTATÍSTICOS	29
4.1 MEDIANA	29
4.2 ASSIMETRIA	29
4.3 CURTOSE.....	30
5 APRENDIZAGEM DE MÁQUINA	32
5.1 APRENDIZADO SUPERVISIONADO	32
5.2 CLASSIFICAÇÃO	34
5.2.1 Algoritmo Knn	36
5.2.2 Algoritmo SVM	39
5.2 APRENDIZADO NÃO SUPERVISIONADO	41
5.3 REDES NEURAS ARTIFICIAIS	42
5.3.1 Rede neural de camada única	43
5.3.2 Rede neural de múltiplas camadas	44
5.3.3 Rede neural recorrente	45
5.3.4 Análise de Componente Principais – PCA	46
6 MATERIAIS E MÉTODOS	47

6.1 BASE DE DADOS.....	47
6.2 SEGMENTAÇÃO E PRÉ-PROCESSAMENTO DOS CICLOS CARDÍACOS	48
6.3 EXTRAÇÃO DE CARACTERÍSTICAS.....	49
6.4 CONFIGURAÇÕES DA MÁQUINA DE APRENDIZADO.....	50
6.5 MÉTRICAS DE AVALIAÇÃO	50
7 RESULTADOS	52
7.1 SAOS EM GRAU LEVE COM PCA	53
7.2 SAOS EM GRAU MODERADO COM PCA.....	57
7.3 SAOS EM GRAU GRAVE COM PCA	61
7.4 AVALIAÇÃO DOS CLASSIFICADORES	65
7.5 CLASSIFICAÇÃO SEM PCA.....	69
8 TRABALHOS FUTUROS	72
9 CONCLUSÃO.....	73
REFERÊNCIAS	74

1 INTRODUÇÃO

O fenômeno do sono é essencial para o processo de regulação do corpo, visto que durante esse intervalo o corpo exerce funções como fortalecimento do sistema imunológico, liberação de hormônios e consolidação da memória, ou seja, é por meio do sono que todo o organismo pode ser renovado (GOLDBERGER *et.al.*, 2000). Esse fenômeno é definido como um processo de redução da atividade metabólica, inativação da musculatura voluntária e diminuição da atividade sensorial, onde se torna reversível a partir de estímulos sensoriais, como estímulos auditivos e táteis (HAGGSTRÄM; ZETTLER; FAM, 2009).

Existem estímulos que ocorrem durante o sono que não causam incômodos no indivíduo a ponto de o manter acordado, quando isso acontece há a presença somente de micro despertares, caracterizado como um despertar transitório que ocasiona contração muscular da faringe e língua, permitindo que haja abertura das vias aéreas e possibilidade de respiração. Segundo Bittencourt, Togeiro e Bagnato (2012) um distúrbio que gera micro despertares frequentes durante o sono é a Síndrome da Apneia Obstrutiva do Sono (SAOS).

A SAOS se caracteriza pelo comprometimento total (apneia) ou parcial (hipopneia) das vias aéreas, por um período maior ou igual a 10 segundos, durante o sono (BITTENCOURT; TOGEIRO; BAGNATO, 2012). Pacientes portadores da SAOS apresentam consequências comuns para essa condição, tais como despertares frequentes, ronco, sono agitado, sonolência diurna e problemas cardiovasculares, como hipertensão arterial e arritmia cardíaca (BITTENCOURT; TOGEIRO; BAGNATO, 2012).

O diagnóstico da SAOS é realizado por meio de análise do quadro clínico do paciente, que engloba aplicação de questionários, escalas de avaliação do sono e exame quantitativo específico, a polissonografia (HAGGSTRÄM; ZETTLER; FAM, 2009; ITO *et. al.*, 2005). O estudo polissonográfico é considerado padrão para diagnóstico e avaliação dos distúrbios respiratórios do sono.

O exame de polissonografia, como o próprio nome sugere, reúne diversos testes que realizam análise da atividade cerebral (Eletroencefalograma - EEG), cardíaca (Eletrocardiograma - ECG), muscular (Eletromiografia - EMG) e ocular (Eletro-oculograma - EOG). Devido a grande quantidade de equipamentos e atendimentos pessoais necessários, o estudo polissonográfico é caracterizado como um exame muito complexo, propõe-se assim a avaliação dos dados de um dos exames presentes na polissonografia, o ECG em conjunto com técnicas de aprendizado de máquinas, para realizar diagnóstico da SAOS de forma mais viável.

Em Queiroz (*et al.*, 2020) é realizada a análise dos batimentos cardíacos através de

métodos estatísticos para extração de características das bases de dados de ECG. A partir de Queiroz (*et al*, 2020) e Silva, Queiroz e Barros (2020) e da metodologia de classificação em diferentes grupos adotada em Sher, Schechtman e Piccirillo (1996), o presente trabalho propõe realizar a separação dos sinais de ECG em dois grupos (SAOS e Saudável), ao utilizar algoritmo de aprendizado de máquinas aliado à análise de componentes principais (PCA) para diagnosticar indivíduos portadores da SAOS em grau leve.

1.1 OBJETIVO GERAL

Utilizar sinais de eletrocardiograma (ECG), em conjunto com técnicas de aprendizado de máquinas e estatísticas de alta ordem, para dá início ao diagnóstico de indivíduos portadores da SAOS.

1.2 OBJETIVOS ESPECÍFICOS

De forma a alcançar o objetivo geral, apresenta-se os objetivos específicos a serem seguidos:

- Realizar pré-processamento das bases de dados por meio da segmentação dos batimentos cardíacos em matrizes;
- Avaliação das características dos indivíduos com apneia e indivíduos saudáveis de cada base de dados por meio de métodos estatísticos;
- Uso de técnicas para redução de dimensionalidade - PCA (*Principal Component Analysis*);
- Realizar classificação e separação entre grupos de indivíduos saudáveis e indivíduos com apnéia em grau leve, moderado e acentuado;

2 DISTÚRBIOS DO SONO

De acordo com Reimão (1996), o sono é uma função biológica de extrema importância para consolidação da memória, visão binocular, termorregulação, conservação e restauração da energia. Em Ferrara e De Gennaro (2001), há ainda afirmação da importância desse fenômeno para restauração do metabolismo energético cerebral.

Dessa forma como o sono é responsável pelas principais funções do corpo humano, perturbações neste processo acarretam mudanças significativas no funcionamento físico, ocupacional, cognitivo e social do indivíduo (MÜLLER; GUIMARÃES, 2007), podendo prejudicar a qualidade de vida consideravelmente. Assim, uma boa qualidade de sono proporciona fatores essenciais para manter corpo e mente saudáveis, visto que durante esse momento há reposição de energias e regulação do metabolismo.

A quantidade adequada de horas de sono varia de acordo com o desenvolvimento e necessidades do corpo humano, onde nos primeiros meses de vida são necessárias 12 a 16 horas (ALMEIDA, 2019). Com o crescimento o tempo de sono da criança diminui 30 minutos por ano até os cinco anos, ao atingir a fase adulta a quantidade de horas decresce mais ainda em função da idade e de fatores externos (MÜLLER; GUIMARÃES, 2007). Dessa forma, tem-se que a partir do avanço da idade, existem perdas na quantidade, constância e qualidade do sono. Outrossim, a sonolência diurna excessiva compromete a segurança pública, pois contribui com o aumento dos casos de acidentes industriais e de tráfego (FERRARA & DE GENNARO, 2001).

2.1 CLASSIFICAÇÃO INTERNACIONAL DOS DISTÚRBIOS DO SONO

Dentre as condições clínicas que afetam a quantidade e a qualidade do sono tem-se dores e o uso de medicações, especialmente entre idosos que são os mais propensos a essas condições. Em Almeida (2019), se explana que 63% da população brasileira relata queixas relacionadas ao distúrbio do sono, o sistema mais utilizado para realizar diagnóstico e classificação desses distúrbios na literatura especializada é a Classificação Internacional dos Distúrbios do Sono (ICSD) (*American Sleep Disorders Association*, 1997).

Dentro da ICSD é realizado classificação dos distúrbios do sono em quatro grupos principais: dissonias, parassonias, distúrbios do sono relacionados a alterações médico-psiquiátricas e distúrbios do sono propostos.

2.1.1 Dissonias

O grupo das dissonias possuem 34 distúrbios que se subdividem em distúrbios extrínsecos, intrínsecos e relacionados ao ritmo circadiano, compondo assim o maior grupo dentre os transtornos do sono. Esse grupo se caracteriza por transtornos primários que estão relacionados à iniciação ou manutenção do sono, e à sonolência excessiva, visto que ocorrem distúrbios na quantidade e regulação do sono (AMERICA PSYCHIATRIC ASSOCIATION, 1994). As características principais das dissonias são queixas de dificuldade para iniciar ou manter o sono e de sono não reparador, ou seja, há histórico de sono “leve” e facilmente perturbado.

2.1.2 Parassonias

Parassonias compõe o segundo maior grupo de transtornos do sono, com 24 distúrbios associados. São definidas como comportamentos e experiências psicológicas anormais, onde ocorrem alterações comportamentais em diferentes momentos do sono (MÜLLER; GUIMARÃES, 2007). Sonambulismo, terror noturno, bruxismo e paralisia do sono são alguns exemplos de parassonias que são tratadas de maneira específica.

2.1.3 Distúrbios do sono relacionados a alterações médico-psiquiátricas

Como o nome sugere, esse grupo é composto por 19 distúrbios que estão relacionados a doenças mentais, neurológicas e afecções médicas. Os transtornos associados a doenças mentais ocorrem em indivíduos portadores de transtornos de ansiedade e psicoses. Para doenças neurológicas tem-se aqueles portadores de doenças degenerativas, demências e mal de Parkinson, enquanto para os casos de enfermidades médicas há os portadores de alterações no sistema cardiovascular, respiratório e digestivo (REIMÃO, 1996).

2.1.4 Distúrbios do sono propostos

Distúrbios do sono propostos ocorrem quando há mudanças nos hábitos de sono. Esse grupo é composto por diferentes síndromes, ou seja, problemas que não possuem causas comuns para que haja definições específicas, engasgue do sono e hiperidrose são alguns

exemplos (MÜLLER; GUIMARÃES, 2007).

2.2 SÍNDROME DA APNEIA OBSTRUTIVA DO SONO (SAOS)

No Brasil é avaliado que entre a população em geral há dez a vinte milhões de portadores de transtornos do sono (MARTINEZ, 1999). Estima-se que a população adulta é a mais acometida por esses distúrbios, com 15% a 27% de prevalência, em Roth (*et. al*, 2002) afirma-se que os distúrbios com maior prevalência são a insônia e a síndrome da apneia obstrutiva do sono (SAOS).

A SAOS é uma síndrome que ocasiona distúrbios frequentes de respiração durante o sono, dessa forma indivíduos portadores desse transtorno sofrem repetidos episódios de apneia e hipopneia enquanto dormem (BALBANI; FORMIGON, 1999). As principais características da SAOS são os esforços inspiratórios ineficazes, altas pressões negativas intratorácicas e aumento da atividade nervosa que são decorrentes das pausas respiratórias.

A gravidade da SAOS é determinada a partir do Índice de Apneia-Hipopneia (IAH), que representa a quantidade da ocorrência dos sintomas da SAOS por hora de sono. O IAH é dividido em três grupos, grau leve para índice maior que 5 e menor que 15 vezes; Grau moderado quando há valores maiores que 15 e menores que 30; e grau acentuado com IAH maior que grau 30 (SBPT, [201-]).

2.2.1 Consequências da SAOS

Dentre os distúrbios do sono a apneia obstrutiva é classificada como uma dissonia, visto que causa alterações e ocasiona casos de insônia ou sonolência excessiva. Para que haja correta interpretação e diagnóstico da SAOS a Academia Americana de Medicina do Sono define que sejam preenchidos alguns critérios, são eles:

- Sonolência diurna excessiva (SDE) que ocasiona estresse psicológico, diminuição da produtividade no trabalho/escola, aumento do risco de acidentes e ainda prejuízos no funcionamento físico (REIMÃO, 1996);
- Engasgos durante o sono;
- Despertares frequentes;
- Sono não reparador com presença de cefaleia matinal, boca seca e dores de garganta;
- Redução da memória.

A ocorrência de cinco ou mais eventos respiratórios obstrutivos por hora de sono (IAH \geq 5/h) caracteriza indistintamente apneias, hipopneias ou algum esforço respiratório relacionado ao despertar (FERRARA & DE GENNARO, 2001). No sono do apneico é notório diversos detalhes importantes para diagnóstico da SAOS, observações que podem ser feitas pelos familiares ou companheiro de quarto, como pausas na respiração, ronco e inquietação no leito. No entanto para que exista diagnóstico preciso da SAOS é necessário a monitorização polissonográfica.

2.2.2 Epidemiologia

A falta de diagnóstico da síndrome da apneia obstrutiva do sono se torna preocupante por conta das comorbidades associadas e principalmente por conta do risco súbito de morte nos portadores (DRAGER, *et al.*, 2002). Os principais fatores de risco para desenvolvimento da SAOS são a obesidade, sexo masculino, hipotireoidismo, diabetes, envelhecimento e aumento da circunferência do pescoço.

De acordo com a literatura a prevalência da SAOS é maior em homens do que em mulheres, onde há uma relação de homem/mulher que varia entre 2:1 e 4:1. O avanço da idade, obesidade e aumento da circunferência do pescoço são fatores agravantes para ocorrência da apneia, visto que com o envelhecimento há aumento do peso e conseqüentemente do comprimento da circunferência do pescoço. Dessa forma, apesar de ser uma síndrome que acomete homens na maioria, mulheres com idade avançada também fazem parte do grupo de risco (MARTINS, *et al.*, 2007).

Por ser uma doença de alta incidência, acometendo 5% da população ocidental (FLEMONS *et al.*, 1999), a falta de diagnóstico e tratamento para a SAOS ocasiona aumento da mortalidade cardiovascular, casos como infarto agudo do miocárdio, insuficiência cardíaca congestiva, arritmias e outras doenças têm sido associadas aos casos de óbitos da SAOS, no entanto essa síndrome possui prevalência mais alta em pacientes com hipertensão arterial sistêmica (LIMA, *et al.*, 2008)

A incidência da síndrome da apnéia obstrutiva é maior em indivíduos com hipertensão arterial devido as condições de risco que ambas as doenças possuem em comum, são elas obesidade, sexo masculino e roncos (DRAGER *et al.*, 2002). De acordo com Silverberg, Oksenberg e Iaina (1997), 40% dos portadores de hipertensão arterial também foram diagnosticados com síndrome da apnéia obstrutiva do sono.

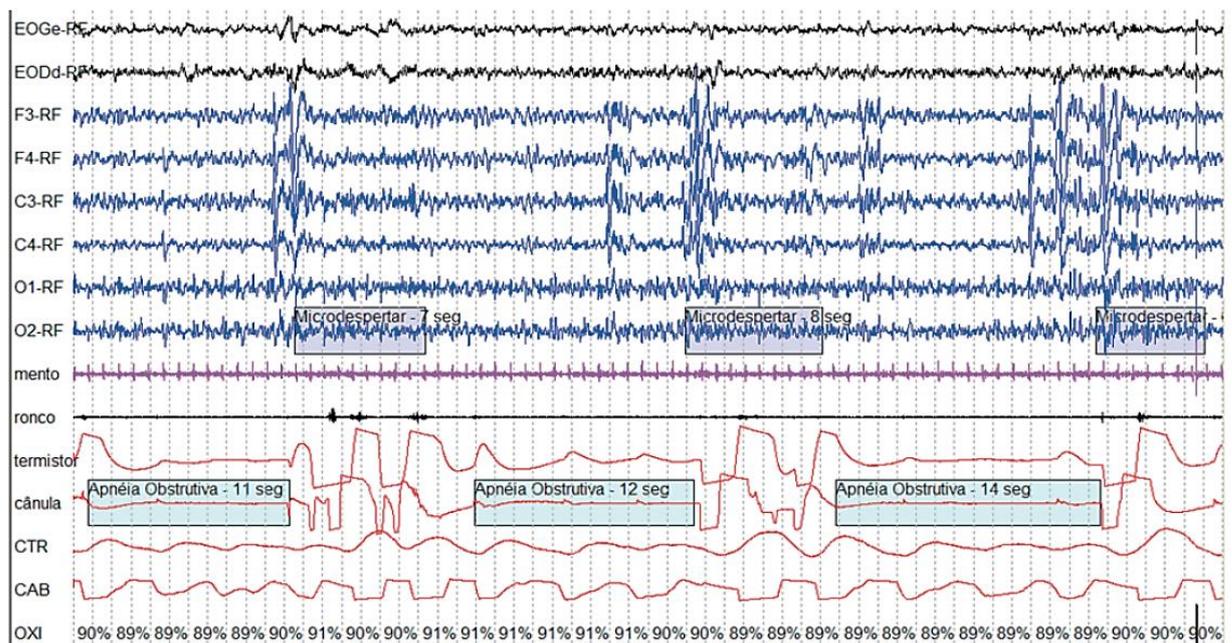
2.3 ESTUDO POLISSONOGRÁFICO

A polissonografia (PSG) é considerada o exame padrão para definir qualquer desordem do sono. Na realização desse procedimento são registrados o estágio e continuidade do sono, os esforços respiratórios realizados e a saturação do oxigênio, eletrocardiograma, eletroencefalograma, posição e movimentos corporais. Para auxiliar o diagnóstico desses distúrbios também são realizados outros exames: ressonância magnética, nasofiboscopia com manobra de Muller, Cefalometria, videofluoroscopia e tomografia computadorizada (FELDMAN; GOLDWASSER, 2004).

Além do estudo polissonográfico, o histórico clínico e exames físicos são fundamentais para que possa ser realizado o diagnóstico da SAOS (GONDIM *et al.*, 2007). Relatos da presença de ronco, dificuldades de respiração durante o sono, insônia, sono agitado e sobrepeso ou obesidade, são informações relevantes para auxiliar a presença dessa síndrome. No entanto, é importante que sempre haja a presença da polissonografia para que exista um diagnóstico mais preciso da enfermidade.

A Figura 1 representa um exame de polissonografia contendo episódios de apneia obstrutiva. A partir da PSG abaixo pode-se visualizar todos os sinais resultantes dos exames supracitados, assim como a possibilidade de caracterização dos pacientes nos diferentes grupos de gravidade da SAOS.

Figura 1- Exame polissonográfico com episódios de apneia obstrutiva



Fonte: Azzari *et al.*, 2014.

Conforme a Figura 1 todos os micros despertares são antecidos de apneias obstrutivas que ocorrem por mais de 10 segundos, caracterizando a redução ou ausência do fluxo aéreo por comprometimento das vias aéreas superiores (VAS) durante o sono. Cada momento de apneia mostrado na PSG representa um índice apneia/hipopneia (IAH), responsável pelo diagnóstico do teor de gravidade do portador de SAOS.

Dessa forma é de suma importância que seja realizado o reconhecimento da SAOS, visto que com um tratamento efetivo as taxas de eventos cardiovasculares, fatais e não fatais, são reduzidas. Porém devido a necessidade de uma grande quantidade de equipamentos e profissionais, o exame polissonográfico exige um alto investimento para ser realizado corretamente, que na maioria das vezes é superior à disponibilidade dos recursos hospitalares.

Por conta da dificuldade para execução do exame de polissonografia, procura-se métodos alternativos para determinação da presença da SAOS, onde esses devem ser mais acessíveis e práticos.

3 ELETROCARDIOGRAMA

O eletrocardiograma (ECG) é um exame complementar no diagnóstico cardiológico, constituído por registos elétricos advindos da atividade cardíaca, seu registo é realizado por meio do aparelho eletrocardiógrafo (RAMOS; SOUSA, 2007).

O eletrocardiógrafo funciona através da medição da diferença de potencial entre dois pontos por meio de eletrodos, esses realizam o registo dos sinais de corrente de pequena intensidade.

3.1 ATIVAÇÃO ELÉTRICA DO CORAÇÃO

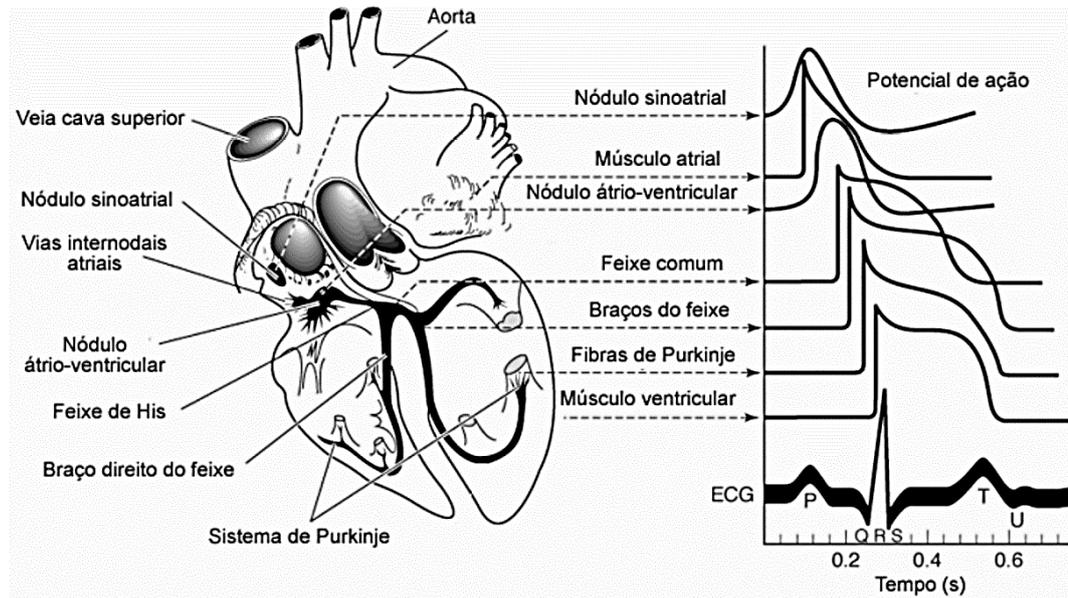
A célula miocárdica quando em repouso é considerada polarizada, visto que nesse estado há uma maior concentração de potássio do que no meio externo, por conta disso a célula apresenta-se negativa. No entanto, ao longo da ativação celular ocorrem trocas iônicas e a célula tende a se tornar positiva enquanto o meio torna-se negativo, essa inversão de polaridade caracteriza uma célula em estado despolarizado (RAMOS; SOUSA, 2007).

A atividade cardíaca desenvolvida durante o processo de ativação do coração desencadeia sinais elétricos originados dos processos de despolarização e repolarização, onde a despolarização é considerada uma onda móvel de cargas positivas (deflexão para cima) pois essas cargas se aproximam do eletrodo (FELDMAN; GOLDWASSER, 2004). Enquanto quando a despolarização se afasta do eletrodo há uma deflexão negativa, ou seja, o ECG apresenta sinal para baixo.

O nódulo sinusal (nódulo de Keith-Flack), localizado no átrio direito, é a origem da despolarização cardíaca, dessa forma o átrio direito é a primeira área do coração a ser ativada. Em sequência ocorre despolarização do átrio esquerdo, nódulo atrioventricular, feixe de His e seus ramos esquerdo e direito, rede de Purkinje e se extingue após alcançar os ventrículos (FELDMAN; GOLDWASSER, 2004), conforme ilustrado na Figura 2.

No lado esquerdo da Figura 2 estão representados os potenciais de ação típicos para cada parte do músculo cardíaco e sua relação com a produção ou não de ondas registráveis no ECG convencional. Como o coração possui batidas em sequência ordenada, visto que primeiramente ocorre a contração (sístole) dos átrios e em seguida a dos ventrículos, essa atividade cardíaca ocasiona ondas no ECG e registo pelo eletrocardiógrafo.

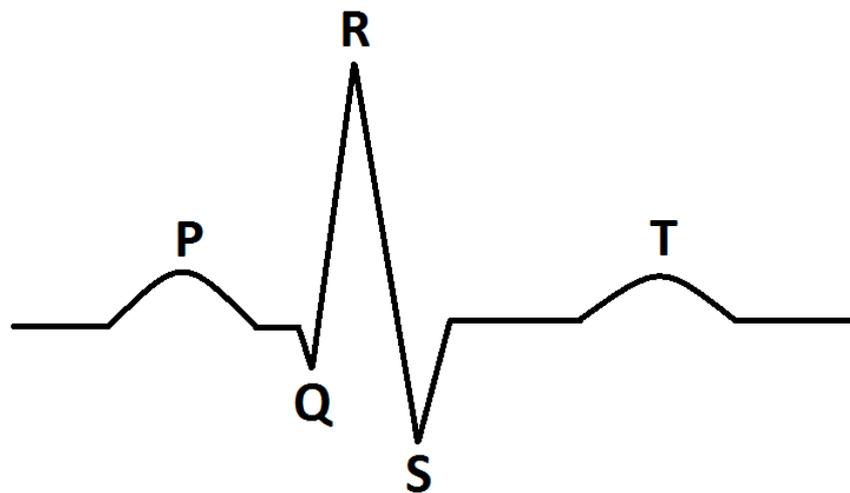
Figura 2- Sistema de condução do coração



Fonte: Schwarz (2009)

A Figura 3 representa as ondas geradas pelas deflexões advindas da ativação cardíaca, dessa forma tem-se que a onda P é resultante da despolarização cardíaca que se inicia no nódulo sinusal (átrio direito) e se difunde para o átrio esquerdo, portanto a onda P se caracteriza como uma onda de contração atrial, pois a atividade elétrica é gerada à medida que a onda de despolarização percorre ambos os átrios.

Figura 3- Ilustração de eletrocardiograma normal, representação das ondas P, T e completo QRS.



Fonte: ResearchGate ([201-]).

Como citado anteriormente a onda de despolarização segue para o nódulo atrioventricular e feixe de His, gerando sinais que o ECG não consegue registrar, dessa forma há um intervalo de tempo até a ativação alcançar os ventrículos e ocasionar seu estímulo. Esse

intervalo, da ativação dos átrios até ativação dos ventrículos, é representado na Figura 3 pelo segmento P-R (SILVERBERG; OKSENBORG; IAINA, 1997).

No início da ativação ventricular ocorrem despolarizações do septo interventricular e das paredes dos ventrículos, como resultado tem-se contração dos ventrículos que ocasionam várias deflexões denominadas complexo QRS.

Ao fim da despolarização nos ventrículos se inicia o processo de repolarização ventricular, inicialmente os sinais elétricos desenvolvidos são de baixa intensidade e como consequência não há deflexões registradas no ECG, nesse intervalo de tempo é registrada uma linha contínua, no segmento S-T, denominada isoeletrica (FELDMAN; GOLDWASSER, 2004). Após esse momento surgem correntes elétricas mais intensas que formam a onda T e possibilitam inscrição no ECG.

3.2 ARRITMIAS CARDÍACAS

Segundo Redline (*et al.*, 2010), indivíduos que apresentam distúrbios respiratórios durante o sono apresentam fatores de risco associados a arritmias complexas com probabilidade até quatro vezes maior do que indivíduos saudáveis. Dessa forma, mesmo que haja ausência de doenças cardíacas, arritmias noturnas podem acontecer em até 50% dos portadores da apneia do sono (WIGGERT *et al.*, 2010).

Para análise do ritmo cardíaco o sistema nervoso autônomo e os estados patológicos são fatores predisponentes, sendo um exemplo de estado patológico a apneia do sono. Para que se possa determinar e analisar o ritmo cardíaco é de suma importância observar o comportamento da onda P e medir o intervalo P-R e o complexo QRS, assim será definido se o ritmo está dentro da faixa de frequência normal ou não (RAMOS; SOUSA, 2007).

Em Cintra (*et al.*, 2014), é realizado um estudo populacional para definir percentualmente a relação entre a SAOS e arritmias cardíacas. Dessa forma, ao utilizar uma amostra de 433 participantes diagnosticados com SAOS em grau leve, moderado ou grave, foi observado que 62,7% dos participantes possuem algum distúrbio do ritmo cardíaco noturno, que varia de arritmia atrial a ventricular. Vale ressaltar que a maior prevalência de arritmias cardíacas ocorreu de forma proporcional ao aumento da gravidade da SAOS.

4 MOMENTOS ESTATÍSTICOS

Em estatística os momentos são de suma importância pois caracterizam distribuições de probabilidade. Dessa forma, considerando o símbolo E como valor esperado e n um momento de ordem n de uma variável aleatória X , tem-se na equação (1) a expressão de esperança.

$$E[X^n], \quad (1)$$

Os momentos estatísticos são medidos a partir de centro de valores, que podem caracterizar funções de densidade de probabilidade e momentos de alta ordem de uma distribuição.

4.1 MEDIANA

Juntamente com a média aritmética e moda, a mediana compõe as medidas de tendência central da estatística. Como o termo *mediana* faz referência a *meio*, tem-se que essa medida corresponde ao valor central de um conjunto de valores ordenados conforme apresentado na equação (2), onde X é um conjunto de variáveis aleatórias, onde no presente trabalho é constituído pela matriz de batimentos cardíacos.

$$M = \frac{X(n+1)}{2} \quad (2)$$

A partir da equação (2) é possível determinar o valor da mediana dentre os dados amostrais. Para que seja calculado a mediana primeiramente o conjunto de dados deve ser colocado em ordem crescente ou decrescente. Se o número de amostras for ímpar a mediana será o valor central, caso seja par a mediana será o resultado da média aritmética calculado entre os dois valores centrais.

4.2 ASSIMETRIA

Assimetria é uma medida de dispersão relativa muito utilizada para realizar análises comparativas, ela expressa os resultados em percentuais visto que mede a relação entre dispersão absoluta e tendência central (ICMC, 2012).

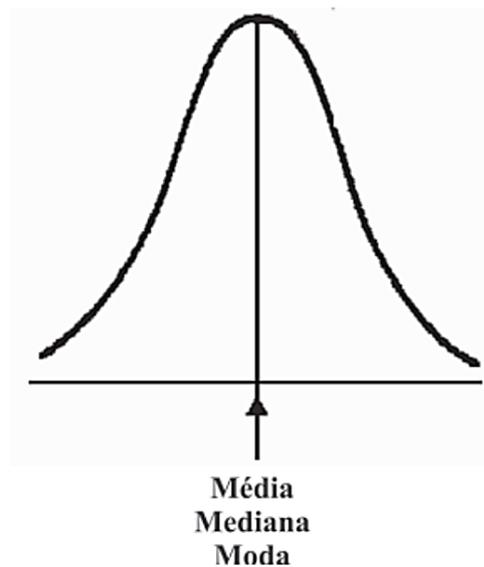
Dessa forma, a assimetria se caracteriza como um grau de desvio da simetria de um conjunto amostral. Esse grau de afastamento é determinado pela equação (3) e por meio da análise das medidas de tendência central, ou seja, quando a média, mediana e moda recaem em pontos diferentes, diz-se que a curva de distribuição é assimétrica.

$$k_X = E[(X - E(X))\sigma^{-1}]^3 \quad (3)$$

Onde X é um vetor de variáveis aleatórias.

A Figura 4 representa uma curva simétrica, pois as medidas de tendência central recaem no mesmo ponto.

Figura 4- Medidas de tendência central (média, mediana e moda) coincidindo no mesmo ponto dentro da curva



Fonte: ICMC, 2012.

4.3 CURTOSE

A curtose também se classifica como uma medida de dispersão relativa, assim como também apresenta seus resultados expressos em porcentagens. É uma medida que faz comparações com a distribuição normal, visto que determina o grau de achatamento de uma distribuição quando relacionada à distribuição normal [37], conforme a equação (4).

$$\lambda_X = E[(X - E(X))\sigma^{-1}]^4 \quad (4)$$

Onde X é um vetor de variáveis aleatórias.

A curtose pode ser do tipo mesocúrtica, leptocúrtica ou platicúrtica, que são nomenclaturas para os diferentes tipos de achatamento apresentado pela curva de distribuição, ilustrados nas Figuras 5, 6 e 7 respectivamente.

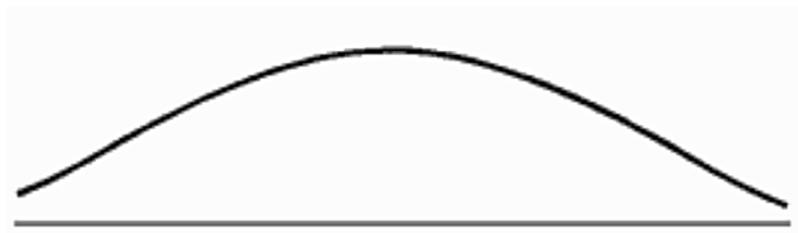
Figura 5- Curtose tipo mesocúrtica – Representa curva com distribuição normal



Figura 6- Curtose tipo leptocúrtica – Representa curva com distribuição anormal pontiaguda



Figura 7- Curtose tipo platicúrtica – Representa curva com distribuição anormal achatada



5 APRENDIZAGEM DE MÁQUINA

Aprendizado de máquina (AM), ou *machine learning*, é uma técnica utilizada na inteligência artificial que objetiva desenvolver tarefas de aprendizado computacional e construir sistemas com capacidade de adquirir conhecimentos em determinada área. Essa área de estudo utiliza algoritmos matemáticos e estatísticos para que assim o sistema computacional consiga realizar inferências, tomada de decisões, através de diversas experiências desenvolvidas como soluções em problemas anteriores (MONARD *et al.*, 2003).

AM visa suprir a necessidade de processamento de dados e consequente obtenção das suas principais informações, haja vista que não é viável a realização do processamento e análise de uma grande quantidade de dados de forma manual, dessa forma faz-se necessário a automatização de tarefas que simulem o comportamento do homem (ESCOVENDO & KOSHIYAMA, 2020).

O aprendizado gera capacidade de modificação e aprimoramento de um comportamento e da forma como ele gera respostas. Assim uma máquina está de fato aprendendo, ou seja, construindo e treinando, quando a máquina procura por padrões dentro do conjunto de dados.

Os sistemas de AM possuem individualidades quanto a linguagem descritiva, modo e forma de aprendizado que são responsáveis pela classificação dos dados. Dessa forma o aprendizado de máquina é dividido entre aprendizagem supervisionada e não supervisionada, constituído por algoritmo e dados de entrada (*dataset*) que são apresentados na saída de forma ordenada.

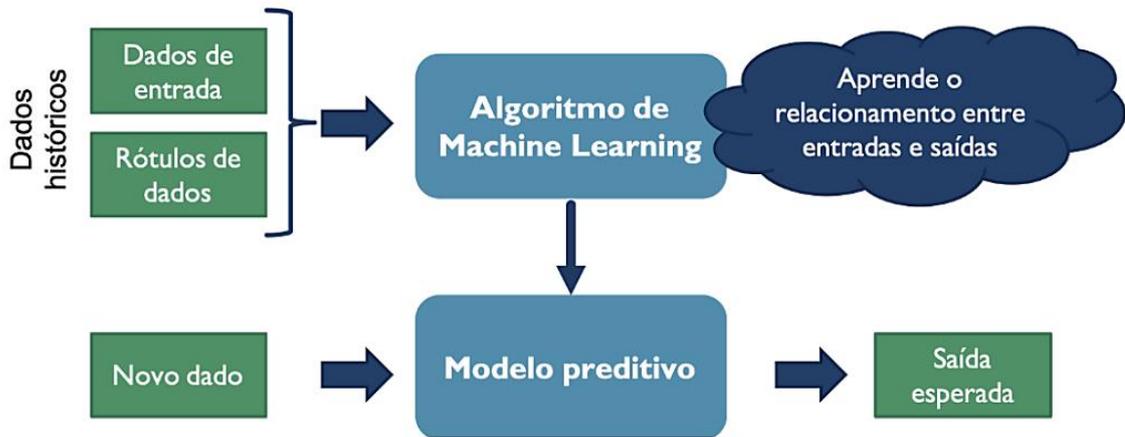
5.1 APRENDIZADO SUPERVISIONADO

Na aprendizagem supervisionada é utilizado dados de entrada rotulados, pois esse conjunto de dados possuem rótulo de classe já associada e conhecida de antemão (MONARD *et al.*, 2003). Dessa forma, o aprendizado acontece ao algoritmo utilizar vetores constituídos por atributos e rótulos de classes associadas para que seja construído um classificador capaz de tomar decisões corretas sobre a classe de novos conjuntos de dados não rotulados que serão utilizados posteriormente (ESCOVENDO & KOSHIYAMA, 2020).

A Figura 8 representa como ocorre o aprendizado supervisionado, onde os dados de entrada são divididos entre dados históricos e novo dado. Como supracitado os dados históricos são utilizados para treinamento, compostos pelos valores de entrada que são o *dataset* e pelos rótulos de dados que são necessários para proporcionar aprendizagem ao algoritmo a partir do

modelo utilizado como exemplo. O bloco de novo dado representa o conjunto teste, ou conjunto de validação, utilizado para verificar o comportamento do modelo aprendido ao utilizar dados novos para adequar possíveis situações não desejadas no aprendizado.

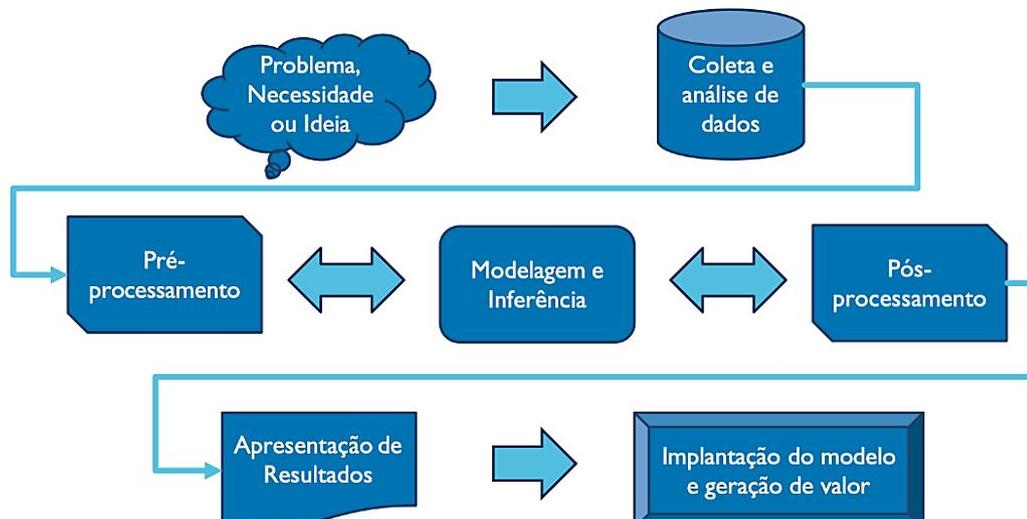
Figura 8- Esquema de modelo de aprendizado supervisionado.



Fonte: Escovendo & Koshiyama, 2020.

Dessa forma para que o modelo desenvolvido pelo algoritmo seja satisfatório ou seja, apresente resultados produtivos, faz-se necessário que o problema e os objetivos a serem alcançados sejam claros, visto que as etapas de coleta e análise de dados adequados serão resultantes de problemas bem definidos anteriormente. Como representado na Figura 9 caso o modelo desenvolvido não seja como esperado, todas as etapas anteriores à implantação do modelo são repetidas, seja para obtenção de mais dados ou mudança de estratégias para construção do aprendizado.

Figura 9- Esquema de aprendizado de máquina



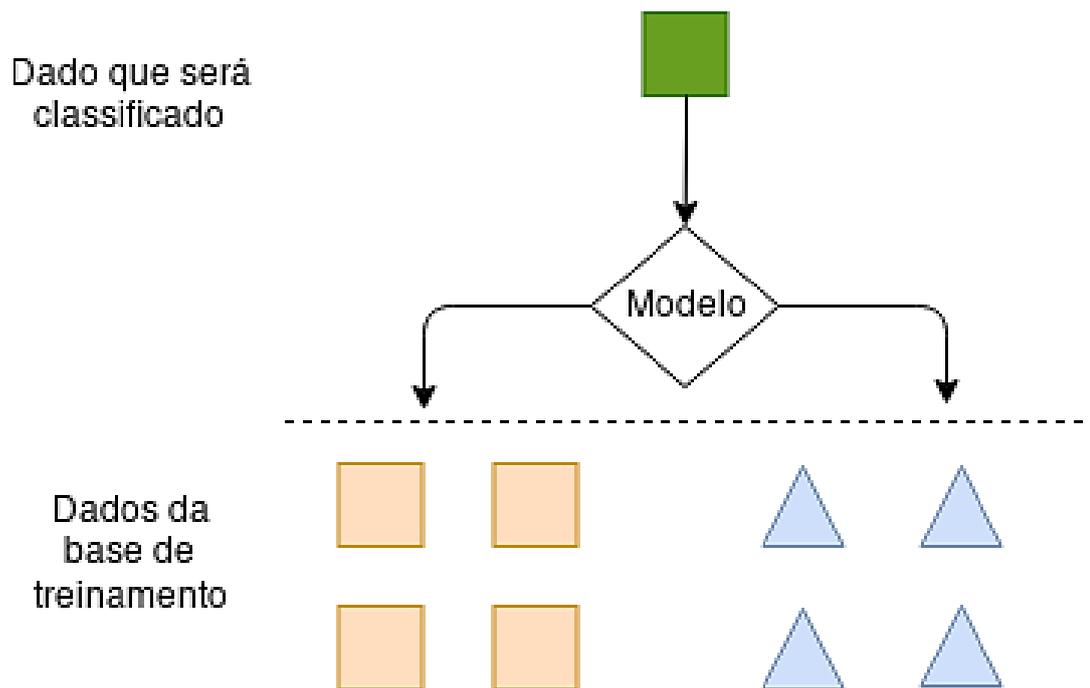
Fonte: Escovendo & Koshiyama, 2020.

Quando utilizado base de dados de treinamento com rótulos de classe discreta, o problema de aprendizagem supervisionada é denominado como classificação, quando há rótulos de classe contínua tem-se problema de regressão (MONARD *et al.*, 2003).

5.2 CLASSIFICAÇÃO

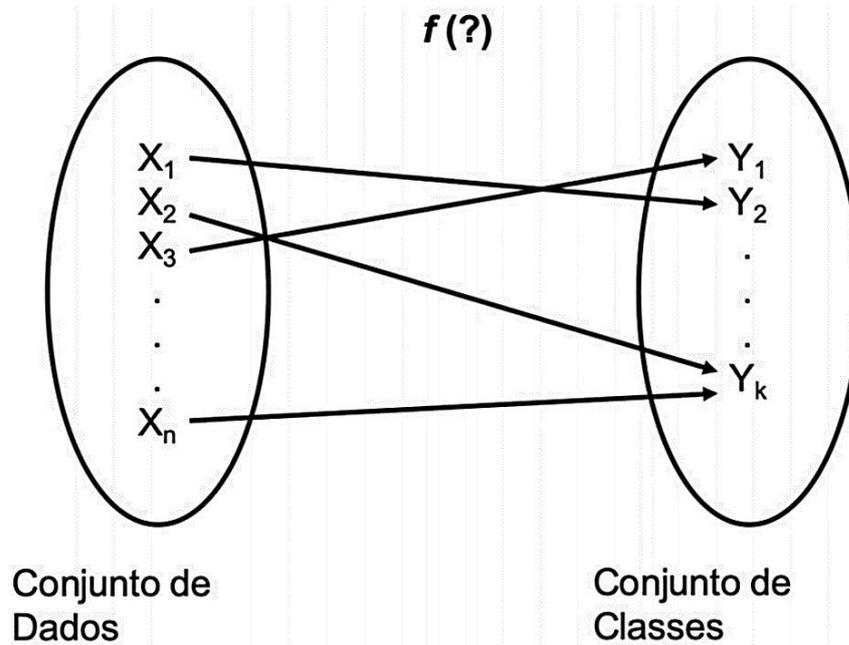
Os problemas de classificação provêm da necessidade de dividir, ou classificar, a base de dados utilizada em diferentes grupos, conforme ilustrado na Figura 10. Dessa forma a classificação é constituída por um algoritmo que visa aprender como associar corretamente todas as entradas em suas respectivas saídas, em Escovendo & Koshiyama (2020) se faz uma definição informal da classificação como a procura por uma função matemática que realize o mapeamento das entradas a uma saída rotulada determinada classe.

Figura 10- Problema de classificação



Fonte: Moura.

Figura 11- Mapeamento das entradas a partir de função matemática



Fonte: Escovendo & Koshiyama, 2020.

A Figura 11 expressa a representação da associação feita entre os dados de entrada e as classes na saída através da função matemática determinada pelo algoritmo durante o aprendizado. Para que exista eficácia na aprendizagem do algoritmo se faz necessário o treinamento do modelo desenvolvido, como citado anteriormente há a base de dados para treinamento e a base de dados novas utilizadas para teste.

A base de treino possibilita que o classificador aprenda os parâmetros utilizados e aprenda como mapear futuras entradas, enquanto a base teste possibilita determinar a qualidade do modelo desenvolvido a partir de medição de erros e acertos ao realizar classificação de dados não vistos. Para estimar a qualidade de um modelo é necessário obter o desempenho da classificação de forma detalhada, onde deve ser determinado para cada saída o número de classificações corretas.

A matriz de confusão é muito utilizada para estimar performance da classificação (LUZ). Essa matriz permite visualizar quantas entradas foram classificadas de forma correta ou errônea, auxiliando a regular parâmetros e entender qual classe está apresentando melhores resultados. A contagem de classificações corretas ou erradas é representada em índices denominados falso positivo (FP), falso negativo (FN), verdadeiro positivo (TP) e verdadeiro negativo (TN), conforme apresentado na Figura 12.

Figura 12- Matriz de confusão

		P R E D I T O	
		POSITIVO	NEGATIVO
R E A L	POSITIVO	 TP verdadeiro positivo	 FN falso negativo
	NEGATIVO	 FP falso positivo	 TN verdadeiro negativo

Fonte: Luz.

Os números de falso positivo (FP) e falso negativo (FN) representam classificação realizada de forma incorreta, visto que ocorrem respectivamente quando o resultado esperado é negativo, mas o modelo classifica como positivo e quando o resultado esperado é positivo, mas o modelo classifica como negativo. Já o verdadeiro positivo (TP) ocorre quando os resultados esperado e predito são positivos, enquanto o verdadeiro negativo (TN) quando os resultados esperado e predito são negativos, indicando classificação correta.

5.2.1 Algoritmo Knn

Por possuir alta aplicabilidade e processo de aprendizado simples, o algoritmo KNN é um dos mais utilizados em AM. Ao ser utilizado para classificação em problemas nas áreas de mercado, saúde e ciências políticas e sociais, esse algoritmo determina em qual grupo cada registro de entrada faz parte, assim é realizada análise a partir de base de dados para gerar a informação desejada (LUZ).

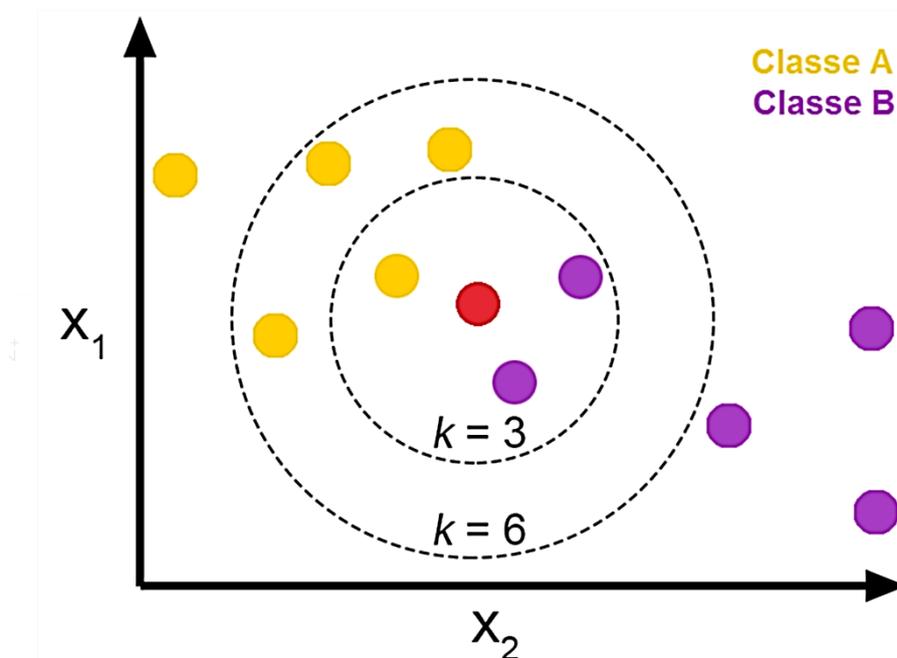
Dessa forma a estrutura do modelo de aprendizado é determinada pela entrada utilizada, onde não há divisão entre dados de treinamento e de teste, visto que o algoritmo não realiza treinamento para aprender e todos os dados de entrada serão utilizados somente na fase de teste. Por possuir essa peculiaridade, o algoritmo KNN é conhecido pelo termo “*lazy*” pois apesar do tempo inicial ser reduzido, por não haver treinamento, o tempo de teste corre de forma muito lenta devido a necessidade de análise e validação com maior atenção acerca dos resultados obtidos.

A sigla KNN provem de *K-Nearest Neighbours* que em português significa K-Vizinhos mais próximos. O significado é devido o algoritmo ser um classificador com aprendizado baseado em analisar a similaridade do *dataset* vizinho, ou seja, quão similar um dado de entrada é do dado vizinho. De acordo com Escovendo & Koshiyama (2020), a expressão popular “diga-me com quem andas e eu te direi quem tu és” ilustra a ideia principal do modelo de aprendizado desenvolvido pelo algoritmo KNN.

O algoritmo KNN funciona a partir do cálculo matemático da distância entre a amostra que se deseja classificar e os demais dados do *dataset* que já estão classificados, onde a quantidade utilizada de dados classificados é representada por K, que são os vizinhos da amostra de estudo.

A Figura 13 demonstra dois exemplos para uso da variável K ao considerar dois grupos de entrada, bolas amarelas e roxas, e um dado não classificado que é a bola vermelha. Quando K é igual a 3, será realizado cálculo da distância somente entre a bola vermelha e as três amostras mais próximas, no entanto quando K é igual a 6, será realizado cálculo da distância entre a bola vermelha e as 6 amostras mais próximas.

Figura 13- Diferentes valores da variável K para classificação de amostra através do algoritmo KNN



Fonte: Luz.

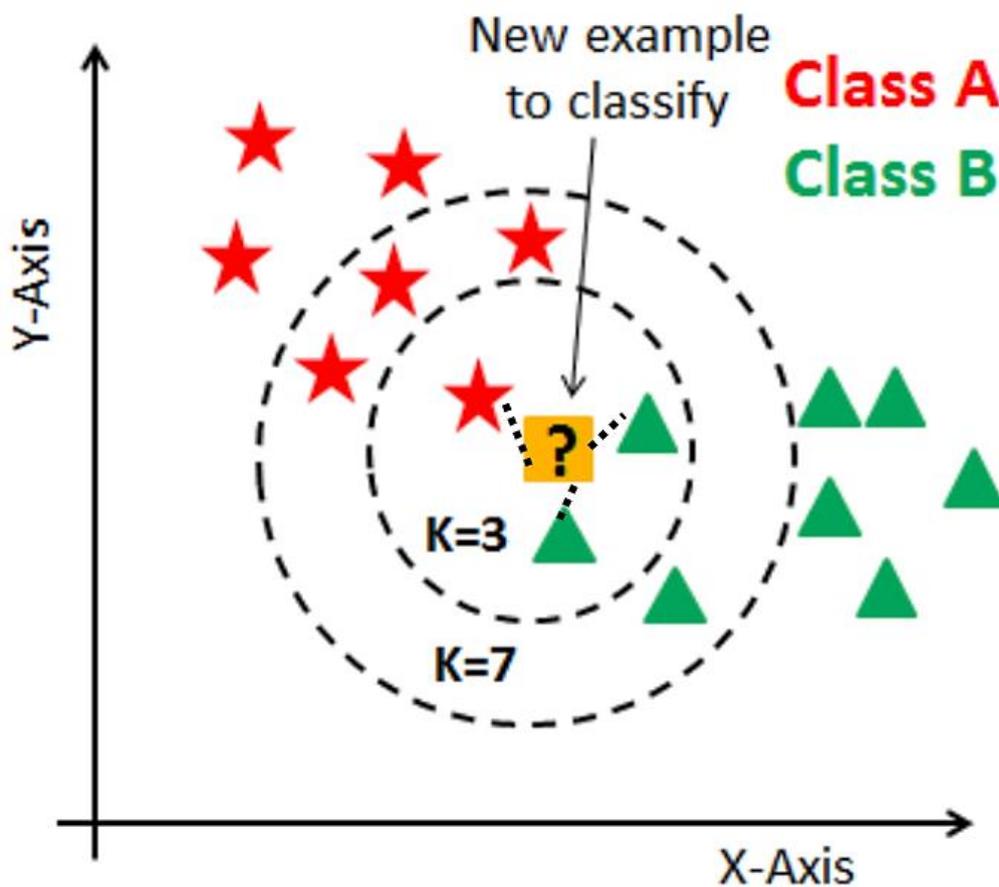
Para o algoritmo KNN realizar o cálculo matemático da distância entre dois pontos pode-se utilizar a distância Euclidiana, de Hamming, Manhattan ou Minkowski. No entanto, a mais utilizada na literatura é a distância euclidiana, representada abaixo pela equação (6).

$$d = \sqrt{\sum_{k=1}^n (p_{ik} - p_{jk})^2}$$

A equação (6) realiza cálculo da distância d entre os pontos p_i e p_j , onde n é a quantidade total de atributos definidos para a variável K .

Dessa forma após definir o valor utilizado na variável K , o algoritmo calcula a distância entre o dado que precisa ser classificado e todos seus dados vizinhos mais próximos que estão dentro do grupo determinado pelo valor de K .

Figura 14- Classificação de dado por algoritmo KNN



Fonte: Luz. (Modificada)

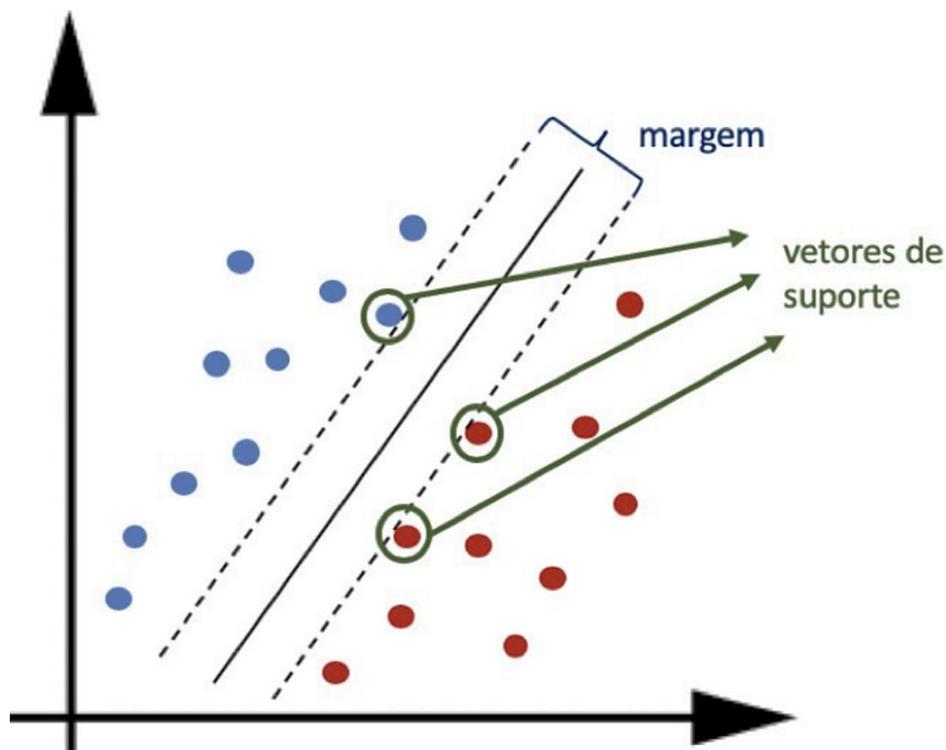
Ao considerar o exemplo da Figura 14 para $K=3$, tem-se que o algoritmo irá calcular a distância entre o dado de estudo e somente três dados do *dataset*, após realizado o cálculo é feita análise de qual classe possuiu distâncias menores do dado de estudo, se por exemplo a classe B possui mais dados com distâncias menores, então o algoritmo classifica o dado como classe B, visto que ele está mais próximo dos dados dessa classe.

5.2.2 Algoritmo SVM

O algoritmo *Support Vector Machine* (SVM), ou Máquina de Vetores de Suporte, é um algoritmo de aprendizado supervisionado utilizado para resolver problemas de classificação. O SVM possui treinamento mais lento pois essa é a etapa de maior foco do algoritmo, por conseguinte ele exige menos ajustes durante a realização dos testes e é caracterizado como uma técnica forte e poderosa para criação de modelos de classificação (ESCOVENDO & KOSHIYAMA, 2020).

O algoritmo SVM utiliza vetores de suporte para determinar um hiperplano de separação entre duas classes diferentes, a classificação é executada com o hiperplano que possui maiores dimensões, ou seja, aquele que mais diferencia as classes. Para que o SVM seja executado corretamente é preciso uso dos vetores de suporte e das margens, conforme apresentado na Figura 15 os vetores de suporte são formados por coordenadas dos dados que melhor segregam as duas classes, enquanto as margens são definidas a partir dos vetores de suporte.

Figura 15- Demonstração dos vetores de suporte e margem utilizadas para algoritmo SVM.

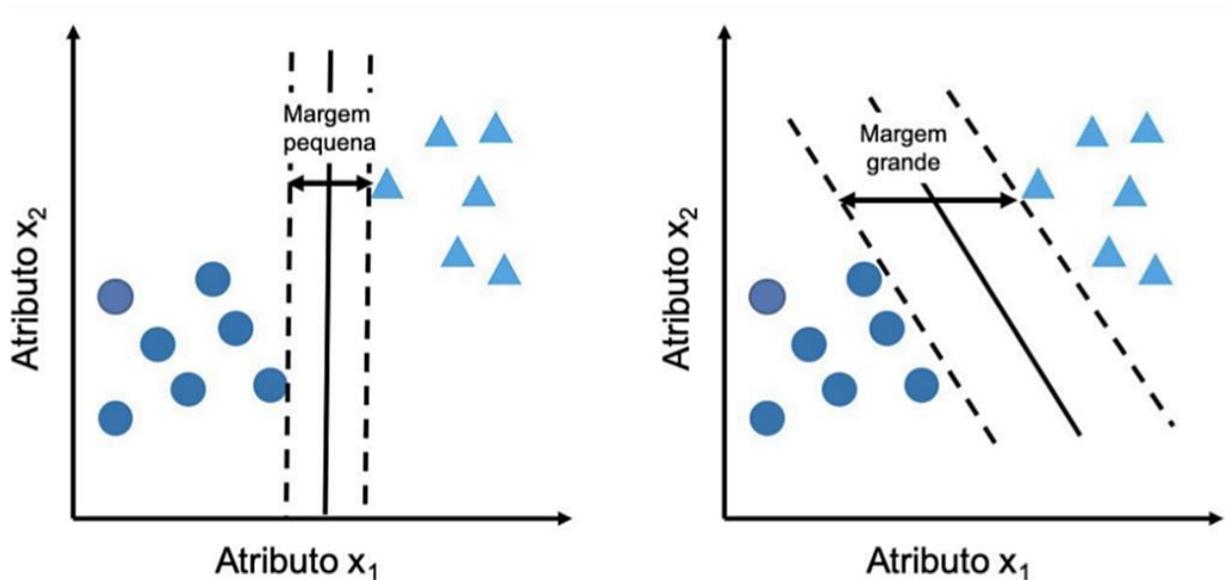


Fonte: Escovendo & Koshiyama (2020)

As linhas pontilhadas representam o local em que estão os vetores de suporte e a margem é a distância entre as duas linhas pontilhadas. Como existem diversas retas que dividem os

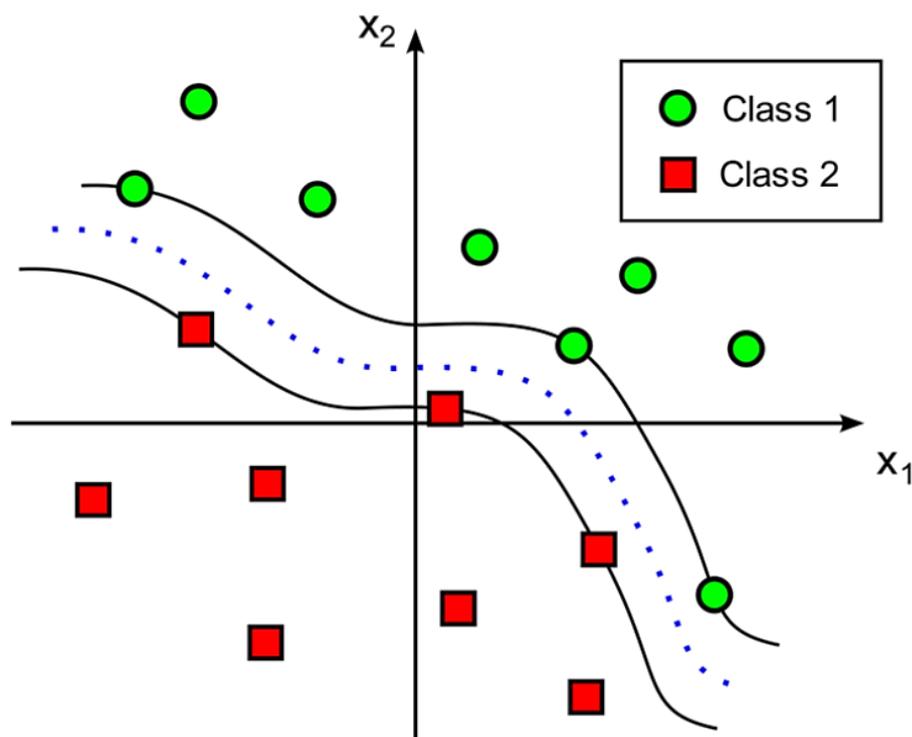
dados em duas classes, cabe ao algoritmo SVM escolher qual reta utilizar, nesse caso conforme a Figura 16, há também diversas margens com tamanhos diferentes. Dessa forma é escolhido o classificador que possui maior margem denominado classificador de margem máxima, que está representado à direita na Figura 16.

Figura 16- Tamanhos diferentes de margem ao considerar o mesmo problema.



Fonte: Escovendo & Koshiyama (2020)

As Figuras 15 e 16 representam o algoritmo SVM realizando classificação linear, visto que os grupos podem ser separados por uma reta. Entretanto existem conjuntos de dados em que não é possível realizar separação de grupos utilizando somente hiperplanos, nesses casos é necessário utilizar SVM não linear para dividir as duas classes. A Figura 17 representa o algoritmo de forma não linear, caracterizado por linhas curvas traçadas de forma a delimitar as classes com máxima margem.

Figura 17- Classificação não linear do algoritmo SVM.

Fonte: Medium.

5.2 APRENDIZADO NÃO SUPERVISIONADO

O aprendizado não supervisionado difere do supervisionado principalmente quanto à intervenção humana no conjunto de dados, visto que o termo supervisão está associado a necessidade de dados já rotulados para realizar treinamento ou seja, os algoritmos de aprendizagem supervisionada só conseguem desenvolver um modelo de aprendizado a partir de treinamento com dados de entradas já pertencentes a uma determinada classe. Assim, como o rótulo é uma informação gerada pelo homem, diz-se que há intervenção ou supervisão externa.

Dessa maneira a aprendizagem não supervisionada busca padrões e realiza agrupamento de dados conforme presença ou ausência de semelhanças, sem saber de fato do que se trata as classes criadas ou os dados utilizados. Como não há dependência de rótulos o aprendizado desenvolvido busca por padrões previamente desconhecidos, conseqüentemente os resultados alcançados possuem análise complexa devido a difícil interpretação sobre o que o algoritmo pretende determinar.

Dessa forma o método não supervisionado é comumente aplicado para realizar pré-processamento de dados rotulados, visto que ao utilizar dados com padrões previamente conhecidos há um aprendizado que auxilia na descoberta de novas relações entre eles,

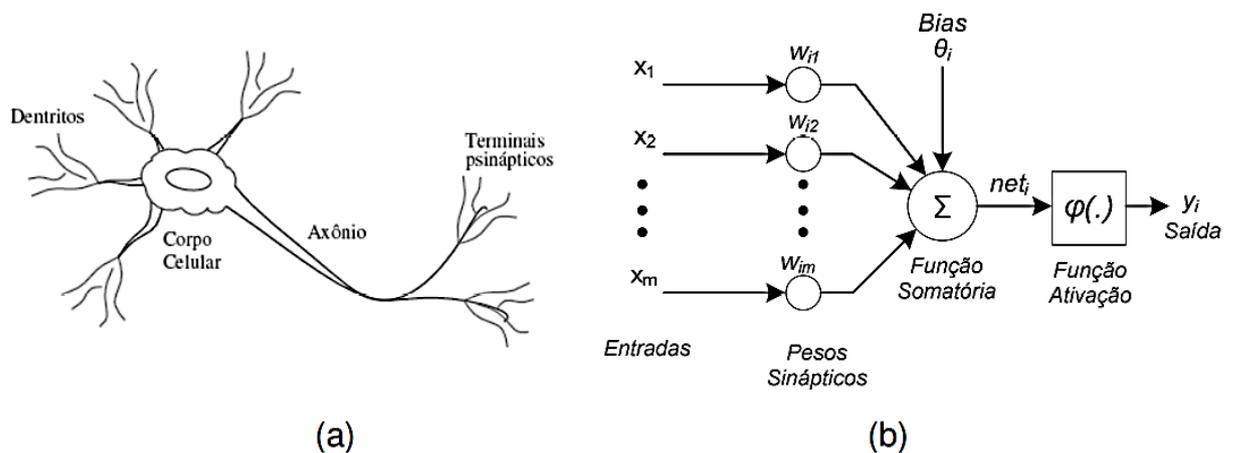
possibilitando uma maior margem para análise do problema em estudo.

5.3 REDES NEURAI ARTIFICIAIS

Redes neurais artificiais (ANNs) ou redes neurais simuladas (SNNs) são estruturas artificiais compostas pela conexão de várias unidades de processamento que refletem o comportamento de neurônios biológicos, em resumo as redes neurais possibilitam que computadores encontrem padrões e solucionem problemas comuns na inteligência artificial, aprendizado de máquina e *deep learning*.

Assim como a sua nomenclatura a estrutura de uma rede neural artificial também é inspirada no cérebro humano pois simula o envio de sinais entre neurônios biológicos, como ilustrado na Figura 18, tem-se que o modelo de neurônio artificial representado em (b) simula o comportamento do neurônio biológico em (a). No neurônio biológico há os dentritos na função de entrada e axônio na saída, no neurônio artificial quem realiza essas funções são o conjunto de entradas X_m e a saída Y_i , respectivamente, já os pesos sinápticos visam armazenamento dos conhecimentos ganhos pela rede.

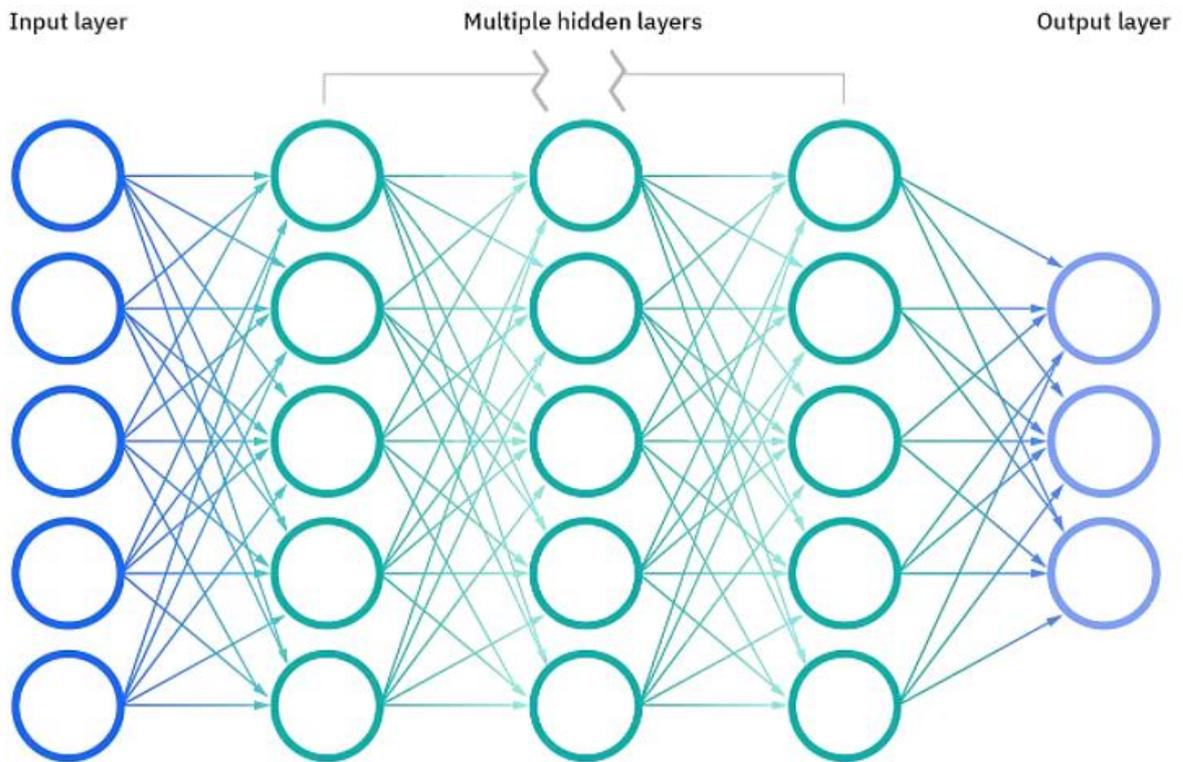
Figura 18- Neurônio biológico em (a) e Neurônio artificial em (b).



Fonte: Rede Neurais Auto-associativas.

Conforme a Figura 19, uma rede neural artificial é um modelo composto resumidamente em camada de entrada, presença ou não de camadas intermediárias e camada de saída, que estão altamente interligadas através de neurônios. Por existir variações na quantidade e em como essas camadas são conectadas, existem diferentes classificações que possuem uma ou múltiplas camadas ocultas visando atender o tipo de aplicação e complexidade do problema abordado.

Figura 19- Estrutura de uma rede neural artificial.

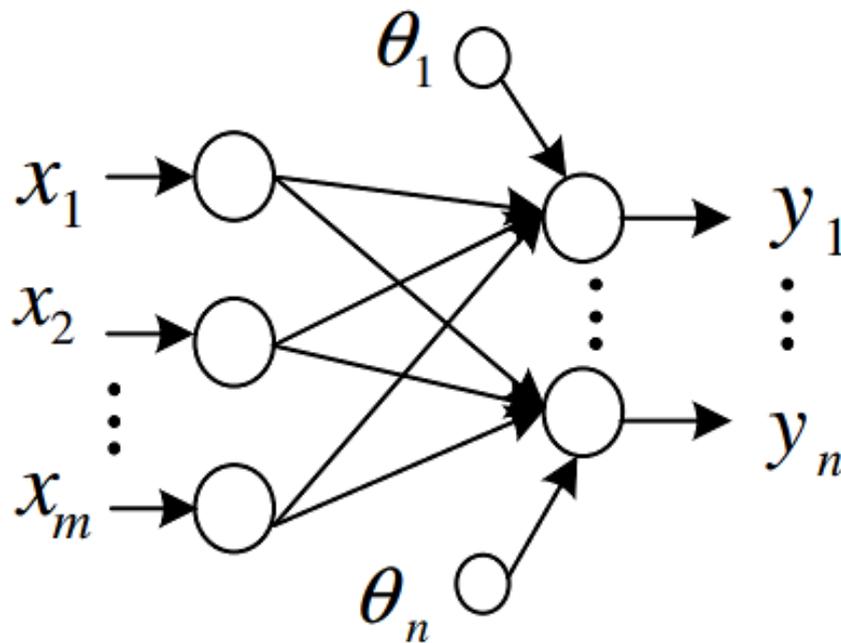


Fonte: IBM Cloud Education.

5.3.1 Rede neural de camada única

É uma rede neural com todos os neurônios organizados em uma única camada, dessa forma cada padrão de treinamento aplicado aos neurônios é associado diretamente aos padrões de saída da rede, conseqüentemente não há representações intermediárias e o sistema se torna incapaz de mapear informações importantes. A Figura 20 apresenta a rede neural de camada única, considerada a rede mais simples pois foi o primeiro modelo de rede neural desenvolvido. Além das limitações citadas tem-se também a incapacidade da rede aprender problemas que não são linearmente separáveis (REDE NEURAIAS AUTO-ASSOCIATIVAS).

Figura 20- Rede neural de camada única.



Fonte: Rede Neurais Auto-Associativas.

5.3.2 Rede neural de múltiplas camadas

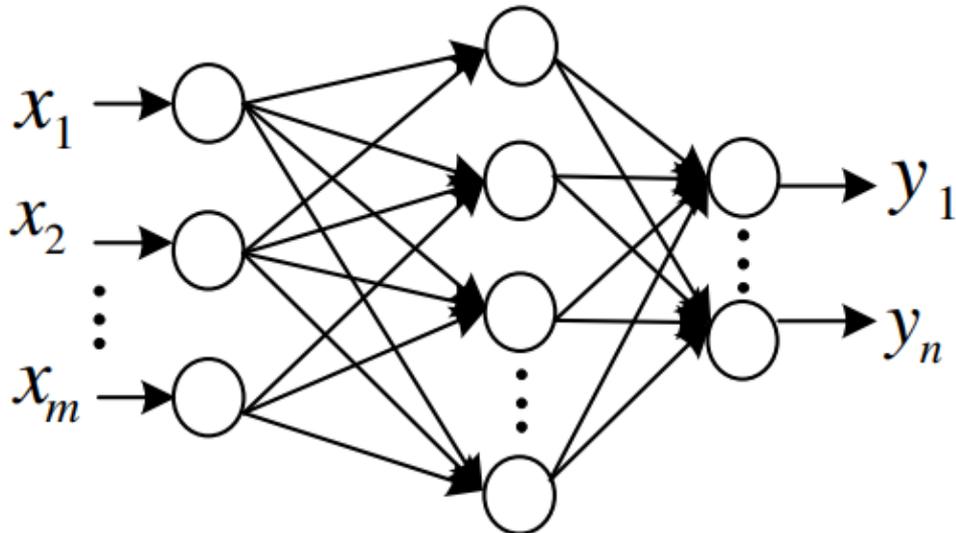
A rede de múltiplas camadas é conhecida como MLP (*Multilayer Perceptron*), esse termo é devido a existência de uma ou mais camadas ocultas, ou intermediárias, entre camada de entrada e de saída (REDE NEURAS AUTO-ASSOCIATIVAS). A camada oculta funciona como sintetizadora de características, visto que é responsável por codificar os atributos apresentadas nos padrões de entrada e enviar para a saída uma representação mais complexa desses padrões, dessa forma a camada de saída recebe esses estímulos e constrói padrões mais complexos e com menor taxa de erro.

O treinamento da rede MLP é realizado pelo algoritmo de retropropagação (*backpropagation*), que funciona por meio de uma sequência de dois passos. Como citado anteriormente o primeiro passo é apresentação de um padrão à camada de entrada, posteriormente o fluxo de estímulos chega nas camadas intermediárias e por fim na camada de saída. O segundo passo é a procura por padrão de saída com erro mínimo, onde a saída adquirida é comparada à saída desejada através do cálculo do erro. O valor do erro é propagado para a camada de entrada ocasionando mudança nos padrões e pesos das conexões nas camadas

ocultas, essas mudanças ocorrem à medida que o valor do erro é retropropagado.

Após a rede MLP encontrar valor de erro satisfatório ela é considerada uma rede treinada e já pode ser utilizada como ferramenta para classificação de novos dados. A Figura 21 apresenta o modelo da rede neural MLP.

Figura 21- Rede neural MLP

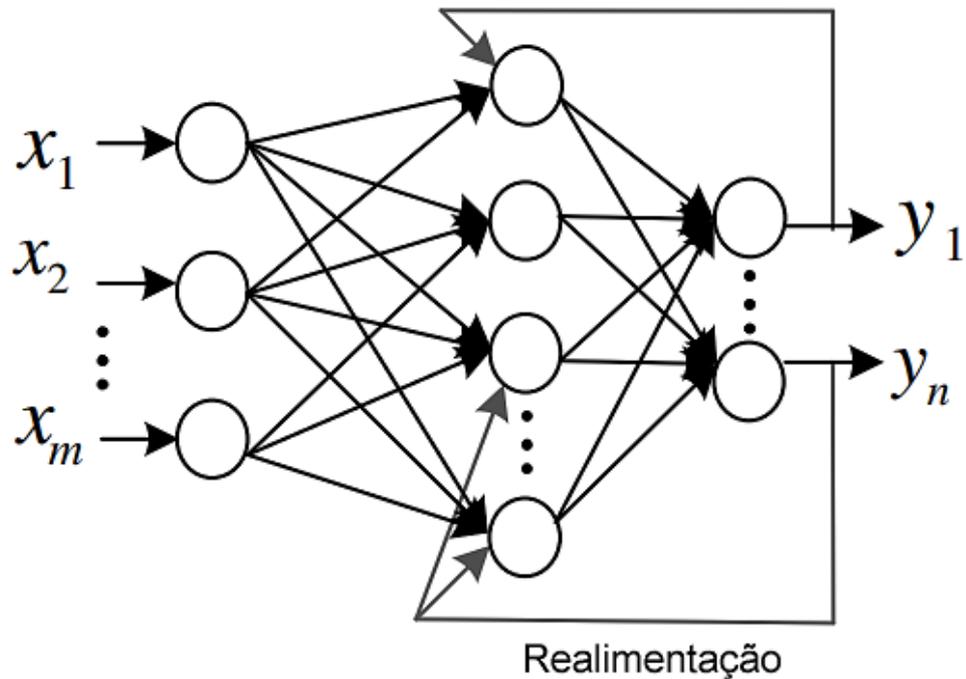


Fonte: Rede Neurais Auto-Associativas.

5.3.3 Rede neural recorrente

Redes neurais recorrentes são também conhecidas como redes realimentadas, visto que são identificadas por possuir saídas com *loops* de *feedback* (IBM CLOUD EDUCATION), ou seja, as camadas de entradas intermediárias dos neurônios são realimentadas pelos sinais de saída da rede, como mostra a Figura 22. Por possuir realimentação, as estruturas neurais recorrentes são qualificadas para uso em processamento dinâmico de informações, já que podem prever estados futuros ao utilizar dados de séries temporais (REDE NEURAIAS AUTO-ASSOCIATIVAS).

Figura 22- Rede neural recorrente



Fonte: Rede Neurais Auto-Associativas.

5.3.4 Análise de Componente Principais – PCA

PCA é um método matemático quantitativamente rigoroso que objetiva diminuir um conjunto de dados, onde o espaço originalmente ocupado por esses dados é transformado em um espaço menor por possui quantidade de dados reduzida (REDE NEURAIAS AUTO-ASSOCIATIVAS). Após essa compactação os dados resultantes desse processo são denominados componentes principais, apesar da redução da quantidade de dados originais e geração de um novo conjunto de variáveis, os componentes principais são capazes de realizar descrição eficaz do problema abordado (MATHWORKS, 2021).

A simplificação formada por componentes principais ocasiona descrição completa do problema pois é composta somente por informações não redundantes. Os componentes são resultantes da combinação e compactação das informações dos dados originais, ou seja, o espaço de características originais passa por uma redução de dimensão.

6 MATERIAIS E MÉTODOS

A metodologia adotada nessa dissertação usa como base a metodologia desenvolvida em Queiroz (*et al.*, 2017) e Silva, Queiroz e Barros (2020). Foram utilizados dois grupos diferentes de indivíduos para realização do exame de eletrocardiograma, há uma classe composta por portadores da SAOS (grau leve, moderado e grave) e outra composta por pessoas saudáveis.

6.1 BASE DE DADOS

As bases de dados utilizadas foram os sinais de eletrocardiograma de indivíduos saudáveis (*MIT-BIH Rhythm Sinus Normal*) e de indivíduos diagnosticados com SAOS (*Apnea-ECG Database*), ambos presentes no trabalho realizado em Goldberger (*et al.*, 2000). A base de dados apresenta gravações de portadores de SAOS com distúrbios respiratórios variando entre 14 e 82 minutos por hora, de acordo com Goldberger (*et al.*, 2000) foi mais apropriado utilizar minutos do que o índice IAH, pois os parâmetros utilizados foram ECG e frequência cardíaca.

As gravações de todos os indivíduos do estudo foram registradas por duas noites seguidas, onde o grupo de indivíduos saudáveis possui 18 gravações enquanto o grupo com SAOS contém número de gravações distribuídos conforme a Tabela 1.

Tabela 1- Número de gravações para cada grupo portador da SAOS

Grupo portador da SAOS	Número de gravações
Apneia em grau leve	20
Apneia em grau moderado	5
Apneia em grau grave	10

Fonte: Autora (2022)

As gravações possuem duração total que varia entre 401 e 578 minutos, dessa forma os grupos foram classificados de acordo com a quantidade de minutos gastos com distúrbios respiratórios.

O grupo C se caracteriza como apneia em grau leve, é composto por gravações com menos de 5 minutos de desordem respiratória, os indivíduos do grupo são seis pessoas do sexo masculino e cinco do sexo feminino com média de idade de 33 anos, variando entre 27 e 42

anos.

O grupo A apresentou distúrbios respiratórios com maior incidência e intervalos de tempo, caracterizando apneia em grau grave, onde todos os indivíduos que compõe esse grupo possuem 100 ou mais minutos de respiração desordenada. Os indivíduos desse grupo são quinze pessoas do sexo masculino e uma do sexo feminino com média de idade de 50 anos, variando entre 29 e 63 anos.

O grupo B é formado por indivíduos com apneia em grau moderado, considerado um grupo limite ou intermediário, apresenta 10 a 96 minutos de distúrbios respiratório e possui maior incidência de desordem que o grupo C e menor que o grupo A. Os indivíduos desse grupo são quatro pessoas do sexo masculino e uma do sexo feminino com média de idade de 46 anos, variando entre 39 e 53 anos.

6.2 SEGMENTAÇÃO E PRÉ-PROCESSAMENTO DOS CICLOS CARDÍACOS

A partir das gravações de ECG obtidas, os sinais foram pré-processados para que fosse possível a criação de matrizes contendo os batimentos cardíacos. Dessa forma os batimentos de cada indivíduo foram retirados das gravações dos sinais de ECG e amostrados em uma frequência de 128 Hz, para ambos os grupos, como proposto em Queiroz (*et al.*, 2017). Para a amostragem em 128 Hz foi utilizada a mesma quantidade de gravações da Tabela 1 para portadores da SAOS e somente 8 gravações para indivíduos saudáveis.

Cada sinal de ECG foi segmentado nas suas respectivas ondas P, Q, R, S e T a fim de se obter cada etapa dos batimentos separada individualmente. Em seguida, foi realizado o procedimento presente na equação (7), onde a média dos batimentos foi subtraída do sinal e esse resultado foi dividido pela entropia de Shannon.

$$\mathbf{b}_{n,m} = \frac{\mathbf{b}_{n,m} - \frac{1}{N} \sum_1^N \mathbf{b}_{n,m}}{-\sum_1^n \mathbf{b}_{n,m} \left(\log_2 \frac{1}{N} \right)}, \quad (7)$$

Os batimentos cardíacos $\mathbf{b}_{n,m}$ dos indivíduos foram concatenados individualmente em matrizes, onde os batimentos dos indivíduos classificados como saudáveis foram agrupados na matriz A, enquanto a classe de indivíduos portadores da SAOS ficou na matriz B. A concatenação dos batimentos cardíacos ocorreu conforme ilustrado na equação (8):

$$B = \begin{bmatrix} \mathbf{b}_{1,1} & \mathbf{b}_{1,2} & \dots & \mathbf{b}_{1,n} \\ \mathbf{b}_{2,1} & \mathbf{b}_{2,2} & \dots & \mathbf{b}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{b}_{m,1} & \mathbf{b}_{m,2} & \dots & \mathbf{b}_{m,n} \end{bmatrix}, \quad (8)$$

Na matriz apresentada acima, o índice n representa o número de batimentos cardíacos e m representa o número total de coluna desses batimentos.

Após obtenção das matrizes de batimentos cardíacos para os indivíduos saudáveis e com SAOS, obtidas através da equação (8), precisou-se realizar uma redução de dimensionalidade das informações contidas nessas matrizes. Dessa forma, utilizou-se o método de análise de componentes principais (PCA) diretamente nas matrizes de batimentos, para ocasionar eliminação de ruídos, redução do conjunto de dados e maior facilidade para posterior classificação dos dados em diferentes grupos.

6.3 EXTRAÇÃO DE CARACTERÍSTICAS

Após aplicado o método de PCA nas matrizes de batimentos cardíacos, as matrizes obtidas como resultados foram aplicadas em funções estatísticas de alta ordem para extração de características. Foram utilizadas as funções mediana, assimetria e curtose, aqui novamente expressas nas equações (9), (10) e (11) respectivamente. A matriz de dados aleatórios é representada abaixo pela variável X .

$$M = \frac{X(n+1)}{2} \quad (9)$$

$$k_X = E[(X - E(X))\sigma^{-1}]^3 \quad (10)$$

$$\lambda_X = E[(X - E(X))\sigma^{-1}]^4 \quad (11)$$

Durante os testes a variância também foi utilizada como método estatístico, no entanto ao analisar os resultados as funções mediana, assimetria e curtose apresentaram melhor espaçamento e classificação das amostras do conjunto de dados, visto que o vetor obtido de cada estatística de alta ordem é utilizado como entrada dos classificadores.

6.4 CONFIGURAÇÕES DA MÁQUINA DE APRENDIZADO

Durante a etapa de classificação a presente dissertação utilizou máquina de aprendizado do tipo supervisionado, onde foram aplicados os algoritmos K-Vizinhos mais Próximos (KNN), Máquina de Vetores de Suporte (SVM) e Rede Neural de Múltiplas Camadas (MLP). A configuração utilizada para o algoritmo KNN foi de índice K igual a 10, para SVM aplicou-se as configurações de Kernel linear, em ambos os algoritmos foi aplicada classificação para doenças cardíacas.

A rede neural artificial MLP (*Multi Layer Perceptron*) utilizada possui duas camadas ocultas, para que ocorresse a otimização iterativa foi utilizado o algoritmo gradiente descendente. Dessa forma o treinamento da máquina de aprendizado conseguiu atingir valores de erro mínimo, na Tabela 2 há as configurações da rede neural usada.

Tabela 2- Configurações da rede neural artificial MLP

<i>Configurações</i>	<i>Valores</i>
Taxa de Aprendizado	0,0001
Número de iterações de treino	1000
Entradas	3
Número de neurônios por camada oculta	100
Número de neurônios na camada de saída	2

Fonte: Autora (2022)

As três entradas aplicadas como parâmetros na rede neural foram mediana, curtose e assimetria, representadas nas equações (9), (10) e (11), enquanto os parâmetros de saídas são duas classes que se dividem em indivíduos classificados como portadores da SAOS e classificados como saudáveis.

6.5 MÉTRICAS DE AVALIAÇÃO

Após a etapa de classificação, foram calculados os valores de acurácia equação (12), sensibilidade equação (13) e especificidade equação (14), utilizados como métricas de avaliação para verificar o desempenho dos classificadores.

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} \times 100 \quad (12)$$

$$Sensibilidade = \frac{VP}{VP+FN} \times 100 \quad (13)$$

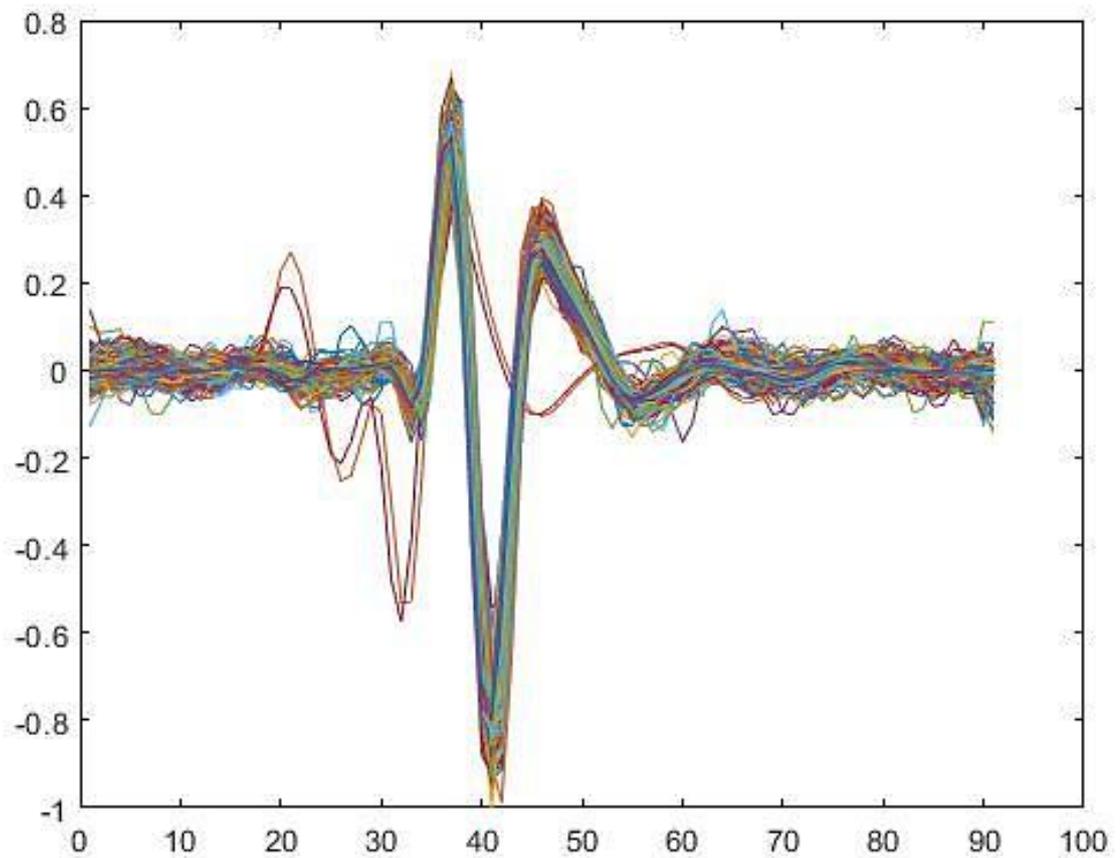
$$Especificidade = \frac{VN}{VN+FP} \times 100 \quad (14)$$

Onde VP representa a quantidade de positivos verdadeiros para presença de portadores da SAOS, VN negativos verdadeiros para ausência de portadores da SAOS, FP falsos positivos para gravações com presença de SAOS em indivíduos saudáveis e FN falsos negativos para classificações de indivíduos portadores da SAOS como saudáveis.

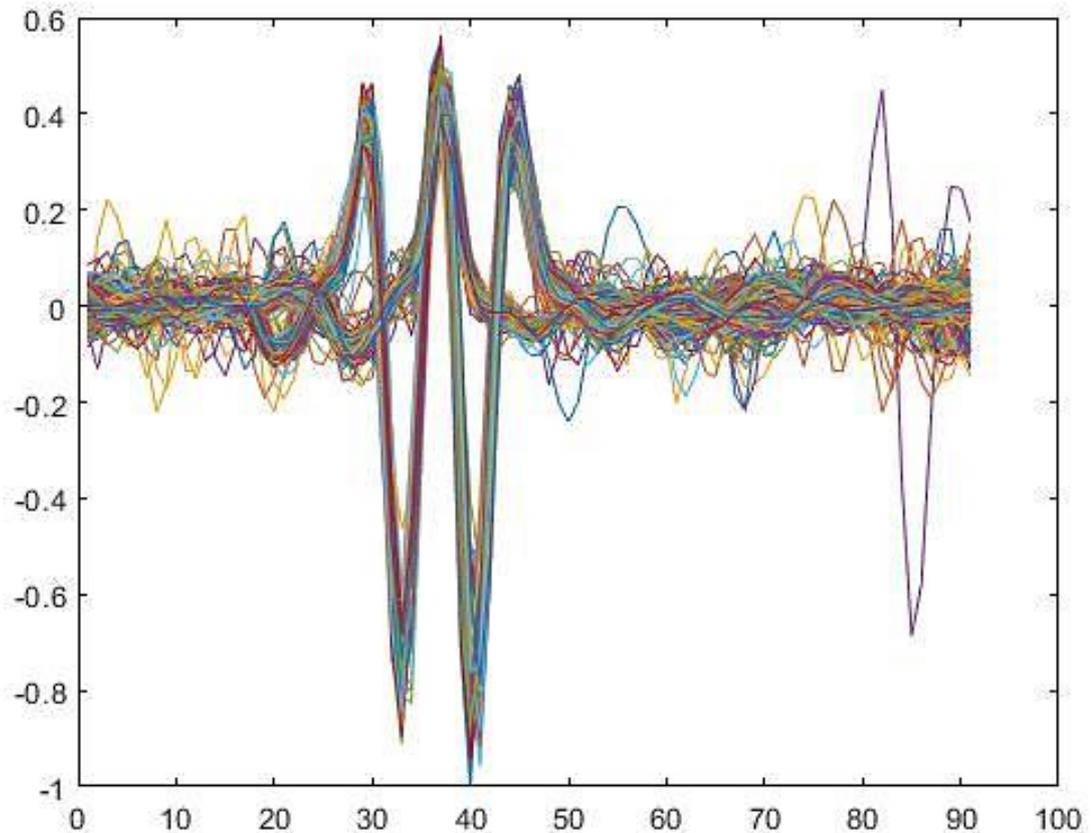
7 RESULTADOS

Os batimentos cardíacos de ambas as classes, indivíduos saudáveis e portadores da SAOS, foram classificados de duas formas: com e sem aplicação do método matemático PCA, a fim de analisar a amostragem dos dados de cada grupo. As Figuras 23 e 24 exibem, respectivamente, os batimentos cardíacos de um indivíduo saudável e de um portador de SAOS.

Figura 23- Batimentos cardíacos de indivíduo saudável



Fonte: Autora.

Figura 24- Batimentos cardíacos de indivíduo com SAOS

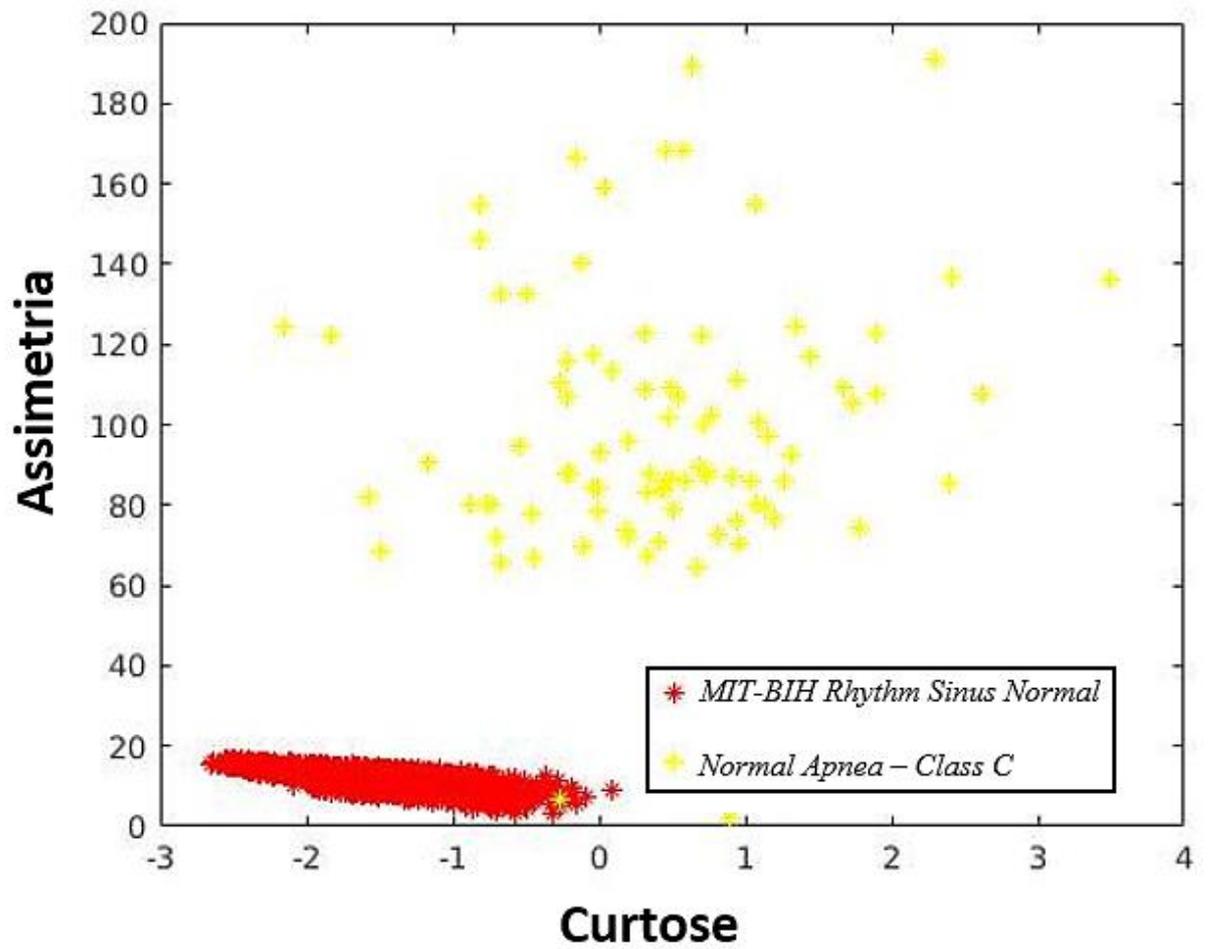
Fonte: Autora.

A classificação dos batimentos cardíacos foi realizada para SAOS em grau leve, moderado e grave, onde foi feita a análise dos gráficos em duas e três dimensões, plotados a partir dos valores estatísticos de mediana, assimetria e curtose.

7.1 SAOS EM GRAU LEVE COM PCA

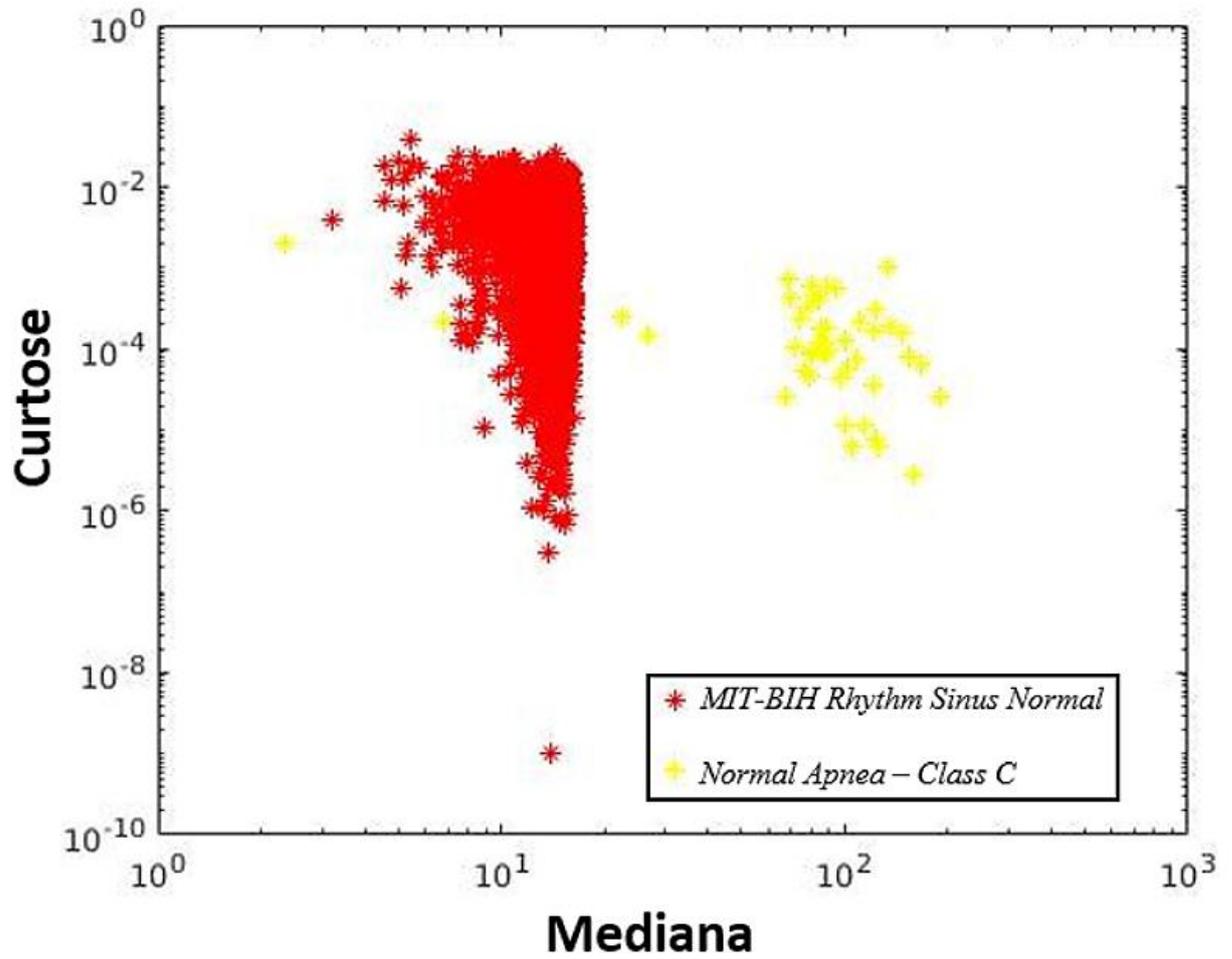
As Figuras 25, 26 e 27 apresentam os gráficos em 2 dimensões para SAOS em grau leve (grupo C). Onde a Figura 25 é um gráfico de dispersão de assimetria e curtose, a Figura 26 apresenta a dispersão de curtose e mediana e a Figura 27 é um gráfico de dispersão de assimetria e mediana.

Figura 25- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau leve



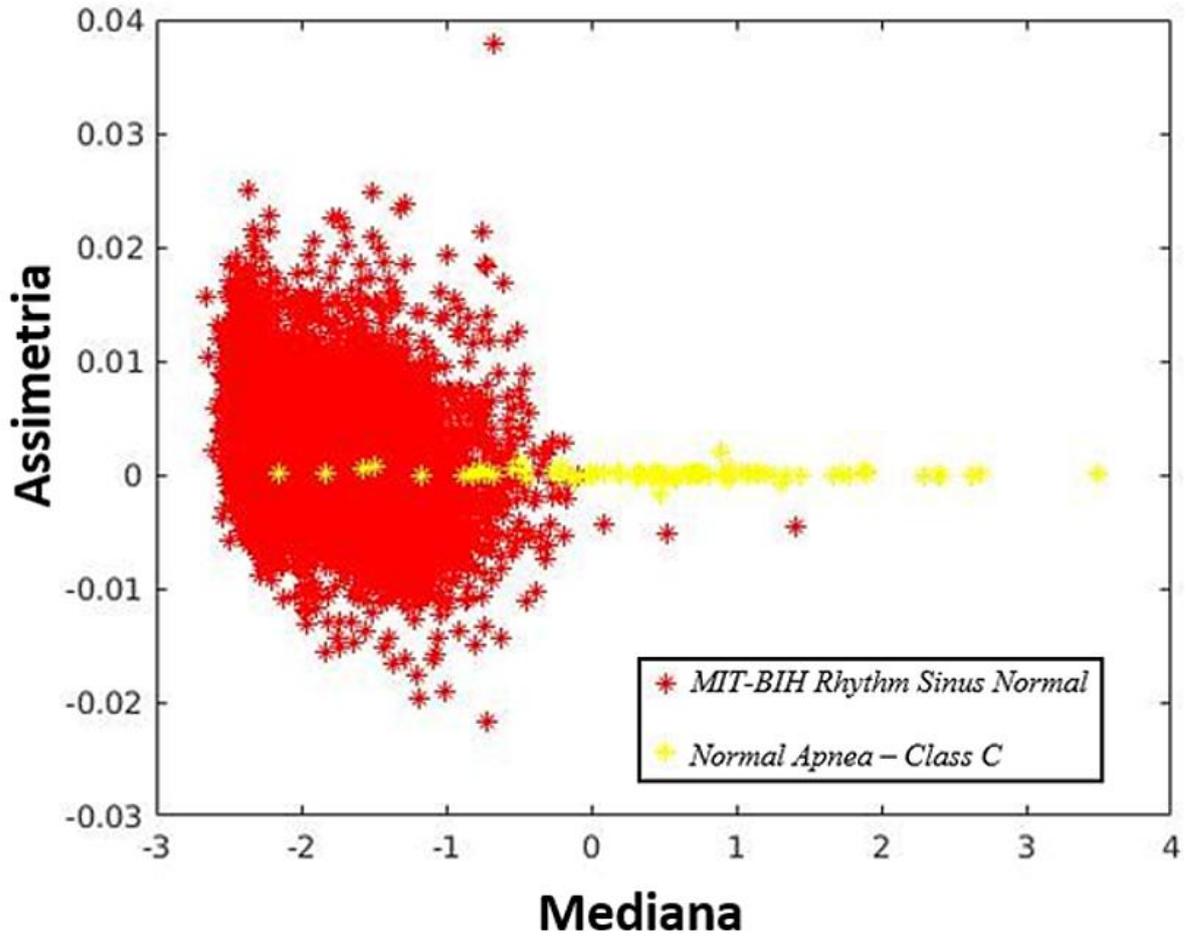
Fonte: Autora.

Figura 26- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve



Fonte: Autora

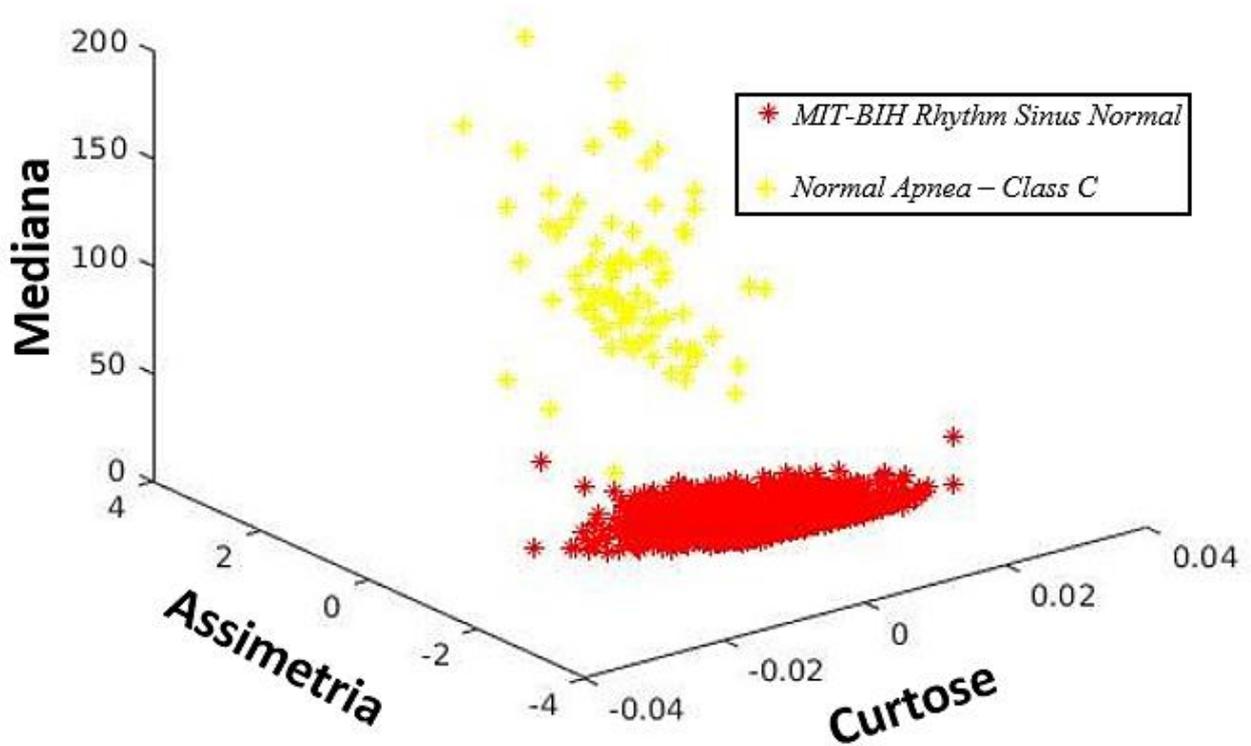
Figura 27- Representação em 2D da dispersão de assimetria e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve



Fonte: Autora

A Figura 28 apresenta o gráfico em três dimensões, utilizando a mediana, assimetria e curtose dos batimentos de cada grupo.

Figura 28- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau leve apresentados em espaço 3D a partir da mediana, assimetria e curtose.

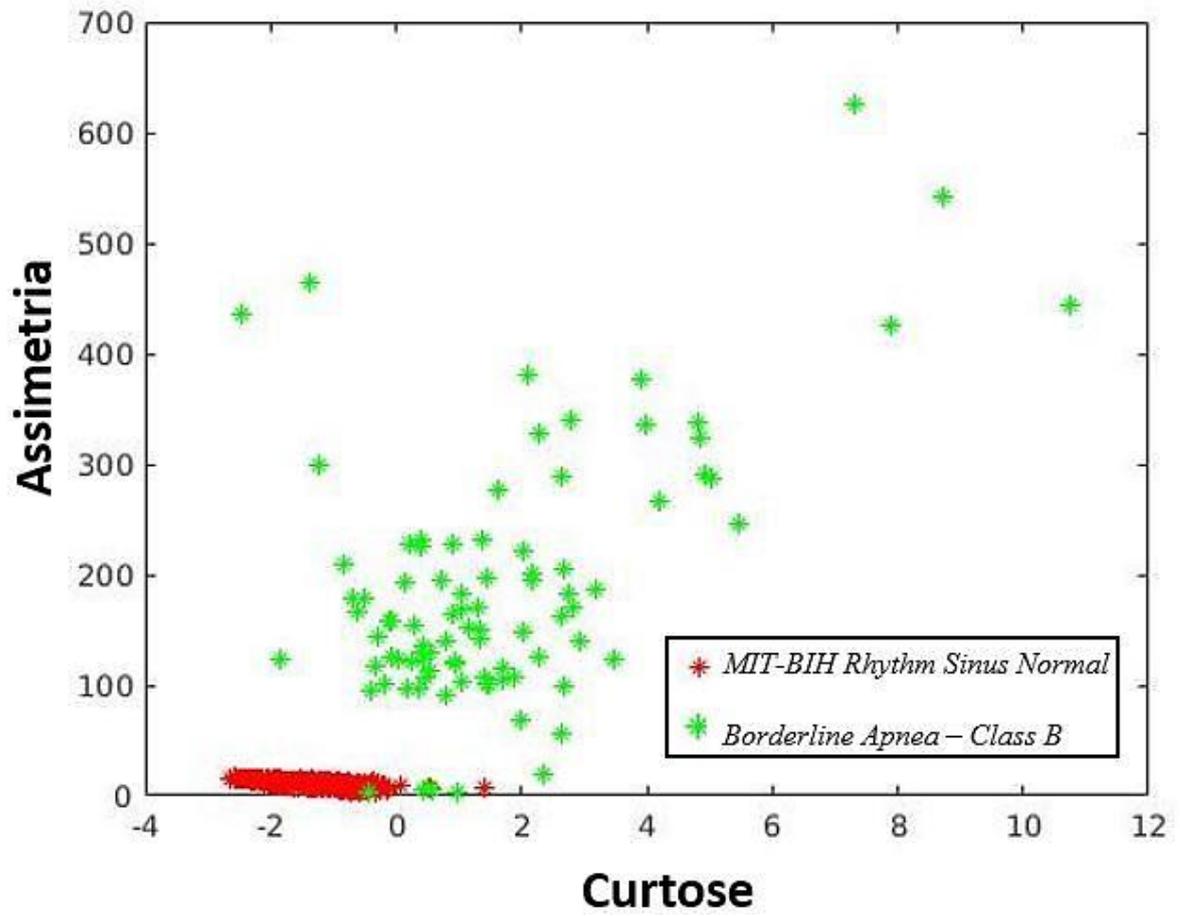


Fonte: Autora.

7.2 SAOS EM GRAU MODERADO COM PCA

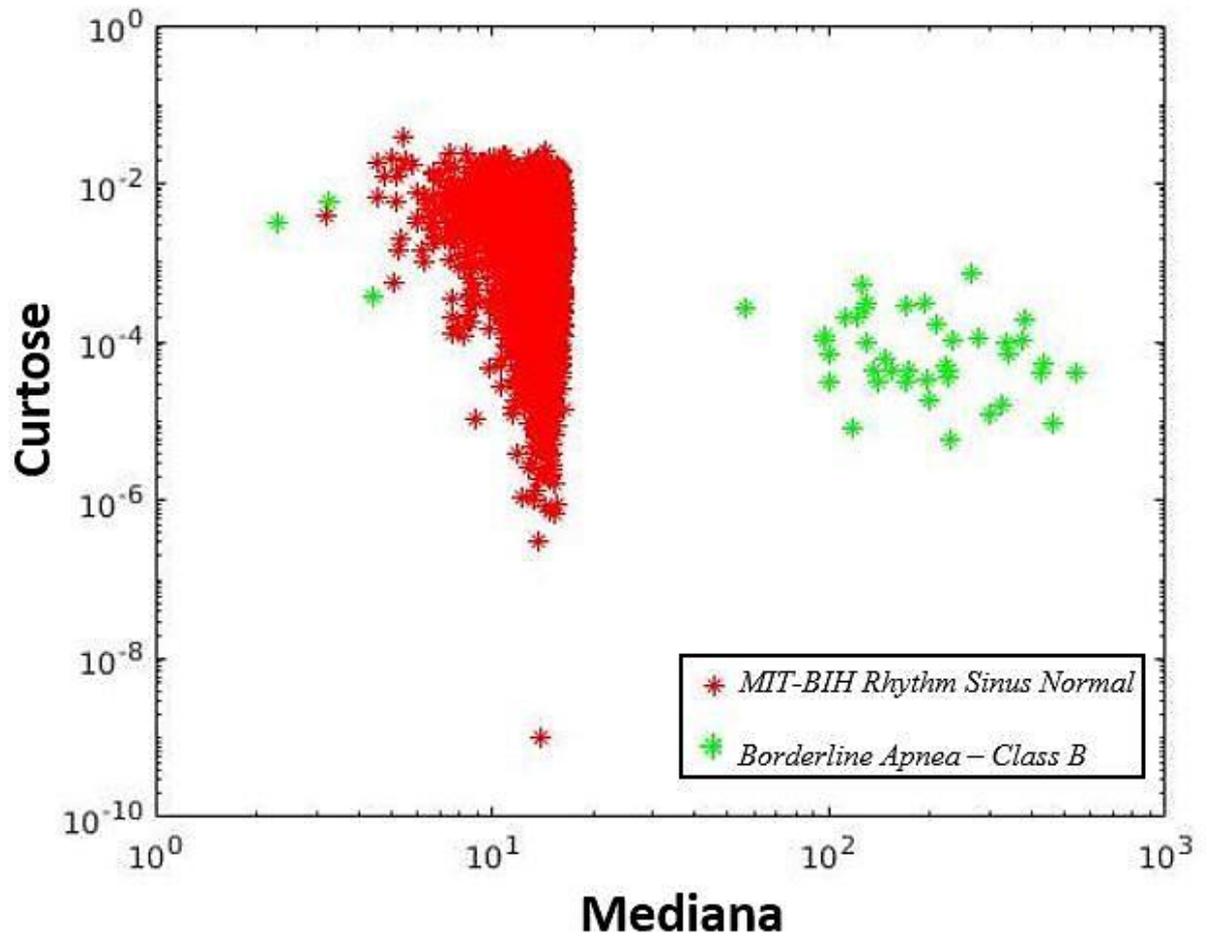
As Figuras 29, 30 e 31 apresentam os gráficos em 2 dimensões para SAOS em grau moderado (grupo B). Onde a Figura 29 é um gráfico de dispersão de assimetria e curtose, a Figura 30 apresenta a dispersão de curtose e mediana e a Figura 31 é um gráfico de dispersão de assimetria e mediana.

Figura 29- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau moderado



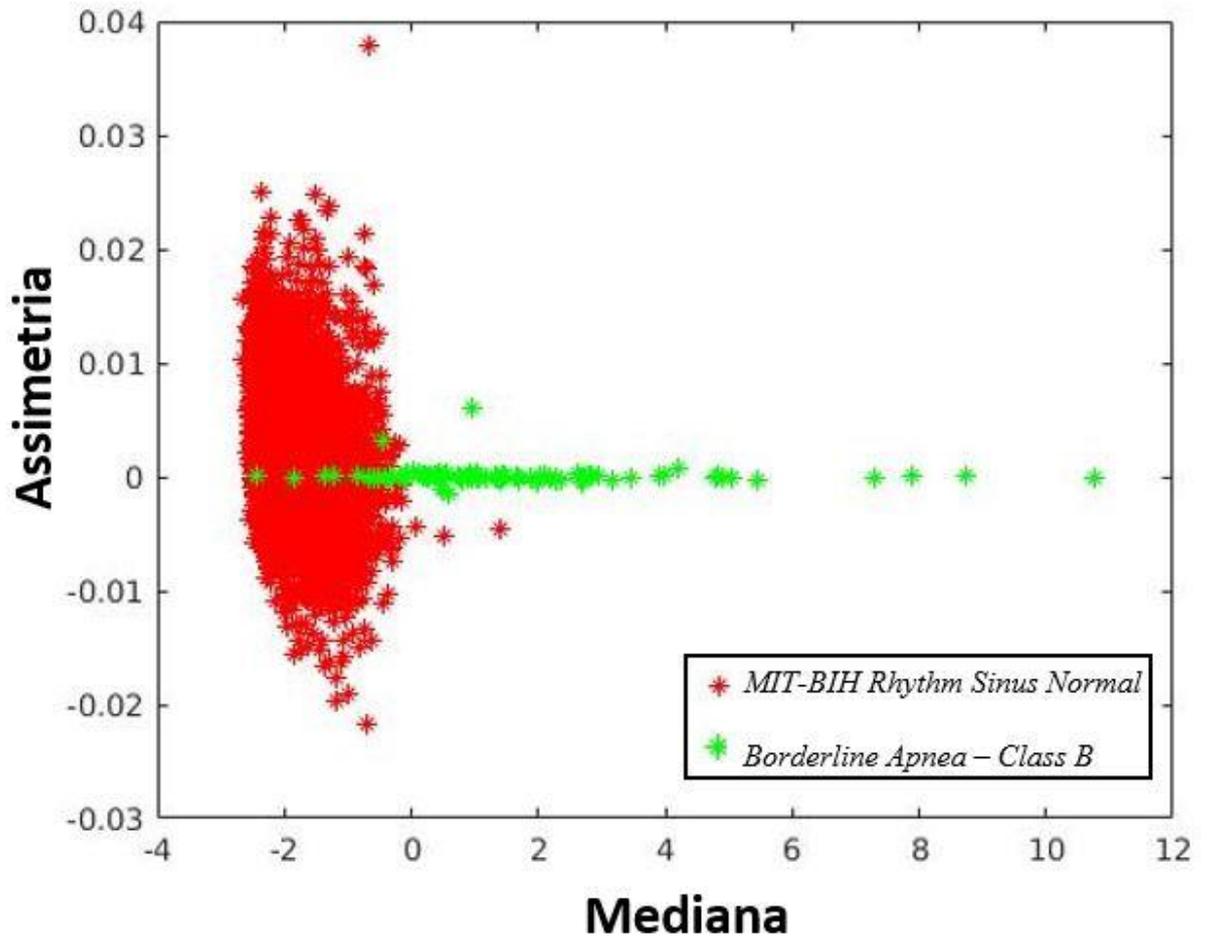
Fonte: Autora

Figura 30- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado



Fonte: Autora

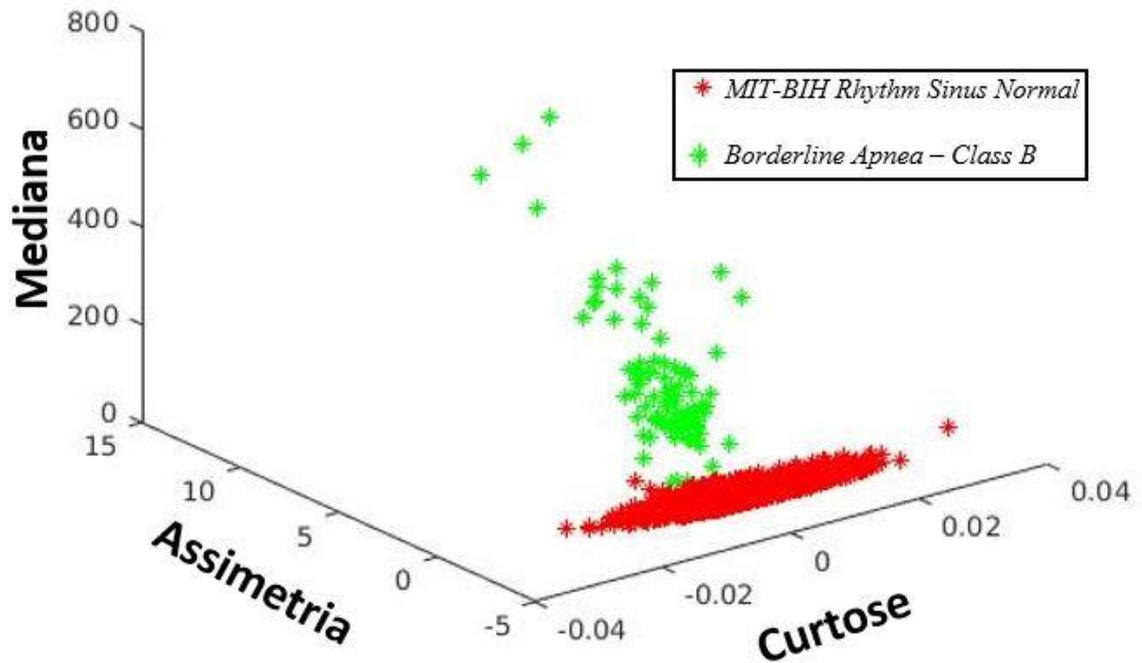
Figura 31- Representação em 2D da dispersão de assimetria e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado



Fonte: Autora.

A Figura 32 apresenta o gráfico em três dimensões, utilizando a mediana, assimetria e curtose dos batimentos de cada grupo.

Figura 32- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau moderado apresentados em espaço 3D a partir da mediana, assimetria e curtose

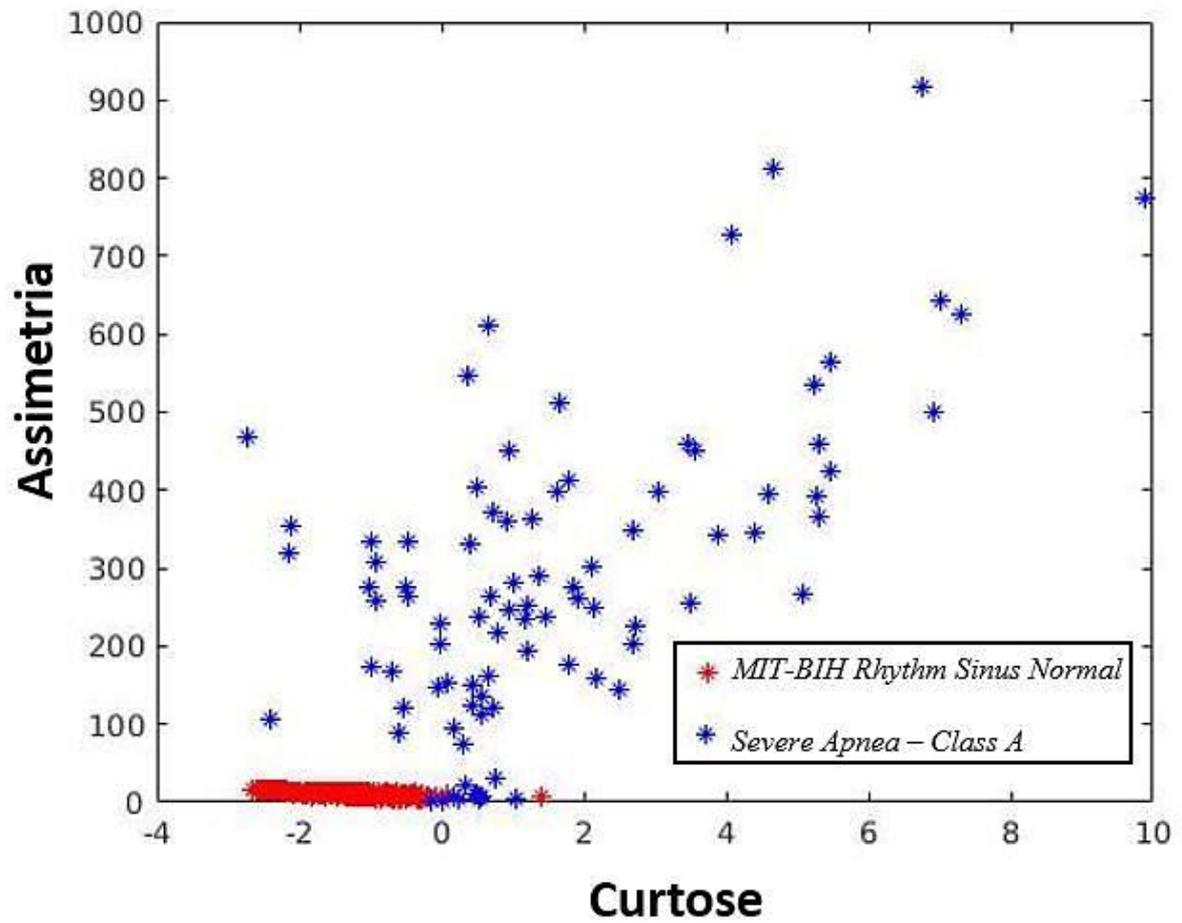


Fonte: Autora.

7.3 SAOS EM GRAU GRAVE COM PCA

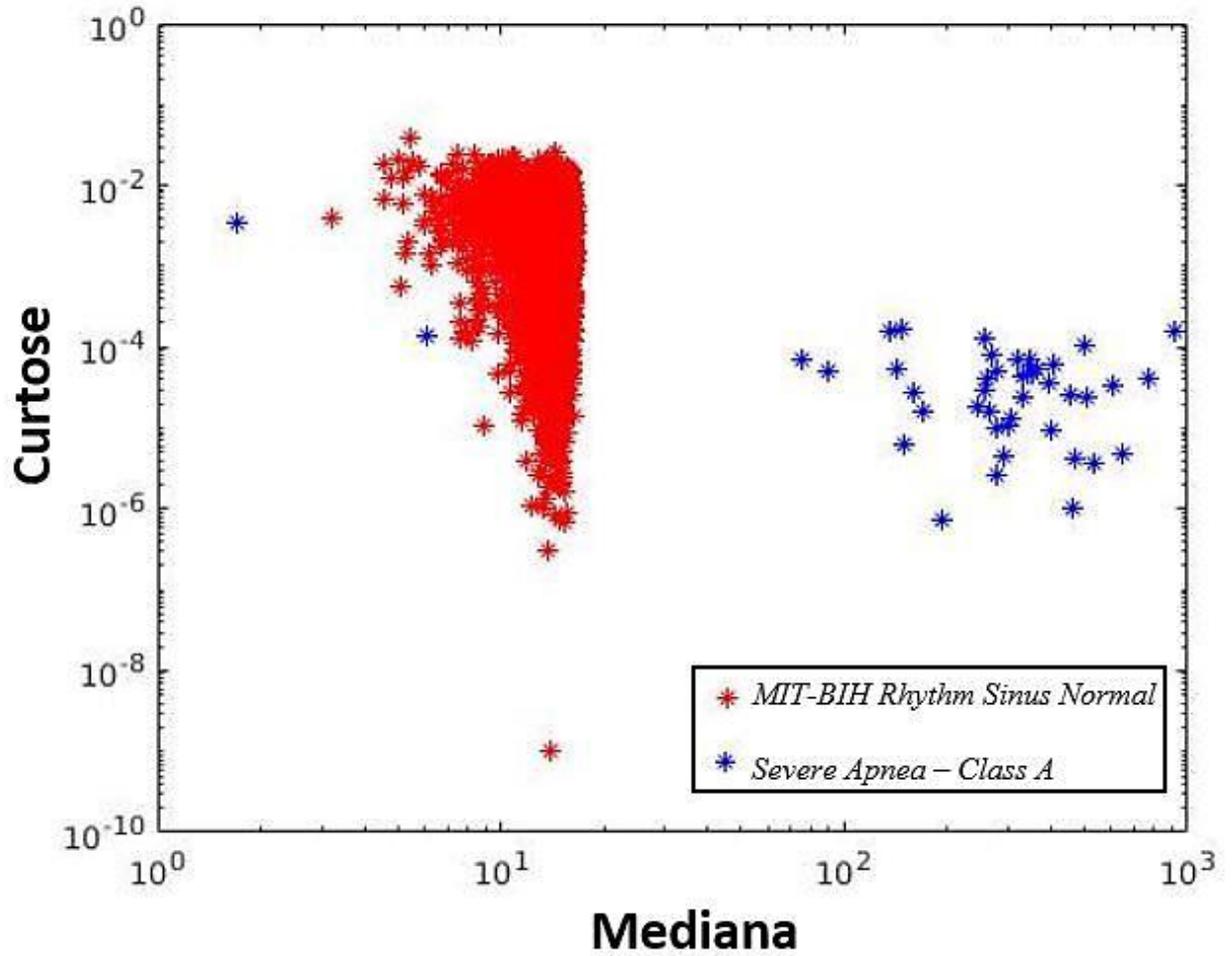
As Figuras 33, 34 e 35 apresentam os gráficos em 2 dimensões para SAOS em grau grave (grupo A). Onde a Figura 33 é um gráfico de dispersão de assimetria e curtose, a Figura 34 apresenta a dispersão de curtose e mediana e a Figura 35 é um gráfico de dispersão de assimetria e mediana.

Figura 33- Representação em 2D da dispersão de assimetria e curtose para os sinais dos indivíduos saudáveis e com SAOS em grau grave



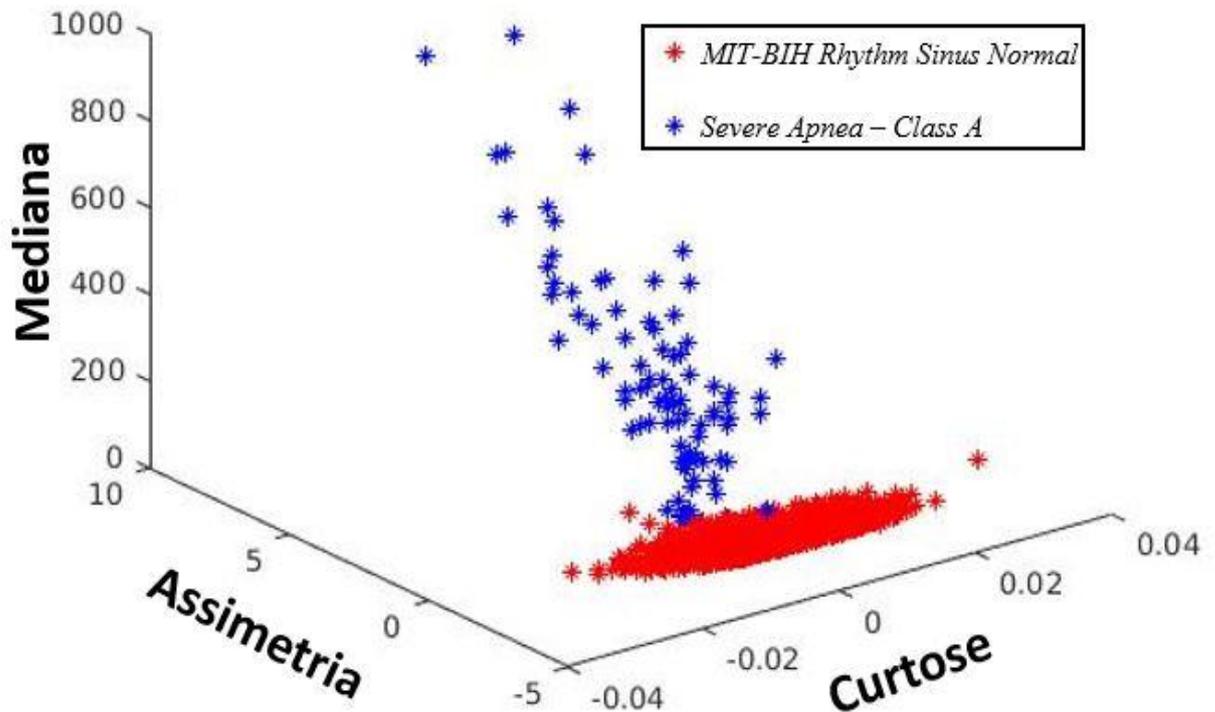
Fonte: Autora

Figura 34- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau grave.



Fonte: Autora.

Figura 36- Os sinais avaliados neste estudo, a saber, saudável e SAOS em grau grave apresentados em espaço 3D a partir da mediana, assimetria e curtose.



Fonte: Autora.

7.4 AVALIAÇÃO DOS CLASSIFICADORES

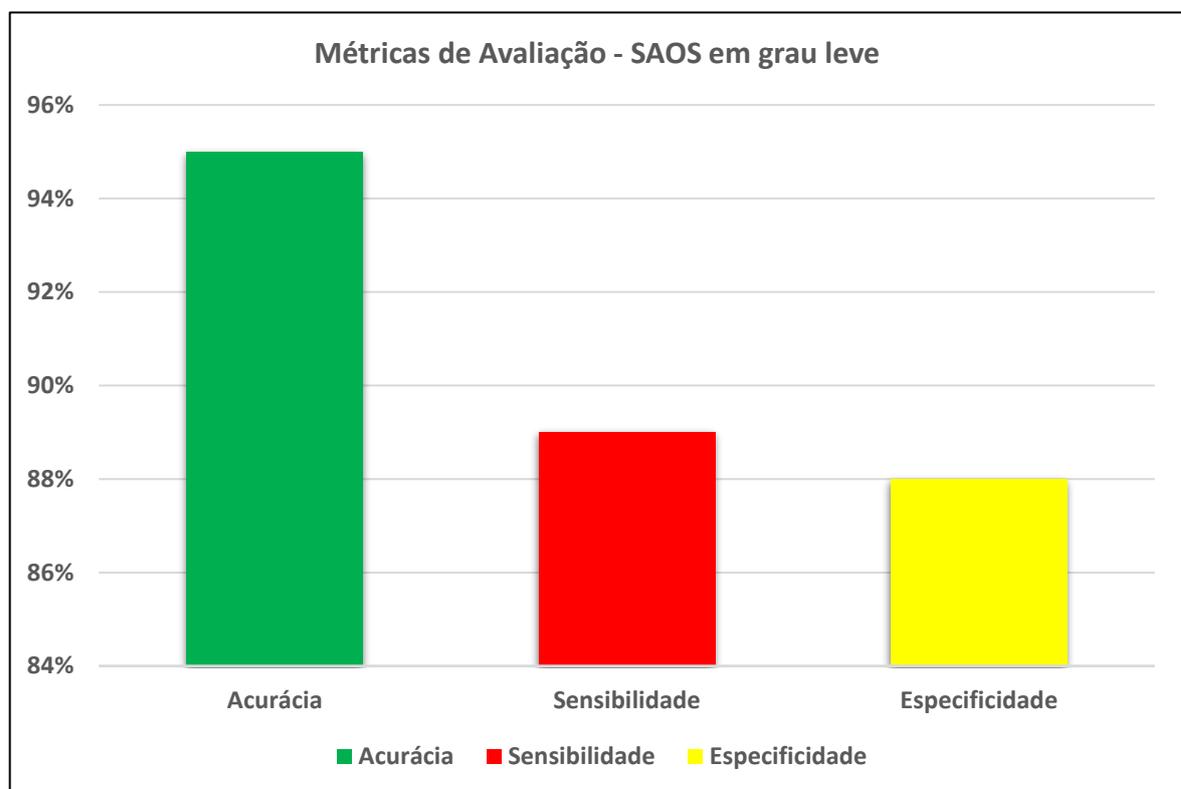
A partir da análise dos gráficos em duas dimensões apresentados nas figuras acima é possível notar uma classificação clara entre indivíduos saudáveis e diagnosticados com SAOS. As figuras plotadas a partir da estatística curtose apresentam uma maior separação entre os grupos, reafirmando os estudos de Queiroz *(et. al., 2017)* indicam a curtose como abordagem mais adequada para análise de sinais de ECG.

As representações tridimensionais proporcionam uma melhor visualização da separação entre ambos os grupos e conseqüentemente melhor classificação de cada indivíduo. Essa representação se torna ainda mais benéfica por ocasionar diagnóstico de forma visual e simples, uma vez que interpretar os resultados da polissonografia é uma tarefa bastante complexa, sendo necessário especialistas qualificados em diferentes especialidades médicas, assim como no

estudo do sono.

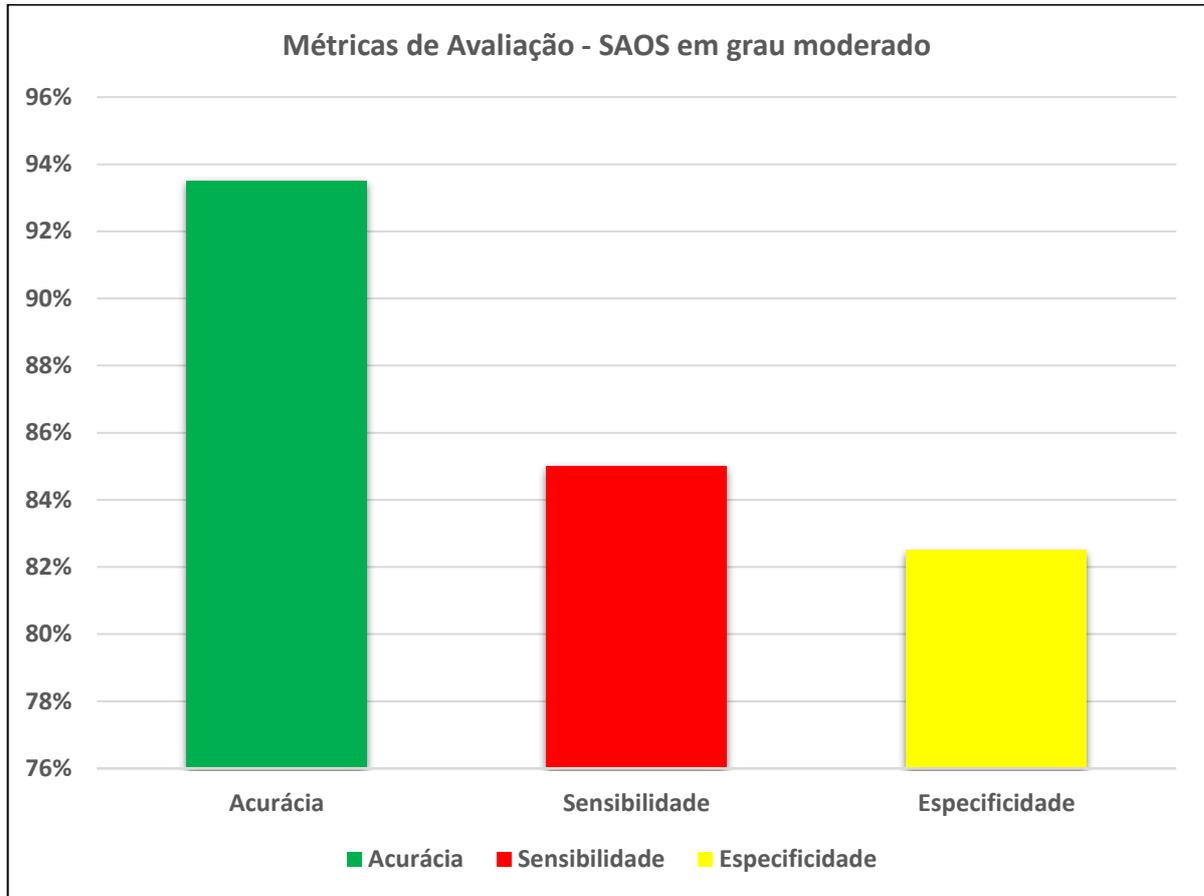
Ao analisar inicialmente as figuras para as três classes diferentes de SAOS, tem-se que as classificações plotadas no grupo de SAOS em grau leve (grupo A) aparentam maior distanciamento entre os dados e uma melhor classificação. Conforme as figuras abaixo, foram utilizadas a acurácia, sensibilidade e especificidade como métricas de avaliação para determinar desempenho das classificações realizadas.

Figura 37- Valores de classificação em porcentagem para SAOS em grau leve para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação



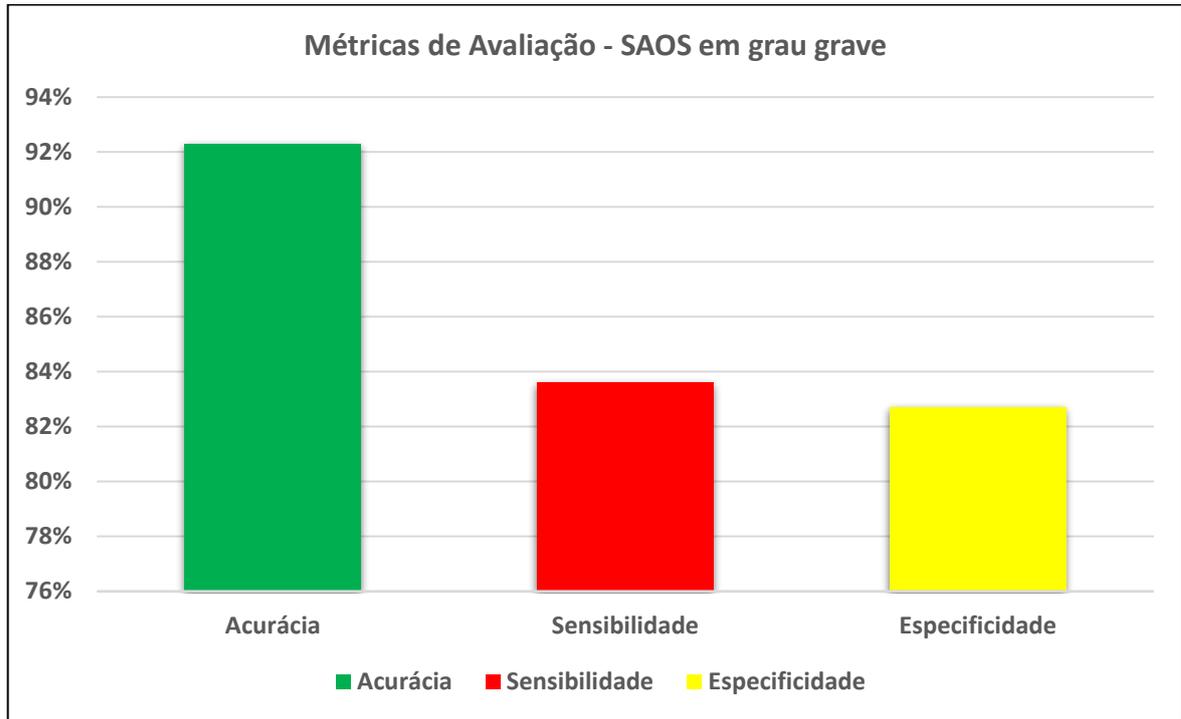
Fonte: Autora.

Figura 38- Valores de classificação em porcentagem para SAOS em grau moderado para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação



Fonte: Autora.

Figura 39- Valores de classificação em porcentagem para SAOS em grau grave para gráfico de duas dimensões (curtose x mediana), obtidos a partir das métricas de avaliação.



Fonte: Autora.

As Figuras 37, 38 e 39 apresentam os maiores valores de métricas de avaliação encontrados a partir da aplicação do algoritmo MLP, haja vista que a execução dos algoritmos KNN e SVM não apresentaram valores satisfatórios. Os melhores resultados correspondem ao gráfico de duas dimensões (curtose x mediana) para SAOS em grau leve apresentado na Figura 26, indicando que esse gráfico apresenta maior separação entre os grupos e conseqüentemente uma melhor performance.

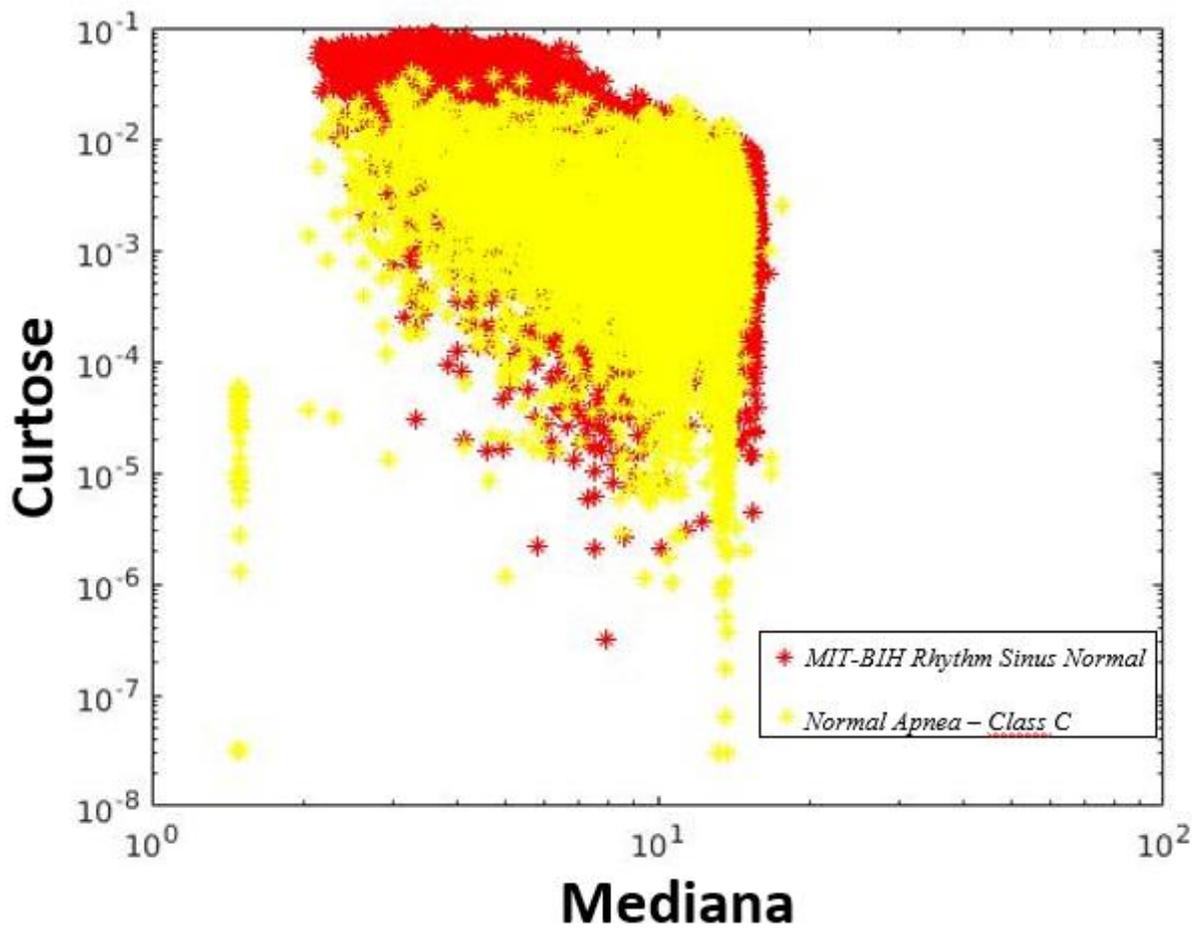
Conforme já discutido, tem-se que a curtose é a melhor estatística para análise de sinais de ECG e que ela juntamente com a mediana apresenta melhor classificação dos batimentos cardíacos, visto que ocasionam maior separação entre os grupos e ambos tiveram seus dados amostrados de forma organizada, pois tiveram maioria dos dados polarizados em lados opostos do gráfico.

A partir dos valores de métricas encontrados, nota-se também que as métricas utilizadas, acurácia, sensibilidade e especificidade, possuem porcentagem igual ou maior que 80%, indicando que o método utilizado ocasiona diagnóstico assertivo para portadores da SAOS em grau grave, moderada e principalmente para indivíduos com baixo índice IAH.

7.5 CLASSIFICAÇÃO SEM PCA

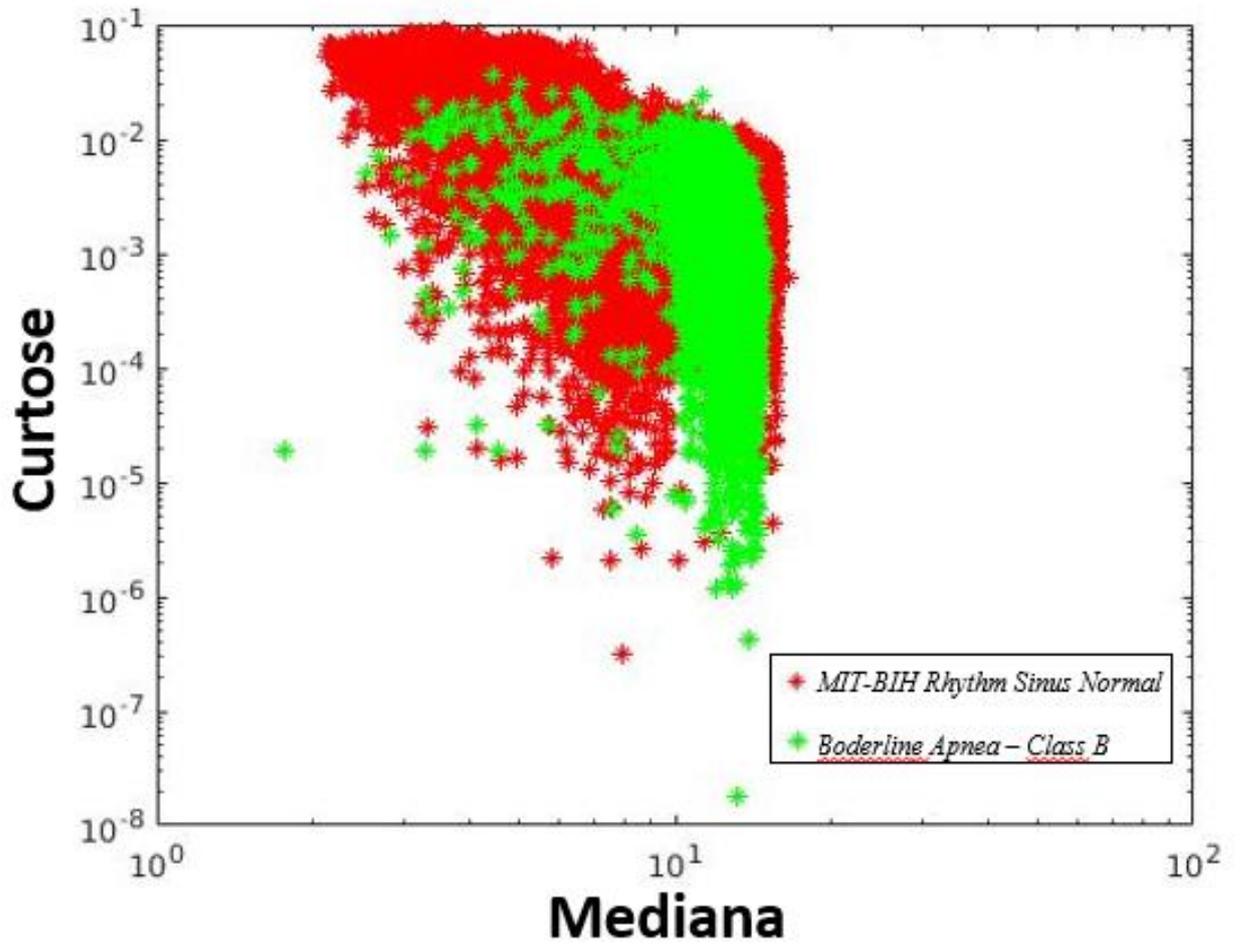
Os resultados obtidos das classificações realizadas sem PCA apresentaram métricas com valores baixos e dificuldade de visualização da separação entre diferentes grupos. Como os maiores valores para classificação com PCA foram para os gráficos bidimensionais entre mediana e curtose. Para fins de comparação, as Figuras 40, 41 e 42 demonstram o comportamento dos gráficos de classificação para mediana e curtose sem aplicação de PCA.

Figura 40- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau leve sem aplicação de PCA.



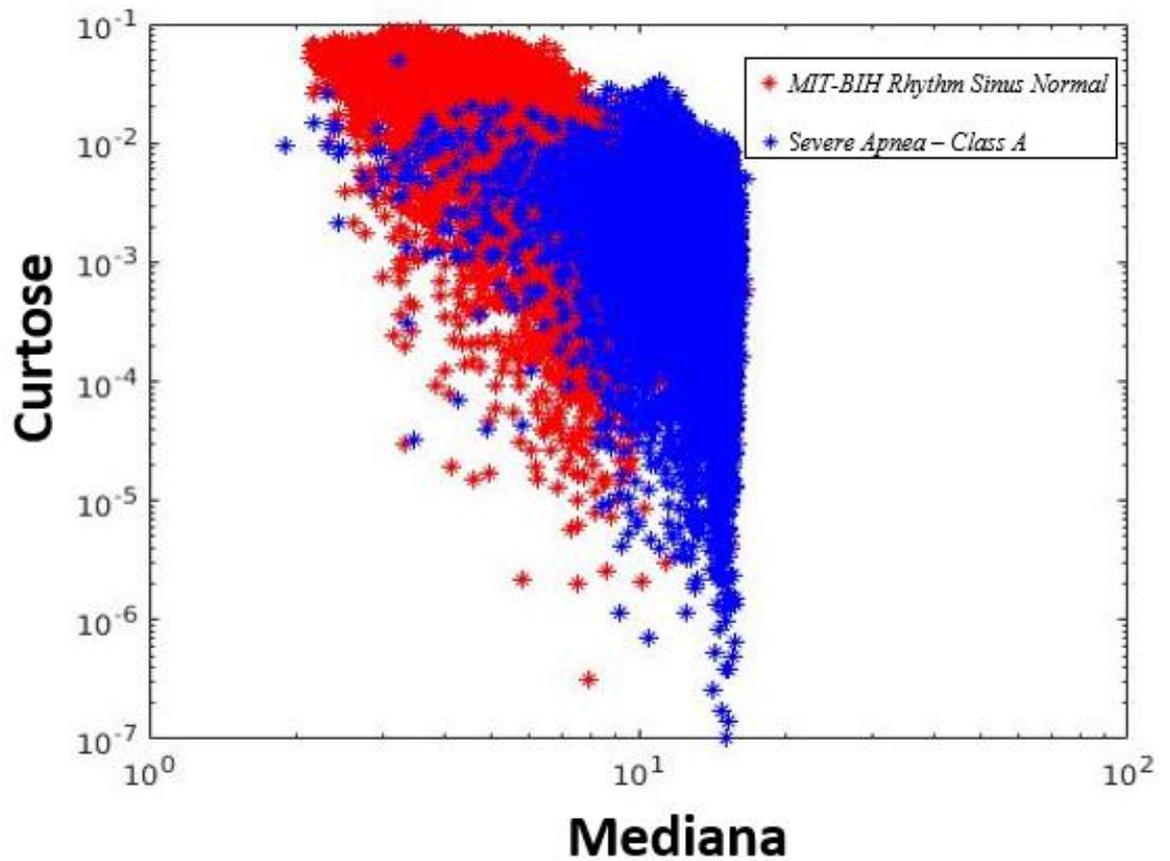
Fonte: Autora.

Figura 41- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau moderado sem aplicação de PCA



Fonte: Autora.

Figura 42- Representação em 2D da dispersão de curtose e mediana para os sinais dos indivíduos saudáveis e com SAOS em grau grave sem aplicação de PCA



Fonte: Autora.

Como observado nas figuras acima, ao realizar uma comparação visual é nítido a diferença entre os gráficos plotados com PCA e sem PCA, visto que nas Figuras 40, 41 e 42 os dados de ambos os grupos se encontram praticamente sobrepostos, não há possibilidade de ser traçada uma reta que separe os dados pois não há uma divisão clara entre as classes. Dessa forma, como esperado os valores das métricas de avaliação não foram considerados adequados, haja vista que todos ficaram abaixo de 70%.

8 TRABALHOS FUTUROS

Como sugestão de trabalhos futuros, tem-se a utilização de outros conjuntos de dados de diferentes fontes, como EEG, mudança de parâmetros dos classificadores objetivando ocasionar maior separação e melhor classificação dos grupos com valores de métricas de avaliação mais próximos de 100%, ocasionando maior assertividade de diagnóstico junto a equipe médica.

Ademais busca-se também aliar o desenvolvimento dos dispositivos de monitoramento biológico ao trabalho desenvolvido nessa dissertação, onde será possível realizar diagnósticos rápidos, pelos próprios pacientes, reduzindo custos e tempo para que se dê início ao tratamento. Dessa forma os pacientes podem procurar diretamente o corpo médico adequado para tomar todas as medidas necessárias para auxílio e melhora do quadro clínico.

9 CONCLUSÃO

A utilização de ferramentas como aprendizado de máquinas e estatísticas de alta ordem foi responsável pela extração de características cardíacas e classificação em grupos de diferentes pacientes. Dessa forma a partir dos sinais de eletrocardiograma há realização do diagnóstico de portadores da SAOS de diferentes níveis de gravidade, possibilitando auxílio da equipe médica para entrega de laudos de forma mais rápida e autônoma.

Portanto, ao associar essa praticidade com a atual evolução tecnológica, o método descrito inova ao desencadear simplicidade de diagnóstico ainda maior. Visto que com o desenvolvimento de dispositivos de monitoramento de sinais biológicos, com maior acessibilidade, robustez e excelente capacidade de processamento, há melhora na área da saúde coletiva e conseqüentemente no bem-estar do corpo humano.

REFERÊNCIAS

ALMEIDA, C. M. O. **Distúrbio do sono**. 2019. Manaus: Copyright: 2019.

AMERICAN PSYCHIATRIC ASSOCIATION. **Manual diagnóstico e estatístico dos transtornos mentais – DSM-IV-TR**. Porto Alegre: Artes Médicas, 1994.

AZZARI, D.; [et al.]. **Síndrome da Apneia Obstrutiva do Sono e Doenças Cardiovasculares**. Universidade Federal de São Paulo – UNIFESP, São Paulo – SP, Brasil. 2014.

BALBANI, A., FORMIGONI, G. (1999). Ronco e Síndrome da apnéia obstrutiva do sono. **Revista Ass Med Brasil**, v. 45, n. 3, p. 273 -278, 1999.

BITTENCOURT L. R., TOGEIRO S. M., BAGNATO M. Diagnóstico da síndrome da apnéia e hipopneia obstrutiva do sono. **Moreira JR Editora**, v. 19, n. 6, p. 584-593, 2012.

CINTRA, F, [et al], **Apneia do Sono e Arritmia Cardíaca Noturna: Estudo Populacional**. Universidade Federal de São Paulo – Unifesp, São Paulo, SP – Brasil. 2014. DOI: 10.5935/abc.20140142.

DRAGER, L., [et al.]. Síndrome da Apnéia Obstrutiva do Sono e sua Relação com a Hipertensão Arterial Sistêmica. Evidências Atuais. **Arq Bras Cardiol**, v 78, n. 5, 531-6, 2002.

ESCOVEDO, T; KOSHIYAMA, **Introdução a Data Science** — Algoritmos de Machine Learning e métodos de análise. São Paulo, Ed. Casa do Código, 2020.

FELDMAN, J.; GOLDWASSER, G. **Eletrocardiograma**: Recomendações para sua interpretação. Universidade Federal do Rio de Janeiro, Faculdade de Medicina Souza Marques, Universidade Gama Filho. Revista da SOCERJ – Out/Nov/Dez 2004.

FERRARA, M., & DE GENNARO, L. How much sleep do we need? **Sleep Medicine**, v. 5, n. 2, p. 155- 179, 2001.

FLEMONS WW, [et al]. The Report of American Academy of Sleep Medicine Task Force. Sleep Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research. **Sleep**, v. 22, p. 667-689, 1999.

GOLDBERGER A. L. [et al.] PhysioBank, PhysioToolkit e PhysioNet: componentes de um novo recurso de pesquisa para sinais fisiológicos complexos. **Circulation**, v. 101, n. 23, p. 215-220, jun., 2000.

GONDIM, L, [et al], Estudo Comparativo da história clínica e da polissonografia na síndrome da apnéia/hipopneia obstrutiva do sono. **Rev. Bras. Otorrinolaringol**, São Paulo, v. 73, n.

6, nov./dec. 2007.

HAGGSTRÄM F., ZETTLER E., FAM C. Apnéia obstrutiva do sono e alterações cardiovasculares. **Scientia Medica**, Porto Alegre, v. 19, n. 3, p. 122-128, jul./set. 2009.

IBM, Cloud Education. **Redes Neurais**. Disponível em: <https://www.ibm.com/br-pt/cloud/learn/neural-networks>. Acesso em: 12 fev. 2022.

ICMC. **Medidas de assimetria e curtose**. Universidade de São Paulo – USP. São Paulo, SP – Brasil. 2012.

ITO, F. A., [et al]. Conduas terapêuticas para tratamento da Síndrome da Apnéia e Hipopnéia Obstrutiva do Sono (SAHOS) e da Síndrome da Resistência das vias Aéreas Superiores (SrvAS) com enfoque no Aparelho Anti-Ronco (AAR-ITO). **R Dental Press Ortodon Ortop Facial**, Maringá, v. 10, n. 4, p.143-156, jul. 2005.

KUNUMI, Blog. **Métricas de avaliação em machine learning: Classificação**. Disponível em: <https://medium.com/kunumi/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-em-machine-learning-classifica>. Acesso em: 31 jan. 2022.

LIMA, Anna Myrna Jaguaribe de [et al]. Contribuição da Apnéia Obstrutiva do Sono para o Estresse Oxidativo da Obesidade. **Arq Bras Endocrinol Metab**, São Paulo, v. 52, n. 4, p.668-676, abr. 2008.

LUZ, F. Data Science for Business, PROVOST. **Algoritmo KNN para classificação**. Inferir. Disponível em: <https://inferir.com.br/artigos/algoritmo-knn-para-classificacao>. Acesso em: 01 fev.2022.

MATHWORKS. **Principal Component Analysis (PCA)**. 2021. Disponível em: <https://www.mathworks.com/help/stats/principal-component-analysis-pca.html>. Acesso em: 21 abr. 2021.

MARTINEZ, D. **Prática da medicina do sono**. São Paulo: BYK, 1999.

MARTINS, A., [et al.]. Síndrome apnéia-hipopnéia obstrutiva do sono. Fisiopatologia. **J Bras Pneumol**. v. 33, n. 1, p. 93-100, 2007.

MEDIUM. **KNN (K-Nearest Neighbors). Como funciona?** Disponível em: <https://medium.com/brasil-ai/knn-k-nearest-neighbors-1-e140c82e9c4e>. Acesso em: 01 fev. 2022.

MONARD, M.; [et al.]. **Conceitos sobre Aprendizado de Máquina**. Sistemas Inteligentes para Engenharia. Capítulo 4. 39-56. USP. 2003.

MOURA, P.; **Conceitos básicos sobre Machine Learning**. [201-]. Disponível em: <https://blog.4linux.com.br/conceitos-basicos-sobre-machine-learning/>. Acesso em: 31 jan. 2022.

MÜLLER, M.; GUIMARÃES, S. Impacto dos transtornos do sono sobre o funcionamento diário e a qualidade de vida. **Estudos de Psicologia**, Campinas, v. 24, n. 4, p. 519-528, out./dez.,

2007.

PACHECO F., ANJOS E.; MAIA A. Síndrome da apnéia/76ornal76ia obstrutiva do sono: Artigo de revisão bibliográfica. **Revista UNILUS Ensino e Pesquisa**, v. 12, n. 29, out./dez. 2015. ISSN 2318-2083 (Eletrônico).

PENZEL T. [et al]. The Apnea – ECG Database. **Computers in Cardiology**. n. 27, p. 255-258, 2000.

PEREIRA A. Síndrome da Apnéia Obstrutiva do Sono: Fisiopatologia, Epidemiologia, Consequências, Diagnóstico e Tratamento. **Arquimed**, Porto, v. 21, n. 5/6, p.159-173, 2007.

QUEIROZ J. A., [et al]. Diagnostic decision support systems for atrial fibrillation based on a novel electrocardiogram approach. **Journal of Electrocardiology**, 2017. Acesso em: 17 jun. 2020.

RAMOS, A., SOUSA, B. **Eletrocardiograma: princípios, conceitos e aplicações**. Centro de Estudos de Fisiologia do Exercício – CEFE, 2007

REDES NEURAIAS AUTO-ASSOCIATIVAS. PUC-Rio – Certificação Digital Nº 09129226/CA. P 44-62.

REIMÃO, R. **Sono: estudo abrangente**. 2ª. ed. São Paulo: Atheneu, 1996

REDLINE, S.; [et al.]. Obstructive Sleep Apnea-Hypopnea and Incident Stroke. The Sleep Heart Health Study. **American Jornal of respiratory and critical care medicine**. v. 182. 2010.

RESEARCHGATE. **Estudo da variabilidade da frequência cardíaca em exercício progressivo máximo**. Disponível em: https://www.researchgate.net/figure/Figura-3-Illustracao-de-Eletrocardiograma-normal-Representacao-das-ondas-P-e-T-e-do_fig1_259841876. Acesso em: 20 dez. 2021.

ROTH, T. [et al.]. (2002). A new questionnaire to detect sleep disorders. **Sleep Medicine**, v. 3 n. 2, p. 99-108, 2002.

SBPT – Sociedade Brasileira de Pneumologia e Tisiologia. **Apnéia do Sono**. Disponível em: <https://sbpt.org.br/portal/espaco-saude-respiratoria-apneia-do-sono>. Acesso em: 29 jan. 2021.

SCHWARZ L. *Artigo de Revisão: Eletrocardiograma*. **Revista Ilha Digital**, ISSN 2177-2649, volume 1, páginas 3-19, 2009.

SHER A. E., SCHECHTMAN K. B., PICCIRILLO J. F. The Efficacy of Surgical Modifications of the Upper Airway in Adults with Obstructive Sleep Apnea Syndrome. **Sleep**, v. 19, n. 2, p. 156-177, 1996.

SILVA L., QUEIROZ J., BARROS A. **Support method for the diagnosis of Atrial Fibrillation using Machine Learning**. XLI CILAMCE, 16 – 19, nov. 2020.

SILVERBERG, D. S; OKSENBERG, A; IAINA A. Sleep related breathing disorders are

common contributing factors to the production of essential hypertension but are neglected, underdiagnosed, and undertreated. **Am J Hypertens**, v. 10, p. 1319-1325, 1997.

WIEGAND L., ZWILICH C. W. Obstructive Sleep Apnea. In Bone RC (Ed) Disease-a-Month., **Mosby Year Book**, St. Louis, v. XL, 4, p. 199-252, 1994.

WIGGERT, G.T, [et al.]. Apnéia obstrutiva do sono e arritmias cardíacas. **Relampa**, v. 23, n. 1, p. 5-11, 2010.