

Universidade Brasil – Campus Itaquera
Instituto Científico e Tecnológico da Universidade Brasil
Programa de Pós-Graduação em Engenharia Biomédica

***Diagnóstico clínico automatizado a partir do uso de métodos de
análise multivariada aplicados a sinais de eletrocardiograma***

ANTONIO CHAVEZ ZENA

Orientador: Prof. Dr. Thiago de Oliveira Mendes

Co-orientadora: Profa. Dra. Laurita dos Santos

São Paulo, SP
2020

***Automated clinical diagnosis by multivariate statistical methods
applied to electrocardiogram signals***

ANTONIO CHAVEZ ZENA

**Tese de doutorado apresentada ao
Programa de Pós-Graduação em
Engenharia Biomédica da Universidade
Brasil, como parte dos requisitos
necessários para obtenção do título de
Doutor em Engenharia Biomédica.**

Orientador: Prof. Dr. Thiago de Oliveira Mendes

Co-orientadora: Profa. Dra. Laurita dos Santos

São Paulo, SP
2020

Ficha catalográfica elaborada pelo Sistema de Bibliotecas da Universidade Brasil,
com os dados fornecidos pelo (a) autor (a).

Z54d ZEÑA, Antonio Chávez

Diagnóstico clínico automatizado a partir do uso de métodos de análise multivariada aplicados a sinais de eletrocardiograma / Antonio Chávez Zeña. -- São Paulo: Universidade Brasil, 2020.

69 p. : il. color.

Tese de Doutorado defendida no Programa de Pós-graduação do Curso de Engenharia Biomédica da Universidade Brasil.

Orientação: Prof. Dr. Thiago de Oliveira Mendes.

Coorientação: Profa. Dra. Laurita dos Santos.

1. Sinais Biomédicos. 2. Variabilidade da frequência cardíaca. 3. Aprendizado de máquina. 4. Diagnóstico. I. Mendes, Thiago de Oliveira. II. Santos, Laurita dos. III. Título.

CDD 620.82

TERMO DE APROVAÇÃO

ANTONIO CHAVEZ ZENA

“DIAGNÓSTICO CLÍNICO AUTOMATIZADO A PARTIR DO USO DE MÉTODOS DE ANÁLISE MULTIVARIADA APLICADOS A SINAIS DE ELETROCARDIOGRAMA”

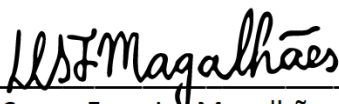
Tese aprovada como requisito parcial para obtenção do título de **Doutor no Programa de Pós-Graduação em Engenharia Biomédica** da Universidade Brasil, pela seguinte banca examinadora:



Prof(a). Dr(a) Thiago de Oliveira Mendes (presidente-orientador)



Prof(a). Dr(a) Lívia Assis Garcia (UNIVERSIDADE BRASIL)



Prof(a). Dr(a) Daniel Souza Ferreira Magalhães (UNIVERSIDADE BRASIL)



Prof(a). Dr(a) Felipe de Oliveira Matos (UNIVERSIDADE ESTADUAL DE MARINGÁ)



Prof(a). Dr(a) Rudinei Martins de Oliveira (UNIVERSIDADE DO ESTADO DE MINAS GERAIS)

São Paulo, 11 de dezembro de 2020.

Presidente da Banca Prof.(a) Dr.(a) Thiago de Oliveira Mendes

Houve alteração do Título: sim () não (X):

Termo de Autorização

Para Publicação de Dissertações e Teses no Formato Eletrônico na Página WWW do Respectivo Programa da Universidade Brasil e no Banco de Teses da CAPES

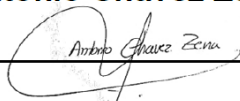
Na qualidade de titular(es) dos direitos de autor da publicação, e de acordo com a Portaria CAPES no. 13, de 15 de fevereiro de 2006, autorizo(amos) a Universidade Brasil a disponibilizar através do site <http://www.universidadebrasil.edu.br>, na página do respectivo Programa de Pós-Graduação *Stricto Sensu*, bem como no Banco de Dissertações e Teses da CAPES, através do site <http://bancodeteses.capes.gov.br>, a versão digital do texto integral da Dissertação/Tese abaixo citada, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira.

A utilização do conteúdo deste texto, exclusivamente para fins acadêmicos e científicos, fica condicionada à citação da fonte.

Título do Trabalho: **“DIAGNÓSTICO CLÍNICO AUTOMATIZADO A PARTIR DO USO DE MÉTODOS DE ANÁLISE MULTIVARIADA APLICADOS A SINAIS DE ELETROCARDIOGRAMA”**

Autor(es):

Discente: **Antonio Chavez Zena**

Assinatura:  _____

Orientador(a): **Prof.(a) Dr.(a) Thiago de Oliveira Mendes**

Assinatura:  _____

Coorientador(a): **Prof.(a) Dr.(a) Laurita dos Santos**

Assinatura:  _____

Houve alteração do Título: sim () não (X):

Data: 11/12/2020

DEDICATÓRIA

Dedico esta tese de doutorado ao meu pai, Antonio Chávez Santamaria (*in memoriam*) que mesmo em outro plano, sempre esteve comigo em todos os momentos da minha vida.

AGRADECIMENTOS

Desejo expressar meu agradecimento ao professor Dr. Thiago de Oliveira Mendes, meu orientador, pela infinita paciência em transmitir seus conhecimentos durante o tempo que foi necessário para elaboração desta tese de doutorado.

Meu agradecimento sincero também a professora Dra. Laurita dos Santos, coorientadora deste trabalho, por sua valiosa ajuda elucidando dúvidas durante este trabalho.

A todos os professores do Instituto Científico e Tecnológico da Universidade Brasil, dos quais tive a oportunidade de ter presenciado as aulas.

A meus pais, Florencia Zena Santisteban e Antonio Chávez Santamaria (*in memoriam*), meus irmãos pelo grande amor e carinho apesar de estarem distantes.

Agradeço a Josefa, minha esposa, pelo apoio e compreensão nos momentos difíceis neste último ano. A minha filha Brenda Florencia, pelos sorrisos e pelas palavras de incentivo mesmo sendo ainda muito jovem: somos uma família.

Aos colegas de trabalho, professoras Claudia e Dora pela amizade e pelo tempo disposto para treinamento na fase de pre-qualificação.

A Deus, por cuidar de minha mãe e família, por guiar nossos caminhos e nos permitir conhecer grandes amigos.

RESUMO

O presente trabalho apresenta um estudo de métodos de reconhecimento de padrões aplicados a parâmetros da análise da variabilidade da frequência cardíaca, a partir de eletrocardiogramas (ECG), para o auxílio diagnóstico clínico automatizado de diferentes doenças associadas ao coração, por meio de métodos estatísticos multivariados e computacionais de aprendizagem de máquina. Para tal propósito, foram obtidos os sinais de eletrocardiogramas de 137 voluntários diagnosticados clinicamente com ritmo sinusal normal (RSN), sendo $n = 54$ indivíduos, que representará o grupo controle, e duas condições clínicas formadas por indivíduos com insuficiência cardíaca congestiva (ICC), com $n = 29$ indivíduos ou supressão de arritmia cardíaca (SAC), com $n = 54$ indivíduos, considerando-se estas duas condições clínicas, como os grupos de caso. Todos estes sinais foram adquiridos proveniente da base de dados de domínio público o *PhysioNet*, que abrange um conjunto de sinais biomédicos reais, software de código aberto e oriundos de estudos consolidados na literatura. Descreveu-se um procedimento de obtenção de variáveis características de tacogramas de sinais de ECG em que as variáveis foram modeladas pelas abordagens de classificação de dados de análise discriminante por regressão por mínimos quadrados parciais (PLS-DA) e redes neurais artificiais (ANN), objetivando o diagnóstico de duas condições clínicas quando comparados com um grupo controle. Foram consideradas matrizes de dados das variáveis associadas ao domínio do tempo, domínio da frequência e obtidas por métodos não-lineares, isoladamente cada uma, e todas essas em uma matriz única de dados, de parâmetros estatísticos associados a variabilidade da frequência cardíaca. As figuras de mérito mostraram que existe um padrão no comportamento dos parâmetros dos tacogramas que podem vir a serem utilizados para o auxílio diagnóstico clínico. Tanto a insuficiência cardíaca congestiva como a classificação e previsão das amostras pertencentes ao estudo de supressão de arritmia cardíaca foram obtidas de forma satisfatória, com área abaixo da curva ROC próxima de 0,9. O modelo PLS-DA demonstrou os melhores resultados de classificação de dados, onde a insuficiência cardíaca congestiva foi diagnosticada com taxas 90,9% de sensibilidade e seletividade de 85,7% e a supressão de arritmia cardíaca foi prevista com taxas de 75,0% de sensibilidade e 100,0% de seletividade, sugerindo que o auxílio diagnóstico clínico em tempo real e um prognóstico personalizado pode se tornar uma realidade que contribuirá positivamente para a prática médica.

Palavras-chave: Sinais Biomédicos. Variabilidade da frequência cardíaca. Aprendizado de máquina. Diagnóstico.

ABSTRACT

This work presents a study of methods applied to pattern recognition of heart rate variability parameters obtained from electrocardiogram (ECG) for the aid of automated clinical diagnosis of various diseases associated with the heart, using multivariate statistical methods and computational machine learning. For this purpose, the electrocardiograms signs of 137 volunteers clinically diagnosed with normal sinus rhythm (NSR), with $n = 54$ individuals, which will represent the control group, and two clinical conditions formed by individuals with congestive heart failure (CHF), with $n = 29$ individuals or suppression of cardiac arrhythmia (CAST), with $n = 54$ individuals, considering these two clinical conditions, such as the case groups. All these signals were obtained from the PhysioNet, which covers a set of real biomedical signals, open source software and from studies consolidated in the literature. A procedure for obtaining characteristic variables of ECG tachograms was described, these variables were modeled by classification approaches of Discriminant Analysis data by Partial Least Squares regression (PLS-DA) and Artificial Neural Networks (ANN), aiming at the diagnosis of two clinical conditions when compared with a control group. Data matrices of variables associated with the time domain, frequency domain and obtained by non-linear methods were considered separately, each one, and all of these in a single data matrix, of statistical parameters associated with heart rate variability. The figures of merit showed that there is a pattern in the behavior of the tachogram parameters that may be used for clinical diagnostic aid. Both congestive heart failure and the classification and prediction of samples belonging to the cardiac arrhythmia suppression were satisfactorily obtained, with an area under the ROC curve close to 0.9. The PLS-DA model demonstrated the best data classification results, where congestive heart failure was diagnosed with rates of 90.9% of sensitivity and selectivity of 85.7% and suppression of cardiac arrhythmia was predicted with rates of 75.0% of sensitivity and 100.0 % of selectivity, suggesting that clinical diagnosis assisting real time and a personalized prognosis can become a reality that will contribute positively to medical practice.

Keywords: Biomedical Signals. Heart rate variability. Machine learning. Diagnosis.

LISTA DE FIGURAS

Figura 1 – a) Derivações: seis em cada plano frontal e horizontal. b) Derivações no plano horizontal observados no ECG.	18
Figura 2 – a) Impulso elétrico do coração b) Vetor resultante do impulso elétrico c) Identificação e visualização do sinal elétrico em forma de ondas.	19
Figura 3 - Perfil típico de um eletrocardiograma.	20
Figura 4 - Características morfológicas das ondas P, Q, R, S e T do ECG.	22
Figura 5 - O tacograma.	23
Figura 6 - Representação esquemática do PLS-DA.	27
Figura 7 - Decomposição de matrizes em modelos PLS.	28
Figura 8 – Esquema de um neurônio (biológico).	29
Figura 9 - Esquema de um neurônio (computacional).	30
Figura 10 - Rede neural perceptron com k-neurônios de entrada.	31
Figura 11 – Exemplo de uma curva ROC.	35
Figura 12 - Fluxo de trabalho.	41
Figura 13- Boxplot: Grupo ritmo sinusal normal (domínio do tempo).	46
Figura 14 - Matriz de dados: Grupo ritmo sinusal normal (domínio do tempo).	46
Figura 15 - Seleção de amostras pelo algoritmo de Kennard-Stone.	50
Figura 16 - Modelo de classificação de dados: Previsão da classe das amostras. ...	51
Figura 17 - Matriz de confusão.	52
Figura 18 - Curva ROC.	52
Figura 19 - Boxplot: Grupo ritmo sinusal normal (domínio da frequência).	62
Figura 20 - Matriz de dados: Grupo ritmo sinusal normal (domínio da frequência) ...	62
Figura 21 - Boxplot: Grupo ritmo sinusal normal (parâmetros não lineares).	63
Figura 22 - Matriz de dados: Grupo ritmo sinusal normal (parâmetros não lineares).	63
Figura 23 - Boxplot: Grupo insuficiência cardíaca congestiva (domínio do tempo). ...	64
Figura 24 - Matriz de dados: Grupo insuficiência cardíaca congestiva (domínio do tempo).	64
Figura 25 - Boxplot: Grupo insuficiência cardíaca congestiva (domínio da frequência).	65

Figura 26 - Matriz de dados: Grupo insuficiência cardíaca congestiva (domínio da frequência).....	65
Figura 27 - Boxplot: Grupo insuficiência cardíaca congestiva (parâmetros não lineares).	66
Figura 28 - Matriz de dados: Grupo insuficiência cardíaca congestiva (parâmetros não lineares).	66
Figura 29 - Boxplot: Grupo supressão de arritmia cardíaca (domínio do tempo).	67
Figura 31 - Boxplot: Grupo supressão de arritmia cardíaca (domínio da frequência).	68
Figura 32 - Matriz de dados: Grupo supressão de arritmia cardíaca (domínio da frequência).....	68
Figura 33 - Boxplot: Grupo supressão de arritmia cardíaca (parâmetros não lineares)	69
Figura 34 - Matriz de dados: Grupo supressão de arritmia cardíaca (parâmetros não lineares).	69

LISTA DE TABELAS

Tabela 1 - Propriedades obtidas a partir de um tacograma.	25
Tabela 2 - Testes de hipóteses univariados.	48
Tabela 3 - Representação da dimensão das matrizes de dados formadas.	49
Tabela 4 - Performance dos modelos de classificação para Insuficiência cardíaca congestiva.	53
Tabela 5 - Performance dos modelos de classificação para supressão de arritmia cardíaca.	54

LISTA DE ABREVIATURAS

AI	Inteligência Artificial (do inglês <i>Artificial Intelligence</i>)
ANN	Rede neural artificial (do inglês <i>Artificial Neural Network</i>)
AUC	Área abaixo da curva (do inglês <i>area under the curve</i>)
ECG	Eletrocardiograma
FN	Falso negativo
FP	Falso positivo
ICC	Insuficiência cardíaca congestiva
PLS	Mínimos quadrados Parciais (do inglês, <i>Partial Least Squares</i>)
PLS-DA	Mínimos quadrados Parciais Análise Discriminante (do inglês <i>Partial Least Squares Discriminant Analysis</i>)
ROC	Curva ROC (do inglês, <i>Receiver Operator Characteristic curve</i>)
RSN	Ritmo sinusal normal
SAC	Supressão de arritmia cardíaca
SNA	Sistema nervoso autônomo
SVM	Máquina de vetor de suporte (do inglês, <i>Support Vector Machine</i>)
VFC	Variabilidade da frequência cardíaca
VP	Verdadeiro positivo
VN	Verdadeiro negativo

SUMÁRIO

1 INTRODUÇÃO	14
2 REVISÃO DE LITERATURA	17
2.1 Obtenção e interpretação do sinal biológico	17
2.1.1 O eletrocardiograma	17
2.1.2 O tacograma RR: a variabilidade da frequência cardíaca	22
2.2 Modelos preditivos baseados em classificadores multivariados ...	25
2.2.1 Mínimos quadrados parciais - análise discriminante (PLS-DA)	26
2.2.2 Redes neurais artificiais (ANN)	29
2.2.3 Performance de modelos preditivos: Figuras de mérito	32
2.3 Bancos de dados	35
2.4 Inteligência artificial na área da saúde	36
3 PROPOSIÇÃO	40
3.1 Objetivo Geral	40
3.2 Objetivos Específicos	40
4 MATERIAIS E MÉTODOS.....	41
4.1 Obtenção dos tacogramas	42
4.2 Pré-processamento de dados	43
4.3 Inspeção dos dados.....	43
4.4 Modelos de classificação supervisionada	44
5 RESULTADOS E DISCUSSÃO	45
6 CONCLUSÕES E PERSPECTIVAS	56
REFERÊNCIAS	57
APÊNDICE: Complemento da inspeção de dados.....	62

1 INTRODUÇÃO

Quando estuda-se o sistema regulatório da pressão arterial cardiovascular e series temporais de sinais biológicos típicas, como a sequência dos intervalos RR, ou quando estima-se a capacidade pulmonar total, afere-se a temperatura corporal, ou obtém-se um eletrocardiograma (ECG) de um indivíduo, realiza-se medições de propriedades físicas que são registradas em forma de um sinal, com uma determinada grandeza e unidade de medida apropriada. Assim o volume de ar (em m³), a temperatura (em graus celsius) e a atividade elétrica (em volts) podem ser utilizados para elucidar informações acerca do organismo que se estuda, estabelecendo, por exemplo, quadros clínicos, efeitos de fármacos e até mesmo o diagnóstico de doenças, (REILLY, 1998).

Um sinal biológico é a medida de uma propriedade física de um organismo vivo, podendo essa propriedade ser de natureza mecânica, térmica, elétrica, ou outra, e com a característica única de rotular o organismo em estudo numa dada condição na qual ele se encontra. Dentro da gama do universo de sinais biológicos que existem, focar-se-á no ECG. O ECG é um sinal de característica elétrica, que faz uma medida das propriedades elétricas associadas ao coração. Este sinal fornece propriedades de interesse médico que estão relacionadas ao quadro clínico do paciente, (ANTONI, 1998).

Embora a avaliação clínica do ECG verse, principalmente, sobre medições relativamente simples dos tempos e amplitudes entre as batidas do coração, uma análise mais aprofundada deste sinal compreende os estudos conhecidos como análise da variabilidade da frequência cardíaca (VFC), (STEIN et al., 1994). A VFC, calculada a partir de eletrocardiogramas de curto e longo prazo, pode inclusive ser utilizada para monitorar o controle da atividade neural, e baseia-se na série temporal dos intervalos, que é a sequência de intervalos entre os pontos fiduciais sucessivos dos picos R dos complexos QRS no ECG. Importante ressaltar que a análise da VFC não mede os ritmos do nó sinoatrial, que é o marcapasso do coração, uma vez que não se baseia nos intervalos PP, (KHANDOKER et al., 2013), (KAMATH; WATANABE; UPTON, 2013).

Esses registros são obtidos de maneira não invasiva ao contrário de outras técnicas disponíveis para avaliação autonômica, e tem uma relação direta com o

quadro clínico do paciente, tornando-se um importante exame associado ao diagnóstico de diferentes enfermidades. O ECG, é um sinal biomédico utilizado para detectar arritmias, aumento de cavidades cardíacas, doenças coronarianas, infarto do miocárdio, entre outros diagnósticos. Entretanto, a avaliação neste tipo de exame é feita sobre o registro da atividade elétrica cardíaca, sendo a tradução do sinal elétrico, obtido a partir do paciente, em um diagnóstico clínico, uma etapa laboriosa que exige extrema dedicação do profissional de saúde e que pode ser de difícil interpretação.

Auxiliando nessa interpretação, métodos estatísticos de análise de dados ganharam espaço em aplicações com sinais de natureza biológica (BLINOWSKA; ZYGIEREWICZ, 2011). O reconhecimento de padrões em exames médicos revelou que é possível promover um diagnóstico automatizado, reduzindo o tempo gasto com interpretação realizada pelo profissional da saúde e propiciando a aplicação de testes clínicos em larga escala de pacientes com resultados em tempo real (FOTIADIS et al., 2006).

No contexto de análise de sinais biológicos, ou biomédicos, desenvolvimentos recentes na área de tecnologia da saúde têm proporcionado o surgimento de equipamentos de diagnóstico cada vez mais instantâneos e confiáveis. A extração automática de um diagnóstico clínico a partir de diferentes conjuntos de dados de sinais biológicos representam problemas complexos, sendo alvos de pesquisas científicas dentro da engenharia biomédica e medicina (SÖRNMO; LAGUNA, 2005).

Assim, o presente estudo de doutoramento é centrado em avaliar o potencial de diferentes técnicas de classificação de dados da área da saúde, tendo como objetivo empregar abordagens estatísticas de reconhecimento de padrões para propiciar um diagnóstico clínico automatizado a partir do uso de métodos de análise multivariada aplicados a dados obtidos do ECG, sem a necessidade da interpretação de um médico especialista. Desta forma, avaliou-se a performance da Regressão por Mínimos Quadrados Parciais – Análise discriminante (PLS-DA, do inglês *discriminant analysis by partial least squares*) e as Redes Neurais Artificiais (ANN, do inglês *artificial neural network*) sobre ensaios provenientes de bancos de dados de ECG amplamente aceitos e estabelecidos pela comunidade médica.

Este texto apresenta na sequência, uma revisão da literatura onde inicialmente aborda a aquisição e interpretação de séries temporais características de tacogramas obtidos a partir de ECG, seguida da fundamentação teórica e revisão da

literatura envolvendo aplicação de métodos de classificação na área da saúde. A seção de materiais e métodos apresenta a sistemática para a obtenção dos dados utilizados no trabalho, e as respectivas figuras de mérito de avaliação da performance dos modelos estatísticos aplicados, quando comparados com os laudos clínicos em duas aplicações distintas. A primeira é sobre a classificação/diagnóstico da insuficiência cardíaca congestiva, e a segunda sobre um conjunto de dados de um ensaio de supressão de arritmia cardíaca. Encerrando as discussões são então apresentadas as conclusões e desdobramentos do presente trabalho.

2 REVISÃO DE LITERATURA

Antes de tentar qualquer processamento de sinal do ECG, seria interessante entender a fisiologia do ECG, convenções de medição e verificar como um médico utiliza o ECG para um diagnóstico de doenças, relacionadas ao sistema cardiovascular e os cuidados da saúde dos pacientes (KUSUMOTO, 2020). Para tal, faz-se necessário o entendimento do funcionamento do coração, um órgão vital composto por um músculo estriado cardíaco (miocárdio), que tende a contrair-se ritmicamente e conduzir a circulação de sangue por todo o corpo. Antes dos batimentos cardíacos normais, ou sístole, uma onda de corrente elétrica passa por todo o coração, o que desencadeia contração do miocárdio (KAMATH; WATANABE; UPTON, 2013). Este impulso elétrico que se inicia no nó sinusal não é aleatório, mas se espalha sobre toda a estrutura do sistema de condução do coração, seguindo um padrão eficaz e coordenado, responsável pela despolarização dos ventrículos. Tal propagação elétrica resulta em um processo mensurável, através da diferença de potencial na superfície corporal, portanto, o sinal amplificado (e filtrado) resultante é conhecido como ECG (MARTINI; TIMMONS; TALLITSCH, 2009). O ECG é afetado por uma série de fatores, incluindo anormalidades de fibras de condução cardíaca, anormalidades metabólicas do miocárdio, e anormalidades macroscópicas da geometria normal do coração e inclusive técnicas, posicionamento do paciente, eletrodos e equipamentos (SAKSENA; CAMM, 2012).

2.1 Obtenção e interpretação do sinal biológico

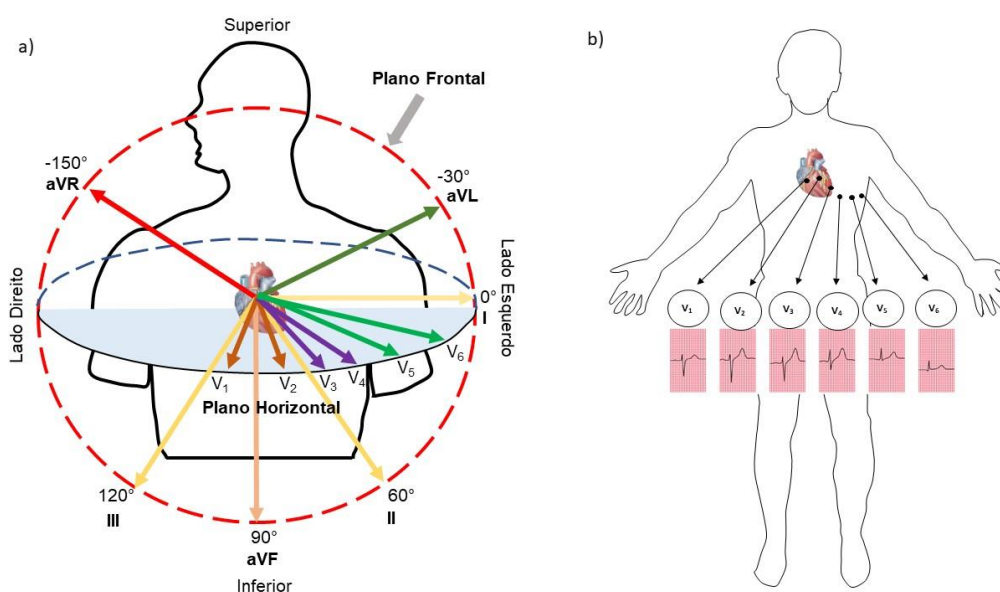
2.1.1 O eletrocardiograma

O ECG é um teste capaz de captar a atividade elétrica do coração, e assim, identificar possíveis arritmias, isquemia ou necrose tecidual. Geralmente é medido por eletrodos colocados na superfície da pele. Isso pode ser feito colocando pelo menos um par de eletrodos diretamente sobre a pele e monitorando a diferença de potencial entre eles, pois, esses sinais elétricos são transmitidos por todo o corpo. As características da forma de onda detectada dependem não apenas da quantidade de

tecido cardíaco envolvido, mas também da orientação relativa dos eletrodos, ou seja, em relação ao dipolo principal do coração. As formas de onda de um ECG parecerão diferentes quando medidas em diferentes posições de eletrodos e, normalmente, uma investigação de ECG é obtida usando uma série de diferentes localizações de eletrodos. Por exemplo, para uma apuração da atividade elétrica do ECG colocam-se eletrodos em pontos específicos do corpo, determinando-se as derivações periféricas: bipolares (DI, DII, DIII) e unipolares (aVR, aVL, aVF) e derivações precordiais (V₁, V₂, V₃, V₄, V₅, V₆) obtendo-se 12 derivações que medem a diferença de potencial, e que foram padronizados por aplicações universais de certas convenções (HOSKINS; LAWFORD; DOYLE, 2017).

Em um ECG consegue-se ver o coração no plano frontal, ou seja, quando o coração é cortado frontalmente e dividido em parte posterior e anterior; e em um plano horizontal quando ele é cortado horizontalmente dividido em partes de superior e inferior. Das 12 derivações observadas no ECG, seis delas são do plano frontal e seis do plano horizontal (IAIZZO, 2015). As seis do plano frontal são DI, DII, DIII e aVR, aVL, aVF como mostra a Figura 1a e as seis derivações do plano horizontal são V₁, V₂, V₃, V₄, V₅ e V₆, como mostra a Figura 1b (KUSUMOTO, 2020).

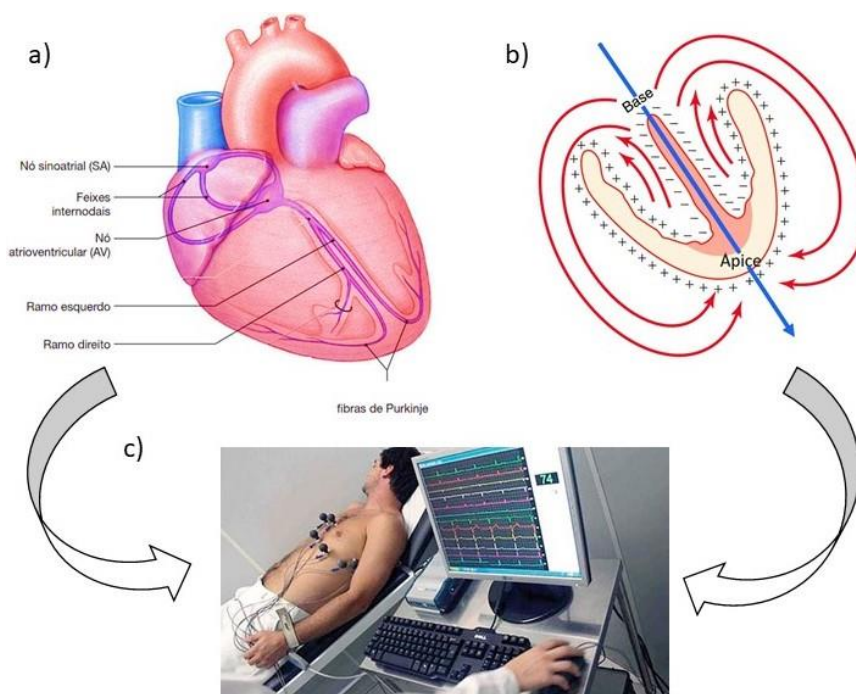
Figura 1 – a) Derivações: seis em cada plano frontal e horizontal. b) Derivações no plano horizontal observados no ECG.



Fonte: Adaptado de (KUSUMOTO, 2020).

Ao observar-se o sistema de condução do coração, verifica-se que o impulso elétrico começa no nó sino atrial, passa pelas vias intermodais, chega ao nó atrioventricular, feixe atrioventricular, ramos direito e esquerdo e fibras de Purkinje, como mostrado na Figura 2a. O coração humano assemelha-se a um formato de cone, é como se ele estivesse apontando de trás pra frente, da esquerda para a direita e de cima para baixo; então ao fazer-se uma soma dos vetores, o vetor resultante vem pra baixo, para frente e para esquerda, como mostrado na Figura 2b (HALL; HALL, 2020); (ABEDIN; CONNER, 2008). Essa atividade elétrica do coração é identificada por eletrodos, que são colados na superfície da pele, e transformam a atividade elétrica em ondas que visualizadas em um monitor, como mostrado na Figura 2c.

Figura 2 – a) Impulso elétrico do coração b) Vetor resultante do impulso elétrico c) Identificação e visualização do sinal elétrico em forma de ondas.

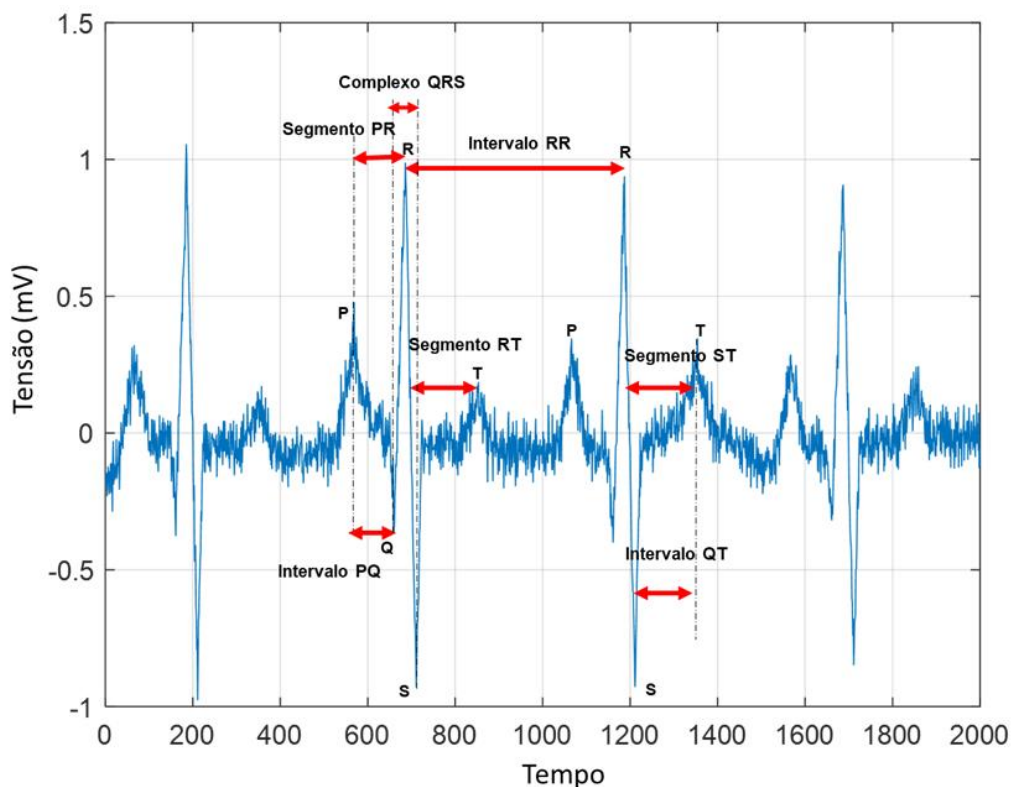


Fonte: Adaptado de (MARTINI; TIMMONS; TALLITSCH, 2009) e (HALL; HALL, 2020).

No ECG observam-se ondas denominadas P, Q, R, S e T representam uma fase da passagem dos impulsos elétricos cardíacos e determinam um ciclo cardíaco (Figura 3). Essas ondas espelham a despolarização e repolarização, e pequenos desvios em milímetros, podendo indicar desde uma isquemia cardíaca até uma lesão tecidual. Todas essas informações podem ser identificadas durante a interpretação de um ECG. Os intervalos PQ, PR, QT, ST e os intervalos RR, também são observados

e fontes de informação para a correta interpretação do ECG (CLIFFORD; AZUAJE; MCSHARRY, 2006).

Figura 3 - Perfil típico de um eletrocardiograma.



Fonte: Adaptado de (CLIFFORD; AZUAJE; MCSHARRY, 2006).

A onda que representa a repolarização dos átrios está escondida atrás do complexo QRS e é chamado de onda T- atrial, porém o impulso elétrico do QRS é tão grande que elas se sobrepõem a onda T-atrinal e geralmente está oculta e não aparece. Lembrando que a despolarização não significa contração, o que se observa em um ECG é uma atividade elétrica e não uma atividade mecânica.

Em eletrocardiograma a onda P e que representa a despolarização dos átrios, tem uma altura de dois a três mm e uma duração de 0,6 a 0,12 s e precede sempre complexo QRS, isso significa que o impulso elétrico está sendo originado no nó sinoatrial. O intervalo PR, descreve a distância entre o começo da despolarização dos átrios e o começo da despolarização dos ventrículos, ou seja, mostra um impulso elétrico passando pelo nó do atrioventricular e pelo feixe de His, essa distância dura entre 0,12 a 0,20 s para ser percorrida. Caso exista um atraso e o intervalo PR se

apresente mais alargado, significa que existe alguma lentificação nesse processo de passagem do impulso elétrico dos átrios para os ventrículos.

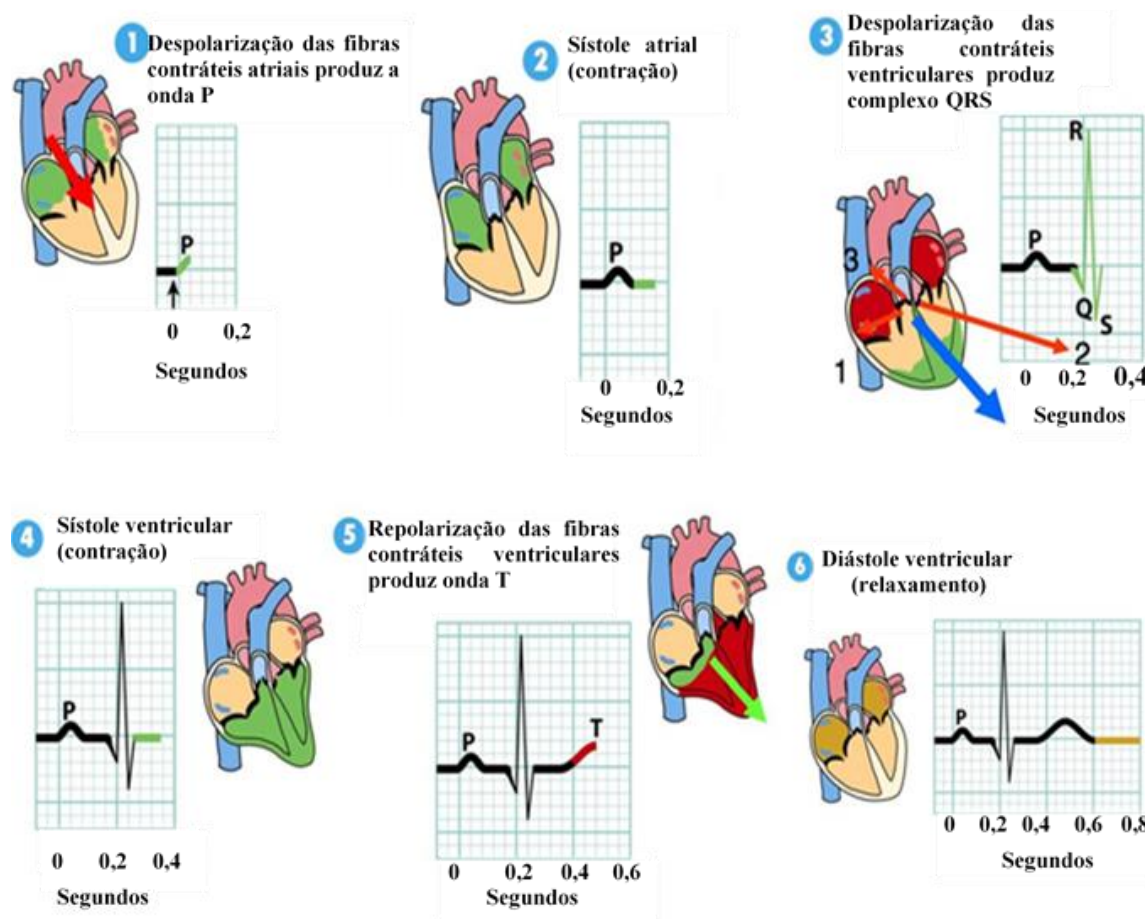
O complexo QRS, que mostra despolarização dos ventrículos tem uma altura de três a cinco mm e uma duração de 0,06 a 0,10 s. Um alargamento do complexo QRS significa que a condução do impulso elétrico pelo ventrículo está lenta e quando o QRS não pode ser observado no eletrocardiograma, significa que existe um bloqueio atrioventricular ou até mesmo uma parada ventricular.

O segmento ST, vai da onda S até onda T e mostra o final da despolarização do ventrículo até a repolarização do ventrículo. Esse segmento geralmente pode ir de - 0,5 a 1,0 mV, podendo estar deslocado para cima ou para baixo. Quando esse segmento está deslocado para baixo, de 2 mm ou mais, têm-se um infra desnivelamento do segmento ST, que reflete uma isquemia cardíaca, ou seja, a célula ainda está funcionando, porém, por alguma obstrução da coronária por exemplo, está chegando menos oxigênio na célula cardíaca. Quando o segmento ST está deslocado para cima, tem-se um supra desnivelamento do segmento ST, o que mostra que o coração apresenta uma lesão na célula cardíaca, já tendo ocorrido uma necrose no tecido.

A última onda T, que representa repolarização do ventrículo está localizada após a onda S e tem uma altura normalmente de 0,5 mm. Essa onda é arredondada e suave, geralmente positiva, sendo negativa em aVR. O intervalo QT, mostra o intervalo entre o começo da despolarização do ventrículo até a repolarização do ventrículo. Esse intervalo varia inversamente com a frequência cardíaca, entre 0,36 e 0,44 s, quando a frequência aumenta, o intervalo diminui, e vice-versa (KEENER; SNEYD, 2009). Estes processos são mostrados na Figura 4.

O intervalo de tempo entre batimentos cardíacos sucessivos é chamado intervalo RR, (ROSENTHAL, 2020) em que uma sequência temporal desses eventos é caracterizada como um tacograma. A análise das variações entre intervalos RR em uma série temporal é chamada variabilidade da frequência cardíaca, (IEVA; PAGANONI; ZANINI, 2013).

Figura 4 - Características morfológicas das ondas P, Q, R, S e T do ECG.



Fonte: Adaptado de (TORTORA, 2018).

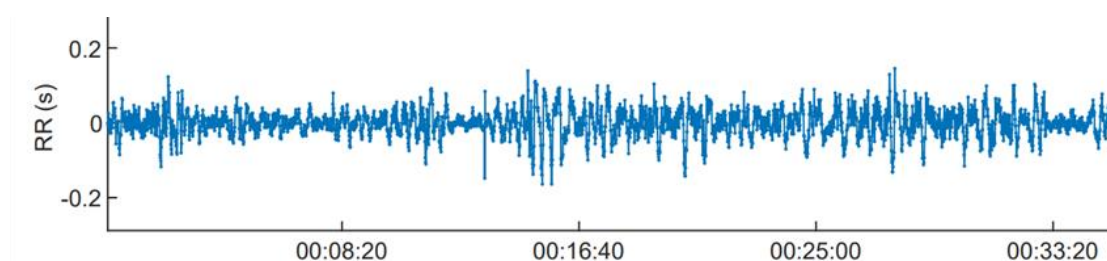
2.1.2 O tacograma RR: a variabilidade da frequência cardíaca

A utilização das séries da variabilidade da frequência cardíaca (VFC) obtidas em registradores de Holter é fundamental como meio auxiliar de diagnóstico. Então a VFC descreve estas variações que ocorre nesses intervalos e, que pode ser utilizado para avaliar a modulação do sistema nervoso autônomo (SNA), que constitui um papel importante na regulação destes processos fisiológicos do ritmo normal do coração, (STEIN et al., 1994), (RAJENDRA ACHARYA et al., 2006). Essas pulsações ou ritmos periódicos são, de fato, a fonte predominante da VFC, proporcionando um meio significativo para analisar a interação entre os sistemas nervoso simpático e parassimpático. Isto mostra que a estrutura em geral no sinal não é apenas linear, mas também envolve contribuições não lineares, (THURASINGHAM, 2006). Estes

registros de longa duração da frequência cardíaca, permitem identificar doenças, nomeadamente arritmias malignas, as quais não são detectadas nos registos de rotina do ECG. As séries longas de variabilidade da frequência cardíaca apresentam características não estacionárias complexas, com ocorrência de pontos de mudança, de valores em falta ou batimentos incorretamente detectados, evidenciando elevados níveis de incerteza, decorrentes da variabilidade das diversas situações da vida diária, (GLADULI et al., 2011).

O tacograma é uma representação da variabilidade resulta do sinal de ECG e representa a evolução temporal dos intervalos RR, como mostrado na Figura 5. A literatura reporta métodos de VFC para a análise de sinais de curta duração, sendo estes: métodos no domínio do tempo, métodos no domínio da frequência, além de métodos não lineares, (HAAKSMA et al., 2003). O primeiro método, refere-se ao cálculo de índices estatísticos do sinal, com a finalidade de interpretar as flutuações na duração dos ciclos cardíacos do ECG, (CLIFFORD; AZUAJE; MCSHARRY, 2006).

Figura 5 - O tacograma.



Fonte: O autor.

Entre estes parâmetros podem-se mencionar, a média, desvio padrão, e valores extremos do intervalo RR. Para o domínio da frequência, o fato do ECG apresentar flutuações, durante períodos curtos ou prolongados, dá origem a um fenômeno ondulatório, onde se define frequência e amplitude. Por fim, os métodos não lineares são baseados na análise de sistemas dinâmicos, em geral evidenciando a não linearidade das séries temporais de intervalos RR, (MCSHARRY et al., 2002).

Para obter um tacograma significativo, a maioria dos algoritmos de detecção de batida está na localização do pico R. A precisão da localização no tempo de cada pico e, portanto, a precisão do valor de cada intervalo entre batimentos que compreende o tacograma, depende, da taxa de amostragem na qual o ECG é digitalizado. Uma adequada frequência deve se manter pelo menos 250 Hz, indivíduos

com variabilidade extremamente baixa no tacograma podem exigir taxas de digitalização maiores, tais como 500 a 1000 Hz, para permitir uma relação sinal/ruído suficientemente alta e assim diferenciar entre as diferenças sutis na variabilidade cardíaca entre estes pacientes (KHANDOKER et al., 2013); (ABBOUD; BARNEA, 1995).

Outro fator importante nos sinais ECG, são as métricas, calculadas a partir de uma análise do tacograma. Sabe-se que a variabilidade na frequência cardíaca ocorre de batimento a batimento, logo o sinal é inerentemente espaçado de maneira desigual ao longo do eixo que representa o tempo. Este fato leva a complicações e erros nas métricas que utilizam interpolação. Existem várias maneiras de avaliar a métrica da variabilidade na frequência cardíaca, uma lista abrangente e detalhada de investigações pode ser observada na literatura (BEHAR et al., 2012). Eles podem ser divididos em três categorias básicas: aquelas que medem as propriedades estatísticas dos dados no domínio do tempo, as métricas no domínio da frequência e estimativas não-lineares.

Os parâmetros obtidos a partir de um tacograma, apresentados na Tabela 1, constituirão as variáveis de entrada para os modelos de classificação supervisionada multivariada que serão investigados nesse trabalho. A saber: um conjunto de onze variáveis relacionadas ao domínio de tempo, onze variáveis no domínio de frequência e doze variáveis obtidas por métodos não lineares.

Tabela 1 - Propriedades obtidas a partir de um tacograma.

	Parâmetro	Unidade	Descrição
Domínio do tempo	mean_RR	[ms]	média dos intervalos RR
	std_RR	[ms]	Desvio padrão dos intervalos RR
	mean_HRV	[batidas/min]	frequência cardíaca média
	std_HRV	[batidas/min]	Desvio padrão da frequência cardíaca
	RMSSD	[ms]	Raiz quadrada das diferenças quadrática média entre intervalos RR sucessivos
	NN50	[batidas]	Número de pares de intervalos RR sucessivos que diferem mais de 50 ms
	pNN50	[%]	NN50 dividido pelo número total de intervalos RR
	HRV_tri_ind	-	A integral do histograma do intervalo RR dividido pela altura do histograma
	TINN	[ms]	Largura da linha de base do histograma do intervalo RR
	SDANN	[ms]	Desvio padrão das médias dos intervalos RR em segmentos de 5 minutos
	SDNN_index	[ms]	Média dos desvios padrão dos intervalos RR no segmento de 5 min
Domínio da frequência	HF_peak	[Hz]	Frequência de pico da banda HF
	VLF_power	[ms ²]	Potências absoluta da banda VLF
	VLF_power_prc	[%]	Potências relativa da banda VLF
	LF_power	[ms ²]	Potências absoluta da banda LF
	LF_power_prc	[%]	Potências relativa da banda LF
	LF_power_nu	[n.u.]	Potências normalizada da banda LF
	HF_power	[ms ²]	Potências absoluta da banda HF
	HF_power_prc	[%]	Potências relativa da banda HF
	HF_power_nu	[n.u.]	Potências normalizada da banda HF
	LF_HF_power	-	Razão entre as potências da banda LF e HF
	tot_power	[ms ²]	potência total da banda
Parâmetros não lineares	Poincare_SD1	[ms]	No gráfico de Poincaré, o desvio padrão perpendicular à linha de identidade
	Poincare_SD2	[ms]	No gráfico de Poincaré, o desvio padrão ao longo da linha de identidade
	ApEn	-	Entropia aproximada
	SampEn	-	Entropia da amostra
	DFA.alpha1	-	Na análise de flutuação sem tendência, inclinação de flutuação de curto prazo
	DFA.alpha2	-	Na análise de flutuação sem tendência, inclinação de flutuação de longo prazo
	RPA.Lmax	[batidas]	Comprimento máximo da linha, gráfico de recorrência
	RPA.Lmean	[batidas]	Comprimento médio da linha, gráfico de recorrência
	RPA.DIV		Divergência
	RPA.REC	[%]	Taxa de recorrência
	RPA.DET	[%]	Determinismo
RPA.ShanEn	-	Entropia de Shannon	

Fonte: O autor.

2.2 Modelos preditivos baseados em classificadores multivariados

O ECG é acessível, não invasivo e amplamente utilizado na prática clínica. Como um registro da atividade elétrica da superfície corporal, ele fornece informações sobre anormalidades do ritmo cardíaco e ajuda a detectar doenças. No entanto, a inspeção visual do ECG fornece recursos discretos, interpretados clinicamente, que

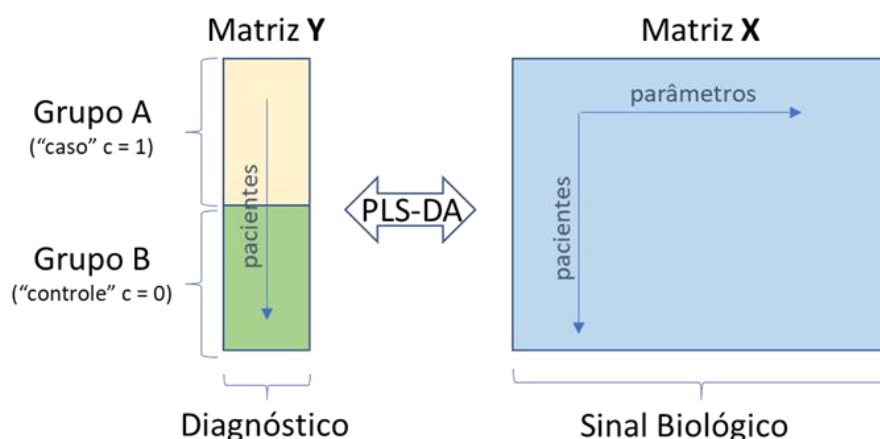
não podem capturar objetivamente a diversidade de anormalidades e morfologias do ECG. Por isso os métodos computacionais são necessários, pois, consegue-se manipular conjuntos de dados complexos multivariados e detectar diferenças que podem ser desafiadoras para o olho humano.

2.2.1 Mínimos quadrados parciais - análise discriminante (PLS-DA)

O método de Mínimos Quadrados Parciais (PLS) é um método multivariado utilizado para avaliar uma relação entre uma matriz de descritora e uma matriz resposta e de previsão, dadas variáveis preditivas (X_1, X_2, \dots, X_p) , com características colineares. Neste algoritmo o foco está em prever a resposta e não em entender o relacionamento entre as variáveis preditoras, isto é, a relação entre variáveis preditivas sem considerar a previsão de resposta final, Y . O PLS cria as combinações lineares das variáveis preditivas, que serão usadas para fins de previsão, (BYLESJÖ et al., 2006).

A estrutura para entendimento de um modelo PLS baseia-se na redução da dimensionalidade de uma matriz, ou seja, é um método de cálculos entre matrizes binárias de dados que descreve as categorias de uma variável categórica em um conjunto de variáveis preditoras. Para tal, duas matrizes são o ponto de partida: a matriz preditora, comumente denotada por X , e matriz predita Y , Figura 6. Esta técnica, procura a covariância máxima entre as duas matrizes X (independente) e Y (dependente), para cada componente, de forma que a matriz preditora contenha todas as variáveis do processo (PÉREZ-ENCISO; TENENHAUS, 2003).

Figura 6 - Representação esquemática do PLS-DA.



Fonte: O autor.

A Figura 6 pode ser representada algebricamente em termos da matriz Y sendo obtida a partir da matriz X ponderada com uma matriz de pesos B , acrescida dos resíduos como (KALIVODOVÁ et al., 2015):

$$Y = XB + E \quad (1)$$

a matriz Y é a matriz resposta contendo a previsão do diagnóstico, como enfermo e sadio, e a matriz B é a matriz de dados contendo o sinal biológico medido em cada paciente, variáveis obtidas a partir do tacograma.

Com as matrizes X e Y , são realizadas as decomposições das duas matrizes, na forma do produto de uma matriz de pontuação por uma matriz de carregamento, acrescida de uma matriz residual. Para a matriz X , a decomposição é expressa por:

$$X = TP^T + E_X \quad (2)$$

sendo T a matriz de *scores* e P a matriz de *loadings* e E é a matriz dos resíduos. Analogamente obtém-se a matriz predita Y :

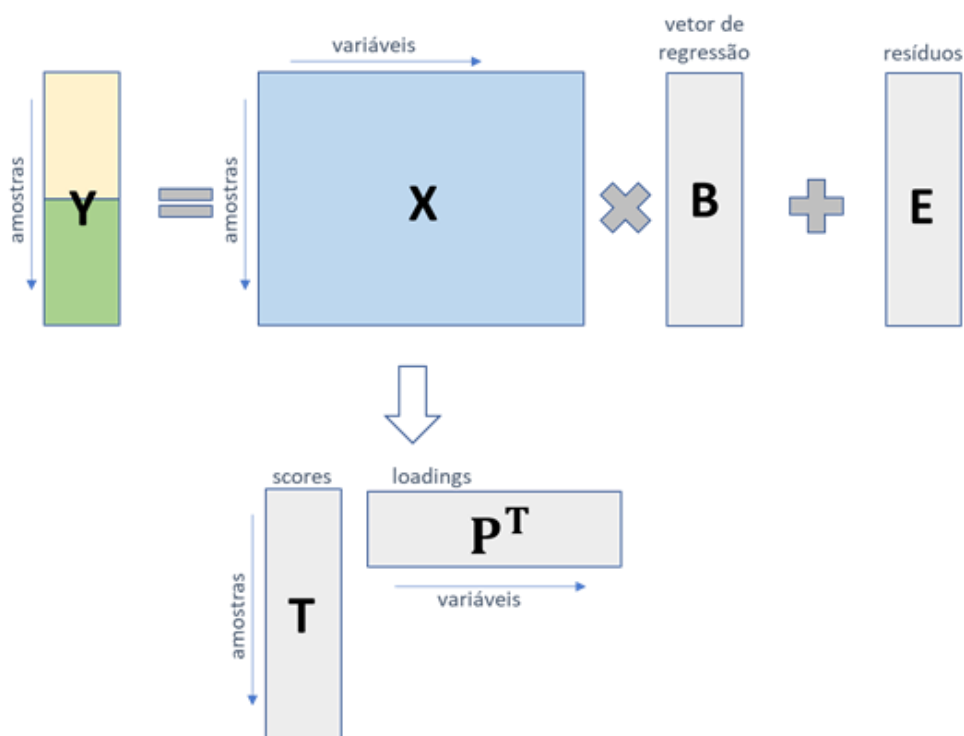
$$Y = UQ^T + E_Y \quad (3)$$

onde U é a matriz de *scores*, Q é a matriz de *loadings* e E é a matriz dos resíduos. Fazendo uma substituição das equações acima obtém-se:

$$UQ^T = TP^TB + E \quad (4)$$

neste ponto inicia-se a resolução de um problema computacional cujo objetivo é apurar esses conjuntos de equações matriciais estabelecendo o máximo de covariância entre X e Y (eq. 1), T e U (eq. 4). A decomposição da matriz X em *scores* e *loadings* é representada na Figura 7. Nesta figura fica ilustrado que a predição Y é obtida pela multiplicação da matriz X, contendo os dados de sinais biológicos, pela matriz B, que representa os vetores de regressão do modelo PLS.

Figura 7 - Decomposição de matrizes em modelos PLS.



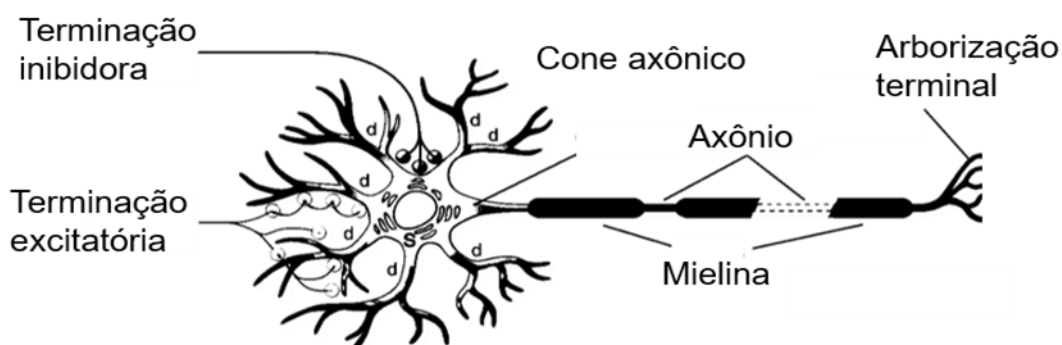
Fonte: O autor.

Os cálculos matriciais para decomposição das matrizes acima citadas e obtenção dos vetores de regressão PLS são realizados por diferentes algoritmos computacionais como Kernel-PLS, (WANG; YAN; FEI, 2015), SIMPLS, (DE JONG, 1993) e NIPLS, (MALTHOUSE; TAMHANE; MAH, 1997). Contudo, softwares de análises de dados como Matlab, R e outros possuem rotinas já estabelecidas para tal.

2.2.2 Redes neurais artificiais (ANN)

Uma rede neural artificial é um sistema de processamento de informação que possui algumas características de desempenho em comum com as redes neurais biológicas. A analogia utilizada para compreender a técnica denominada ANN é cunhada na compreensão da função de um neurônio. No cérebro humano, cada neurônio é tipicamente conectado a milhares de outros neurônios. Um neurônio coleta sinais de outros através de uma série de estruturas finas chamadas dendritos. O neurônio envia picos de atividade elétrica através de uma longa e estrutura conhecida como axônio, que se divide em vários ramos conhecidas como terminais axônios ou terminais nervosos. No final de cada ramo, uma estrutura chamada sinapse, converte a atividade do axônio em efeitos elétricos, que inibem ou excitam a atividade do axônio em efeitos elétricos, excitam a atividade nos neurônios conectados (CARTWRIGHT (EDITOR), 2020). Quando um neurônio, recebe uma entrada excitatória que é suficientemente grande em comparação com sua entrada inibitória, ele envia um pico de atividade elétrica para baixo de seu axônio, veja Figura 8.

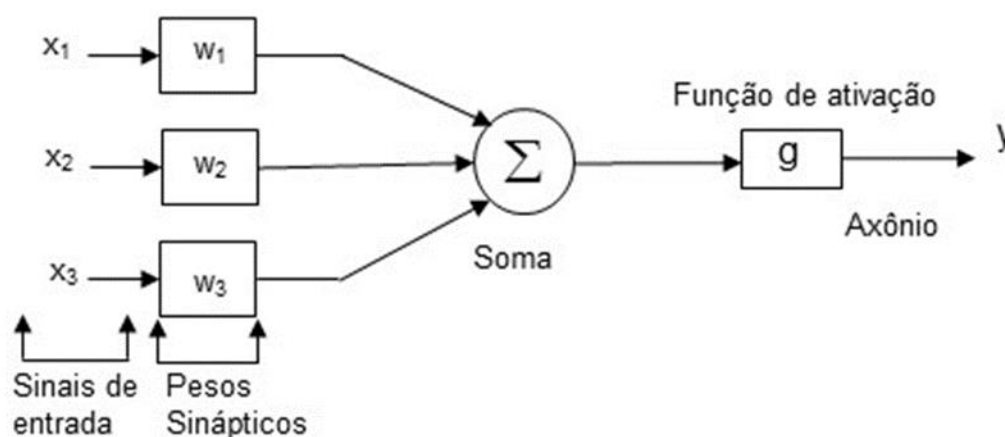
Figura 8 – Esquema de um neurônio (biológico).



Fonte: Adaptado de (HUDSON; COHEN, 1999).

As redes neurais artificiais (ANN), tornaram-se um método popular de inteligência artificial, devido ao sucesso recente no aprendizado de máquina, assim como reconhecimento de padrões. Uma ANN, simula o mecanismo da rede neural humana, usando pesos adaptativos entre as camadas da rede, para modelar sistemas muito complicados, Figura 9.

Figura 9 - Esquema de um neurônio (computacional).



Fonte: Adaptado de (HAYKIN, 2008).

A aprendizagem em sistemas biológicos, envolve ajustes nas conexões sinápticas, que existem entre os neurônios. Uma ANN, apresenta aprendizado adaptável (a capacidade de aprender a executar tarefas, com base nos dados fornecidos para treinamento) e auto-organização (criando sua própria organização ou representação das informações que recebe na fase de aprendizado). O modelo matemático simples de um neurônio foi introduzido por McCulloch e Pitts em 1943. O modelo neural inclui:

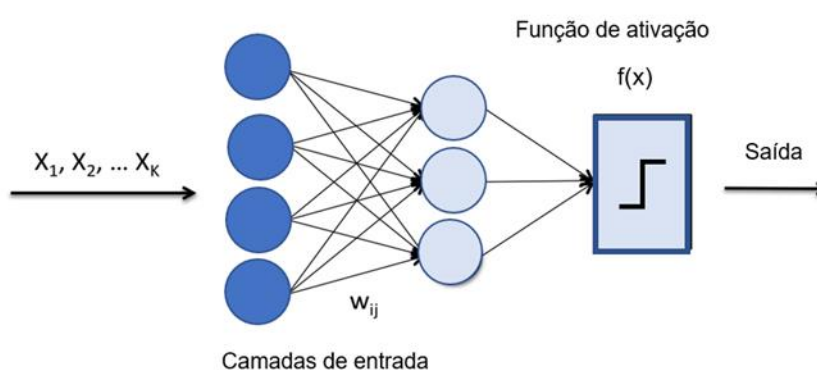
1. Um conjunto de entradas ponderadas $w_i (i = 1, \dots, k)$, que correspondem às sinapses. Um w_i positivo (negativo) corresponde a conexões excitatórias (inibitórias), que tornam os neurônios mais (menos) propensos a disparar.
2. Um somador que soma os sinais de entrada (equivalente à membrana da célula que coleta carga elétrica).
3. Uma função θ de ativação (limiar) que decide se o neurônio dispara (picos) para as entradas atuais, (MCCULLOCH; PITTS, 1943).

O modelo dos neurônios McCulloch e Pitts, mostrado na Figura 9, é um modelo linear. Sua função é muito limitada, pois, o modelo de neurônio é ativado (saída 1) ou inativo (saída 0). Neurônios reais, são muito mais complicados, uma diferença significativa, é que o neurônio real não emite uma única resposta de saída, mas emite uma sequência de respostas (ou seja, uma sequência de pulso), e é essa sequência de pico que codifica a informação. Isso significa que o neurônio, não responderá como um dispositivo de limiar, mas produzirá continuamente uma saída hierárquica, ou seja, os limites que eles disparam mudarão com o tempo. Como os

neurônios são dispositivos bioquímicos, o número de neurotransmissores (entre outros fatores que afetam a carga necessária para emitir) muda de acordo com o estado atual do organismo. Além disso, os neurônios são atualizados de forma assíncrona ou aleatória.

Entre os parâmetros no modelo de neurônio McCulloch-Pitts: pesos, soma de funções e limiares, pesos (modelagem de sinapses), são os mais interessantes, porque podem mudar durante o processo de aprendizado. Ou seja, o aprendizado é basicamente modelado por mudanças nos pesos. Para uma determinada rede neural, a função soma e o limiar são geralmente escolhidos por simplicidade. O perceptron é uma extensão do modelo de neurônio McCulloch-Pitts, para múltiplos neurônios, Figura 10. Como cada neurônio tem duas saídas possíveis, n neurônios juntos têm 2^n saídas/respostas diferentes e cada combinação possui um código binário de n bits que pode ser atribuído a significados diferentes.

Figura 10 - Rede neural perceptron com k -neurônios de entrada.



Fonte: Adaptado de (HAYKIN, 2008).

Até este ponto mostrou-se que uma ANN é constituída por uma camada de entrada (*input*) de dados, uma camada de cálculos intermediários destes dados (conhecida como camada oculta (*hidden*), e uma camada de saída (*output*), que são os resultados. Cada nódulo nesta camada é chamado então de neurônio e a arquitetura de uma rede neural pode conter de alguns poucos a dezenas de neurônios por cada camada.

A função matemática que transforma a informação contida num neurônio ao passar por cada camada é a função de ativação. Geralmente é uma função sigmoideal (ou logística), ou tangente hiperbólica, ou função exponencial ou outra, como uma

função de comportamento linear ou novas implementações como a função Leaky ReLU (AGARWAL et al., 2020). Assim, os neurônios da camada de entrada, cada neurônio representa uma variável biológica medida, é interligado ao neurônio da camada subsequente, no qual a informação propaga mudando seus pesos conforme passa por cada neurônio contendo uma determinada função de ativação. A complexidade destas redes de neurônios aumenta conforme aumentam os números de neurônios e o número de camadas. Redes perceptron, mostradas anteriormente, são muito populares e podem conter algoritmos de retropropagação (do inglês *backpropagation*) (HAYKIN, 2008). Nesta abordagem, o erro obtido ao final de uma estimativa volta à camada anterior, propagando-se da direita para esquerda, sendo os pesos de cada conexão recalculados, otimizando a rede neural para minimizar os erros em suas determinações finais.

2.2.3 Performance de modelos preditivos: Figuras de mérito

A determinação da acurácia do modelo matemático de classificação de dados é realizada comparando-se os resultados obtidos pelo classificador com o laudo clínico, padrão ouro, fornecido pelo médico responsável pela análise de rotina. Estes parâmetros que indicam a qualidade ou não destes modelos são chamados de figuras de mérito.

Dentre o conjunto de figuras de mérito utilizadas em abordagens qualitativas e quantitativas somente serão utilizadas aqui aquelas relacionadas à avaliação de modelos de classificação de dados. Assim, para uma classificação binária (1 ou 0), como é o caso de um diagnóstico do tipo enfermo ou sadio, as figuras de mérito são apresentadas por um conjunto de equações baseadas na contagem das determinações corretas e incorretas das duas classes em questão.

Diz-se um verdadeiro positivo (VP) quando uma estimativa feita pelo classificador matemático que está sendo testado faz uma previsão correta para a classe 1 (caso, ou enfermo), ou seja, um indivíduo que possui uma enfermidade é classificado como enfermo pelo algoritmo matemático. Diz-se um verdadeiro negativo (VN) quando este classificador faz uma previsão correta para a classe 0 (controle, ou sadio), ou seja, um indivíduo que está sadio é diagnosticado como indivíduo normal

pelo classificador matemático. Assim, falso positivo (FP) é referente a uma previsão incorreta para classe caso, ou seja, um indivíduo que está sadio é erroneamente classificado como doente pelo classificador matemático e um falso negativo (FN) refere-se a uma estimativa incorreta para a classe controle, ou seja, um indivíduo enfermo é incorretamente classificado como sadio.

A partir da contagem do número de VP, VN, FP e FN as seguintes figuras de mérito podem ser estimadas: sensibilidade, seletividade, taxa de falso positivo, taxa de falso negativo, eficiência, índice de Younden's (FLUSS; FARAGGI; REISER, 2005) e coeficiente de correlação de Matthews (BALLABIO; GRISONI; TODESCHINI, 2018).

A sensibilidade é interpretada como sendo a capacidade do sistema em prever corretamente a condição (indivíduo doente) para casos que realmente a têm.

$$\frac{VP}{VP + FN} \quad (5)$$

A seletividade é a capacidade do sistema em prever corretamente a ausência da condição (indivíduo sadio) para casos que realmente não a têm.

$$\frac{VN}{VN + FP} \quad (6)$$

A taxa de falso positivo refere-se as determinações incorretas de indivíduos sadios classificados como enfermos.

$$\frac{FP}{VN + FP} \quad (7)$$

A taxa de falso negativo refere-se as determinações incorretas de indivíduos doentes classificados como sadios.

$$\frac{FN}{VP + FN} \quad (8)$$

A eficiência é um estimador que pondera as determinações verdadeiras de caso e controle sobre o universo de dados. Este parâmetro é interessante pois na

maior parte de aplicações na área da saúde a sensibilidade e a seletividade caminham em direções opostas, ou seja, quando um método é muito sensível a positivos, tende a gerar muitos falso-positivos, e vice-versa. Assim a eficiência agrupa, em partes, sensibilidade e seletividade. De forma análoga pode ser interpretado o também utilizado índice de Younden's, (FLUSS; FARAGGI; REISER, 2005).

$$\frac{VN + VP}{VN + FP + VP + FN} \quad (9)$$

$$(\text{Sensibilidade} + \text{Seletividade} - 1) \quad (10)$$

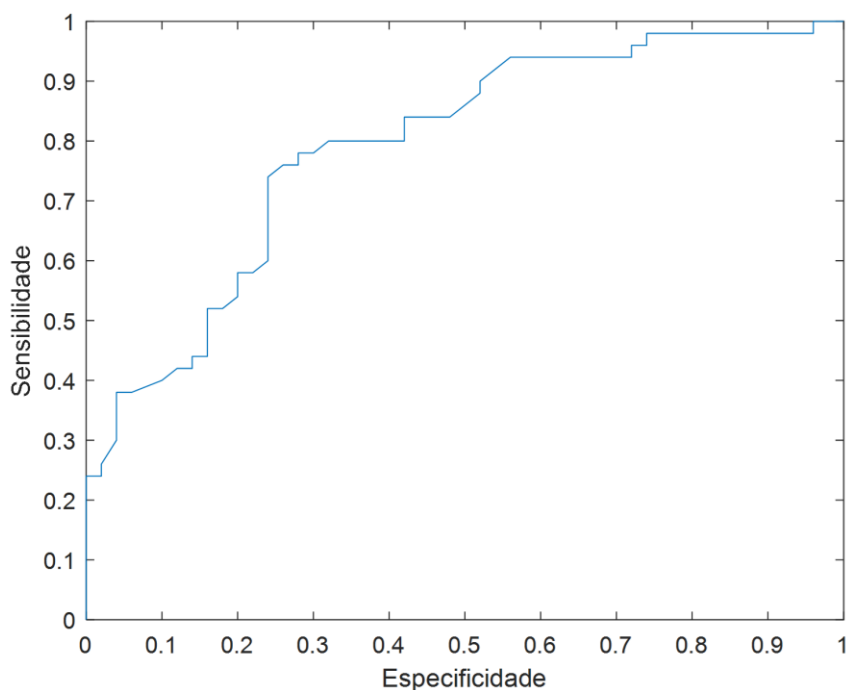
O coeficiente de correlação de Matthews é uma medida de qualidade do classificador, retornando valores compreendidos entre (-1) e (+1), na qual um coeficiente +1 representa uma predição perfeita, 0 representa uma predição aleatória e -1 uma predição inversa.

$$\frac{VP \times VN - FP \times FN}{\sqrt{(VP + FP)(VP + FN)(VN + FP)(VN + FN)}} \quad (11)$$

Estimadores como a acurácia, valor preditivo positivo e valor preditivo negativo não serão utilizados pois tendem a induzir conclusões equivocadas quando os grupos testados estão desbalanceados.

Além destas figuras de mérito existe uma informação gráfica muito útil para avaliar a performance de classificadores, a denominada curva ROC (do inglês *Receiver Operating Characteristic*). A Figura 11, representa um exemplo de curva ROC, classificando a especificidade e Sensibilidade (HAND; TILL, 2001).

Figura 11 – Exemplo de uma curva ROC.



Fonte: O autor.

A curva ROC é esboçada no plano bidimensional, como no plano cartesiano, com eixos representado a taxa de verdadeiros positivos (sensibilidade) no eixo vertical e a taxa de falsos positivos ($1 -$ seletividade) no eixo horizontal. Assim, visualmente é possível apreciar tanto sensibilidade e seletividade do classificador, além disso, os lados de uma curva ROC possuem valor máximo 1 (sensibilidade ou seletividade igual a 100%), de modo que a área abaixo da curva ROC apresentara valor máximo igual a 1. Esta informação é útil pois a área abaixo da curva ROC próxima de 0,9 é obtida para um classificador que distingue muito bem as classes caso e controle. Área próxima de 0,8 tem-se uma separação razoável, 0,7 uma separação ruim e 0,5 as estimativas são totalmente aleatórias.

2.3 Bancos de dados

A digitalização de dados do setor de saúde está acontecendo vertiginosamente. Um dos principais resultados dessa transformação são os registros em mídias eletrônicas é a proliferação de dados de saúde. E com isso, é claro, surgem os bancos de dados de saúde. Um banco de dados é qualquer coleção de dados organizados

para armazenamento, acessibilidade e recuperação. Assim, a divulgação de informações de estudos clínicos acelerou uma parcela importante de trabalhos acadêmicos na área de desenvolvimento de abordagens de inteligência artificial para reconhecimento de padrões e diagnósticos na saúde (TOPOL; J, 2019).

Alguns desses bancos de dados fornecem informações cruciais para o desenvolvimento de modelos de diagnóstico que futuramente farão parte de análises de rotina da prática médica além de caminharem juntamente com o desenvolvimento de novos equipamentos dotados de ferramentas de diagnóstico automatizado para a área da saúde. Assim, bases de dados como Human Mortality Database (WILMOTH; SHKOLNIKOV; BARBIERI, 2020), National Cancer Institute Database (SHARPLESS; LOWY, 2020), World in Data COVID-19 dataset (ROSER et al., 2020), PhysioBank (GOLDBERGER et al., 2000), Lung Image Database Consortium Data Collection (MCNITT-GRAY et al., 2007) e muitas outras são cruciais para desenvolvimento de aplicações inovadoras de aprendizado de máquina na análise da saúde. Esta tese também fará uso de banco de dados de sinais de ECG de acesso público, *PhysioNet*.

2.4 Inteligência artificial na área da saúde

A aplicação dos métodos de aprendizagem de máquina e inteligência artificial (IA), junto à geração de *big data*, tornou-se uma tendência importante para prevenção, detecção de doenças e na identificação dos melhores métodos de tratamento. A Big data é caracterizada pela grande variedade de tipos de dados e fontes de dados, bem como seus aplicativos e no setor da saúde, uma grande variedade de dados é adquirida, incluindo novos tipos do que não existia antes do surgimento dos dispositivos inteligentes. Com o avanço da tecnologia, e com a existência de ferramentas de diagnósticos por meio de *softwares*, observa-se, até mesmo, uma redução na realização de cirurgias invasivas e de alto risco. Algoritmos de aprendizagem de máquina tem proporcionado a profissionais que atuam em conjunto com a área da saúde, processar grandes quantidades de dados de saúde de forma automatizada e instantânea, tornando a identificação das doenças mais rápida e confiável (BOHR; MEMARZADEH, 2020).

Então, a medicina preditiva em colaboração com a Big data de dados da saúde com os avanços da tecnologia da inteligência Artificial, comporão além dos *softwares*, transformar a prática atual do atendimento clínico, então registros eletrônicos de dados clínicos passaram a formar grandes bases de dados digitais dos pacientes. Estas fontes de dados, alimentadas deste o preenchimento do prontuário de cada paciente, auxiliam na triagem dos mesmos, relacionam com informações de histórico familiar do indivíduo, e já classificam como pertencente a determinado grupo de prioridade ou não. Muitos algoritmos para detecção automática e de classificação de padrões de batimentos cardíacos de ECG foram apresentados na literatura que demonstram o sucesso nos dados na área da saúde.

Aplicações de inteligência artificial (LI GANG et al., 2000) e redes neurais (SILIPO; MARCHESI, 1998 e VAN GILS *et al.*, 1997) para análise automatizada de ECG foram descritas. Van Gils *et al*, mostrou que as técnicas de redes neurais, podem ser utilizadas para fins de classificação de ECG , Barro *et al*, (BARRO et al., 1998) usou a teoria de ressonância adaptativa multicanal para classificar os padrões de ECG com uma precisão de 90%. Frenkel *et al*, utilizou uma rede neural treinada em estudo de isquemia (FRENKEL; NADAL, 1999).

Redes neurais artificiais (ANNs) são ferramentas poderosas para reconhecimentos de padrões pois, têm a capacidade de aprender e resolver problemas complexos e não lineares entre diferentes classes e, isto pode ser a chave para reconhecimento e classificação de batimentos de ECG. Assim, Soowhan fez uma classificação de formas de onda de ECG, usando as redes neurais e algoritmos genéticos, obtendo-se ótimos resultados (HAM; SOOWHAN HAN, 1996). Engin, 2004, utilizou a rede neural híbrida-difusa e Transformadas Wavelet, para classificação de batimento de ECG (ENGIN, 2004). Em 2005 GÜLER descreveu um modelo de rede neural para classificação de batimentos de ECG, obtendo-se uma precisão de diagnóstico de 96,94% (GÜLER; ÜBEYLI, 2005). Em 2012 Sarkaleh, propôs um sistema para classificação de arritmias de ECG, utilizando a Transformação Wavelet e a rede neural Multi-Layer Perceptron, com precisão da classificação de 96,5% (SARKALEH, 2012), (ELSAYAD; ELSALAMONY, 2013). Li, *et al*, 2014, desenvolveram um algoritmo de classificação da qualidade do sinal de ECG quanto ao ruído. Para tal fim, uma máquina de vetores de suporte (SVM) foi treinada para

realizar a classificação, obtendo-se uma precisão de 89,46% (LI; RAJAGOPALAN; CLIFFORD, 2014), (ZHAO; ZHANG, 2005), (ZHANG; LUO, 2014).

Em 2014, com avanços das tecnologias de telecomunicações, novas abordagens foram feitas, tais como um sistema suporte e decisão clínica, com reconhecimento automático do ECG em tempo real, com implementação de máquina de vetores de suporte (SVM), e algoritmos de aprendizagem automática. Estes modelos de classificação automática, obtiveram 88,4% de sensibilidade e especificidade de 85,9% (HO et al., 2014). Em 2017, Kshirsagar *et al.*, apresentaram uma aplicação de redes neurais probabilísticas, para a classificação e detecção de ECG, com a ajuda de Wavelet (KSHIRSAGAR; AKOJWAR; DHANORIYA, 2017). Em 2018, Lyon, fez uma análise de classificação de ECG utilizando técnicas de aprendizado de máquina e modelagem computacional, *clustering* e simulação, para abordar a análise de dados médicos (LYON et al., 2018). Avanços recentes em dispositivos sensores de ECG oferecem oportunidades para o autogerenciamento do usuário, serviços de autodiagnóstico e monitoramento pela internet. No trabalho de Guo *et al.*, propuseram um modelo de aprendizado profundo, utilizando a rede neural convolucional densamente conectada, para abordar a classificação de ECG, com resultados satisfatórios (GUO; SIM; MATUSZEWSKI, 2019). Nesse ano de 2020, Hua et al., estabelece um método de classificação de sinal de ECG com base em uma nova estratégia de segmentação via redes neurais convolucionais, foi realizado, para auxiliar na classificação de sinais de ECG e obtendo-se a melhor classificação, com boa sensibilidade e área sob a curva AUC (HUA et al., 2020).

Por outro lado, o PLS-DA junto com a técnica de árvore de decisão, o PLS-DA, demonstrou grande sucesso na análise de classificação de tecidos mamários maligno e benigno (RASHID et al., 2020); (PÉREZ-ENCISO; TENENHAUS, 2003b); (SZI-WEN CHEN, 2000) e autenticação de produtos em análises de alimentos (BALLABIO; CONSONNI, 2013).

Como descrito acima, algumas abordagens envolvendo aprendizado de máquina com métodos ANN foram aplicados em tacogramas RR. Nestes casos as redes neurais eram redes que ajustavam parâmetros sobre as próprias séries temporais de intervalos RR. A modelagem ANN que vamos discutir é realizada sobre parâmetros estatísticos obtidos a partir destes tacogramas. Esses parâmetros são conhecidos e bem estabelecidos em estudos de variabilidade da frequência cardíaca.

Além disso, a abordagem PLS-DA figura como um diferencial desta tese para aplicações em dados como estes, não sendo encontrado qualquer relato na literatura de modelagem PLS-DA sobre variáveis associadas a variabilidade da frequência cardíaca.

3 PROPOSIÇÃO

3.1 Objetivo Geral

Demonstrar o potencial do uso de técnicas de classificação multivariada aplicadas a dados biológicos, em especial a variabilidade da frequência cardíaca, com o fim de auxiliar profissionais da saúde proporcionando um auxílio diagnóstico automatizado de enfermidades por meio de métodos computacionais de aprendizagem de máquina.

3.2 Objetivos Específicos

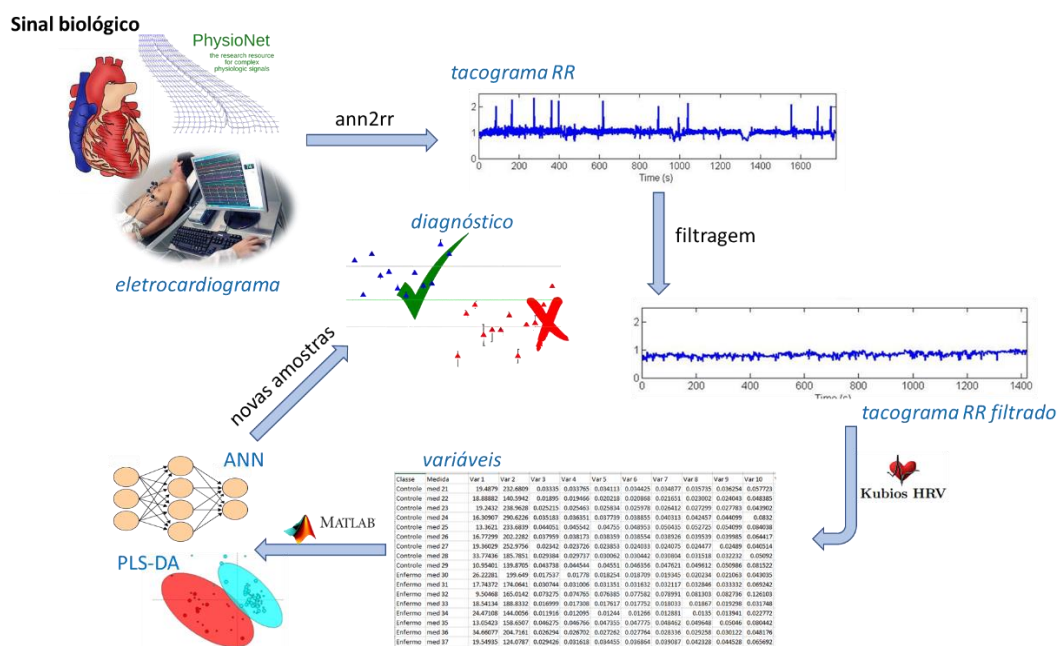
- Obtenção de parâmetros característicos de tacogramas obtidos a partir de sinais reais do tipo eletrocardiogramas com a utilização de ferramentas de filtragem de sinais e utilização do *software* Kubios.
- Ajuste de modelos de classificação supervisionada por métodos estatísticos multivariados, a saber: PLS-DA e ANN, sobre matrizes de dados separadas por variáveis características do domínio de tempo, domínio da frequência, parâmetros não-lineares e a fusão de todos estes supracitados, característicos de tacogramas.
- Aplicar os modelos de classificação a um conjunto de dados reais obtidos a partir de banco de dados público. Dois estudos de casos serão investigados: insuficiência cardíaca congestiva e supressão de arritmia cardíaca.
- Determinação das figuras de mérito dos modelos de classificação através dos cálculos de um conjunto de parâmetros que avaliam a performance destes modelos.

4 MATERIAIS E MÉTODOS

As etapas de obtenção dos dados, filtragem, transformação dos dados, ajuste de modelos de classificação assim como a avaliação das previsões obtidas será apresentada nas subseções que seguem. A Figura 12 organiza sequencialmente esse processo ilustrando graficamente as etapas que serão executadas.

A primeira etapa é a obtenção de dados do tipo ECG distribuídos livremente na base de dados *PhysioNet*, na sequência, a aplicação *ann2rr* (do próprio *PhysioNet*) extrai os tacogramas RR do sinal ECG. Os tacogramas são então filtrados através de um filtro adaptativo antes de serem importados para o software Kubios. O software Kubios é utilizado para extrair os parâmetros estatísticos característicos de cada tacograma individual, esses parâmetros constituem o conjunto de variáveis que serão utilizadas nos modelos multivariados. Modelos de previsão ANN e PLS-DA são ajustados sobre a matriz de dados com tais variáveis. Finalmente, amostras que não participaram do ajuste destes modelos são utilizadas para previsão e comparação com os laudos clínicos associados a cada indivíduo, obtendo-se então um modelo diagnóstico no qual figuras de mérito foram calculadas para avaliação da performance dos mesmos.

Figura 12 - Fluxo de trabalho.



Fonte: O autor.

4.1 Obtenção dos tacogramas

PhysioNet, é um banco de acesso gratuito e é suportado por servidores da Web hospedada em vários locais ao redor do mundo, para prover acesso confiável à comunidade de pesquisa. O *PhysioNet* está composto pelo *PhysioBank*, disponibilizando bancos de dados clínicos de sinais fisiológicos e o *PhysioToolkit*, proporcionando as ferramentas necessárias como softwares de código aberto. Não há nenhum custo para usar este banco de dados, e pode-se iniciar sua exploração por o *PhysioBank*, que contém dicas sobre amostras dos bancos de dados disponíveis e sugestões para visualizá-los.

Os dados utilizados nesta tese foram obtidos diretamente do banco de dados *PhysioNet*, todos disponibilizados gratuitamente para *download*. Trata-se de três conjuntos de sinais biológicos do tipo ECG de longa duração, de indivíduos humanos, agrupados por quadros clínicos específicos. Foram selecionados dados provenientes de três grupos de pacientes: um grupo de indivíduos com ritmo sinusal normal, outro com indivíduos diagnosticados com insuficiência cardíaca congestiva e um ensaio de supressão de arritmia cardíaca.

O banco de dados com indivíduos com ritmo sinusal normal, denominada *Normal Sinus Rhythm RR Interval Database*, compreende ECG de longo prazo de 30 homens, com idades entre 28,5 e 76 anos, e 24 mulheres, com idades entre 58 e 73 anos. Aqui, este conjunto de dados será denominado como grupo controle nas aplicações que serão discutidas nas seções subsequentes.

O banco de dados com indivíduos diagnosticados com insuficiência cardíaca congestiva, denominado *Congestive Heart Failure RR Interval Database*, inclui 29 gravações de ECG de indivíduos com idade entre 34 e 79 anos, com insuficiência cardíaca congestiva (classes I, II e III da NYHA). Os sujeitos incluíram 8 homens e 2 mulheres; gênero não é conhecido para os 19 restantes. Este conjunto de dados será utilizado na primeira comparação caso versus controle que será discutida nesse trabalho.

O banco de dados do estudo de supressão de arritmia cardíaca, denominado *The CAST RR Interval Sub-Study Database*, foi um estudo de grande escala compreendendo mais de 3500 voluntários. Trata-se de um estudo com indivíduos que sobreviveram ao infarto do miocárdio sendo então três medicamentos testados

objetivando verificar a melhora na qualidade de vida pós infarto. Para todos as avaliações de eficácia ou não dos fármacos testados foi também registrado o ECG no indivíduo antes da medicação. Cinquenta e quatro eletrocardiogramas deste grupo de dados, sem medicação, foi o selecionado para a segunda comparação do tipo caso versus controle que será discutida nesse trabalho.

4.2 Pré-processamento de dados

Os arquivos disponíveis para *download* diretamente do banco de dados *PhysioNet* foram baixados, extraídas as anotações referentes aos registros de eletrocardiograma. A partir de cada eletrocardiograma a ferramenta *ann2rr* disponibilizada pelo *Physionet* foi aplicada aos sinais do ECG e obtidos os intervalos RR, que constituem assim o tacograma. Cada tacograma foi então filtrado pela aplicação de filtragem adaptativa desenvolvida por (DOS SANTOS et al., 2016).

O *software* Kubios (TARVAINEN et al., 2014) foi utilizado para obtenção dos trinta e quatro parâmetros característicos de cada tacograma. Assim, para cada tacograma foram obtidas onze variáveis relacionadas ao domínio de tempo, onze variáveis no domínio de frequência e doze variáveis obtidas por métodos não lineares. Esse conjunto de parâmetros serão as variáveis a serem modeladas pelas técnicas de classificação e previsão PLS-DA e ANN. Cada base de dados obtidas a partir do banco *PhysioNet* gerou uma matriz de dados de dimensão $n \times m$, sendo n o número de indivíduos e m as 34 variáveis descritas na Tabela 1.

4.3 Inspeção dos dados

Os dados pré-processados foram separados por grupos: ritmo sinusal normal (54 indivíduos), insuficiência cardíaca congestiva (29 indivíduos) e supressão de arritmia cardíaca (54 indivíduos); e por origem do sinal: variáveis relacionadas ao domínio de tempo (onze), variáveis no domínio de frequência (onze) e variáveis obtidas por métodos não lineares (doze). Para cada subconjunto formado foi realizada uma inspeção dos dados através da análise do gráfico boxplot, objetivando verificar a

largura de dispersão dos dados, por grupo, e a possibilidade de presença de valores anômalos, ou *outliers*.

Nesta análise de triagem foi realizada uma série de testes de hipóteses, saber: teste de Lilliefors, teste-F, teste-*t* e Kruskal-Wallis, ao nível de 95% de confiança, objetivando conhecer características dos dados, como dispersão e característica de grupo. A análise estatística dos dados foi realizada no software Matlab.

4.4 Modelos de classificação supervisionada

Em cada grupo de amostras utilizado nos modelos de previsão, tanto PLS-DA como ANN, os dados foram divididos em conjuntos de treinamento e de teste. Para essa divisão foi utilizado o algoritmo de Kennard-Stone (KENNARD; STONE, 1969) e considerou-se a proporção de 2/3 das amostras para treinamento e 1/3 para teste, de acordo com a norma ASTM E166563.

Modelos PLS-DA e ANN foram ajustados e otimizados explorando condições como número de variáveis latentes, para a abordagem PLS-DA, e número de neurônios e camadas, para a abordagem ANN. Para cada modelo testado foi obtido um valor limiar de separação entre as classes (*threshold*) e a partir dele realizada a contagem de VP, VN, FP e FN sobre os dados de teste. Os modelos foram avaliados por um conjunto de figuras de mérito: sensibilidade, seletividade, taxa de falso positivo, taxa de falso negativo, eficiência, índice de Younden's, coeficiente de correlação de Matthews e área baixo da curva ROC. Todos os ajustes de modelos de classificação assim como a obtenção dos parâmetros de desempenho foram realizadas no *software* Matlab.

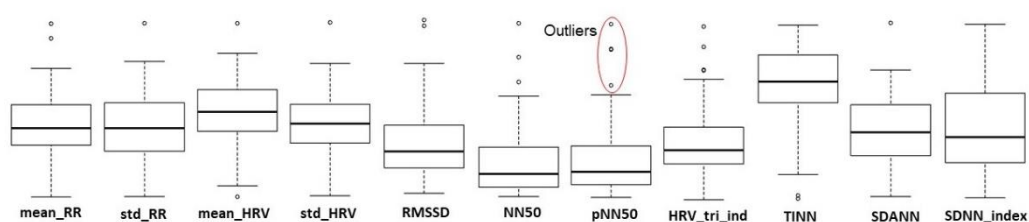
5 RESULTADOS E DISCUSSÃO

Após o *download* dos dados do tipo ECG diretamente do banco de dados *PhysioNet*, obtenção dos correspondentes tacogramas pela aplicação da ferramenta *an2rr*-(esta ferramenta utiliza-se para obter séries de intervalo RR de arquivos de anotação de ECG), filtragem dos dados e conversão dos mesmo no *software* Kubios para obtenção das variáveis que foram modeladas nos classificadores, cada uma das 34 variáveis foi investigada objetivando realizar uma inspeção dos dados.

Essa etapa de filtragem dos dados é importante pois permite a remoção dos possíveis artefatos no sinal e manutenção dos intervalos RR de origem sinusal. O filtro adaptativo consiste em basicamente três etapas: a) remoção dos intervalos RR maiores de 1200 ms e menores 350 ms; b) elaboração de uma série binomial baseado na média e desvio padrão adaptativo e c) substituição dos intervalos que sejam superiores a 20 % dos seus vizinhos por um valor entre a média e o desvio padrão adaptativo. Após a etapa de filtragem o tacograma é então exportado para a obtenção dos parâmetros a serem modelados.

Amostras anômalas são aquelas com características muito diferentes do grupo na qual está inserida. Dentre os métodos de detecção de amostras anômalas, também chamadas de outliers, foi selecionado o boxplot. O boxplot é uma forma de visualmente conhecer como é a dispersão dos dados, uma a uma variável, e a presença de candidatos a outliers. Assim, um total de 102 gráficos boxplot foram obtidos. Eles são apresentados no apêndice, correspondentes aos conjuntos de variáveis pertencentes ao domínio de tempo, frequência e aquelas obtidas por métodos não-lineares. A Figura 13 apresenta o conjunto de gráficos boxplot para as onze variáveis do domínio de tempo para o grupo de amostras com ritmo sinusal normal. Nesta figura cada ponto destacado num boxplot refere-se a uma amostra candidata a *outlier*. Para avaliar corretamente se uma amostra é ou não um candidato a *outlier* ela deve apresentar um valor atípico num grande número de variáveis. Se não for assim, não há argumentos para remover a amostra do conjunto de dados. A figura 14 mostra a matriz de dados que originou a figura 13, nesta matriz cada célula representada em amarelo corresponde a um valor apontado como *outlier* na análise boxplot.

Figura 13- Boxplot: Grupo ritmo sinusal normal (domínio do tempo).



Fonte: O autor.

Figura 14 - Matriz de dados: Grupo ritmo sinusal normal (domínio do tempo).

amostra	mean_RR	std_RR	mean_HRV	std_HRV	RMSSD	NN50	pNN50	HRV_tri_ind	TINN	SDANN	SDNN_index
nsr001	0.75806	0.16509	82.73187	16.80444	0.02754	7508	7.07908	18.38128	0.525	0.15224	0.05505
nsr0010	0.84936	0.17938	74.41368	18.51278	0.02744	6097	6.44933	31.98173	0.61	0.16437	0.06793
nsr0011	0.73751	0.11722	83.35261	12.7325	0.02028	2250	1.94279	16.36484	0.53	0.11182	0.03356
nsr0012	0.86436	0.16602	72.24976	15.1446	0.01959	2467	2.52386	24.01081	0.58	0.15885	0.04379
nsr0013	0.72901	0.15287	86.38431	20.18902	0.0176	1375	1.18794	22.82098	0.6	0.14754	0.03761
nsr0014	0.75992	0.13354	81.39181	14.20659	0.04788	22886	20.69165	19.5279	0.52	0.11916	0.05525
nsr0015	0.79372	0.09921	76.82693	9.97842	0.01549	371	0.36937	16.93206	0.465	0.09393	0.02997
nsr0016	0.78119	0.11141	78.42344	11.4871	0.01641	770	0.81882	18.15074	0.465	0.10325	0.03799
nsr0017	0.84415	0.14235	73.18876	12.76733	0.03697	13328	13.46385	30.18969	0.48	0.12703	0.05826
nsr0018	0.77408	0.1237	79.38674	11.96186	0.02505	4851	4.5317	20.32795	0.48	0.11444	0.04342
nsr0019	0.69999	0.11689	88.12742	14.6717	0.01344	413	0.33628	13.29156	0.5	0.11204	0.03084
nsr002	0.75435	0.12867	81.84124	13.70558	0.01661	1053	0.9463	20.59522	0.5	0.11987	0.04413
nsr0020	0.61313	0.10083	100.22942	14.76121	0.0192	2510	1.83995	11.56379	0.59	0.09212	0.03724
nsr0021	0.82675	0.13415	74.36909	11.25074	0.02622	5247	5.0678	16.44489	0.425	0.11985	0.05512
nsr0022	1.01262	0.10328	59.91022	6.54227	0.0346	8915	12.28114	16.38276	0.325	0.0839	0.05446
nsr0023	0.76727	0.11256	79.91182	11.8298	0.01844	1827	1.63025	19.17037	0.525	0.1033	0.04053
nsr0024	0.76059	0.14679	81.6866	14.87105	0.04901	15176	17.71097	20.25242	0.515	0.1314	0.05921
nsr0025	0.65491	0.08881	93.27736	12.49027	0.02764	7054	5.48254	11.583	0.54	0.07456	0.04306
nsr0026	0.73732	0.14471	84.27521	15.08749	0.01319	146	0.12538	14.17468	0.515	0.13949	0.03505
nsr0027	0.74037	0.15476	84.77371	18.16366	0.0153	671	0.5957	19.86263	0.56	0.14802	0.04212
nsr0028	0.78523	0.1284	78.60197	13.80702	0.02066	2769	2.5647	19.31086	0.545	0.11612	0.04975
nsr0029	0.68733	0.16041	91.67511	19.46368	0.01912	2388	2.05642	19.06188	0.6	0.151	0.04994
nsr003	0.86714	0.10028	70.21871	8.95158	0.02665	5680	5.86771	15.23481	0.4	0.08961	0.04195
nsr0030	0.80178	0.14094	77.27697	14.18383	0.01722	661	0.63098	30.5598	0.485	0.13338	0.04189
nsr0031	0.7345	0.11643	83.66912	12.70812	0.0206	3104	2.81961	27.82082	0.495	0.10657	0.04526
nsr0032	0.81107	0.12456	75.80509	12.07589	0.01586	800	0.7823	22.50506	0.465	0.11503	0.04176
nsr0033	0.97856	0.1211	62.3215	8.23765	0.02576	3186	5.10945	24.57864	0.32	0.0986	0.06532
nsr0034	0.91017	0.09848	66.70994	7.34719	0.01941	1237	1.72347	17.19986	0.36	0.08272	0.04979
nsr0035	0.71773	0.1002	85.11106	11.00845	0.02002	2829	2.40977	11.18183	0.465	0.09306	0.03296
nsr0036	0.72709	0.09732	84.12861	12.17661	0.02495	6110	5.37474	17.79881	0.515	0.08588	0.04175
nsr0037	0.87036	0.14464	70.94636	12.38093	0.01747	758	0.8252	18.26203	0.435	0.13799	0.03839
nsr0038	0.73144	0.13976	85.03327	16.03407	0.02136	2934	2.88748	19.12516	0.56	0.1319	0.04381
nsr0039	0.83648	0.10906	73.02677	10.11813	0.0155	381	0.38728	18.3203	0.51	0.10135	0.03729
nsr004	0.84769	0.1695	73.8376	15.68118	0.02721	6690	6.89933	30.31166	0.525	0.15995	0.0531
nsr0040	0.72046	0.11861	85.65431	14.74464	0.02236	5873	4.99256	15.65141	0.52	0.11362	0.03203
nsr0041	0.75865	0.08526	80.12014	9.27417	0.0202	2188	2.15583	15.6456	0.455	0.07753	0.03228
nsr0042	0.79195	0.12897	77.90813	13.39704	0.02649	6094	5.6484	20.4492	0.51	0.11019	0.06191
nsr0043	0.76809	0.13591	80.87572	16.06554	0.01546	439	0.42734	19.60103	0.51	0.12761	0.04111
nsr0044	0.73513	0.12768	84.00925	14.07601	0.0227	3745	3.39855	17.37269	0.515	0.12163	0.03676
nsr0045	0.82529	0.13008	74.53708	11.82166	0.02929	8014	8.57186	22.55017	0.49	0.12038	0.04645
nsr0046	0.85818	0.16847	73.03015	16.41046	0.02526	5336	5.63928	35.80136	0.545	0.15921	0.04921
nsr0047	0.76348	0.13147	80.82429	13.35123	0.02997	9232	8.20097	18.54885	0.5	0.11076	0.06578
nsr0048	0.72424	0.1413	85.867	15.8185	0.02579	5987	5.16085	18.4434	0.57	0.12225	0.06654
nsr0049	0.81562	0.13667	75.71563	13.22137	0.02827	7103	6.94799	19.58467	0.545	0.12233	0.05737
nsr005	0.7209	0.20582	89.97522	24.38218	0.02297	4073	3.57799	32.05745	0.61	0.1977	0.05045
nsr0050	0.67715	0.11986	91.36656	16.00023	0.02015	3055	2.40976	15.01741	0.575	0.10953	0.04494
nsr0051	0.7967	0.1767	79.11121	17.80976	0.04008	18465	17.67781	23.41493	0.56	0.15926	0.07153
nsr0052	0.68819	0.15038	91.34076	19.63761	0.0194	2906	2.44877	16.32141	0.595	0.13961	0.05045
nsr0053	0.81831	0.15813	76.09256	14.72937	0.03131	9253	9.41848	21.70179	0.54	0.14112	0.0671
nsr0054	0.68947	0.14121	90.5577	17.58983	0.02622	7441	6.20849	18.31215	0.565	0.13013	0.05303
nsr006	0.81034	0.16452	77.73837	18.65549	0.01892	1847	1.82382	24.68844	0.57	0.15716	0.04282
nsr007	0.78986	0.0984	77.24785	10.49271	0.01664	466	0.43684	25.47206	0.495	0.09255	0.03146
nsr008	0.79541	0.14345	77.83932	13.55646	0.02369	4817	4.47822	19.21164	0.51	0.13259	0.0511
nsr009	0.82731	0.15857	75.21821	14.43295	0.03295	8886	8.80438	39.04371	0.48	0.14606	0.05552

Fonte: O autor.

Como pode-se notar, não foi destacado em amarelo toda uma linha horizontal. Ou seja, não há indícios suficientes para remover qualquer amostra deste conjunto de dados. A mesma inspeção foi realizada para todos os conjuntos de dados. A análise de boxplot e apresentação das respectivas matrizes de dados é apresentado na seção de Apêndices.

Assim, não se encontrou amostras a serem removidas do conjunto total de dados selecionados para *download* e filtragem. Complementando a análise univariada, uma sequência de testes de hipóteses foi realizada tais como, teste de Lilliefors, teste-F, teste-*t* e Kruskal-Wallis, ao nível de 95% de confiança, objetivando explorar os parâmetros selecionados nesse estudo. Estes parâmetros correspondem as entradas nos modelos de classificação multivariada. Assim, se for detectada uma significância estatística que discrimine os grupos de amostras estudados, há fortes indícios que os modelos poderão ser utilizados como modelos de previsão de classe, uma vez que os dados que os originam possuem tal característica de grupo.

A Tabela 2 apresenta o resultado após uma completa análise por testes de hipóteses. Inicialmente foi verificada a normalidade dos dados pelo teste de Lilliefors, na sequência foi observada a dispersão dos dados pelo teste-F e então aplicando um teste-*t*, todos ao nível de 95% de confiança. A análise dos *p*-valores, tendo como valor crítico $p = 0,05$, revelou que apenas cinco variáveis não apresentaram diferenças significativas entre as comparações de grupos de Insuficiência cardíaca congestiva *versus* Ritmo sinusal normal e a supressão de arritmia cardíaca *versus* Ritmo sinusal normal. São elas: VLF_power_prc, LF_power_prc, SampEn, RPA.Lmean e RPA.ShanEn. As demais 29 variáveis mostraram fortes indícios de característica de grupo. Esta observação é um forte indício de que os modelos de classificação utilizando estas variáveis podem resultar em bons modelos de classificação.

Tabela 2 - Testes de hipóteses univariados.

Parâmetro	Significância estatística* (<i>p</i> -valor)		
	Insuficiência cardíaca congestiva (ICC) <i>versus</i> Ritmo sinusal normal	Supressão de arritmia cardíaca (SAC) <i>versus</i> Ritmo sinusal normal	
Domínio do tempo	mean_RR	3.05933E-05	1.58218E-06
	std_RR	4.23483E-19	1.77979E-29
	mean_HRV	9.77551E-05	0.000262583
	std_HRV	2.78247E-08	1.04616E-20
	RMSSD	1.50058E-05	4.2076E-09
	NN50	2.65483E-05	8.75617E-07
	pNN50	0.000226191	1.28098E-07
	HRV_tri_ind	2.52077E-12	3.29894E-19
	TINN	1.8739E-06	2.35614E-21
	SDANN	7.39622E-17	5.32703E-28
SDNN_index	1.97927E-11	6.42264E-22	
Domínio da frequência	HF_peak	4.68685E-05	3.39408E-05
	VLF_power	9.68398E-22	1.78879E-22
	VLF_power_prc **	0.921760032	0.88861277
	LF_power	4.52676E-13	1.45649E-12
	LF_power_prc **	0.044528199	0.056431647
	LF_power_nu	2.46631E-10	6.5846E-09
	HF_power	2.31743E-06	4.87546E-10
	HF_power_prc	0.00049678	0.000219937
	HF_power_nu	2.84299E-10	7.65564E-09
	LF_HF_power	1.39096E-09	2.0549E-08
tot_power	1.25001E-21	1.39391E-22	
Parâmetros não lineares	Poincare_SD1	1.50231E-05	4.21613E-09
	Poincare_SD2	4.36612E-19	1.83551E-29
	ApEn	0.001772531	2.87065E-11
	SampEn **	0.076315022	8.61736E-09
	DFA.alpha1	3.78963E-09	5.44384E-15
	DFA.alpha2	0.000517007	0.000816523
	RPA.Lmax	4.08925E-06	8.25063E-12
	RPA.Lmean **	0.640543001	0.520625855
	RPA.DIV	5.1347E-05	9.68932E-11
	RPA.REC	0.004978879	0.000107125
	RPA.DET	0.000399044	6.36861E-09
	RPA.ShanEn **	0.344393143	0.053008612

*Ao nível de 95% de confiança. Após verificação dos pressupostos da normalidade dos dados e distribuição das variâncias.

**Variáveis que não apresentaram diferenças significativas.

Fonte: O autor.

Para a abordagem multivariada, os dados foram agrupados em 4 conjuntos. Um contendo 11 variáveis do domínio de tempo, outro contendo 11 variáveis do domínio de frequência, outro contendo 12 variáveis obtidas por métodos não-lineares e um quarto grupo contendo 34 variáveis, agrupando todos os subconjuntos descritos anteriormente num único *dataset*.

Deste modo, o grupo de amostras com pacientes diagnosticados com ritmo sinusal normal, denominado grupo controle, foi subdividido em 4 subconjuntos. Cada um destes conjuntos contém 54 indivíduos. O grupo de indivíduos com insuficiência cardíaca congestiva também foi dividido em 4 subconjuntos, com 29 indivíduos e o grupo com supressão de arritmia cardíaca, também dividido em 4 subconjuntos, com 54 indivíduos.

A primeira etapa ao iniciar a análise multivariada objetivando construir classificadores do tipo caso *versus* controle é dividir o conjunto de dados em conjunto de treinamento e conjunto de teste. As amostras do conjunto de treinamento, ou calibração, são aquelas com os indivíduos que serão utilizados para ajustar o modelo de classificação, responsáveis pelos pesos de cada coeficiente presente no modelo de previsão. As amostras do conjunto de teste, ou validação, são aquelas que não participam do ajuste, ou seja, não tem qualquer influência sobre os modelos que estão sendo implementados, sendo então utilizadas para testar a eficiência das previsões destes modelos. Em todos os modelos que serão implementados a proporção utilizada entre conjunto de treinamento e teste foi de 70% do total de voluntários para treinamento e os 30% restante utilizados para teste. A tabela 3 ilustra como fica o tamanho de cada um destes conjuntos de dados.

Tabela 3 - Representação da dimensão das matrizes de dados formadas.

	Grupo	Conjunto	Domínio Tempo	Domínio Frequência	Não-Linear	Todas Variáveis
1º Caso (N = 83)	ICC (29)	Treinamento (N = 58)	58x11	58x11	58x12	58x34
	vs RSN (54)	Teste (N = 25)	25x11	25x11	25x12	25x34
2º Caso (N = 108)	SAC (54)	Treinamento (N = 76)	76x11	76x11	76x12	76x34
	vs RSN (54)	Teste (N = 32)	32x11	32x11	32x12	32x34

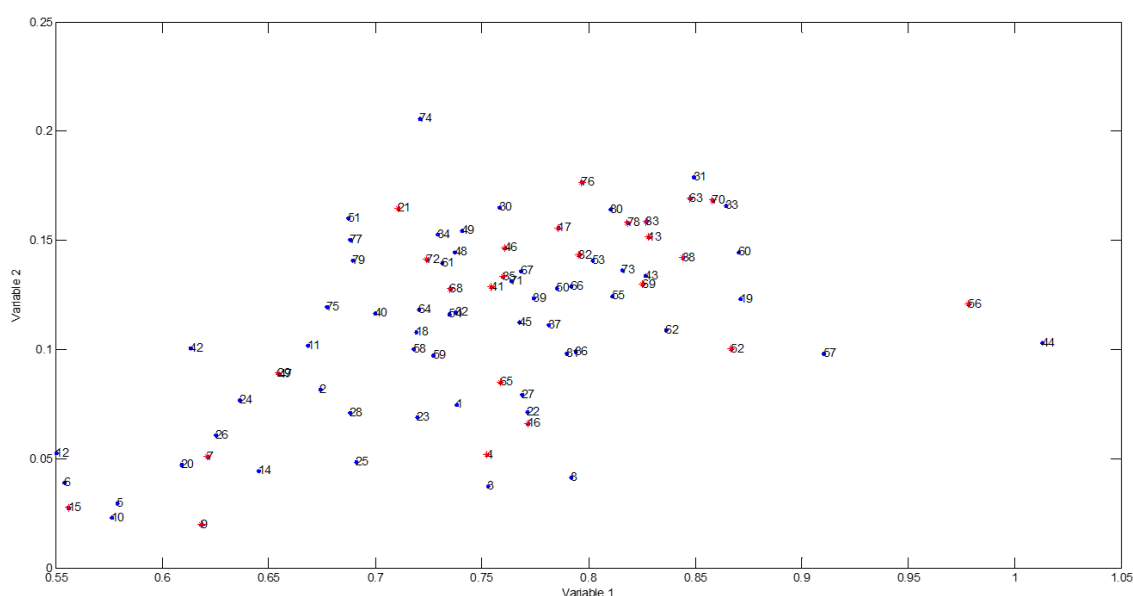
Fonte: O autor.

A divisão dos dados em conjunto de treinamento e conjunto de teste foi realizada pelo algoritmo de Kennard-Stone (para a abordagem PLS-Da) e por uma seleção aleatória interna a rotina de redes neurais (para a abordagem ANN). Neste ponto vale reforçar que embora o tamanho dos conjuntos de dados sejam os mesmos nas abordagens de classificadores PLS-DA e ANN os indivíduos dentro de cada grupo

não necessariamente são os mesmos, devido a característica das rotinas computacionais utilizadas.

A seguir é mostrada a figura 15 que representa a divisão de dados em treinamento e teste através do algoritmo de Kennard-Stone. Esta figura representa a seleção de amostras de treinamento, em azul, e amostras de teste, em vermelho, para o conjunto de variáveis do domínio de tempo da investigação insuficiência cardíaca congestiva versus ritmo sinusal normal.

Figura 15 - Seleção de amostras pelo algoritmo de Kennard-Stone.



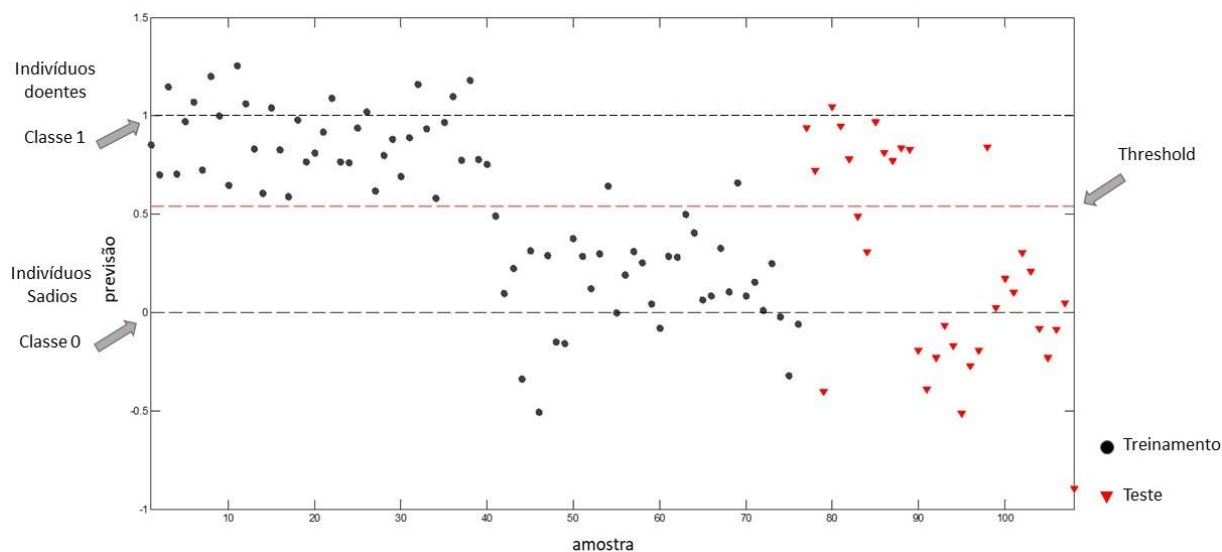
Fonte: O autor.

Neste algoritmo é determinado o centroide do conjunto de dados e então sucessivamente é selecionado uma amostra mais próxima e uma afastada, contemplando assim a seleção de amostras que representa todo o conjunto de dados. Ao final tem-se separadas as amostras em 70% para treinamento e 30% para teste. O mesmo procedimento foi realizado com os outros subconjuntos de variáveis e para o segundo caso em estudo.

Após o ajuste de um modelo de previsão, tanto PLS-DA como ANN, obtém-se um valor associado a cada indivíduo como pertencente a uma classe ou outra. O limiar de separação entre a classe denominada caso (classe igual a 1) e a classe denominada controle (classe igual a 0) é definido por um valor denominado *threshold*. A figura 16 ilustra a predição do modelo PLS-DA, com o conjunto de todas as

variáveis, para o caso de estudo supressão de arritmia cardíaca *versus* ritmo sinusal normal.

Figura 16 - Modelo de classificação de dados: Previsão da classe das amostras.



Fonte: O autor.

Nesta figura, o eixo horizontal representa as amostras presente em todo o conjunto de dados, inicialmente aquelas pertencentes ao conjunto de treinamento, em preto, seguidas das amostras de teste, em vermelho. A linha tracejada em vermelho representa o valor de *threshold*, ou seja, o limiar de separação entre as duas classes para um diagnóstico clínico. Para cada modelo multivariado ajustado foi obtida uma figura similar, contudo, a matriz de confusão foi obtida, para cada um destes modelos, a fim de não incluir erros com contagens de pontos pertencentes a uma classe ou outra.

A Figura 17 representa a matriz de confusão para a mesma determinação representada acima. Nesta matriz tem-se representado 14 como o número de verdadeiros negativos, 3 falsos negativos, 5 falsos positivos e 10 verdadeiros positivos. Essa contagem é realizada apenas sobre os dados de teste. Este conjunto, de teste, é que fornece os valores de VP, VN, FP e FN utilizados para estimar as figuras de mérito de cada modelo.

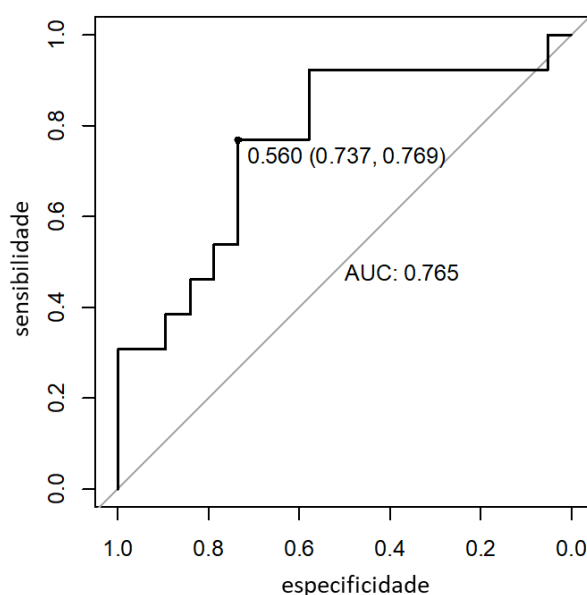
Figura 17 - Matriz de confusão.

```
> confusion.matrix(w, p, threshold = 0.560)
      obs
pred 0 1
0   14 3
1    5 10
```

Fonte: O autor.

Também, ilustrando para um modelo, e extrapolando para os demais, é mostrado na figura 18 a representação da curva ROC, característica de cada modelo de classificação. Esta Figura trata-se da curva ROC característica do modelo PLS-DA supracitado. Para cada estimativa de curva ROC foi obtido o valor de área abaixo da curva ROC, denotado por AUC, e registrado esse valor para ser discutido a seguir.

Figura 18 - Curva ROC.



Fonte: O autor.

Finalmente, foram ajustados 8 modelos de classificação multivariada para o caso em estudo Insuficiência Cardíaca Congestiva versus Ritmo sinusal normal (Controle). Os modelos PLS-DA foram ajustados com 3, 2, 3 e 5 variáveis latentes, respectivamente (Domínio do Tempo, Domínio da Frequência, Não-Linear e Todas Variáveis). Os modelos ANN foram ajustados com uma rede do tipo *feed-forward* com duas camadas treinada com algoritmo Levenberg-Marquardt *backpropagation* e 10 neurônios, onde 30% dos dados foram separados para teste e os outros 70% sendo

55% treinamento e 15% validação. Tanto o número de variáveis latentes (PLS-DA) como os parâmetros ajustáveis ANN foram obtidos na condição *default* do *toolbox* Matlab.

A Tabela 4 reúne o conjunto de estimativas de figuras de mérito para esses classificadores. Esses parâmetros dizem respeito a performance de cada modelo ajustado, podendo ser interpretado como um bom classificador para diagnóstico clínico ou não. Todas as estimativas de performance foram obtidas a partir do conjunto de dados de teste.

Tabela 4 - Performance dos modelos de classificação para Insuficiência cardíaca congestiva.

Parâmetro	Modelo de Classificação – Insuficiência Cardíaca Congestiva (ICC) <i>versus</i> Controle (RSN)							
	Domínio Tempo		Domínio Frequência		Não-Linear		Todas Variáveis	
	PLS-DA	ANN	PLS-DA	ANN	PLS-DA	ANN	PLS-DA	ANN
Sensibilidade	66.7	58.3	44.4	54.5	84.6	62.5	90.9	54.5
Seletividade	93.8	69.2	87.5	92.9	75.0	88.2	85.7	35.7
Taxa de Falso Positivo	6.3	30.8	12.5	7.1	25.0	11.8	14.3	64.3
Taxa de Falso Negativo	33.3	41.7	55.6	45.5	15.4	37.5	9.1	45.5
Eficiência	84.0	64.0	72.0	76.0	80.0	80.0	88.0	44.0
Índice de Younden	0.604	0.276	0.319	0.474	0.596	0.507	0.766	-0.097
Coefficiente de Matthews	0.646	0.277	0.359	0.524	0.600	0.527	0.761	-0.099
AUC curva ROC	0.806	0.609	0.632	0.724	0.846	0.684	0.942	0.422
<i>threshold</i>	0.298	0.322	0.663	0.795	0.276	0.829	0.142	0.143
VN	15	9	14	13	9	15	12	5
FN	3	5	5	5	2	3	1	5
FP	1	4	2	1	3	2	2	9
VP	6	7	4	6	11	5	10	6

Fonte: O autor.

Analisando os resultados da Tabela 4, representa a determinação do modelo de classificação da insuficiência cardíaca congestiva *versus* ritmo sinusal normal, se fez uma análise dos parâmetros de cada modelo através dos classificadores PLS-DA e ANN. Então, pode-se afirmar que em relação as variáveis, Domínio Tempo, Não-Linear e Todas as Variáveis o classificador PLS-DA foi um melhor classificador de sensibilidade, com respeito ao ANN, portanto, o PLS-DA mostrou uma boa classificação dos indivíduos doentes. Quanto ao parâmetro de seletividade, que representa a capacidade do sistema em prever corretamente a ausência de indivíduo sadio, para casos que realmente não existem, o PLS-DA mostrou-se mais seletivo nas variáveis Domínio Tempo e Todas as Variáveis, por outro lado o classificador ANN, foi maior no Domínio Frequência e a variável Não-Linear. O parâmetro Taxa de Falso Positivo, que se refere às determinações incorretas de indivíduos sadios classificados como enfermos, o PLS-DA teve uma classificação maior nas variáveis Domínio Frequência e Não-Linear, e o ANN foi maior nas variáveis

Domínio Tempo e em Todas as Variáveis. Em relação a Taxa de Falso Negativo, o PLS-DA, representa um bom classificador na variável Domínio Frequência e Não-Linear, e o ANN foi melhor nas outras variáveis. No parâmetro Eficiência, que representa um estimador que pondera as determinações verdadeiras, teve uma pequena variação pelos classificadores PLS-DA e ANN, exceto na variável Todas as Variáveis. O Índice de Younden, em termos gerais teve pouca variação entre os classificadores PLS-DA e ANN, exceto em Todas as Variáveis. Por outro lado, o Coeficiente de Matthews, mostra que os dois classificadores PLS-DA e ANN, estão tendendo para +1, que significa uma predição perfeita, e o parâmetro curva ROC, confirma uma ótima classificação realizada pelos classificadores PLS-DA e ANN.

Para o caso em estudo supressão de arritmia cardíaca *versus* ritmo sinusal normal (controle), foram ajustados 8 modelos de classificação binária. Os ajustes PLS-DA foram obtidos com 4, 1, 6 e 6 variáveis latentes, respectivamente (Domínio do Tempo, Domínio da Frequência, Não-Linear e Todas Variáveis). O ajuste de cada rede neural seguiu o mesmo padrão descrito na primeira abordagem. Na tabela 5 são apresentadas as estimativas de desempenho de cada um destes modelos de classificação.

Tabela 5 - Performance dos modelos de classificação para supressão de arritmia cardíaca.

Parâmetro	Modelo de Classificação – Supressão de arritmia cardíaca (SAC) <i>versus</i> controle (RSN)							
	Domínio Tempo		Domínio Frequência		Não-Linear		Todas Variáveis	
	PLS-DA	ANN	PLS-DA	ANN	PLS-DA	ANN	PLS-DA	ANN
Sensibilidade	75.0	66.7	84.2	70.6	81.3	56.3	76.9	64.3
Seletividade	100.0	64.7	69.2	93.3	56.3	87.5	73.7	77.8
Taxa de Falso Positivo	0.0	35.3	30.8	6.7	43.8	12.5	26.3	22.2
Taxa de Falso Negativo	25.0	33.3	15.8	29.4	18.8	43.8	23.1	35.7
Eficiência	90.6	65.6	78.1	81.3	68.8	71.9	75.0	71.9
Índice de Younden	0.750	0.314	0.534	0.639	0.375	0.438	0.506	0.421
Coeficiente de Matthews	0.808	0.313	0.542	0.649	0.387	0.461	0.498	0.425
AUC curva ROC	0.892	0.537	0.721	0.804	0.660	0.703	0.765	0.635
<i>threshold</i>	0.718	0.748	0.346	0.806	0.314	0.713	0.560	0.374
VN	20	11	9	14	9	14	14	14
FN	3	5	3	5	3	7	3	5
FP	0	6	4	1	7	2	5	4
VP	9	10	16	12	13	9	10	9

Fonte: O autor.

Na Tabela 5, representa a performance dos modelos de classificação para supressão de arritmia cardíaca em comparação ritmo sinusal normal, também foram analisados os parâmetros de cada modelo através dos classificadores PLS-DA e ANN. Então, como observa-se nesta tabela, pode-se afirmar que com respeito a todas as variáveis, o classificador PLS-DA obteve-se uma melhor classificação de

sensibilidade, dos indivíduos doentes comparado ao classificador ANN. Quanto ao parâmetro de seletividade, o PLS-DA mostrou-se mais seletivo nas variáveis Domínio Tempo, e o ANN, foi maior em todas as outras variáveis. O parâmetro Taxa de Falso Positivo, o PLS-DA teve uma classificação de 0% na variável Domínio Tempo, e o ANN foi maior em todas as outras variáveis. No que se refere a Taxa de Falso Negativo, o ANN fez uma maior classificação em todas as variáveis. No parâmetro Eficiência, o PLS-DA foi maior na variável Domínio Tempo e Não- Linear nas outras variáveis o ANN fez uma melhor classificação. O Índice de Younden, em termos gerais teve pouca variação entre os classificadores PLS-DA e ANN. Por outro lado, o Coeficiente de Matthews, mostra que os dois classificadores PLS-DA e ANN, estão tendendo para +1, que significa uma predição perfeita. O parâmetro de AUC curva ROC, confirma uma ótima classificação realizada pelos classificadores PLS-DA e ANN.

É importante destacar que nos artigos encontrados na literatura não utilizam o PLS-DA como classificador para análise automatizada de ECGs, por outro lado mostra-se que as redes neurais ANN foram utilizadas para fins de classificação (VAN GILS et al., 1997). Com avanços das tecnologias, foram feitas abordagens com reconhecimento automático do ECG em tempo real, para isto, utilizaram implementações de máquinas de vetores e algoritmos de aprendizagem automática. Estes modelos de classificação obtiveram 88,4% de sensibilidade e especificidade de 85,9% (HO et al., 2014), comprando-se com nosso trabalho foi melhor com relação a especificidade maior de 90%.

6 CONCLUSÕES E PERSPECTIVAS

As figuras de mérito mostraram que existe um padrão no comportamento dos parâmetros dos tacogramas que podem ser utilizados para diagnóstico clínico. Assim tanto a insuficiência cardíaca congestiva como a classificação e previsão das amostras pertencentes ao estudo de supressão de arritmia cardíaca foram obtidas de forma satisfatória, ou seja, com área abaixo da curva ROC próxima de 0,9. O modelo estatístico que apresentou o melhor poder de classificação para o estudo insuficiência cardíaca congestiva *versus* controle foi o ajuste PLS-DA formado a partir do conjunto de todas variáveis, domínio do tempo, frequência e não-lineares simultaneamente, apresentando área abaixo da curva ROC de 0,942, sensibilidade de 90,9% e seletividade de 85,7%. Para o estudo supressão de arritmia cardíaca *versus* controle o melhor ajuste de classificação foi obtido com o modelo PLS-DA com o conjunto de variáveis pertencentes ao domínio do tempo, com área abaixo da curva ROC de 0,892, seletividade de 100,0% e sensibilidade de 75,0%. Em vista disso, entende-se que o PLS-DA mostrou-se mais preciso e apresentou uma robustez suficiente quanto a classificação das doenças de insuficiência cardíaca congestiva (ICC) e supressão de arritmia cardíaca (SAC), quando comparados ao grupo controle denominado de ritmo sinusal normal (RSN). Vale ressaltar que as amostras pertencentes aos grupos de treinamento e teste, dos modelos PLS-DA quando comparadas com ANN, não são necessariamente as mesmas, visto que os algoritmos de seleção de amostras são diferentes.

Como perspectivas futuras, além de reunir os resultados dos achados em formato de artigo para publicação, as discussões levantadas ao longo da elaboração dessa tese abriram possibilidades de colaborações acadêmicas para captar dados de estudos clínicos de enfermidades e/ou condições de interesse biomédico cujo modelo de automatização de classificadores possa ser utilizado para diagnóstico em diversas aplicações, unindo conhecimento médico, matemático e estatístico/computacional.

REFERÊNCIAS

- ABBOUD, S.; BARNEA, O. **Errors due to Sampling Frequency of the Electrocardiogram in Spectral Analysis of Heart rate Signals with low Variability.** Vienna, Austria: IEEE, 1995.
- ABEDIN, Z.; CONNER, R. **ECG Interpretation The Self-Assessment Approach.** 2nd. ed. Massachusetts, USA: El Passo, TX, 2008.
- AGARWAL, B. et al. **Deep Learning Techniques for Biomedical and Health Informatics.** 1. ed. United States: ELSEVIER-Academic Press, 2020.
- ANTONI, H. **Electrical Properties of the Heart.** In: REILLY, J. P. Applied Bioelectricity. New York, NY: Springer New York, p. 148–193, 1998.
- BALLABIO, D.; CONSONNI, V. **Classification tools in Chemistry. Part 1: linear models. PLS-DA.** Analytical Methods, v. 5, n. 16, p. 3790, 2013.
- BALLABIO, D.; GRISONI, F.; TODESCHINI, R. **Multivariate Comparison of Classification Performance Measures.** Chemometrics and Intelligent Laboratory Systems, v. 174, p. 33–44, mar. 2018.
- BARRO, S. et al. **Classifying Multichannel ECG Patterns with an Adaptive Neural Network.** IEEE Engineering in Medicine and Biology Magazine, v. 17, n. 1, p. 45–55, fev. 1998.
- BEHAR, J. et al. **A Single Channel ECG Quality Metric.** Computing in Cardiology. 1 jan. 2012.
- BLINOWSKA, K. J.; ZYGIEREWICZ, J. **Practical Biomedical Signal Analysis Using MATLAB® -Series in Medical Physics and Biomedical Engineering.** 1. ed. New York, NY: CRC Press, 2011.
- BOHR, A.; MEMARZADEH, K. **Artificial Intelligence in Healthcare.** 1. ed. United States: Academic Press/Elsevier, 2020.
- BYLESJÖ, M. et al. **OPLS Discriminant Analysis: Combining the Strengths of PLS-DA and SIMCA Classification.** Journal of Chemometrics, v. 20, n. 8–10, p. 341–351, ago. 2006.
- CARTWRIGHT (EDITOR), H. **Artificial Neural Networks.** 3. ed. New York, NY: Springer, 2020.
- CLIFFORD, G. D.; AZUAJE, F.; MCSHARRY, P. **Advanced Methods And Tools for ECG Data Analysis.** 1. ed. Londres: ARTECH HOUSE, 2006.
- DE JONG, S. SIMPLS: An alternative approach to partial least squares regression. **Chemometrics and Intelligent Laboratory Systems**, v. 18, n. 3, p. 251–263, mar. 1993.
- DOS SANTOS, L. et al. **T-RR Filter: Ferramenta Computacional Gráfica no uso do Filtro Adaptativo para Análise da Variabilidade da Frequência Cardíaca.** In:

Proceeding Series of the Brazilian Society of Applied and Computational Mathematics. Natal-RN, 9 ago. 2016.

ELSAYAD, A.; ELSALAMONY, H. **Diagnosis of Breast Cancer using Decision Tree Models and SVM**. International Journal of Computer Applications, v. 83, p. 19–29, 1 dez. 2013.

ENGIN, M. **ECG Beat Classification Using Neuro-fuzzy Network**. Pattern Recognition Letters, v. 25, n. 15, p. 1715–1722, nov. 2004.

FLUSS, R.; FARAGGI, D.; REISER, B. **Estimation of the Youden Index and its Associated Cutoff Point**. Biometrical Journal, v. 47, n. 4, p. 458–472, ago. 2005.

FOTIADIS, D. et al. In: **Electrocardiogram (ECG): Automated Diagnosis**. 1. ed. Greece: Wiley-Interscience, v. 6, p. 17, 2006.

FRENKEL, D.; NADAL, J. **Ischemic Episode Detection Using an Artificial Neural Network Trained with Isolated ST-T Segments**. Hannover, Germany: IEEE, 1999.

GLADULI, A. et al. **Poincaré Plots and Tachograms Reveal Beat Patterning in Sick Sinus Syndrome with Supraventricular Tachycardia and Varying AV Nodal Block**. Journal of Veterinary Cardiology, v. 13, n. 1, p. 63–70, mar. 2011.

GOLDBERGER, A. L. et al. **PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals**. Circulation, v. 101, n. 23, 13 jun. 2000.

GÜLER, İ.; ÜBEYLI, E. D. **ECG beat Classifier Designed by Combined Neural Network Model**. Pattern Recognition, v. 38, n. 2, p. 199–208, fev. 2005.

GUO, L.; SIM, G.; MATUSZEWSKI, B. **Inter-patient ECG Classification with Convolutional and Recurrent Neural Networks**. Biocybernetics and Biomedical Engineering, v. 39, n. 3, p. 868–879, jul. 2019.

HAAKSMA, J. et al. **Applicability and Performance of Heart Rate Variability Methods**. Computers in Cardiology. Anais Thessaloniki Chalkidiki, Greece: IEEE, 2003.

HALL, J.; HALL, M. **Guyton and Hall Textbook of Medical Physiology**. 14th Edition ed. Philadelphia: Elsevier, 2020.

HAM, F. M.; SOOWHAN HAN. **Classification of Cardiac Arrhythmias Using Fuzzy ARTMAP**. IEEE Transactions on Biomedical Engineering, v. 43, n. 4, p. 425–429, abr. 1996.

HAND, D. J.; TILL, R. J. **A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems**. Machine Learning, v. 45, n. 2, p. 171–186, 2001.

HAYKIN, S. O. **Neural Networks and Learning Machines**. 3rd Edition. Canada: Prentice Hall, 2008.

- HO, T.-W. et al. **A clinical Decision and Support System with Automatically ECG Classification in Telehealthcare**. Taiwan, IEEE, out. 2014.
- HOSKINS, P. R.; LAWFORD, P. V.; DOYLE, B. J. **Cardiovascular Biomechanics**. 1. ed. Switzerland: Springer International Publishing, 2017.
- HUA, X. et al. **A Novel Method for ECG Signal Classification via One-Dimensional Convolutional Neural Network**. 20 jun. 2020.
- HUDSON, D. L.; COHEN, M. E. **Neural Networks and Artificial Intelligence for Biomedical Engineering**. 1. ed. New York: Wiley-IEEE Press, 1999.
- IAZZO (EDS.), P. A. **Handbook of Cardiac Anatomy, Physiology, and Devices**. 3. ed. Springer International Publishing, 2015.
- IEVA, F.; PAGANONI, A. M.; ZANINI, P. **Detection of Structural Changes in Tachogram Series for the Diagnosis of Atrial Fibrillation Events**. *Computational and Mathematical Methods in Medicine*, v. 2013, p. 1–11, 2013.
- KALIVODOVÁ, A. et al. **PLS-DA for Compositional Data with Application to Metabolomics: PLS-DA for Compositional Data**. *Journal of Chemometrics*, v. 29, n. 1, p. 21–28, jan. 2015.
- KAMATH, M. V.; WATANABE, M.; UPTON, A. **Heart Rate Variability (HRV) Signal Analysis: Clinical Applications**. New York, NY: CRC Press, 2013.
- KEENER, J.; SNEYD, J. **Mathematical Physiology: II: Systems Physiology**. 2. ed. USA: Springer-Verlag New York, 2009.
- KENNARD, R. W.; STONE, L. A. **Computer Aided Design of Experiments**. *Technometrics*, v. 11, n. 1, p. 137–148, fev. 1969.
- KHANDOKER, A. H. et al. **Poincaré Plot Methods for Heart Rate Variability Analysis**. 1. ed. [s.l.] Springer US, 2013.
- KSHIRSAGAR, P.; AKOJWAR, S.; DHANORIYA, R. **Classification of ECG-signals using Artificial Neural Networks**. p. 1–5, 2 mar. 2017.
- KUSUMOTO, F. **ECG Interpretation: From Pathophysiology to Clinical Application**. 2. ed. Switzerland: Springer Nature, 2020.
- LI GANG et al. **An artificial-intelligence approach to ECG analysis**. p. 95–100, abr. 2000.
- LI, Q.; RAJAGOPALAN, C.; CLIFFORD, G. D. **A Machine Learning Approach to Multi-level ECG Signal Quality Classification**. *Computer Methods and Programs in Biomedicine*, v. 117, n. 3, p. 435–447, 1 dez. 2014.
- LYON, A. et al. **Computational Techniques for ECG Analysis and Interpretation in Light of their Contribution to Medical Advances**. *Journal of The Royal Society Interface*, v. 15, n. 138, p. 20170821, 31 jan. 2018.

MALTHOUSE, E. C.; TAMHANE, A. C.; MAH, R. S. H. **Nonlinear Partial Least Squares**. *Computers & Chemical Engineering*, v. 21, n. 8, p. 875–890, abr. 1997.

MARTINI, F. H.; TIMMONS, M. J.; TALLITSCH, R. B. **Anatomia Humana**. 6ª edição ed. Porto Alegre: Artmed, 2009.

MCCULLOCH, W. S.; PITTS, W. **A Logical Calculus of the Ideas Immanent in Nervous Activity**. *The Bulletin of Mathematical Biophysics*, v. 5, n. 4, p. 115–133, dez. 1943.

MCNITT-GRAY, M. F. et al. **The Lung Image Database Consortium (LIDC) Data Collection Process for Nodule Detection and Annotation**. *Academic radiology*, v. 14, n. 12, p. 1464–1474, dez. 2007.

MCSHARRY, P. E. et al. **Method for Generating an Artificial RR Tachogram of a Typical Healthy Human Over 24-hours**. *Computers in Cardiology. Anais In: Computers in Cardiology*. Memphis, TN, USA: IEEE, 2002.

PÉREZ-ENCISO, M.; TENENHAUS, M. **Prediction of Clinical Outcome with Microarray Data: a Partial Least Squares Discriminant Analysis (PLS-DA) Approach**. *Human Genetics*, v. 112, n. 5, p. 581–592, 1 maio 2003.

RAJENDRA ACHARYA, U. et al. **Heart Rate Variability: a Review**. *Medical & Biological Engineering & Computing*, v. 44, n. 12, p. 1031–1051, dez. 2006.

RASHID, N. A. et al. **Comparison Analysis: Large Data Classification Using PLS-DA and Decision Trees**. *Mathematics and Statistics*, v. 8, n. 2, p. 100–105, mar. 2020.

REILLY (AUTH.), J. P. **Applied Bioelectricity: From Electrical Stimulation to Electropathology**. 1. ed. Springer-Verlag New York, 1998.

ROSENTHAL, L. **Normal Electrocardiography (ECG) Intervals: Normal Electrocardiography Intervals**. 10 jun. 2020.

ROSER, M. et al. **Coronavirus Pandemic (COVID-19)**. *Our World in Data*. Page, disponível em: <<https://ourworldindata.org/coronavirus-data>>. Acesso em: 12 out. 2020.

SAKSENA, S.; CAMM, A. J. **Electrophysiological Disorders of the Heart**. 2nd. ed. Philadelphia, PA: Elsevier/Saunders, 2012.

SARKALEH, M. **Classification Of Ecg Arrhythmias Using Discrete Wavelet Transform and Neural Networks**. *International Journal of Computer Science, Engineering and Applications*, v. 2, p. 1–13, 29 fev. 2012.

SHARPLESS, N. E.; LOWY, D. R. **National Cancer Institute**. Page, disponível em: <<https://www.cancer.gov/research/resource>>. Acesso em: 12 out. 2020.

SILIPO, R.; MARCHESI, C. **Artificial Neural Networks for Automatic ECG analysis**. *IEEE Transactions on Signal Processing*, v. 46, n. 5, p. 1417–1425, maio 1998.

SÖRNMO, L.; LAGUNA, P. **Bioelectrical Signal Processing in Cardiac and Neurological Applications**. 1. ed. Sweden: Elsevier Academic Press, 2005.

STEIN, P. K. et al. **Heart Rate Variability: A Measure of Cardiac Autonomic Tone**. American Heart Journal, v. 127, n. 5, p. 1376–1381, maio 1994.

SZI-WEN CHEN. **A two-stage Discrimination of Cardiac Arrhythmias Using a Total Least Squares-Based Prony Modeling Algorithm**. IEEE Transactions on Biomedical Engineering, v. 47, n. 10, p. 1317–1327, out. 2000.

TARVAINEN, M. P. et al. **Kubios HRV – Heart Rate Variability Analysis Software**. Computer Methods and Programs in Biomedicine, v. 113, n. 1, p. 210–220, jan. 2014.

THURASINGHAM, R. A. **Preprocessing RR Interval Time Series for Heart Rate Variability Analysis and Estimates of Standard Deviation of RR Intervals**. Computer Methods and Programs in Biomedicine, v. 83, n. 1, p. 78–82, jul. 2006.

TOPOL; J, E. **Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again**. First edition ed. New York: Basic Books, 2019.

TORTORA, G. J. **Principles of Anatomy and Physiology**. 15. ed. United States of America: John Wiley, 2018.

VAN GILS, M. et al. **Using artificial Neural Networks for classifying ICU Patient States**. IEEE Engineering in Medicine and Biology Magazine, v. 16, n. 6, p. 41–47, nov. 1997.

WANG, M.; YAN, G.; FEI, Z. **Kernel PLS Based Prediction Model Construction and Timulation on Theoretical Cases**. Neurocomputing, v. 165, p. 389–394, 1 out. 2015.

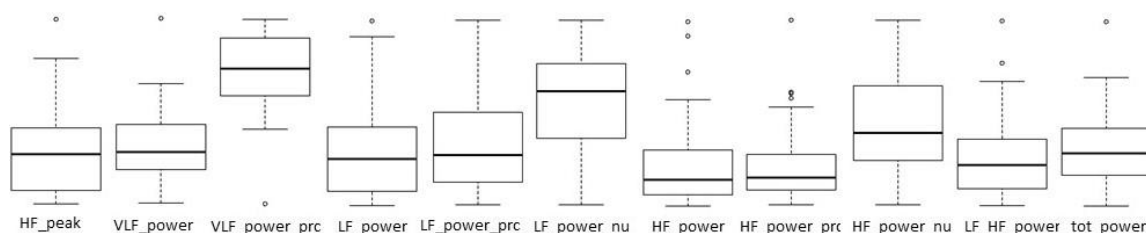
WILMOTH, J.; SHKOLNIKOV, V.; BARBIERI, M. **Human Mortality Database**. Disponível em: <<https://www.mortality.org/>>. Acesso em: 12 out. 2020.

ZHANG, Z.; LUO, X. **Heartbeat Classification Using Decision Level Fusion**. Biomedical Engineering Letters, v. 4, n. 4, p. 388–395, dez. 2014.

ZHAO, Q.; ZHANG, L. **ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines**. China: IEEE, 2005.

APÊNDICE: Complemento da inspeção de dados

Figura 19 - Boxplot: Grupo ritmo sinusal normal (domínio da frequência).



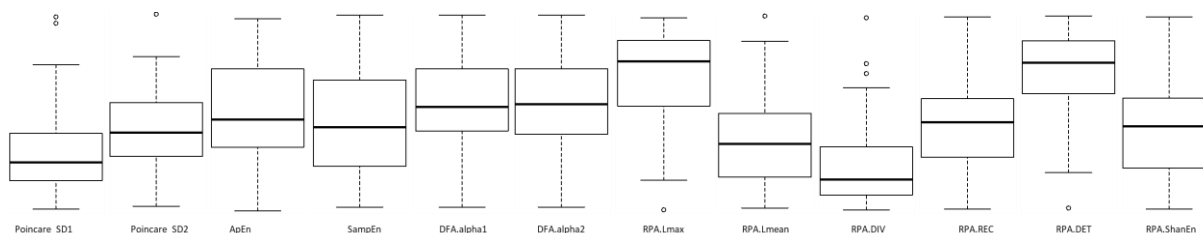
Fonte: O autor.

Figura 20 - Matriz de dados: Grupo ritmo sinusal normal (domínio da frequência)

amostra	HF_peak	VLF_power	VLF_power_prc	LF_power	LF_power_prc	LF_power_nu	HF_power	HF_power_prc	HF_power_nu	LF_HF_power	tot_power
nsr001	0.22656	0.01987	95.72813	0.00057	2.72371	63.75914	0.00032	1.54566	36.18238	1.76216	0.02076
nsr0010	0.15234	0.01902	91.96736	0.00147	7.13123	88.77812	0.00019	0.89913	11.19344	7.93127	0.02068
nsr0011	0.25	0.00964	97.21038	0.0002	2.04526	73.31665	0.00007	0.74033	26.53883	2.76262	0.00991
nsr0012	0.28516	0.01769	97.24081	0.00041	2.26984	82.26478	0.00009	0.48749	17.66786	4.65618	0.01819
nsr0013	0.19922	0.01472	96.45714	0.00045	2.97717	84.0329	0.00009	0.56394	15.9178	5.27918	0.01526
nsr0014	0.16016	0.01195	91.07004	0.00065	4.91612	55.05201	0.00052	3.98898	44.66957	1.23243	0.01312
nsr0015	0.23438	0.00641	96.34466	0.00019	2.83946	77.67992	0.00005	0.81202	22.21465	3.49679	0.00666
nsr0016	0.23438	0.00839	97.10985	0.00018	2.02907	70.20625	0.00007	0.85826	29.69615	2.36415	0.00864
nsr0017	0.24219	0.0133	92.68438	0.00069	4.79867	65.59484	0.00036	2.50456	34.23578	1.91597	0.01435
nsr0018	0.27734	0.01112	93.64802	0.00061	5.17108	81.40884	0.00014	1.17665	18.52416	4.39474	0.01187
nsr0019	0.25781	0.00933	97.66593	0.00018	1.84624	79.09983	0.00005	0.48609	20.82602	3.79812	0.00955
nsr002	0.15234	0.01159	95.84849	0.00046	3.81007	91.77534	0.00004	0.33893	8.16401	11.24145	0.01209
nsr0020	0.15234	0.00775	91.65391	0.00058	6.87142	82.33101	0.00012	1.47116	17.62694	4.67075	0.00846
nsr0021	0.25	0.01339	95.21879	0.00045	3.19817	66.89025	0.00022	1.5778	32.99993	2.02698	0.01406
nsr0022	0.15234	0.00688	89.16376	0.00051	6.63248	61.20648	0.00032	4.18922	38.65938	1.58322	0.00771
nsr0023	0.24219	0.00871	96.35007	0.00022	2.47735	67.87381	0.00011	1.16926	32.03524	2.11872	0.00904
nsr0024	0.16016	0.01524	90.59866	0.00088	5.21624	55.48396	0.0007	4.16267	44.27736	1.2531	0.01683
nsr0025	0.16797	0.00501	80.92684	0.00076	12.31587	64.57176	0.00042	6.74705	35.37459	1.82537	0.00619
nsr0026	0.17578	0.01596	99.07601	0.00011	0.70526	76.32768	0.00004	0.21781	23.57277	3.23796	0.01611
nsr0027	0.17188	0.01583	97.75071	0.00025	1.51409	67.31425	0.00012	0.73409	32.63653	2.06254	0.01619
nsr0028	0.23828	0.01067	92.91738	0.00069	5.9971	84.67349	0.00012	1.08246	15.2833	5.54026	0.01148
nsr0029	0.15234	0.01977	95.6632	0.00078	3.79006	87.39312	0.00011	0.54552	12.57875	6.94768	0.02067
nsr003	0.27344	0.00624	94.20445	0.00021	3.10547	53.58368	0.00018	2.67527	46.16067	1.16081	0.00663
nsr0030	0.17969	0.01334	98.17646	0.00017	1.28457	70.44336	0.00007	0.53621	29.40471	2.39565	0.01359
nsr0031	0.19141	0.00935	89.80966	0.00093	8.93153	87.64705	0.00013	1.25557	12.32122	7.1135	0.01041
nsr0032	0.24219	0.01038	97.55288	0.00018	1.67321	68.37478	0.00008	0.77197	31.54606	2.16746	0.01064
nsr0033	0.15234	0.00966	90.25204	0.00091	8.50467	87.24562	0.00013	1.23874	12.70768	6.86558	0.0107
nsr0034	0.26172	0.00668	91.81672	0.00051	7.01958	85.77955	0.00008	1.15904	14.16348	6.05639	0.00727
nsr0035	0.21094	0.00724	93.19148	0.00037	4.8035	70.55123	0.00016	1.99982	29.37229	2.40197	0.00777
nsr0036	0.22656	0.0059	91.98823	0.00028	4.33249	54.07548	0.00024	3.6682	45.78509	1.1811	0.00641
nsr0037	0.29688	0.01394	98.00137	0.00022	1.57887	78.99742	0.00006	0.41745	20.88658	3.78221	0.01422
nsr0038	0.21094	0.01327	93.30778	0.00079	5.54164	82.80707	0.00016	1.14821	17.15732	4.82634	0.01422
nsr0039	0.26563	0.00788	97.49783	0.00016	1.98857	79.47404	0.00004	0.50972	20.37099	3.90133	0.00808
nsr004	0.21875	0.01862	95.19081	0.00066	3.3787	70.25502	0.00028	1.42815	29.69618	2.36579	0.01956
nsr0040	0.30078	0.00933	97.8766	0.00013	1.38825	65.37875	0.00007	0.72845	34.30563	1.90577	0.00953
nsr0041	0.19922	0.00477	93.15601	0.00024	4.60095	67.22618	0.00011	2.23459	32.65041	2.05897	0.00512
nsr0042	0.23438	0.01079	91.02217	0.00084	7.12386	79.34953	0.00022	1.84922	20.59766	3.85236	0.01186
nsr0043	0.23047	0.01177	97.71047	0.00023	1.92465	84.06347	0.00004	0.36281	15.84647	5.30487	0.01204
nsr0044	0.20313	0.01165	96.28858	0.00028	2.31836	62.46568	0.00017	1.38753	37.38551	1.67085	0.0121
nsr0045	0.21094	0.01131	93.95447	0.00043	3.56474	58.96498	0.0003	2.47394	40.92187	1.44092	0.01204
nsr0046	0.32422	0.01759	97.19188	0.00036	1.98834	70.80671	0.00015	0.81599	29.05813	2.43673	0.0181
nsr0047	0.15234	0.01247	89.34404	0.00116	8.31861	78.06536	0.00033	2.33312	21.89494	3.56545	0.01396
nsr0048	0.15234	0.01408	88.26607	0.00168	10.54979	89.90836	0.00019	1.18188	10.07229	8.92631	0.01595
nsr0049	0.15234	0.01206	88.78554	0.00131	9.65073	86.05611	0.00021	1.55981	13.90892	6.18712	0.01358
nsr005	0.15234	0.03044	97.48065	0.00066	2.11311	83.87493	0.00013	0.40484	16.06902	5.21967	0.03122
nsr0050	0.23438	0.00997	93.46311	0.00057	5.37714	82.25843	0.00012	1.15652	17.69219	4.64942	0.01067
nsr0051	0.23828	0.02138	91.50283	0.00134	5.72088	67.32682	0.00065	2.77117	32.61283	2.06443	0.02336
nsr0052	0.1875	0.0161	95.72854	0.00056	3.34426	78.29307	0.00016	0.92565	21.67051	3.61289	0.01682
nsr0053	0.18359	0.01696	91.25888	0.00122	6.57628	75.23381	0.0004	2.16194	24.73303	3.04184	0.01858
nsr0054	0.17188	0.01384	91.235	0.00099	6.51322	74.30939	0.00034	2.24886	25.65728	2.89623	0.01516
nsr006	0.21094	0.01568	96.95766	0.0004	2.45078	80.55549	0.0001	0.58948	19.37581	4.15753	0.01617
nsr007	0.37109	0.00619	97.35838	0.00013	1.99136	75.38383	0.00004	0.64463	24.40296	3.08913	0.00636
nsr008	0.21875	0.01439	94.21538	0.0007	4.58844	79.32145	0.00018	1.19299	20.62354	3.84616	0.01527
nsr009	0.15234	0.01667	89.88983	0.00155	8.36396	82.72817	0.00032	1.74225	17.2326	4.80068	0.01854

Fonte: O autor.

Figura 21 - Boxplot: Grupo ritmo sinusal normal (parâmetros não lineares).



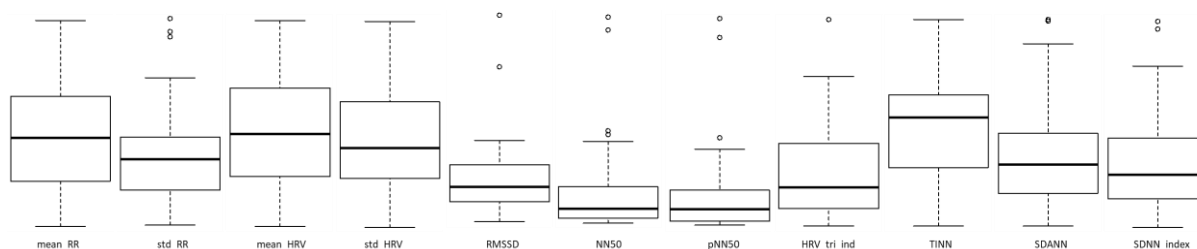
Fonte: O autor.

Figura 22 - Matriz de dados: Grupo ritmo sinusal normal (parâmetros não lineares).

amostra	Poincare_SD1	Poincare_SD2	ApEn	SampEn	DFA.alpha1	DFA.alpha2	RPA.Lmax	RPA.Lmean	RPA.DIV	RPA.REC	RPA.DET	RPA.ShanEn
nsr001	0.01947	0.23266	1.09008	1.11924	1.26873	1.11406	856.87736	33.9718	0.00169	0.48247	0.9941	4.08739
nsr0010	0.01941	0.25294	1.04643	1.02486	1.51697	0.98758	887.24468	23.90868	0.00122	0.46429	0.99585	3.8751
nsr0011	0.01434	0.16516	1.24977	1.44549	1.11871	1.09859	709.38261	27.84359	0.00199	0.46238	0.9916	3.94667
nsr0012	0.01385	0.23438	0.9946	1.01238	1.36801	1.15736	891.1134	38.65657	0.00129	0.50265	0.99573	4.21363
nsr0013	0.01245	0.21583	1.16416	1.26834	1.31353	1.03499	896.94783	30.53937	0.00126	0.45555	0.99446	3.9295
nsr0014	0.03385	0.1858	1.2445	1.26987	0.92507	1.13402	425.27273	22.96037	0.00401	0.40765	0.98524	3.75798
nsr0015	0.01096	0.13987	1.26214	1.41775	1.19848	1.08202	809.73	29.00359	0.00169	0.47294	0.993	3.97417
nsr0016	0.01161	0.15713	1.07818	1.14301	1.22528	1.20537	809.7234	29.7119	0.0018	0.47328	0.99349	4.0668
nsr0017	0.02614	0.19961	1.18864	1.20968	1.07558	1.06605	565.69388	28.40773	0.00255	0.46367	0.99212	3.96995
nsr0018	0.01772	0.17403	1.20441	1.30431	1.34913	1.02368	743.09346	21.23488	0.00177	0.44431	0.99242	3.81223
nsr0019	0.0095	0.16503	1.18162	1.27988	1.30096	1.1062	813.91803	28.30603	0.00154	0.48622	0.99451	3.9868
nsr002	0.01175	0.18158	1.01338	1.00784	1.54153	1.08903	911.01802	23.55468	0.00117	0.47531	0.99652	3.95937
nsr0020	0.01358	0.14194	1.25569	1.40053	1.37054	0.95401	724.79412	20.13765	0.00161	0.44932	0.99366	3.75734
nsr0021	0.01854	0.18881	1.18524	1.23461	1.20132	1.14469	841.61165	25.83138	0.00156	0.47007	0.99346	3.95951
nsr0022	0.02446	0.144	1.27229	1.43197	1.05485	1.10686	618.04167	31.39994	0.00295	0.46496	0.98735	3.90564
nsr0023	0.01304	0.15865	1.09884	1.19751	1.20546	1.1903	839.08929	34.09406	0.00151	0.48483	0.9945	4.11514
nsr0024	0.03466	0.20469	1.21625	1.20568	1.02842	1.0483	505.95294	24.52411	0.00331	0.42049	0.9889	3.79408
nsr0025	0.01954	0.12406	1.24092	1.33682	1.15538	0.98238	626.07031	21.32108	0.00231	0.41224	0.98988	3.64329
nsr0026	0.00933	0.20444	1.06298	1.10366	1.23499	1.26565	949.96552	43.7141	0.00114	0.5418	0.99664	4.0107
nsr0027	0.01082	0.21859	0.93458	0.96365	1.33769	1.1831	943.79464	40.31447	0.00109	0.51534	0.99665	4.25869
nsr0028	0.01461	0.18099	1.11339	1.14779	1.49528	1.017	889.79439	21.21617	0.00124	0.44856	0.99497	3.79832
nsr0029	0.01352	0.22645	1.00893	1.00194	1.51751	0.99565	916.71552	24.86208	0.00115	0.46966	0.99653	3.93292
nsr003	0.01885	0.14056	1.25515	1.39599	0.9999	1.16825	639.33333	27.65796	0.00247	0.46331	0.9901	3.93619
nsr0030	0.01218	0.19895	1.01692	1.02162	1.22813	1.24393	926.91346	48.14693	0.00123	0.54429	0.99675	4.40059
nsr0031	0.01457	0.16401	1.05494	1.01902	1.56546	0.88642	800.34545	17.3908	0.00141	0.43169	0.99521	3.64041
nsr0032	0.01121	0.1758	1.0209	1.0399	1.2442	1.1956	844.54902	41.3041	0.00151	0.50963	0.99514	4.2529
nsr0033	0.01821	0.17029	0.98528	0.93093	1.47551	1.12897	936.16129	27.2867	0.00114	0.48894	0.99669	4.03774
nsr0034	0.01372	0.13859	1.03857	1.03682	1.51177	1.11248	905.43662	24.38644	0.00115	0.48864	0.99683	3.99199
nsr0035	0.01416	0.14099	1.28631	1.43508	1.24601	0.99188	651.8547	19.28427	0.00221	0.42842	0.99082	3.70129
nsr0036	0.01764	0.13649	1.19466	1.28502	1.07692	1.06063	606.23894	27.71276	0.00316	0.44242	0.98848	3.82165
nsr0037	0.01235	0.20417	1.09996	1.14582	1.27375	1.15342	872.0989	36.99992	0.00151	0.5046	0.99433	4.15278
nsr0038	0.0151	0.19707	1.16191	1.21367	1.45948	0.8839	765.37624	21.07675	0.00164	0.44195	0.99373	3.68729
nsr0039	0.01096	0.15385	1.07402	1.1446	1.26597	1.18338	870.58163	39.91074	0.00139	0.49884	0.99553	4.23118
nsr004	0.01924	0.23894	1.14197	1.16432	1.27481	1.0495	843.53125	27.8216	0.00147	0.47993	0.99412	3.95876
nsr0040	0.01581	0.167	1.17672	1.24159	0.93138	1.17733	689.47009	32.33685	0.00222	0.46905	0.99119	4.05022
nsr0041	0.01428	0.11973	1.26168	1.39263	1.10799	1.07912	617.14851	29.47664	0.00294	0.47839	0.9891	3.90217
nsr0042	0.01873	0.18142	0.98137	0.95645	1.38987	1.11728	875.78505	32.48923	0.00139	0.46921	0.99476	4.06788
nsr0043	0.01093	0.1919	1.00645	0.98511	1.31398	1.17857	852.17647	41.78949	0.00161	0.51468	0.99516	4.23639
nsr0044	0.01605	0.17985	1.26708	1.39574	1.13041	1.05334	635.95455	25.40611	0.00222	0.44784	0.99186	3.83339
nsr0045	0.02071	0.18279	1.29913	1.41604	1.09369	1.01787	639.04301	20.20017	0.00225	0.42868	0.98922	3.70502
nsr0046	0.01786	0.23759	1.06721	1.13584	1.22612	1.12764	741.03191	42.78558	0.00204	0.48715	0.99278	4.17586
nsr0047	0.02119	0.18471	1.13467	1.13457	1.4406	1.02348	896.80357	20.72137	0.00138	0.4428	0.99397	3.76235
nsr0048	0.01823	0.199	1.0337	0.97955	1.59639	0.95244	910.58621	19.09624	0.00116	0.45079	0.9959	3.72994
nsr0049	0.01999	0.19224	1.1263	1.12221	1.47262	0.86585	740.95098	20.23435	0.00184	0.4273	0.99265	3.67315
nsr005	0.01624	0.29062	1.05564	1.06591	1.19057	1.16367	845.78761	44.07549	0.00131	0.49894	0.99532	4.2865
nsr0050	0.01425	0.1689	1.07749	1.08819	1.44648	0.98667	827.48413	24.93165	0.00145	0.46352	0.9948	3.89195
nsr0051	0.02834	0.24829	1.22292	1.22996	1.24642	1.04686	832.65385	20.91102	0.00147	0.42843	0.99146	3.69332
nsr0052	0.01372	0.21222	0.95769	0.933	1.45199	1.11561	887.66102	30.41444	0.00136	0.49705	0.9965	4.06834
nsr0053	0.02214	0.22253	1.13408	1.12513	1.36476	0.97576	861.77551	22.37775	0.00131	0.45104	0.99388	3.78136
nsr0054	0.01854	0.19884	1.09823	1.08979	1.39588	0.93364	753.69748	17.43771	0.00173	0.42037	0.99242	3.61386
nsr006	0.01338	0.23229	1.07671	1.11501	1.33806	1.07299	816.9901	32.86537	0.00165	0.48069	0.99452	4.04121
nsr007	0.01176	0.13866	1.18044	1.27187	1.13654	1.17926	777.92453	29.43442	0.00193	0.48288	0.9929	4.08675
nsr008	0.01675	0.20217	1.06483	1.0569	1.40964	1.06029	894.16822	25.84268	0.00123	0.47039	0.99593	3.98172
nsr009	0.0233	0.22304	1.07862	1.04063	1.37696	0.95627	678.8	25.25556	0.00222	0.43815	0.99236	3.75349

Fonte: O autor.

Figura 23 - Boxplot: Grupo insuficiência cardíaca congestiva (domínio do tempo).



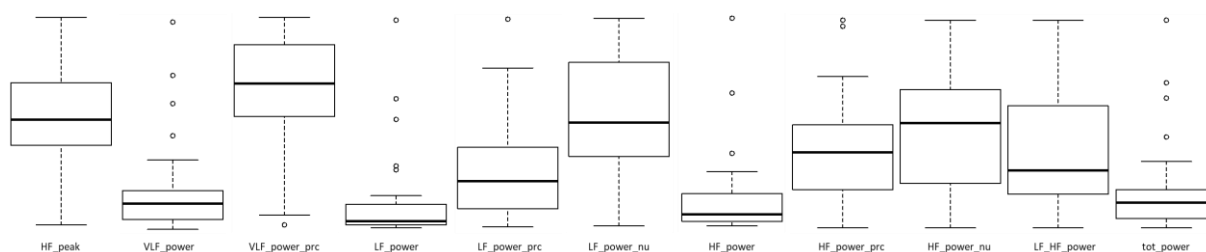
Fonte: O autor.

Figura 24 - Matriz de dados: Grupo insuficiência cardíaca congestiva (domínio do tempo).

amostra	mean_RR	std_RR	mean_HRV	std_HRV	RMSSD	NN50	pNN50	HRV_tri_ind	TINN	SDANN	SDNN_index
chf201	0.73803	0.07481	82.11969	8.21005	0.01903	1964	1.75101	11.50882	0.435	0.05343	0.04784
chf202	0.6742	0.08186	90.35135	11.38874	0.02153	3126	2.86597	11.66567	0.52	0.07412	0.03132
chf203	0.75278	0.03765	79.91369	4.18907	0.01033	161	0.16271	5.71932	0.2	0.03124	0.01976
chf204	0.75228	0.05196	80.15299	5.76614	0.02011	2654	2.69526	8.00634	0.325	0.04028	0.03117
chf205	0.57876	0.02973	103.94423	5.36771	0.01193	603	0.44989	6.27701	0.26	0.0264	0.01324
chf206	0.55407	0.03921	108.91921	8.96782	0.00815	166	0.12263	4.17352	0.22	0.03737	0.00932
chf207	0.62112	0.05106	97.23087	7.74824	0.02082	4321	3.96957	8.38952	0.38	0.04532	0.02153
chf208	0.7917	0.04158	75.99173	3.91591	0.01516	1169	1.08289	7.29708	0.295	0.03737	0.01663
chf209	0.61813	0.01994	97.16893	3.16949	0.00699	10	0.00915	5.75042	0.12	0.01796	0.00836
chf210	0.57621	0.02336	104.30425	4.33576	0.01046	146	0.10324	5.41129	0.22	0.01803	0.01415
chf211	0.66821	0.1019	91.90172	14.13232	0.02719	8009	6.30486	18.09544	0.565	0.08282	0.05529
chf212	0.55007	0.0528	110.16924	11.47909	0.02113	2583	2.54105	6.34096	0.335	0.04915	0.01888
chf213	0.82855	0.15182	75.18945	15.5679	0.05819	17827	17.10319	29.70447	0.615	0.13318	0.06824
chf214	0.645	0.04458	93.46933	6.48141	0.01508	1206	1.28691	7.76936	0.245	0.04225	0.01392
chf215	0.55589	0.02771	108.2093	5.52468	0.011	471	0.32131	4.3838	0.215	0.02424	0.01244
chf216	0.77181	0.06636	78.3057	6.63936	0.01708	1639	1.61511	9.66476	0.395	0.05659	0.03175
chf217	0.78613	0.15562	79.34535	15.68855	0.02714	7670	7.23708	18.29185	0.56	0.14827	0.04452
chf218	0.71932	0.10807	85.39483	13.44368	0.01555	906	0.87132	16.8363	0.565	0.10164	0.03428
chf219	0.8714	0.1232	70.26683	10.10062	0.02341	3407	3.77097	22.69505	0.415	0.10983	0.05184
chf220	0.60924	0.0474	99.0807	7.73131	0.01282	964	0.69994	6.2561	0.355	0.0384	0.02527
chf221	0.71092	0.16484	88.16786	16.7711	0.04536	16701	15.56448	14.56536	0.55	0.14755	0.06605
chf222	0.77112	0.07158	78.56653	8.23922	0.01494	1148	1.13602	8.62105	0.375	0.06255	0.02991
chf223	0.71953	0.06925	84.23164	8.84009	0.01774	2064	1.83447	9.79055	0.395	0.0597	0.03074
chf224	0.63622	0.07697	95.72181	11.84481	0.01124	58	0.04306	12.4235	0.4	0.07273	0.02356
chf225	0.69107	0.0488	87.27083	6.38636	0.01208	260	0.2876	8.30614	0.315	0.04157	0.02303
chf226	0.62482	0.06115	96.9074	9.03391	0.00783	280	0.21358	8.90639	0.26	0.05938	0.01264
chf227	0.7688	0.07955	78.9075	8.4049	0.01622	1760	1.65201	14.3776	0.43	0.07637	0.02083
chf228	0.68785	0.07134	88.21919	9.64035	0.01524	764	0.64249	12.48048	0.41	0.06637	0.02383
chf229	0.6545	0.08929	93.46755	13.33375	0.0256	7103	5.8249	16.19214	0.475	0.08154	0.03492

Fonte: O autor.

Figura 25 - Boxplot: Grupo insuficiência cardíaca congestiva (domínio da frequência).



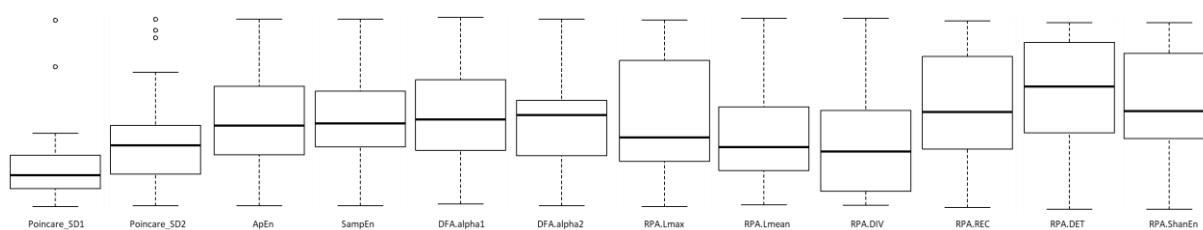
Fonte: O autor.

Figura 26 - Matriz de dados: Grupo insuficiência cardíaca congestiva (domínio da frequência).

amostra	HF_peak	VLF_power	VLF_power_prc	LF_power	LF_power_prc	LF_power_nu	HF_power	HF_power_prc	HF_power_nu	LF_HF_power	tot_power
chf201	0.23828	0.00417	88.22215	0.00042	8.85239	75.16135	0.00014	2.92062	24.79756	3.031	0.00473
chf202	0.25781	0.00446	93.21497	0.00018	3.68563	54.32003	0.00015	3.08887	45.52472	1.1932	0.00478
chf203	0.32031	0.00099	96.6619	0.00001	1.19392	35.76643	0.00002	2.13046	63.82265	0.5604	0.00102
chf204	0.16406	0.00179	91.44119	0.00009	4.69989	54.91284	0.00008	3.83342	44.78913	1.22603	0.00196
chf205	0.30859	0.00057	92.22501	0.00003	4.73961	60.95967	0.00002	3.01736	38.80849	1.57078	0.00061
chf206	0.39063	0.00091	98.86573	0.00001	0.60242	53.11068	0	0.52665	46.43101	1.14386	0.00092
chf207	0.15234	0.00163	85.25636	0.00017	8.63588	58.57359	0.00012	6.07403	41.1976	1.42177	0.00191
chf208	0.35938	0.00117	94.73435	0.00002	1.84061	34.95501	0.00004	3.396	64.49341	0.54199	0.00124
chf209	0.35547	0.00026	96.03219	0.00001	1.91771	48.33172	0.00001	2.02851	51.12416	0.94538	0.00027
chf210	0.33594	0.00035	88.20512	0.00003	7.24354	61.41255	0.00002	4.51159	38.25043	1.60554	0.0004
chf211	0.27734	0.00686	84.59635	0.00093	11.43615	74.24313	0.00032	3.95773	25.69348	2.88957	0.0081
chf212	0.26563	0.00164	94.31872	0.00004	2.22929	39.23928	0.00006	3.4293	60.36137	0.65007	0.00174
chf213	0.30078	0.01379	89.02695	0.00078	5.02969	45.83676	0.00092	5.91528	53.90736	0.85029	0.01549
chf214	0.38281	0.0013	97.50484	0.00001	0.88635	35.5229	0.00002	1.59514	63.9293	0.55566	0.00133
chf215	0.24609	0.00048	94.58805	0.00001	2.58023	47.67659	0.00001	2.81354	51.98753	0.91708	0.00051
chf216	0.25	0.00315	94.299	0.0001	3.12712	54.85213	0.00009	2.56429	44.97955	1.21949	0.00334
chf217	0.28125	0.01684	97.2366	0.00023	1.34589	48.70396	0.00024	1.41317	51.13864	0.95239	0.01732
chf218	0.28906	0.00776	97.72562	0.00013	1.60068	70.37898	0.00005	0.67004	29.46035	2.38894	0.00794
chf219	0.25	0.01038	94.53116	0.00045	4.11342	75.21558	0.00015	1.35223	24.72609	3.04195	0.01098
chf220	0.16406	0.00163	93.99602	0.00007	4.00321	66.67589	0.00003	1.98856	33.12064	2.01312	0.00173
chf221	0.16016	0.02261	91.56094	0.00149	6.02342	71.3755	0.00059	2.40718	28.52432	2.50227	0.0247
chf222	0.22656	0.00321	97.0559	0.00005	1.57705	53.56656	0.00004	1.3595	46.17721	1.16002	0.00331
chf223	0.26172	0.00303	94.50473	0.00009	2.93228	53.36007	0.00008	2.55206	46.44103	1.14899	0.00321
chf224	0.33203	0.00402	98.01173	0.00006	1.46877	73.87169	0.00002	0.51577	25.9406	2.84772	0.0041
chf225	0.28516	0.00162	94.96226	0.00006	3.65677	72.58749	0.00002	1.37259	27.24617	2.66414	0.00171
chf226	0.26563	0.00264	99.17446	0.00001	0.51632	62.54333	0.00001	0.30649	37.12604	1.68462	0.00266
chf227	0.18359	0.00411	97.30925	0.00005	1.25299	46.56636	0.00006	1.42801	53.07087	0.87744	0.00422
chf228	0.39844	0.00334	97.91174	0.00004	1.14949	55.04523	0.00003	0.91588	43.85844	1.25507	0.00341
chf229	0.29297	0.00508	92.38012	0.00024	4.43421	58.19265	0.00017	3.17311	41.64254	1.39743	0.0055

Fonte: O autor.

Figura 27 - Boxplot: Grupo insuficiência cardíaca congestiva (parâmetros não lineares).



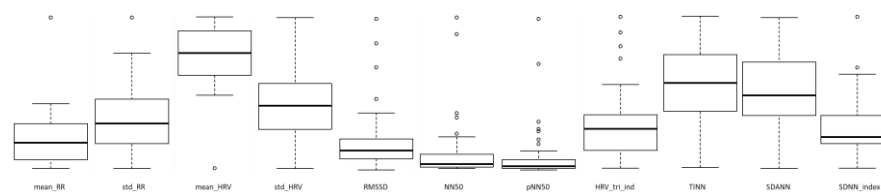
Fonte: O autor.

Figura 28 - Matriz de dados: Grupo insuficiência cardíaca congestiva (parâmetros não lineares).

amostra	Poincare_SD1	Poincare_SD2	ApEn	SampEn	DFA.alpha1	DFA.alpha2	RPA.Lmax	RPA.Lmean	RPA.DIV	RPA.REC	RPA.DET	RPA.ShanEn
chf201	0.01345	0.10494	1.046	1.03801	1.35017	1.2268	904.57143	22.92389	0.00123	0.46087	0.99591	3.9375
chf202	0.01522	0.11476	1.33195	1.48441	0.89086	1.21734	480.44037	23.04631	0.00272	0.44208	0.98914	3.7784
chf203	0.0073	0.05274	1.20153	1.20324	0.76751	1.56108	751.68367	36.68473	0.00159	0.48117	0.99482	4.26983
chf204	0.01422	0.0721	1.25984	1.4174	0.95291	1.35342	443.39796	23.78375	0.00315	0.43635	0.99111	3.87528
chf205	0.00844	0.04119	1.33187	1.33847	0.81218	1.09694	261.68657	16.81287	0.00479	0.40366	0.98393	3.57078
chf206	0.00576	0.05515	1.0623	0.92865	0.53974	1.22319	466.55556	44.59526	0.00346	0.44234	0.98542	3.98024
chf207	0.01472	0.07069	1.30346	1.37052	0.85086	0.90625	274.62037	19.44389	0.00613	0.37901	0.98643	3.53645
chf208	0.01072	0.05782	1.34677	1.40693	0.64785	1.27631	420.04673	28.61558	0.00326	0.43534	0.98687	3.87405
chf209	0.00495	0.02777	1.06689	0.96325	0.62266	1.26299	413.65138	25.38538	0.0039	0.41735	0.98554	3.84939
chf210	0.00739	0.0322	1.33266	1.35783	0.8208	1.2398	395.12766	18.80887	0.00402	0.40543	0.98508	3.60855
chf211	0.01923	0.14282	1.05595	1.02913	1.36994	1.01314	770.95276	21.84335	0.00149	0.45123	0.99466	3.82675
chf212	0.01494	0.07317	1.45538	1.64857	0.68961	0.91466	233.87129	17.56765	0.0065	0.37885	0.97987	3.43341
chf213	0.04114	0.21072	1.13086	1.09292	0.91928	1.0652	325.26923	26.62181	0.00485	0.43817	0.9878	3.74992
chf214	0.01067	0.06214	1.28648	1.27178	0.5422	1.00518	317.02151	24.63057	0.0044	0.38535	0.98341	3.70559
chf215	0.00778	0.0384	1.25122	1.22368	0.70793	1.17975	395.15068	23.6694	0.00385	0.42494	0.98605	3.75932
chf216	0.01208	0.09306	1.19415	1.26545	1.18211	1.23997	777.9604	24.46602	0.00148	0.4841	0.99514	3.97609
chf217	0.01919	0.21924	1.09251	1.13807	1.09859	1.20007	806.12381	36.85562	0.00158	0.47808	0.99297	4.16072
chf218	0.01099	0.15243	1.12849	1.20503	1.15697	1.24976	809.37864	31.3586	0.00142	0.50283	0.9956	4.18094
chf219	0.01656	0.17345	1.1604	1.24515	1.37834	1.13228	873.04444	21.89204	0.00148	0.4591	0.9942	3.86011
chf220	0.00907	0.06642	1.11504	1.08359	1.10204	1.38786	747.75912	30.40002	0.00179	0.48913	0.99551	4.13248
chf221	0.03207	0.2309	1.17712	1.13412	1.12216	0.97102	356.78505	17.87746	0.0039	0.41772	0.99042	3.63515
chf222	0.01057	0.10067	1.11648	1.15046	0.94399	1.35579	676.11881	41.75354	0.00204	0.49458	0.99345	4.24297
chf223	0.01255	0.09713	1.19023	1.24893	0.92397	1.19884	583.09821	32.45534	0.00271	0.46347	0.99192	4.0373
chf224	0.00795	0.10856	1.1839	1.23019	1.0732	1.29663	891.33582	31.37759	0.00121	0.48463	0.99498	4.18008
chf225	0.00854	0.06848	1.18954	1.21965	1.08	1.24464	758.95556	29.68618	0.00171	0.48901	0.99417	4.11697
chf226	0.00554	0.0863	0.98912	0.89887	0.83089	1.31107	691.41985	36.63674	0.0023	0.47911	0.99263	4.21572
chf227	0.01147	0.11192	1.28561	1.36631	0.83909	1.04564	401.57547	27.7575	0.00435	0.43418	0.98889	3.83599
chf228	0.01078	0.10031	1.23413	1.34055	0.77969	1.31247	681.62712	35.86033	0.00202	0.45335	0.99031	4.14408
chf229	0.0181	0.12497	1.31272	1.52304	0.96066	1.08183	481.67769	21.93473	0.00314	0.41488	0.98704	3.72141

Fonte: O autor.

Figura 29 - Boxplot: Grupo supressão de arritmia cardíaca (domínio do tempo).



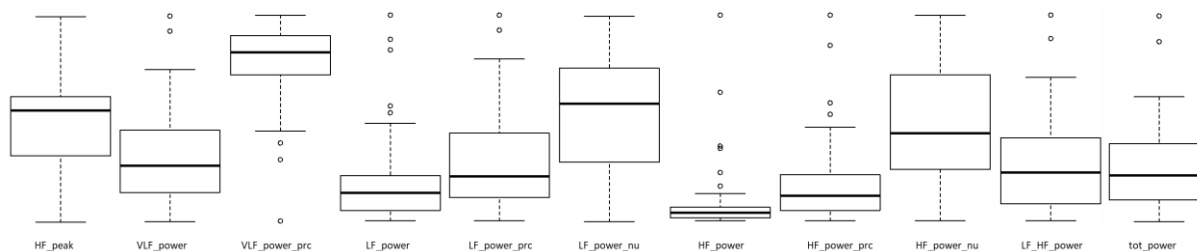
Fonte: O autor.

Figura 30 - Matriz de dados: Grupo supressão de arritmia cardíaca (domínio do tempo).

amostra	mean_RR	std_RR	mean_HRV	std_HRV	RMSSD	NN50	pNN50	HRV_tri_ind	TINN	SDANN	SDNN_index
m006a	0.73081	0.07167	82.91342	8.33963	0.01616	942	0.89123	11.83098	0.48	0.0679	0.02158
m012a	0.71814	0.06167	84.19281	7.54394	0.01705	1364	1.17213	11.40882	0.355	0.05685	0.02305
m014a	0.70793	0.0707	85.60648	8.58092	0.01337	225	0.21993	10.80752	0.365	0.06563	0.02465
m022a	0.78174	0.06861	77.33703	6.68927	0.02024	4273	3.97392	10.44966	0.34	0.06472	0.02109
m028a	0.73526	0.04673	81.95269	5.52625	0.01453	711	0.62389	6.59966	0.345	0.03159	0.03014
m032a	0.67892	0.07304	89.41651	9.75852	0.01455	1178	1.06861	22.11394	0.48	0.0674	0.02601
m044a	0.96933	0.10532	62.73508	7.90331	0.05988	27247	32.80676	14.91631	0.45	0.08191	0.063
m058a	0.72584	0.05957	83.23077	6.98245	0.02832	6363	8.92915	8.10348	0.38	0.05012	0.02906
m063a	0.71922	0.06192	84.07853	7.65343	0.01561	1231	1.01184	9.625	0.385	0.04844	0.03657
m065a	0.70054	0.04786	86.06741	6.17961	0.01873	2155	1.82327	6.9913	0.335	0.03803	0.02628
m068a	0.71303	0.07608	85.08687	8.89414	0.01167	640	0.61542	12.5009	0.495	0.07138	0.02303
m069a	0.67635	0.04585	89.13783	6.30725	0.01282	320	0.25868	7.50312	0.25	0.04022	0.02079
m072a	0.76501	0.06162	78.97534	6.83046	0.01848	1705	1.51094	8.31884	0.435	0.05531	0.02423
m074a	0.79197	0.06451	76.27907	6.39783	0.02307	4189	4.08054	23.95217	0.36	0.05602	0.02996
m089a	0.68678	0.05557	87.949	7.32749	0.02129	5094	4.16868	9.69671	0.38	0.04957	0.02135
m091a	0.68099	0.05182	88.65259	7.20562	0.01011	145	0.11581	8.02255	0.26	0.04892	0.01464
m094a	0.6918	0.05503	87.30107	7.2389	0.01292	635	0.57422	8.3949	0.305	0.04929	0.02134
m100a	0.72655	0.05775	83.09628	6.52322	0.02409	6361	5.63788	9.50763	0.43	0.04588	0.03023
m103a	0.78265	0.08373	77.51709	8.04361	0.0158	862	0.87542	12.50546	0.41	0.08012	0.02246
m107a	0.67265	0.04797	89.65069	6.34593	0.01204	109	0.07902	8.57184	0.33	0.04317	0.0198
m127a	0.71767	0.05778	84.19559	7.40399	0.01561	689	0.81225	14.69884	0.38	0.04973	0.02698
m135a	0.75844	0.07903	79.99272	8.56375	0.02249	3404	3.16701	13.454	0.475	0.0654	0.04156
m136a	0.66939	0.05551	90.2844	7.89538	0.01964	2873	2.26149	8.36346	0.31	0.05014	0.02315
m141a	0.71741	0.05561	84.17971	7.08054	0.00917	109	0.09385	6.92393	0.28	0.05113	0.01969
m145a	0.72267	0.06591	83.72711	7.74494	0.01263	397	0.37858	19.0598	0.345	0.06188	0.02042
m151a	0.72449	0.05742	83.34772	6.736	0.01909	2613	2.24344	20.49877	0.395	0.04634	0.03168
m158a	0.75473	0.07532	80.25156	7.60884	0.01288	435	0.37066	11.3742	0.385	0.07421	0.01269
m169a	0.79601	0.09128	76.27214	7.9311	0.04369	9996	10.44831	15.895	0.415	0.07502	0.04395
m172a	0.68316	0.05237	88.35367	6.90527	0.01447	652	0.54184	8.50516	0.305	0.04761	0.02071
m183a	0.68139	0.07599	89.21446	10.53628	0.01582	651	0.54655	12.50772	0.385	0.06746	0.03132
m185a	0.67693	0.07553	89.79024	10.45329	0.01649	1481	1.21126	11.10435	0.4	0.06724	0.03083
m186a	0.66596	0.05568	90.75751	7.97239	0.01047	90	0.06982	8.11782	0.325	0.05007	0.02195
m188a	0.75699	0.07566	80.06253	8.10257	0.05177	24259	23.00151	13.6652	0.465	0.05788	0.04624
m192a	0.70945	0.04934	85.00482	6.24372	0.01333	395	0.57957	6.31767	0.24	0.04542	0.01857
m205a	0.77739	0.06899	77.87358	7.89341	0.01226	213	0.22023	10.59564	0.39	0.0635	0.02308
m210a	0.76317	0.05673	79.06108	5.96276	0.01281	161	0.14871	10.22787	0.285	0.0506	0.02426
m211a	0.73714	0.06622	82.10955	8.07296	0.02675	5782	6.60355	10.4988	0.42	0.05577	0.03321
m213a	0.78051	0.06232	77.34231	5.92003	0.01947	2487	2.53866	14.38984	0.405	0.05606	0.02295
m215a	0.67113	0.06062	90.20024	8.93092	0.01095	202	0.16479	7.7021	0.335	0.05578	0.02165
m217a	0.66853	0.07399	90.88644	10.34857	0.01267	523	0.43447	14.41123	0.435	0.07009	0.0218
m218a	0.71008	0.06706	85.25798	8.10764	0.02056	3602	3.06997	11.09198	0.43	0.05977	0.02728
m229a	0.70351	0.05313	85.71828	6.00431	0.01876	372	1.05263	8.05401	0.445	0.03993	0.02578
m242a	0.71307	0.05816	84.69655	6.82708	0.01961	1756	1.61308	10.94961	0.385	0.05234	0.02307
m245a	0.72685	0.06882	83.4401	9.568	0.01322	903	0.75266	9.66379	0.335	0.06715	0.01447
m257a	0.73793	0.07769	82.24793	9.00473	0.0131	305	0.29264	13.69027	0.44	0.06968	0.03285
m260a	0.68089	0.05753	88.75884	7.59758	0.01584	800	0.67377	11.11855	0.415	0.05337	0.02077
m263a	0.70224	0.05221	85.91024	6.33977	0.01366	311	0.25188	8.23263	0.375	0.04817	0.01853
m267a	0.7565	0.06934	79.99615	7.4961	0.01678	1607	1.63971	10.68418	0.35	0.06518	0.02251
m269a	0.66871	0.04942	90.2353	6.93155	0.01309	141	0.11199	12.41392	0.295	0.04607	0.01716
m273a	0.78072	0.06675	77.4478	7.01571	0.01262	327	0.31825	16.02245	0.35	0.06443	0.01575
m275a	0.71679	0.07092	84.48616	7.98268	0.033	9267	8.40788	10.8708	0.47	0.05583	0.03713
m277a	0.72151	0.07339	84.01756	8.52476	0.01821	1785	1.52606	12.21481	0.495	0.06577	0.03037
m282a	0.68098	0.07317	89.17404	9.97831	0.00969	56	0.04467	11.04918	0.33	0.06939	0.02073
m293a	0.77565	0.08367	78.24792	8.37061	0.01362	523	0.47554	11.77275	0.46	0.07552	0.03231

Fonte: O autor.

Figura 30 - Boxplot: Grupo supressão de arritmia cardíaca (domínio da frequência).



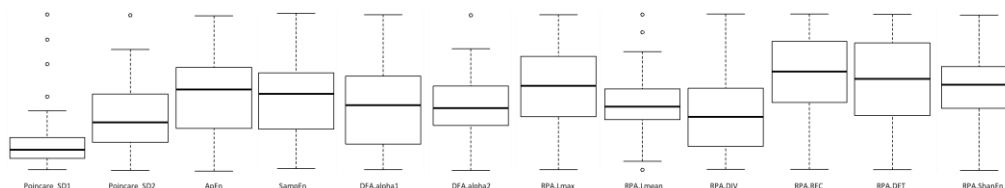
Fonte: O autor.

Figura 31 - Matriz de dados: Grupo supressão de arritmia cardíaca (domínio da frequência).

amostra	HF_peak	VLF_power	VLF_power_prc	LF_power	LF_power_prc	LF_power_nu	HF_power	HF_power_prc	HF_power_nu	LF_HF_power	tot_power
m006a	0.28906	0.00337	96.98208	0.00004	1.24373	41.21127	0.00006	1.76208	58.38724	0.70583	0.00348
m012a	0.15625	0.00239	93.10314	0.00006	2.243	32.52203	0.00012	4.6423	67.31024	0.48317	0.00257
m014a	0.35547	0.00335	95.41771	0.00012	3.27398	71.44858	0.00005	1.30032	28.37705	2.51783	0.00351
m022a	0.30078	0.00316	96.75283	0.00004	1.182	36.40093	0.00007	2.04975	63.12413	0.57666	0.00327
m028a	0.32813	0.00152	88.66575	0.00015	8.60796	75.94642	0.00005	2.70729	23.8859	3.17955	0.00171
m032a	0.15234	0.00357	94.95399	0.00015	4.00038	79.27804	0.00004	1.0391	20.59258	3.84984	0.00376
m044a	0.32813	0.00622	82.51226	0.00053	6.97188	39.86722	0.00079	10.44915	59.75126	0.66722	0.00754
m058a	0.28906	0.00213	84.38887	0.00025	9.78264	62.66451	0.00015	5.78881	37.08127	1.68992	0.00253
m063a	0.15234	0.00241	86.22798	0.00032	11.44772	83.12297	0.00006	2.31409	16.80283	4.94696	0.0028
m065a	0.28906	0.00151	93.49541	0.00004	2.78535	42.82129	0.00006	3.68099	56.59064	0.75669	0.00161
m068a	0.15234	0.00401	96.31031	0.00013	3.12014	84.56366	0.00002	0.56596	15.33888	5.51303	0.00416
m069a	0.25	0.00133	91.19853	0.00008	5.80865	65.9963	0.00004	2.98194	33.88004	1.94794	0.00146
m072a	0.21094	0.00238	92.88518	0.00009	3.56989	50.1754	0.00009	3.52598	49.55821	1.01245	0.00256
m074a	0.26953	0.00268	93.00845	0.00008	2.86351	40.95677	0.00012	4.09671	58.59515	0.69898	0.00288
m089a	0.34766	0.002	94.46218	0.00005	2.47609	44.7124	0.00006	3.03303	54.76938	0.81638	0.00211
m091a	0.28516	0.00171	97.78719	0.00001	0.78104	35.29636	0.00002	1.425	64.39763	0.5481	0.00175
m094a	0.375	0.00198	96.45666	0.00005	2.24243	63.28589	0.00003	1.28674	36.31434	1.74272	0.00205
m100a	0.36328	0.00234	92.2114	0.00012	4.71102	60.48612	0.00008	3.04852	39.14076	1.54535	0.00254
m103a	0.15625	0.00487	97.2647	0.00009	1.71566	62.72292	0.00005	1.01349	37.05236	1.69282	0.00501
m107a	0.27734	0.00154	93.02104	0.00006	3.77032	54.02412	0.00005	3.1997	45.84785	1.17834	0.00166
m127a	0.32813	0.00215	94.24162	0.00009	4.03069	69.99683	0.00004	1.71604	29.80078	2.34883	0.00228
m135a	0.15234	0.00416	90.68279	0.00034	7.39918	79.4142	0.00009	1.90789	20.47704	3.87821	0.00459
m136a	0.23828	0.00192	93.05504	0.00009	4.46372	64.27274	0.00005	2.46063	35.43041	1.81406	0.00207
m141a	0.28906	0.00202	97.90027	0.00003	1.22352	58.2702	0.00002	0.87152	41.50599	1.4039	0.00206
m145a	0.33984	0.00293	97.63771	0.00004	1.20424	50.97779	0.00003	1.1491	48.64327	1.04799	0.003
m151a	0.23438	0.00219	89.62456	0.00016	6.49532	62.60286	0.00009	3.86463	37.2479	1.68071	0.00245
m158a	0.15234	0.00402	99.12318	0.00002	0.42243	48.17785	0.00002	0.45023	51.34841	0.93825	0.00406
m169a	0.15234	0.00587	86.68482	0.0006	8.79847	66.07847	0.0003	4.48043	33.64903	1.96376	0.00677
m172a	0.24609	0.0018	94.216	0.00006	3.37743	58.39265	0.00005	2.3943	41.39529	1.41061	0.00191
m183a	0.22656	0.00374	94.7888	0.00014	3.63628	69.7781	0.00006	1.56896	30.10755	2.31763	0.00395
m185a	0.30078	0.00378	95.9094	0.0001	2.5342	61.95187	0.00006	1.54797	37.84207	1.63712	0.00394
m186a	0.30078	0.00204	97.38599	0.00004	1.8774	71.82062	0.00002	0.7308	27.95712	2.56896	0.0021
m188a	0.15234	0.00308	75.43403	0.0005	12.33838	50.22551	0.0005	12.14902	49.45469	1.01559	0.00408
m192a	0.28906	0.00155	95.28742	0.00004	2.71672	57.64815	0.00003	1.98243	42.06685	1.37039	0.00163
m205a	0.38281	0.00291	97.3382	0.00006	1.90014	71.38541	0.00002	0.75156	28.23519	2.52824	0.00298
m210a	0.26172	0.00215	94.1397	0.00009	3.90133	66.57211	0.00004	1.95218	33.31196	1.99844	0.00228
m211a	0.16406	0.00262	87.5257	0.00018	5.92708	47.51433	0.0002	6.50955	52.18369	0.91052	0.003
m213a	0.27734	0.00269	95.31742	0.00007	2.40669	51.39659	0.00006	2.25855	48.23305	1.06559	0.00283
m215a	0.28516	0.00235	97.66833	0.00004	1.56853	67.27084	0.00002	0.75719	32.47412	2.07152	0.0024
m217a	0.23828	0.00355	96.7131	0.00009	2.43622	74.11911	0.00003	0.84687	25.7651	2.87673	0.00367
m218a	0.26953	0.00302	95.02892	0.00007	2.34733	47.21971	0.00008	2.60841	52.47169	0.89991	0.00318
m229a	0.33203	0.00143	91.12898	0.00011	6.90395	77.82586	0.00003	1.95406	22.02749	3.53312	0.00157
m242a	0.28906	0.00224	93.81558	0.0001	4.31953	69.84547	0.00004	1.85267	29.95704	2.33152	0.00239
m245a	0.28906	0.00268	98.46382	0.00002	0.59951	39.02623	0.00003	0.92769	60.38936	0.64624	0.00272
m257a	0.28906	0.00401	94.98922	0.00017	4.00495	79.92654	0.00004	1.00202	19.9973	3.99687	0.00422
m260a	0.28906	0.00212	92.69414	0.00013	5.49903	75.26873	0.00004	1.79599	24.58287	3.06184	0.00229
m263a	0.29297	0.00186	97.05899	0.00003	1.40798	47.87405	0.00003	1.51917	51.65455	0.92681	0.00192
m267a	0.28906	0.00318	96.78317	0.00004	1.34799	41.90412	0.00006	1.85166	57.56174	0.72799	0.00329
m269a	0.25391	0.00156	95.01093	0.00003	1.78433	35.76474	0.00005	3.1943	64.02604	0.5586	0.00164
m273a	0.23828	0.00287	98.32796	0.00002	0.67437	40.33231	0.00003	0.98938	59.1719	0.68161	0.00292
m275a	0.15234	0.00349	85.81211	0.00029	7.01231	49.42461	0.00029	7.14281	50.34441	0.98173	0.00407
m277a	0.15234	0.00359	91.60865	0.00026	6.68201	79.62973	0.00007	1.6994	20.25185	3.93197	0.00392
m282a	0.26563	0.0035	98.50736	0.00004	1.03233	69.1617	0.00002	0.45727	30.63477	2.25762	0.00356
m293a	0.26563	0.00495	96.33882	0.00014	2.81065	76.7689	0.00004	0.8467	23.12631	3.31955	0.00514

Fonte: O autor.

Figura 32 - Boxplot: Grupo supressão de arritmia cardíaca (parâmetros não lineares)



Fonte: O autor.

Figura 33 - Matriz de dados: Grupo supressão de arritmia cardíaca (parâmetros não lineares).

amostra	Poincare_SD1	Poincare_SD2	ApEn	SampEn	DFA.alpha1	DFA.alpha2	RPA.Lmax	RPA.Lmean	RPA.DIV	RPA.REC	RPA.DET	RPA.ShanEn
m006a	0.01143	0.10071	1.3394	1.49392	0.72187	1.18788	552.54286	32.80275	0.00294	0.44411	0.98838	3.99396
m012a	0.01206	0.08638	1.4202	1.64169	1.04423	0.96518	513.64655	23.32311	0.00353	0.37443	0.98306	3.62283
m014a	0.00945	0.09954	1.2571	1.36145	1.18217	1.10933	803.89216	25.78903	0.00205	0.44818	0.991	3.91085
m022a	0.01431	0.09596	1.30061	1.40584	0.69919	1.11138	358.2243	26.02881	0.00578	0.38254	0.98077	3.61532
m028a	0.01027	0.06529	1.20568	1.28328	1.15954	1.21858	663.95575	28.36182	0.00186	0.47875	0.99348	4.01071
m032a	0.01029	0.10278	1.16444	1.19415	1.22698	1.11916	670.93636	24.18597	0.00176	0.45864	0.99467	3.91579
m044a	0.04234	0.14279	1.34509	1.48981	0.79471	1.0371	338.19277	24.88	0.00515	0.38801	0.98027	3.65017
m058a	0.02003	0.08183	1.35227	1.53406	0.8571	0.95451	298.8169	17.58956	0.00468	0.38024	0.98226	3.49897
m063a	0.01104	0.08687	1.08404	1.08854	1.40536	1.04655	831.86777	25.13168	0.00136	0.47645	0.99643	3.96822
m065a	0.01324	0.06638	1.2971	1.43848	0.73233	1.27462	587.36441	30.61644	0.00291	0.43709	0.98721	3.94899
m068a	0.00825	0.10728	1.13129	1.12793	1.32963	1.1383	646.5534	27.67942	0.0019	0.49367	0.99554	4.00968
m069a	0.00907	0.06421	1.33159	1.44774	1.02925	1.06331	636.56098	27.3288	0.00249	0.42334	0.98843	3.77369
m072a	0.01307	0.08616	1.36228	1.53778	0.89373	1.07686	438.91071	32.05561	0.00507	0.44298	0.98253	3.78184
m074a	0.01632	0.08976	1.31395	1.55448	0.81057	1.14728	562.15686	26.05848	0.00358	0.41744	0.98471	3.77928
m089a	0.01505	0.07712	1.29295	1.37431	0.69291	1.09105	377.95902	26.88766	0.00446	0.41003	0.98366	3.70602
m091a	0.00715	0.07293	1.19143	1.15782	0.71555	1.24588	559.64	46.04482	0.00302	0.46556	0.98805	4.06978
m094a	0.00914	0.07729	1.22368	1.27274	0.93095	1.25637	598.52727	37.32862	0.00259	0.48256	0.9911	4.12806
m100a	0.01703	0.07988	1.11341	1.10014	0.79282	1.22394	436.36607	25.99637	0.00358	0.4274	0.98875	3.87854
m103a	0.01118	0.11788	1.34705	1.4888	1.03397	1.04731	535.14286	24.03744	0.00414	0.42069	0.98528	3.72778
m107a	0.00851	0.06731	1.35622	1.45407	1.03043	1.09613	692.37226	19.66066	0.00198	0.40439	0.98896	3.67359
m127a	0.01104	0.08096	1.22897	1.32977	1.09565	1.28215	647.33333	24.30554	0.00256	0.4608	0.99201	3.91476
m135a	0.0159	0.11062	1.21167	1.32588	1.23891	1.08381	725.45794	23.85038	0.00192	0.46506	0.99248	3.90436
m136a	0.01389	0.07727	1.41037	1.68977	0.89153	1.09054	371.90551	17.02049	0.00399	0.39274	0.98362	3.56097
m141a	0.00649	0.07837	1.16313	1.13792	1.03451	1.46118	800.46552	29.17246	0.00185	0.47527	0.99412	4.05914
m145a	0.00893	0.09279	1.2063	1.25387	0.8668	1.27138	759.09615	38.17759	0.00219	0.48459	0.99165	4.17381
m151a	0.0135	0.08008	1.26349	1.43615	1.06076	1.20578	588.21552	23.14886	0.00231	0.44988	0.99199	3.85867
m158a	0.00911	0.10613	1.30916	1.33018	0.67328	1.08283	459	27.6706	0.00339	0.4278	0.98431	3.80832
m169a	0.03089	0.12534	1.21295	1.23105	1.02832	1.09736	569.71579	26.60166	0.00276	0.43887	0.98919	3.90995
m172a	0.01023	0.07335	1.39209	1.54618	0.91154	1.09646	469.61667	23.28359	0.00372	0.40285	0.98596	3.6763
m183a	0.01118	0.10688	1.27427	1.446	1.12481	1.10761	699.37815	29.30199	0.00276	0.45928	0.98964	3.90607
m185a	0.01166	0.10617	1.17779	1.23733	0.98909	1.30159	711.78689	32.83336	0.00202	0.4731	0.99287	4.09988
m186a	0.00741	0.0784	1.14763	1.1691	1.02214	1.33288	804.5625	36.11879	0.00165	0.49434	0.99429	4.23898
m188a	0.03661	0.10054	1.30839	1.36974	0.83023	0.92941	230.13333	15.64182	0.00623	0.34343	0.97808	3.41325
m192a	0.00943	0.06914	1.30007	1.35209	0.944	1.14471	587.39706	27.4723	0.00322	0.43101	0.98853	3.81106
m205a	0.00867	0.09719	1.1863	1.23842	1.09036	1.22001	800.64583	35.19717	0.00153	0.49508	0.99428	4.14836
m210a	0.00906	0.07972	1.29708	1.42235	1.11461	1.11566	802.83333	28.47435	0.0016	0.46794	0.99341	4.00624
m211a	0.01892	0.09172	1.15381	1.1577	0.89443	1.02496	272.11494	24.50564	0.00516	0.41683	0.99084	3.76496
m213a	0.01377	0.08705	1.31469	1.42165	0.74082	1.16177	419.02062	24.89651	0.00359	0.43506	0.98804	3.8868
m215a	0.00774	0.08539	1.13538	1.15588	0.96333	1.34651	811.2541	38.34543	0.00164	0.50082	0.99443	4.32147
m217a	0.00896	0.10426	1.26889	1.33742	1.13272	1.05781	699.14167	24.78141	0.00201	0.45843	0.99157	3.85906
m218a	0.01454	0.09372	1.28259	1.4088	0.79252	1.16384	537.4188	35.61353	0.00309	0.46297	0.98816	3.97725
m229a	0.01327	0.07395	1.14143	1.18149	1.23353	1.17947	719.82857	33.53031	0.00168	0.48198	0.99455	3.98234
m242a	0.01386	0.08107	1.35944	1.60415	0.99985	1.03768	446.46296	21.02968	0.00348	0.39103	0.98086	3.57646
m245a	0.00934	0.09688	1.32861	1.37672	0.67509	1.1295	394.52941	26.11407	0.00439	0.40555	0.98334	3.75008
m257a	0.00926	0.10948	1.10903	1.15923	1.32453	1.19175	887.47115	29.27895	0.00125	0.48857	0.99655	4.10829
m260a	0.0112	0.08058	1.4548	1.6887	1.12508	0.93262	455.66949	13.90295	0.00386	0.36988	0.98172	3.36995
m263a	0.00966	0.07321	1.3108	1.39529	0.70404	1.2078	581.7561	35.88517	0.00328	0.44074	0.98684	3.98245
m267a	0.01187	0.09735	1.3433	1.48318	0.75852	1.19324	528.96939	28.44702	0.00347	0.44188	0.98948	3.94311
m269a	0.00926	0.06928	1.44651	1.59065	0.70376	1.18513	488.064	27.17849	0.00385	0.40488	0.98336	3.69395
m273a	0.00892	0.09398	1.32028	1.37217	0.67915	1.1645	486.53922	29.0152	0.00415	0.41427	0.9845	3.84261
m275a	0.02334	0.09754	1.27354	1.37239	0.81661	1.1509	412.91818	25.32196	0.00387	0.41109	0.98741	3.76557
m277a	0.01288	0.10299	1.24519	1.34694	1.27835	1.04553	663.18966	21.93484	0.00204	0.46804	0.99331	3.86132
m282a	0.00685	0.10324	1.0886	1.07616	1.02802	1.23652	775.84	42.42037	0.00208	0.48422	0.99318	4.21446
m293a	0.00963	0.11794	1.10485	1.13407	1.25541	1.2493	828.22018	27.35868	0.0014	0.49216	0.99561	4.10074

Fonte: O autor.