



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CAMPUS QUIXADÁ**  
**CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO**

**STEFANE RIBEIRO DOS SANTOS**

**O USO DE APRENDIZADO PROFUNDO PARA PREDIÇÃO DE DIAGNÓSTICOS DE  
DOENÇAS CARDIOVASCULARES**

**QUIXADÁ**

**2023**

STEFANE RIBEIRO DOS SANTOS

O USO DE APRENDIZADO PROFUNDO PARA PREDIÇÃO DE DIAGNÓSTICOS DE  
DOENÇAS CARDIOVASCULARES

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Campus Quixadá da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Emanuel Ferreira Coutinho

Coorientador: Prof. Dr. Victor Aguiar Evangelista de Farias

QUIXADÁ

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- S238u Santos, Stefane Ribeiro dos.  
O uso de aprendizado profundo para predição de diagnósticos de doenças cardiovasculares / Stefane Ribeiro dos Santos. – 2023.  
51 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Engenharia de Computação, Quixadá, 2023.  
Orientação: Prof. Dr. Emanuel Ferreira Coutinho.  
Coorientação: Prof. Dr. Victor Aguiar Evangelista de Farias.
1. Doenças cardiovasculares. 2. Aprendizado profundo. 3. Eletrocardiograma. 4. Redes neurais. 5. Classificação. I. Título.

CDD 621.39

---

STEFANE RIBEIRO DOS SANTOS

O USO DE APRENDIZADO PROFUNDO PARA PREDIÇÃO DE DIAGNÓSTICOS DE  
DOENÇAS CARDIOVASCULARES

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Campus Quixadá da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Aprovada em: 07 de Dezembro de 2023

BANCA EXAMINADORA

---

Prof. Dr. Emanuel Ferreira Coutinho (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Victor Aguiar Evangelista de  
Farias (Coorientador)  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Lívia Almada Cruz  
Universidade Federal do Ceará (UFC)

À minha família, por sua capacidade de acreditar em mim e me apoiar nesta caminhada. Mãe, seu cuidado e dedicação foi o que deram a esperança e a força necessária para seguir.

## AGRADECIMENTOS

A Deus, que permitiu que eu tivesse determinação e me conduziu nesta trajetória.

À minha mãe, Ednilce Ribeiro da Cruz, que me ensinou a importância do esforço e da dedicação. Seu exemplo de vida é minha fonte de inspiração e motivação para buscar sempre o melhor.

À minha família, pelo apoio que sempre me deram ao longo de toda a minha vida.

Aos meus amigos, que sempre me encorajaram a buscar meus objetivos e me ajudaram a manter a motivação.

Ao Prof. Dr. Emanuel Ferreira Coutinho, meu orientador, e ao Prof. Dr. Victor Aguiar Evangelista de Farias, meu coorientador, por me orientarem em meu Trabalho de Conclusão de Curso.

À Universidade Federal do Ceará e aos seus docentes, pela excelência da qualidade de ensino oferecida.

A todos os que, de alguma forma, colaboraram para a realização deste Trabalho de Conclusão de Curso, seja fornecendo informações ou materiais, seja simplesmente ouvindo minhas ideias.

Ao Doutor em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, aluno de graduação em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da Universidade Federal do Ceará (UFC).

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.”

(Madre Teresa de Calcutá)

## RESUMO

As Doenças Cardiovasculares (DCV) são um grupo de doenças do coração e dos vasos sanguíneos. Essas doenças configuram a maior causa de morte no mundo. As DCV vem afetando cada vez mais jovens adultos, devido à prática de hábitos não saudáveis, como não realizar exercícios físicos, excesso de peso, obesidade, tabagismo, estresse, depressão e alimentação com alto teor de gordura e industrializados. Nesse contexto, a detecção precoce de DCV desempenha um papel essencial para que o manejo com aconselhamento e medicamentos possa ter início. A maior parte das adversidades cardíacas podem ser diagnosticadas por meio de sinais de Eletrocardiograma (ECG). A classificação do ECG apresenta uma relevância bastante alta atualmente, visto que, há diversas aplicações médicas em que este problema pode ser evidenciado. Na segunda metade do século 20, o Aprendizado de Máquina (AM) evoluiu como um subcampo da Inteligência Artificial (IA), sendo capaz de desenvolver algoritmos de autoaprendizagem que formam conhecimento de dados para poder fazer previsões. Embora haja muitos métodos de AM que apresentam soluções capazes de auxiliar na interpretação do ECG, a predição de DCV configura um dos impasses mais complexos na área de análise de dados clínicos. Dito isso, realizar análises com métodos ainda pouco explorados para este fim pode trazer contribuições significativas para a área da saúde. Nessa perspectiva, este trabalho propõe o uso de Aprendizado Profundo (AP) para classificação de diagnósticos de DCV, desenvolvendo modelos de redes neurais *Transformer* e *Long Short-Term Memory* (LSTM), e avaliando-os por meio das métricas de precisão, revocação e *f1-score*, com o intuito de comparar o desempenho dos métodos que foram utilizados para realizar a predição de diagnósticos de DCV. Os resultados dos experimentos mostraram valores de precisão de, aproximadamente, 70% a 84% para a rede neural LSTM e valores de precisão igual ou inferiores a 48% para a rede neural *Transformer*. Para a métrica de revocação, o *Transformer* obteve valores que alcançaram até 62% e a LSTM obteve resultados de até 77%. Para a métrica *f1-score*, a LSTM apresentou valores entre 41% a 72% e o *Transformer* obteve valores menores, de até 36%.

**Palavras-chave:** doenças cardiovasculares; aprendizado profundo; eletrocardiograma; redes neurais; classificação.

## ABSTRACT

Cardiovascular diseases (CVD) are a group of diseases of the heart and blood vessels. These diseases are the leading cause of death worldwide. CVD is affecting more and more young adults, due to unhealthy habits such as not exercising, being overweight, obesity, smoking, stress, depression and eating a diet high in fat and processed foods. In this context, early detection of CVD plays an essential role so that management with counseling and medication can begin. Most heart conditions can be diagnosed using electrocardiogram (ECG) signals. The classification of the ECG is highly relevant today, as there are many medical applications in which this problem can be evidenced. In the second half of the 20th century, Machine Learning (ML) evolved as a sub-field of Artificial Intelligence (AI), capable of developing self-learning algorithms that form knowledge from data in order to make predictions. Although there are many AM methods that present solutions capable of helping with ECG interpretation, predicting CVD is one of the most complex impasses in the field of clinical data analysis. That said, carrying out analyses with methods that have not yet been explored for this purpose could make significant contributions to the health field. With this in mind, this work proposes the use of deep learning to classify CVD diagnoses, developing Transformer and LSTM neural network models and evaluating them using the metrics of accuracy, recall and f1-score, in order to compare the performance of the methods that have been used to predict cardiovascular disease diagnoses. The results of the experiments showed accuracy values of approximately 70% to 84% for the LSTM neural network and accuracy values of 48% or less for the Transformer neural network. For the recall metric, Transformer obtained values of up to 62% and LSTM obtained results of up to 77%. For the f1-score metric, LSTM had values between 41% and 72% and Transformer had lower values of up to 36%.

**Keywords:** cardiovascular diseases; deep learning; electrocardiogram; neural networks; classification.

## LISTA DE FIGURAS

Figura 1 – Exemplos de Eletrocardiogramas para cada superclasse de diagnóstico . . .	18
Figura 2 – O Aprendizado de Máquina pode ajudar no ensino dos humanos . . . . .	19
Figura 3 – Um conjunto de treinamento rotulado para aprendizado supervisionado, por exemplo, classificação de spam . . . . .	20
Figura 4 – Perceptron Multicamadas . . . . .	23
Figura 5 – Célula LSTM . . . . .	24
Figura 6 – Arquitetura do modelo <i>Transformer</i> . . . . .	27
Figura 7 – Fluxograma de procedimentos metodológicos . . . . .	35
Figura 8 – Conjunto de dados PTB-XL em termos de superclasses e subclasses de diagnóstico representado graficamente . . . . .	42

## LISTA DE TABELAS

Tabela 1 – Análise comparativa entre os trabalhos relacionados e este trabalho . . . . .	34
Tabela 2 – Informações sobre a base de dados PTB-XL . . . . .	38
Tabela 3 – Descrição dos atributos da base de dados . . . . .	39
Tabela 4 – Visão geral da declaração de diagnóstico . . . . .	40
Tabela 5 – Hiperparâmetros da rede neural <i>Transformer</i> . . . . .	44
Tabela 6 – Desempenho do modelo 1 da rede neural <i>Transformer</i> . . . . .	44
Tabela 7 – Desempenho do modelo 2 da rede neural <i>Transformer</i> . . . . .	45
Tabela 8 – Desempenho do modelo 3 da rede neural <i>Transformer</i> . . . . .	45
Tabela 9 – Desempenho da rede neural LSTM . . . . .	45

## LISTA DE ABREVIATURAS E SIGLAS

AFib	Fibrilação Atrial
AFlut	Flutter Atrial
AM	Aprendizado de Máquina
AP	Aprendizado Profundo
APS	Atenção Primária à Saúde
AVC	Acidente Vascular Cerebral
CNN	Redes Neurais Convolucionais
DCV	Doenças Cardiovasculares
ECG	Eletrocardiograma
HAS	Hipertensão Arterial Sistêmica
IA	Inteligência Artificial
LSTM	<i>Long Short-Term Memory</i>
PACS	<i>Picture Archiving and Communication System</i>
PLN	Processamento de Linguagem Natural
PTB	<i>Physikalisch-Technische Bundesanstalt</i>
RNA	Redes Neurais Artificiais

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	14
<b>1.1</b>	<b>Objetivos</b>	15
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	16
<b>2.1</b>	<b>Doenças Cardiovasculares</b>	16
<i>2.1.1</i>	<i>Eletrocardiograma</i>	17
<i>2.1.2</i>	<i>Séries Temporais</i>	18
<b>2.2</b>	<b>Aprendizado de Máquina</b>	19
<i>2.2.1</i>	<i>Aprendizado supervisionado</i>	19
<i>2.2.2</i>	<i>Classificação</i>	20
<i>2.2.3</i>	<i>Métricas de Avaliação</i>	21
<i>2.2.3.1</i>	<i>Precisão</i>	21
<i>2.2.3.2</i>	<i>Revocação</i>	21
<i>2.2.3.3</i>	<i>F1-score</i>	22
<b>2.3</b>	<b>Redes Neurais e Aprendizado Profundo</b>	22
<i>2.3.1</i>	<i>Redes Neurais Recorrentes</i>	23
<i>2.3.1.1</i>	<i>LSTM</i>	24
<i>2.3.2</i>	<i>Transformer</i>	25
<i>2.3.2.1</i>	<i>Arquitetura do Modelo</i>	26
<i>2.3.2.2</i>	<i>Mecanismo de Atenção</i>	26
<i>2.3.2.2.1</i>	<i>Pilhas de Codificador e Decodificador</i>	27
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	29
<b>3.1</b>	<i>Deep Learning for ECG Classification</i>	29
<b>3.2</b>	<i>Towards Understanding ECG Rhythm Classification Using Convolutional Neural Networks and Attention Mappings</i>	30
<b>3.3</b>	<b>Aprendizado de Máquina para Predição de Diagnósticos de Doenças Cardiovasculares</b>	31
<b>3.4</b>	<i>Deep Learning Approach for Detection of Atrial Fibrillation and Atrial Flutter Based on ECG Images</i>	32
<b>3.5</b>	<b>Análise Comparativa</b>	33
<b>4</b>	<b>PROCEDIMENTOS METODOLÓGICOS</b>	35

<b>4.1</b>	<b>Obtenção da base de dados</b>	35
<b>4.2</b>	<b>Pré-processamento dos dados</b>	36
<b>4.3</b>	<b>Treinamento e otimização dos modelos</b>	36
<b>4.4</b>	<b>Avaliação dos modelos</b>	36
<b>5</b>	<b>RESULTADOS E DISCUSSÕES</b>	38
<b>5.1</b>	<b>Obtenção da base de dados</b>	38
<b>5.1.1</b>	<i>Base de dados PTB-XL</i>	38
<b>5.1.2</b>	<i>Pré-processamento dos dados</i>	40
<b>5.1.2.1</b>	<i>Remoção de dados ausentes</i>	42
<b>5.1.2.2</b>	<i>Codificação de Variáveis Categóricas</i>	42
<b>5.1.3</b>	<i>Treinamento dos Modelos</i>	42
<b>5.1.3.1</b>	<i>Arquitetura da Rede Neural LSTM</i>	43
<b>5.1.3.2</b>	<i>Arquitetura da Rede Neural Transformer</i>	43
<b>5.1.3.3</b>	<i>Hiperparâmetros</i>	43
<b>5.1.4</b>	<i>Avaliação dos Modelos</i>	44
<b>6</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS</b>	47
<b>6.1</b>	<b>Considerações Finais</b>	47
<b>6.2</b>	<b>Trabalhos futuros</b>	47
	<b>REFERÊNCIAS</b>	49
	<b>ANEXOS</b>	50
	<b>ANEXO A – Pré-processamento dos dados fornecido pela Physionet</b>	51

## 1 INTRODUÇÃO

As Doenças Cardiovasculares (DCV) são um grupo de doenças do coração e dos vasos sanguíneos (WHO, 2023). Ataques cardíacos e derrames são, geralmente, eventos agudos e são ocasionados, principalmente, por um bloqueio que impossibilita o fluxo de sangue para o coração ou cérebro (WHO, 2023). A causa mais comum para que isso ocorra é o acúmulo de depósitos de gordura nas paredes internas dos vasos sanguíneos que irrigam o coração ou o cérebro. Além disso, a causa dos derrames pode está relacionada ao sangramento de um vaso sanguíneo no cérebro ou por coágulos sanguíneos (WHO, 2023).

Os fatores de risco comportamentais mais relevantes para DCV e Acidente Vascular Cerebral (AVC) são alimentação pouco saudável, inatividade física, tabagismo e consumo nocivo de álcool. Os efeitos desses fatores de risco podem ser percebidos em pessoas como aumento da pressão arterial, aumento da glicemia, aumento dos lipídios no sangue, sobrepeso e obesidade (WHO, 2023). Em 2021, a Atenção Primária à Saúde (APS) realizou 18,1 milhões de atendimentos a indivíduos com hipertensão e 2,3 milhões a indivíduos com obesidade, sendo estes fatores de risco para o coração (MS, 2021).

A detecção precoce de doenças cardiovasculares desempenha um papel muito importante para que o manejo com aconselhamento e medicamentos possa ter início (WHO, 2023). A maioria das adversidades cardíacas, podem ser diagnosticadas por meio de sinais de ECG, sendo apontada uma estimativa de centenas de milhões de ECG anualmente registrados (PYAKILLYA *et al.*, 2017). Os ECG se apresentam na forma de séries temporais (FILHO, 2022).

A classificação do ECG apresenta uma relevância bastante alta atualmente, visto que, há várias aplicações médicas nas quais este problema pode ser evidenciado (PYAKILLYA *et al.*, 2017). Na segunda metade do século 20, o Aprendizado de Máquina (AM) teve sua evolução como um subcampo da Inteligência Artificial (IA), sendo capaz de desenvolver algoritmos de autoaprendizagem que formam conhecimento de dados para poder realizar previsões (RASCHKA; MIRJALILI, 2019). Atualmente, há uma variedade de métodos de AM apresentando soluções que podem ser utilizadas para examinar e classificar dados de ECG (PYAKILLYA *et al.*, 2017).

Entretanto, embora haja diversos métodos de AM que apresentam soluções capazes de auxiliar na interpretação do ECG, de acordo com Mohan *et al.* (2019), a predição de DCV configura um dos impasses mais complicados na área de análise de dados clínicos. À vista disso, realizar análises com métodos ainda pouco explorados para este fim pode resultar em contribuições significativas para a área da saúde. Dessa forma, este trabalho busca realizar a

predição de diagnósticos de DCV utilizando AM através de arquiteturas de Aprendizado Profundo (AP), com o intuito de contribuir, de maneira geral, na tomada de decisões de pesquisadores relacionados à área da saúde, médicos e gestores de saúde.

Portanto, o presente trabalho deseja contribuir com a precisão, eficiência e personalização do cuidado médico, proporcionando benefícios tanto para os profissionais da área da saúde, quanto para os pacientes, por meio da detecção automática de diagnósticos de DCV, auxiliando na tomada de decisão dos profissionais e possibilitando intervenções precoces e tratamentos adequados para as doenças cardiovasculares.

## **1.1 Objetivos**

Dessa forma, o objetivo geral deste trabalho é utilizar Aprendizado Profundo para realizar a predição de diagnósticos de doenças cardiovasculares, como normal, infarto do miocárdio, distúrbio de condução, alteração ST/T e hipertrofia, de forma automática.

Como objetivos específicos estão:

- Levantar métodos de aprendizado profundo que ainda foram pouco explorados para realizar a predição de diagnósticos de doenças cardiovasculares.
- Treinar modelos de aprendizado profundo para a predição de doenças cardiovasculares.
- Comparar o desempenho dos métodos que foram utilizados para realizar a predição de diagnósticos de doenças cardiovasculares.

O restante do trabalho está organizado da seguinte maneira. No Capítulo 2 são apresentados os conceitos necessários para o entendimento do trabalho. O Capítulo 3 apresenta trabalhos relacionados ao presente trabalho. Em seguida, no Capítulo 4 estão os procedimentos metodológicos a serem empregados para o desenvolvimento deste trabalho. No Capítulo 5 estão os resultados obtidos neste trabalho. Por fim, as considerações finais e trabalhos futuros são apresentados no Capítulo 6.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão apresentados os principais conceitos necessários para o entendimento e desenvolvimento do projeto proposto neste trabalho.

### 2.1 Doenças Cardiovasculares

As DCV consistem em um grupo de problemas que atingem o funcionamento do coração, o qual é responsável pela condução de oxigênio e nutrientes das células do organismo (BARBOSA *et al.*, 2020). As DCV configuram a maior causa de morte no mundo (WHO, 2023). Estima-se que 17,9 milhões de pessoas morreram devido a essas doenças em 2019, representando uma porcentagem de 32% de todas as mortes globais, sendo 85% dessas mortes causadas por ataque cardíaco e AVC (WHO, 2023).

Segundo Barbosa *et al.* (2020), há diversos fatores para que se desenvolva DCV, esses fatores podem ser classificados em modificáveis e não modificáveis. Os fatores modificáveis incluem a hiperlipidemia, o tabagismo, etilismo, hiperglicemia, obesidade, sedentarismo, má alimentação e uso de contraceptivos. Os fatores não modificáveis, por sua vez, consistem no histórico familiar de DCV, idade, sexo e a raça. Além disso, a Hipertensão Arterial Sistêmica (HAS) também se configura como outro fator de risco para o desenvolvimento dessas doenças, que atinge cerca de 60% dos idosos e é mais comum em homens negros. Este fato está relacionado a fatores socioeconômicos, como, por exemplo, a baixa renda, este sendo um fator que está diretamente ligado à adesão ao tratamento, tendo em vista que quanto menor a renda, mais difícil é o acesso à medicação, alimentação saudável, entre outros (BARBOSA *et al.*, 2020).

Nesse contexto, cabe também ressaltar que de acordo com Maceno e Garcia (2022), as DCV afetam cada vez mais jovens adultos, esse fato estando relacionado à prática de hábitos não saudáveis, como não realizar exercícios físicos, excesso de peso, obesidade, tabagismo, estresse, depressão e alimentação com alto teor de gordura e industrializados. Além disso, as pessoas tem adotado hábitos prejudiciais de vida desde a sua infância e adolescência, o que tem desencadeado doenças cardiovasculares em estágios iniciais de suas vidas (MACENO; GARCIA, 2022). É essencial realizar a detecção de doenças cardiovasculares o mais cedo possível para que o tratamento com aconselhamento e medicamentos possa ser iniciado (WHO, 2023).

### **2.1.1 Eletrocardiograma**

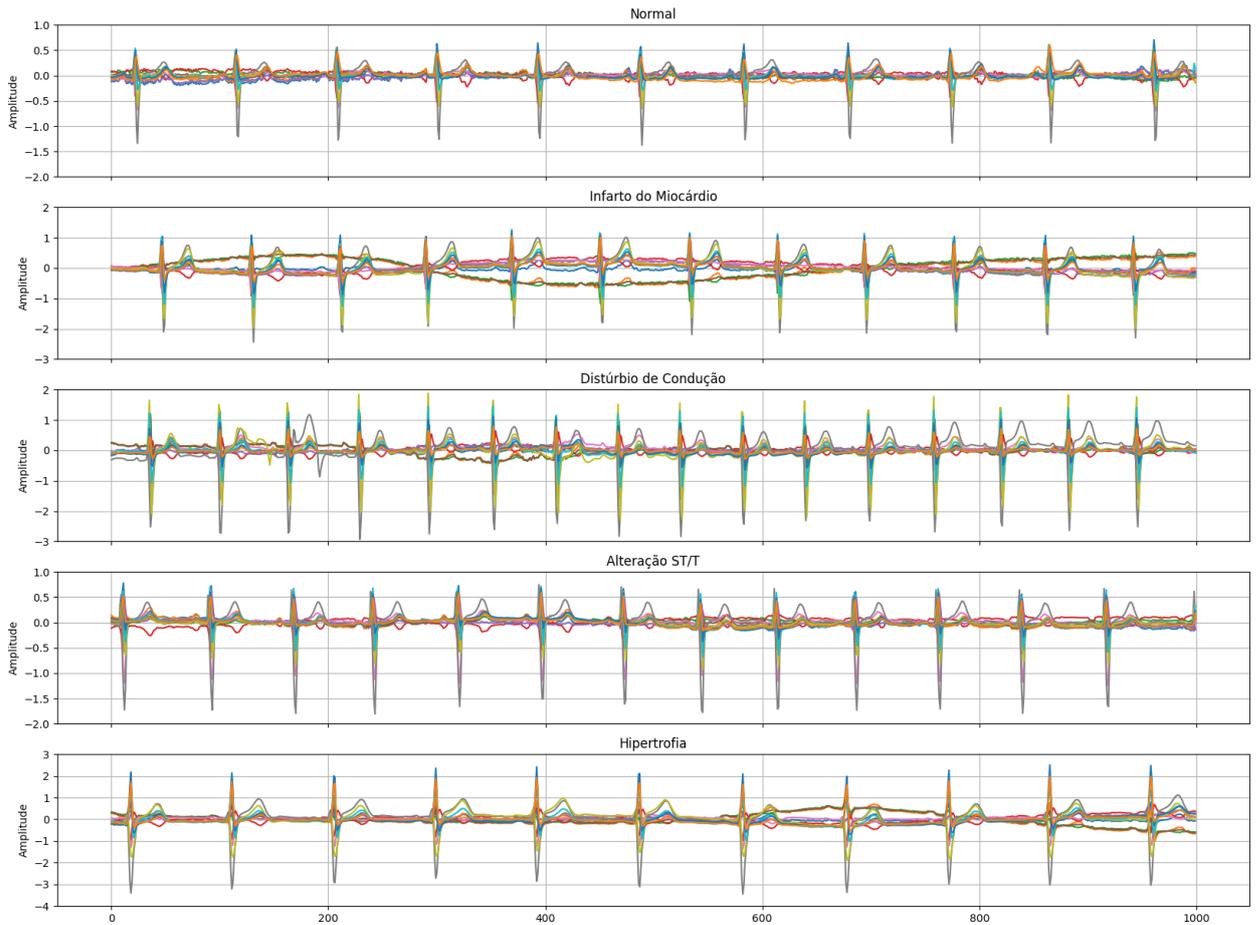
O ECG consiste no registro dos eventos elétricos originados no decurso da atividade cardíaca através de um aparelho que é capaz de medir pequenas intensidades de corrente, as quais são obtidas por meio de dois eletrodos, ou seja, placas de metal pequenas que são ligadas a um fio condutor, localizados em certos pontos do corpo humano (RAMOS; SOUSA, 2007).

A célula miocárdica em repouso é chamada de polarizada e possui uma grande concentração de potássio, apresentando-se negativa em relação ao meio externo, à medida que este possui grande concentração de sódio, além disso, ao ser despolarizada, a célula fica com sua polaridade invertida (RAMOS; SOUSA, 2007). Uma onda progressiva de despolarização pode ser observada como uma onda de cargas positivas em movimento (RAMOS; SOUSA, 2007). Quando uma onda positiva de despolarização se movimenta em direção a um eletrodo na pele, ou seja, um eletrodo positivo, é registrado no ECG uma deflexão positiva, isto é, para cima, em contrapartida, quando ela vai se distanciando do eletrodo, obtêm-se uma deflexão negativa e, por fim, quando não há atividade elétrica, a linha é isoeétrica (RAMOS; SOUSA, 2007)

As derivações, que são representadas por um eletrodo positivo e um eletrodo negativo, registram diferentes pontos de vista de uma mesma atividade cardíaca e, de acordo com a posição dos eletrodos exploradores, são formadas derivações periféricas, com eletrodos no plano frontal, e derivações precordiais, com eletrodos no plano horizontal (RAMOS; SOUSA, 2007). O sinal resultante da atividade elétrica cardíaca pode ser compreendido por meio da análise da série temporal que é formada por ele (FERREIRA, 2010).

Na análise do ECG, áreas essenciais para serem consideradas são frequência cardíaca, ritmo, eixo, hipertrofia e infarto (RAMOS; SOUSA, 2007). O ritmo sinusal é caracterizado pela apresentação de uma sequência de ciclos cardíacos com uma sucessão regular, ou seja, ritmada entre 60 e 100 bpm. Em contrapartida, existem as arritmias cardíacas, sendo exemplos delas arritmia sinusal, marcapasso migratório, fibrilação atrial, extra-sístole atrial, extra-sístole nodal, extra-sístole ventricular, parada sinusal, taquicardia paroxística, flutter atrial, flutter ventricular e bloqueio sinusal. Com relação a hipertrofia, tem-se como exemplo hipertrofia atrial direita/esquerda, e relacionado ao infarto, exemplos são isquemia, lesão e infarto do miocárdio (RAMOS; SOUSA, 2007). A Figura 1 ilustra exemplos de ECG com rótulos de superclasses da base de dados utilizada neste trabalho (Normal, Infarto do Miocárdio, Distúrbio de Condução, Alteração ST/T e Hipertrofia).

Figura 1 – Exemplos de Eletrocardiogramas para cada superclasse de diagnóstico



Fonte: Elaborado pela autora (2023)

### 2.1.2 Séries Temporais

De acordo com Ehlers (2007), uma série temporal é um conjunto de observações realizadas sequencialmente ao longo do tempo. A característica essencial desse tipo de dados é que as observações vizinhas são dependentes e o ponto de interesse é analisar e modelar esta dependência. Dados de séries temporais surgem em diversos campos do conhecimento, como na Medicina (eletrocardiograma, eletroencefalograma). O objetivo para se estudar séries temporais pode ser realizar previsões de valores futuros ou, ainda, pode se ter como o interesse principal conhecer a estrutura da série ou sua relação com outras séries (EHLERS, 2007).

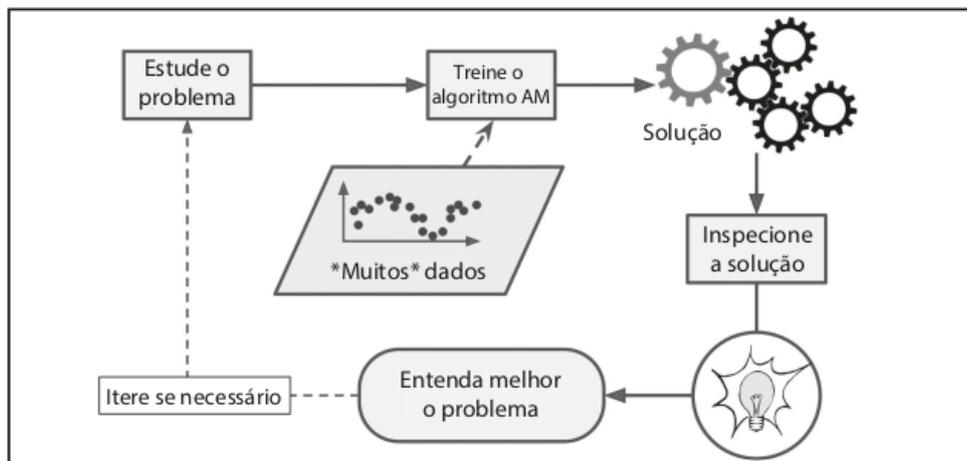
Dessa forma, toda vez que o tempo é importante em seus dados (ou a ideia de ordem de sequência), faz sentido o armazenar em um tensor 3D com um eixo de tempo explícito. Desse modo, toda amostra pode ser codificada como uma sequência de vetores, ou seja, um tensor 2D, e um lote de dados será codificado como um tensor 3D (CHOLLET, 2018). Um tensor é

um contêiner para dados, quase sempre esses dados são numéricos. De maneira semelhante as matrizes, as quais são tensores 2D. À vista disso, tensores são uma generalização de matrizes para uma quantidade arbitrária de dimensões (CHOLLET, 2018).

## 2.2 Aprendizado de Máquina

O AM é a ciência e, também, a arte da programação de computadores de forma a possibilitar que eles sejam capazes de aprender com os dados. Além disso, o AM pode auxiliar os seres humanos a aprender, visto que, algoritmos de AM podem ser inspecionados para que possa-se visualizar o que eles aprenderam (GÉRON, 2019). A Figura 2 ilustra esse processo, onde, ao observá-la, é possível notar e interpretar que o AM possibilita estudar o problema de forma mais detalhada, realizar o treinamento de modelos de AM para que eles aprendam com os dados e possam fornecer soluções que podem ser eficientes para resolver determinados problemas. Essas soluções devem ser avaliadas e após isso, podem proporcionar um melhor entendimento do problema.

Figura 2 – O Aprendizado de Máquina pode ajudar no ensino dos humanos



Fonte: Géron (2019)

### 2.2.1 Aprendizado supervisionado

Segundo Géron (2019), dividir os dados em dois conjuntos: o conjunto de treinamento e o conjunto de teste, ou seja, treinar seu modelo usando o conjunto de treinamento e o testar usando o conjunto de teste, é uma melhor opção do que colocar um modelo em produção e monitorar a qualidade do desempenho dele, sendo comum utilizar 80% dos dados

para treinamento e 20% dos dados para teste. Além disso, os sistemas de AM permitem que sua classificação seja feita de acordo com a quantidade e o tipo de supervisão que lhes é dada durante o treinamento (GÉRON, 2019).

Aprendizagem é uma investigação através do espaço de hipóteses com grandes chances de acontecer por aquele que terá um bom desempenho, mesmo em exemplos que não fazem parte do conjunto de treinamento (RUSSELL, 2013). Para calcular a precisão de uma hipótese, é fornecido um conjunto de teste de exemplos diferentes do conjunto de treinamento. Dizemos que uma hipótese generaliza bem se prevê de forma correta o valor de uma saída  $y$  para novos exemplos (RUSSELL, 2013).

Na aprendizagem supervisionada, o agente considera alguns exemplos de pares de entrada e saída, e aprende uma função que realiza o mapeamento de uma entrada para uma saída (RUSSELL, 2013). Assim, no aprendizado supervisionado, é fornecido ao algoritmo, por meio do conjunto de treinamento, as soluções desejadas, chamadas de rótulos (GÉRON, 2019). A Figura 3 ilustra esse processo.

Figura 3 – Um conjunto de treinamento rotulado para aprendizado supervisionado, por exemplo, classificação de spam



Fonte: Géron (2019)

Dessa forma, o aprendizado supervisionado tem como objetivo geralmente encontrar padrões nos dados e usá-los para realizar previsões (GÉRON, 2019). A classificação é uma tarefa típica desse tipo de aprendizado.

### 2.2.2 Classificação

Na tarefa de classificação, o programa de computador é solicitado a especificar a qual das categorias alguma entrada é pertencente (GOODFELLOW *et al.*, 2016). Há diferentes formas de categorizar a classificação, entre elas temos a classificação binária e a classificação multiclasse (RUSSELL, 2013). Na ocasião em que a saída  $y$  for de um conjunto finito de valores como,

por exemplo, ensolarado, nublado ou chuvoso, o problema de aprendizado será denominado de classificação, e será intitulado de classificação booleana ou binária se existir somente dois valores (RUSSELL, 2013). Enquanto os classificadores binários se discriminam entre duas classes, os classificadores multiclases, também denominados classificadores multinomiais, podem discriminar-se entre mais de duas classes (GÉRON, 2019).

### 2.2.3 Métricas de Avaliação

Os modelos de AM podem ser avaliados por diversas métricas, nesse trabalho, os modelos são avaliados pelas métricas precisão, revocação e *f1-score*. Para o melhor entendimento dessas métricas será, primeiramente, apresentado as definições de Verdadeiros Positivos (VP), Verdadeiros Negativos (VN), Falsos Positivos (FP) e Falsos Negativos (FN).

Nesse contexto, de acordo com Géron (2019), temos que:

- VP: classificado corretamente como a classe positiva.
- VN: classificado corretamente como a classe negativa.
- FP: classificado erroneamente como a classe positiva.
- FN: classificado erroneamente como a classe negativa.

#### 2.2.3.1 Precisão

A precisão do classificador é a acurácia das previsões positivas (GÉRON, 2019). O conceito de acurácia está relacionado à proporção de exemplos para os quais um modelo fornece a saída corretamente (GOODFELLOW *et al.*, 2016). A equação 2.1 mostra como é feito o cálculo da precisão.

$$\frac{VP}{VP + FP} \quad (2.1)$$

#### 2.2.3.2 Revocação

Também conhecida como sensibilidade ou taxa de verdadeiros positivos, revocação é a taxa de instâncias positivas que são detectadas de forma correta pelo classificador (GÉRON, 2019). A equação 2.2 mostra como é feito o cálculo da revocação.

$$\frac{VP}{VP + FN} \quad (2.2)$$

### 2.2.3.3 *F1-score*

Diversas vezes, é apropriado combinar precisão e revocação em uma única métrica denominada *f1-score*, essa métrica é a média harmônica da precisão e revocação. Ela pode ser útil, principalmente, se for preciso uma forma simples de comparar dois classificadores (GÉRON, 2019). A equação 2.3 mostra como é feito o cálculo da *f1-score*.

$$\frac{VP}{VP + \frac{FN+FP}{2}} \quad (2.3)$$

## 2.3 Redes Neurais e Aprendizado Profundo

Segundo Géron (2019), como a natureza inspirou diversas invenções, parece lógico, então, observar a arquitetura do cérebro para adquirir a inspiração sobre como criar uma máquina inteligente, esta sendo a ideia-chave que inspirou as redes neurais artificiais (RNA). Russell (2013) mostra resultados básicos da neurociência — em particular, a hipótese de que a atividade mental constitui-se basicamente na atividade eletroquímica em redes de células cerebrais que são denominadas de neurônios. Entretanto, as RNA gradualmente têm se tornado muito diferentes de seus primos biológicos (GÉRON, 2019).

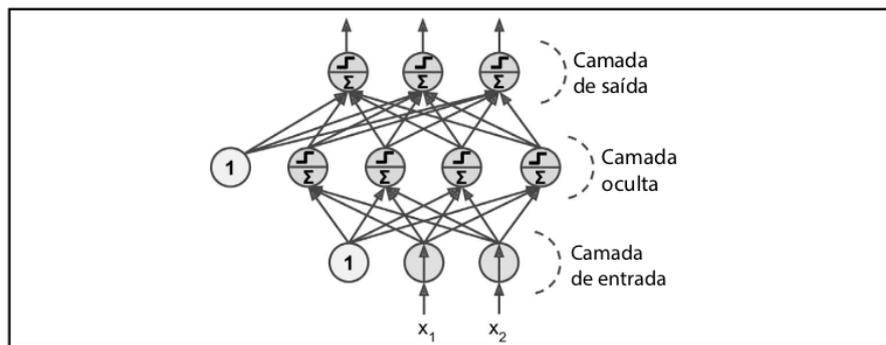
Uma parte dos trabalhos mais antigos de IA tiveram o objetivo de construir Redes Neurais Artificiais (RNA), outros nomes do campo que pode-se usar são conexionismo, processamento distribuído em paralelo e computação neural (RUSSELL, 2013). Um modelo simples de neurônio biológico, que tempos depois ficou conhecido como neurônio artificial, foi proposto por Warren McCulloch e Walter Pitts, esse neurônio contendo uma ou mais entradas binárias (ligado/desligado) e uma saída binária (GÉRON, 2019). McCulloch e Pitts mostraram que, embora com um modelo tão simplificado, é possível criar uma rede de neurônios artificiais capaz de calcular qualquer proposta lógica desejada (GÉRON, 2019). Uma rede neural é somente uma coleção de unidades conectadas; as propriedades da rede são determinadas por meio de sua topologia e propriedades dos “neurônios” (RUSSELL, 2013).

Segundo Géron (2019), as RNA estão no cerne do AP, elas são versáteis, poderosas e escaláveis, o que faz com que sejam uma opção ideal para lidar com grandes tarefas altamente complexas do AM, como classificar bilhões de imagens (por exemplo, as do *Google*), alimentar serviços de reconhecimento da fala (por exemplo, a Siri da *Apple*), fazer recomendações dos melhores vídeos para centenas de milhões de usuários (como o *YouTube*), ou aprender a derrotar

o campeão mundial no jogo Go, analisando milhões de jogos passados e depois jogando contra si mesmo (*AlphaGo* da *DeepMind*) (GÉRON, 2019).

O AP consiste em uma nova abordagem no aprendizado de representações a partir de dados que ressalta o aprendizado de camadas sucessivas de representações cada vez mais significativas (CHOLLET, 2018). A camada é o elemento central das redes neurais, ela é um módulo de processamento de dados que pode ser considerado como um filtro de dados. Especificamente, as camadas extraem representações dos dados que são fornecidos a elas – esperançosamente, representações que tenham maior relevância para o problema em questão (CHOLLET, 2018). Um modelo de AP pode ser comparado a uma peneira para o processamento de dados, composta de uma sucessão de filtros de dados cada vez mais refinados, ou seja, as camadas (CHOLLET, 2018). A arquitetura Perceptron Multicamadas é um exemplo de RNA que pode possuir muitas camadas, a Figura 4 ilustra essa arquitetura.

Figura 4 – Perceptron Multicamadas



Fonte: Géron (2019)

A Perceptron Multicamadas é composta de uma camada de entrada, uma ou mais camadas ocultas e uma camada final chamada de camada de saída (GÉRON, 2019). Por meio da Figura 4, é possível observar uma camada de entrada, que consiste na primeira camada da rede neural; uma camada oculta, que consiste na camada que está entre a camada de entrada e a camada de saída da rede neural; e uma camada de saída, que consiste na última camada da rede neural, a qual produz os resultados ou as previsões do modelo baseado nas operações realizadas nas camadas anteriores. De acordo com Géron (2019), no caso em que uma RNA possui duas ou mais camadas ocultas, ela é denominada de rede neural profunda.

### 2.3.1 Redes Neurais Recorrentes

As informações contidas nesta seção são derivadas de Géron (2019).

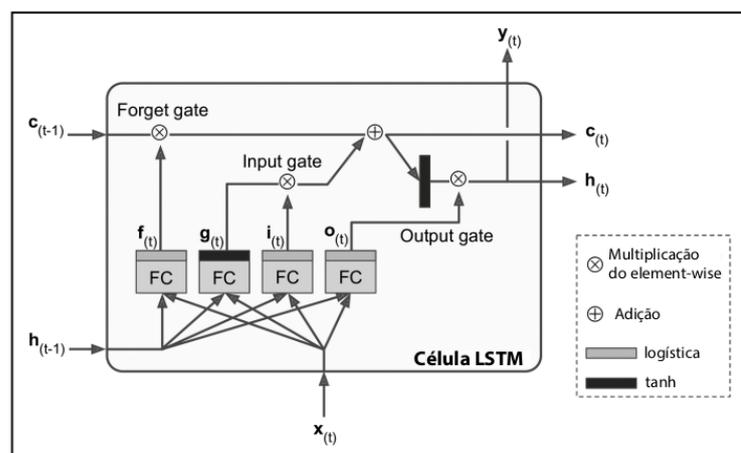
Prever o futuro é algo comum em nosso dia a dia, seja terminando a frase de um amigo seja antecipando o cheiro do café da manhã. As Redes Neurais Recorrentes são uma classe de redes capaz de prever o futuro (isso até certo ponto). Essas redes neurais conseguem analisar dados de séries temporais, como, por exemplo, preços de ações, e informar quando comprar ou vender, além disso, são capazes de antecipar trajetórias de carros, podendo, assim, auxiliar na prevenção de acidentes em sistemas autônomos de direção. Elas são extremamente úteis para sistemas de Processamento de Linguagem Natural (PLN) como a tradução automática, voz para texto ou análise comportamental (por exemplo, lendo resenhas de filmes e extraindo a impressão do avaliador quanto ao filme), visto que, podem trabalhar em sequências de comprimentos arbitrários.

### 2.3.1.1 LSTM

As informações contidas nesta seção, assim como na anterior, são derivadas de Géron (2019).

A célula Memória Longa de Curto Prazo (LSTM, do inglês *Long Short-Term Memory*) é uma célula com memória de longo prazo que tem como ideia-chave que a rede possa aprender o que armazenar no estado de longo prazo, o que jogar fora e o que ler a partir dele. Elas têm sido bem-sucedidas em identificar padrões de longo prazo em séries temporais, textos extensos, gravações de áudio e em muitas outras tarefas. A arquitetura de uma célula LSTM é mostrada na Figura 5.

Figura 5 – Célula LSTM



Fonte: Géron (2019)

O estado da célula LSTM é dividido em dois vetores:  $h_{(t)}$  e  $c_{(t)}$  (“c” se refere a

“célula”), onde  $h_{(t)}$  é o estado de curto prazo e  $c_{(t)}$  é o estado de longo prazo. Além disso, pela Figura 5, é possível observar o vetor de entrada atual  $x_{(t)}$  e o estado de curto prazo anterior  $h_{(t-1)}$ , eles são fornecidos para quatro camadas diferentes totalmente conectadas, sendo elas a camada principal e outras três camadas *gate controllers*, denominadas *forget gate*, *input gate* e *output gate*.

A camada principal é responsável por produzir  $g_{(t)}$  e, geralmente, tem o papel de analisar  $x_{(t)}$  e o  $h_{(t-1)}$ . Na célula LSTM, a saída dessa camada não se extingue diretamente, ela é armazenada parcialmente no estado de longo prazo. Nas camadas *gate controllers* as saídas variam de 0 a 1 e, como é mostrado na Figura 5, as saídas são fornecidas para operações de multiplicação elemento por elemento, logo, se elas produzem 0s, fecham o *gate* e, se produzem 1s, abrem os *gates*. O *forget gate* é responsável por controlar quais partes do estado de longo prazo devem ser esquecidas; o *input gate* é responsável por controlar quais partes de  $g_{(t)}$  devem ser inseridas ao estado de longo prazo; por fim, o *output gate* é responsável por controlar quais partes do estado de longo prazo devem ser lidas e exibidas neste intervalo de tempo (ambos para  $h_{(t)}$  e  $y_{(t)}$ , onde  $y_{(t)}$  é a saída da célula).

Portanto, uma célula LSTM pode aprender a reconhecer uma entrada relevante, este sendo o papel do *input gate*, armazenar essa entrada no estado de longo prazo, aprender a preservar a entrada pelo tempo que for preciso, este sendo o papel do *forget gate*, e aprender a extraí-la sempre que for necessário.

### 2.3.2 Transformer

*Transformer* é uma arquitetura de rede simples baseada exclusivamente em mecanismos de atenção, dispensando o uso de Redes Neurais Recorrentes ou Redes Neurais Convolucionais (VASWANI *et al.*, 2017). Da mesma forma que as Redes Neurais Recorrentes, os *Transformers* são projetados para lidar com dados de entrada sequenciais, como linguagem natural, para tarefas como tradução e resumo de texto. Entretanto, os *Transformers* não exigem que os dados sequenciais sejam processados de maneira ordenada, visto que, a operação de atenção identifica o contexto para qualquer posição na sequência de entrada (ACADEMY, 2022). Além disso, os *Transformers* demonstraram capacidade elevada de modelagem para dependências e interações de longo alcance em dados sequenciais e, portanto, são atraentes para a modelagem de séries temporais (WEN *et al.*, 2022).

### 2.3.2.1 Arquitetura do Modelo

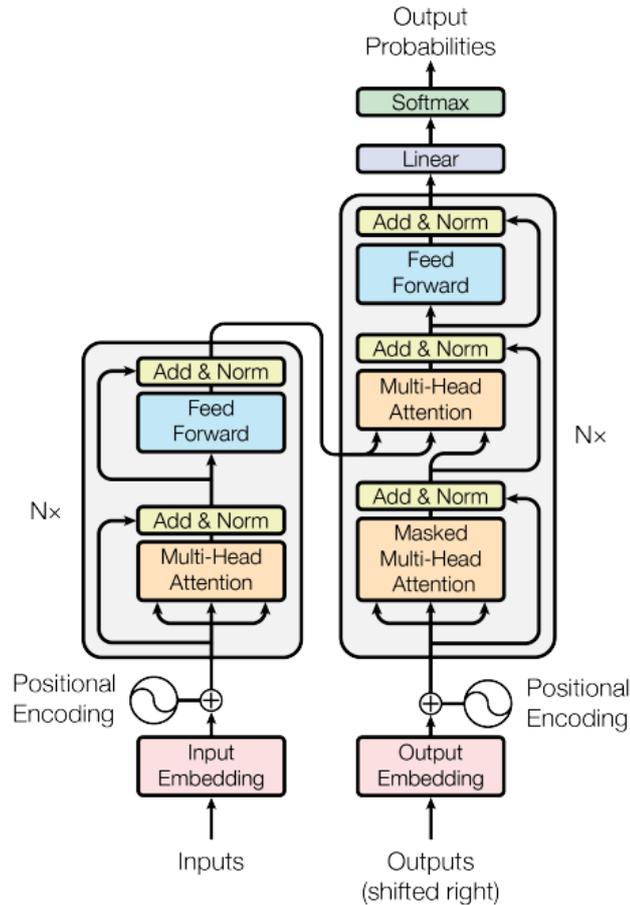
O modelo proposto por Vaswani *et al.* (2017) possui uma estrutura codificador-decodificador, onde o codificador mapeia uma sequência de entrada de representações de símbolos  $(x_1, \dots, x_n)$  para uma sequência de representações contínuas  $z = (z_1, \dots, z_n)$ . Dado a sequência  $z$ , o decodificador produz uma sequência de saída  $(y_1, \dots, y_m)$  de símbolos um elemento de cada vez (VASWANI *et al.*, 2017). O *Transformer*, em cada passo, é um modelo auto-regressivo e consome os símbolos gerados anteriormente como entrada adicional ao gerar o próximo (VASWANI *et al.*, 2017). No modelo são utilizados *embeddings* aprendidos para converter unidades individuais que compõem uma sequência de texto de entrada e as unidades individuais que compõem uma sequência de texto de saída em vetores de dimensão  $d_{modelo}$  (VASWANI *et al.*, 2017). Segundo Géron (2019), para chegar a uma representação significativa, a solução mais frequente é representar cada palavra no vocabulário com o uso de um vetor razoavelmente pequeno e denso, denominado *embedding*. O *Transformer* segue uma arquitetura geral usando camadas empilhadas de autoatenção e camadas pontuais totalmente conectadas para o codificador e o decodificador, que estão ilustradas nas metades esquerda e direita da Figura 6 (VASWANI *et al.*, 2017).

### 2.3.2.2 Mecanismo de Atenção

Pode-se descrever uma função de atenção como o mapeamento de uma consulta e um conjunto de pares chave-valor para uma saída, em que a consulta, as chaves, os valores e a saída são vetores (VASWANI *et al.*, 2017). Para calcular a saída é realizada uma soma ponderada dos valores, em que o peso dado a cada valor é calculado por meio de uma função de compatibilidade da consulta com a chave correspondente (VASWANI *et al.*, 2017). Para a Atenção *Multi-head*, em vez de executar somente uma função de atenção foram projetadas linearmente as consultas, chaves e valores  $h$  vezes com diferentes projeções lineares (VASWANI *et al.*, 2017). Em cada uma dessas versões projetadas de consultas, chaves e valores, foi realizada a função de atenção em paralelo, gerando valores de saída dimensionais, esses valores são concatenados e projetados novamente, resultando nos valores finais (VASWANI *et al.*, 2017). De acordo com Wen *et al.* (2022), o módulo de autoatenção é essencial para o *Transformer* e ele pode ser visto como uma camada totalmente conectada com pesos que são gerados dinamicamente com base na semelhança entre pares de padrões inseridos. Como resultado, ele compartilha

o mesmo comprimento máximo de caminho que as camadas totalmente conectadas, mas com um número muito menor de parâmetros, o que o torna adequado para modelar dependências de longo prazo (WEN *et al.*, 2022).

Figura 6 – Arquitetura do modelo *Transformer*



Fonte: Vaswani *et al.* (2017)

### 2.3.2.2.1 Pilhas de Codificador e Decodificador

No trabalho de Vaswani *et al.* (2017), o codificador é constituído por uma pilha de  $N = 6$  camadas idênticas, onde cada camada possui duas subcamadas. A primeira é um mecanismo de auto-atenção *Multi-Head*, e a segunda é uma rede *feed-forward* - a informação se movimenta da entrada para a saída e não possui um *loop* de *feedback* - inteiramente conectada, que é aplicada a cada posição separadamente e de forma idêntica (VASWANI *et al.*, 2017). É empregada uma conexão residual em torno de cada uma das duas subcamadas, seguida de normalização de camada. Para diminuir a complexidade dessas conexões residuais, todas as subcamadas do modelo e as camadas de *embedding*, fornecem saídas de dimensão  $d_{modelo} = 512$

(VASWANI *et al.*, 2017).

O decodificador também é constituído por uma pilha de  $N = 6$  camadas idênticas. Ele é semelhante ao codificador em relação às subcamadas, porém insere uma terceira subcamada, que faz a atenção *Multi-Head* sobre a saída da pilha do codificador (VASWANI *et al.*, 2017). Assim como o codificador, são empregadas conexões residuais em torno de cada uma das subcamadas, seguidas de normalização de camada. Além disso, a subcamada de autoatenção na pilha do decodificador é modificada para evitar que as posições atendam às posições subsequentes (VASWANI *et al.*, 2017). Este mascaramento, combinado com o fato de que os *embeddings* de saída são compensados por uma posição, faz com que as previsões para a posição  $i$  dependam somente das saídas conhecidas em posições inferiores a  $i$  (VASWANI *et al.*, 2017). Além disso, a transformação linear aprendida usual e a função *softmax* são utilizadas para converter a saída do decodificador em probabilidades previstas do próximo símbolo (VASWANI *et al.*, 2017).

Dessa forma, o codificador consiste em um conjunto de camadas de codificação capazes de processar a entrada iterativamente uma camada depois da outra e o decodificador consiste em um conjunto de camadas de decodificação que realizam o mesmo processo com a saída do codificador (ACADEMY, 2022).

### 3 TRABALHOS RELACIONADOS

Neste capítulo, serão apresentados trabalhos relacionados com o projeto proposto neste trabalho.

#### 3.1 *Deep Learning for ECG Classification*

Em Pyakillya *et al.* (2017) é ressaltado que a maioria dos problemas cardíacos podem ser diagnosticados por meio de sinais de ECG com uma estimativa de centenas de milhões de ECG registrados por ano. Além disso, a classificação do ECG tem uma alta importância agora, por causa das várias aplicações médicas atuais onde o problema pode ser declarado. Os autores afirmam que existem muitas soluções de AM que possibilitam a análise e classificação de dados de ECG.

Entretanto, são citadas como principais desvantagens dos resultados de AM o uso de recursos heurísticos feitos a mão ou projetados com arquiteturas de aprendizado que apresentam recursos superficiais. Como solução, eles apoiam o uso de arquiteturas de AP, onde as primeiras camadas de neurônios convolucionais funcionam como extratores de recursos e no final camadas totalmente conectadas são utilizadas para tomar a decisão em relação ao ECG.

Em seu trabalho, Pyakillya *et al.* (2017) apresenta a arquitetura de AP com camadas convolucionais 1D e camadas totalmente conectadas (FCN) para classificação de ECG. Os autores descrevem a tarefa que será realizada como uma tarefa para decidir a qual classe o ECG do paciente pode ser atribuído, onde o número de classes são quatro: ritmo sinusal normal, arritmico, outro tipo de ritmo e muito ruidoso.

Além disso, os dados não são balanceados, sendo 60% dos dados pertencente a classe ritmo sinusal normal. Para balancear os dados, foi utilizado operações como a multiplicação do ECG existente para alguma classe através de deslocamento de valores de tempo e uma tentativa de criar algum comprimento unificado de ECG através de valores de séries temporais de duplicação.

A arquitetura da rede neural convolucional 1D é constituída por 7 camadas convolucionais com largura de filtro 5 e 128 neurônios, *max-pooling* e *dropout* após cada camada, *GlobalAveragePooling* e 3 camadas FCN com (256/128/64) neurônios, *dropout* após cada camada e camada *softmax* com 4 saídas. Além disso, para o pré-processamento foram utilizados procedimentos como subtração da média dos valores das séries temporais e divisão pelo valor do

desvio padrão com tamanho do lote igual a 256 e como acurácia dos melhores resultados nos dados de validação foi obtido o valor de, aproximadamente, 86%.

A principal diferença entre o trabalho de Pyakillya *et al.* (2017) e este trabalho, é que neste são utilizadas as arquiteturas de AP *Transformer* e LSTM para realizar a classificação de diagnósticos de ECG, enquanto que Pyakillya *et al.* (2017) utiliza Redes Neurais Convolucionais (CNN).

### **3.2 *Towards Understanding ECG Rhythm Classification Using Convolutional Neural Networks and Attention Mappings***

Goodfellow *et al.* (2018) afirma que muitos pesquisadores utilizaram diversas técnicas de AM para classificação/deteção de ritmo cardíaco a partir de formas de onda de ECG usando uma das duas estratégias de aprendizado de máquina.

A primeira estratégia é o AM passo a passo, que é uma abordagem mais comum e envolve processamento de sinal, recursos de engenharia manual e classificação. A segunda abordagem é o AP de ponta a ponta, o qual usa redes neurais profundas, que utilizam formas de onda de ECG brutas como entrada e saída de uma classificação, dessa forma, eliminando a necessidade de mão de engenharia de atributos.

O estudo feito por Goodfellow *et al.* (2018) extrai as formulações derivadas para mapeamento de atenção de rede neural convolucional de imagens e as aplica à classificação de séries temporais para poder observar a região de uma forma de onda de ECG que recebe mais atenção durante a classificação do ritmo.

O objetivo do trabalho foi construir um classificador de ritmo de ECG de CNN que tenha um desempenho comparável ao da apresentação superior do Desafio *Physionet* 2017 e contenha uma camada de *pool* de média global antes do classificador *softmax*, possibilitando que mapeamentos de ativação de classe, ou seja, mapeamentos de atenção, sejam extraídos. Para isso, é treinada uma rede neural convolucional profunda para realizar a classificação de formas de onda de ECG de derivação única em três classes: Ritmo Sinusal Normal, Fibrilação Atrial e Outro Ritmo. O conjunto de dados usado é do Desafio *Physionet* de 2017 composto por 12.186 formas de onda rotuladas doadas pela *AliveCor*, sendo retirada a classe Ruído deste conjunto de dados para realizar o estudo.

O pré-processamento foi feito da seguinte maneira: as formas de onda completas primeiramente foram filtradas utilizando um filtro passa-banda de resposta ao impulso finito

com um limite inferior de 3 Hz e um limite superior de 45 Hz. Após isso, os picos R foram escolhidos usando o algoritmo de Hamilton-Tompkins que retornou uma matriz de tempos de pico R escolhidos como linhas verticais cinzas. Por fim, cada forma de onda de ECG foi normalizada para a amplitude mediana do pico R. O modelo apresentou as pontuações médias a seguir, em todas as classes de ritmo, no conjunto de dados de validação: precisão igual a 0,84; revocação igual a 0,85;  $F_1$  igual a 0,84 e acurácia igual a 0,88.

A principal diferença entre o trabalho de Goodfellow *et al.* (2018) e este trabalho está, mais uma vez, na arquitetura de AP utilizada para realizar a classificação de ECG, onde Goodfellow *et al.* (2018), faz uso de rede neural convolucional profunda.

### 3.3 Aprendizado de Máquina para Predição de Diagnósticos de Doenças Cardiovasculares

Em Filho e Coutinho (2022), as DCV são definidas como adversidades que afetam o coração e os vasos sanguíneos. Além disso, trazem dados que estimam que em 2019, as DCV tenham sido a causa da morte de 17,9 milhões de pessoas, o que representa 32% das mortes daquele ano. No trabalho é destacado que com a influência maciça de dados disponíveis em várias formas, o papel da análise de dados em informática médica tem sido elevado rapidamente na última década, o que também acarretou um interesse progressivo na geração de modelos analíticos orientados por dados com base em AM. Além disso, é ressaltado que a classificação usando AM desempenha um papel relevante na previsão de doenças cardíacas e na investigação de dados, para amenizar os impactos no coração e combater uma possível morte prematura.

O objetivo do trabalho é treinar e avaliar modelos de AM e AP para predição de diagnósticos de doenças cardiovasculares. Para isso, o trabalho utiliza algoritmos que se enquadram como: AM tradicional - possui um único modelo preditivo - onde os algoritmos escolhidos previamente e que serão usados são regressão logística e árvore de decisão; *Ensemble Learning* - combina múltiplos modelos preditivos - onde os modelos utilizados foram floresta aleatória e *XGBoost*, especificamente *Gradient Boosting*; e redes neurais de AP, onde foi implementada a rede neural do tipo LSTM.

Os resultados do trabalho mostram que o melhor modelo para a classificação de 5 classes foi a rede neural LSTM apresentando 77% de acurácia, 43% de precisão, 68% de revocação e 52% de *F1-score*. Considerando a sua superioridade, somente a rede neural foi usada para realizar a predição das 23 subclasses de diagnósticos e os 44 diagnósticos, tendo o desempenho de 94% de acurácia, 21% de precisão, 70% de revocação e 30% de *F1-score* na

classificação das 23 subclasses. O desempenho da rede neural LSTM na classificação dos 44 diagnósticos foi 95% de acurácia, 35% de precisão, 51% de revocação e 39% de *F1-score*. Dessa forma, o trabalho realça como contribuições o uso de redes neurais LSTM para séries temporais ao avaliar e comparar com outras abordagens.

O presente trabalho, assim como em Filho e Coutinho (2022), tem o intuito de realizar um estudo no domínio das DCV utilizando a base de dados PTB-XL. Contudo, usando modelos de AP *Transformer* para predição de diagnósticos dessas doenças para verificar o seu desempenho em comparação a rede neural LSTM.

### **3.4 *Deep Learning Approach for Detection of Atrial Fibrillation and Atrial Flutter Based on ECG Images***

O estudo realizado por Ribeiro *et al.* (2023) explora a aplicação de técnicas de AP com base em imagens para diferenciar entre Fibrilação Atrial (AFib) e Flutter Atrial (AFlut), para isso são utilizadas imagens de exames padrão ECG de 12 derivações de um banco de dados privado coletados de 2017 a 2020 do *Picture Archiving and Communication System* (PACS) de um hospital terciário de referência no Brasil especializado em cardiologia (*Heart Institute Hospital*).

A definição dada em Ribeiro *et al.* (2023) dessas doenças é que AFib e AFlut são doenças cardiovasculares que tem como característica ritmos cardíacos irregulares dos átrios. Além disso, pessoas idosas são mais propensas a serem afetadas por elas. É destacado que o desenvolvimento e a validação dos métodos de detecção automática de AFib e AFlut são capazes de auxiliar na rapidez do seu diagnóstico para fornecer o tratamento necessário.

Para realizar a avaliação do desempenho da classificação de AFib e AFlut baseada em imagens, foi usada uma rede neural convolucional bidimensional tradicional e amplamente utilizada chamada *MobileNet*. Sua arquitetura é a seguinte: as camadas totalmente conectadas eram um perceptron personalizado de 3 camadas com uma taxa de *dropout* de 30%, uma função de ativação ReLU nas camadas intermediárias e uma função sigmóide na última camada.

Para avaliação dos modelos empregados, foram abordadas cinco métricas estatísticas, incluindo Sensibilidade (Se), Especificidade (Spe), F1-score (F1), Área sob Curva de Recebimento Operacional (AUROC) e Acurácia (Acc). Além disso, foram selecionados dois métodos para auxiliar no entendimento do modelo. O método *Gradient-weighted Class Activation* (Grad-CAM) e o método *Local Interpretable Model-Agnostic Explanations* (LIME), que permitiram

identificar regiões significativas nas imagens de ECG, que colaboraram para a classificação. O modelo conseguiu um desempenho de Acc 95%, AUROC 97.6%, F1 83.2%, Spe 99.6% e Se 72.7%, com AFlut como classe positiva.

Os autores ressaltam que, embora os resultados sejam promissores, existem algumas limitações no estudo, como o conjunto de dados ter sido adquirido de um único hospital, destacando que este fato pode limitar a generalização dos resultados para outras populações e ambientes. Além disso, há também a questão que o modelo pode não ser tão assertivo para outros tipos de arritmias ou na presença de sinais de ECG ruidosos. Suas pesquisas futuras devem focar na expansão do conjunto de dados para incluir uma população de pacientes mais diversificada e na investigação do desempenho do modelo para diversos tipos de arritmias.

O presente trabalho, assim como em Ribeiro *et al.* (2023), se dedica a realizar um estudo no domínio das DCV utilizando AP. Entretanto, usando os modelos de AP *Transformer* e LSTM, para compará-los em termos de predição de diagnósticos das doenças cardiovasculares, diferentemente de Ribeiro *et al.* (2023), que faz uso de uma rede neural convolucional em seu estudo.

### 3.5 Análise Comparativa

Assim como realizado em Pyakillya *et al.* (2017), Goodfellow *et al.* (2018) e Ribeiro *et al.* (2023), este trabalho pretende realizar a classificação de ECG no domínio de doenças cardiovasculares utilizando aprendizado profundo. Contudo, esses trabalhos utilizam arquitetura CNN, diferentemente do presente trabalho, que utiliza as arquiteturas *Transformer* e LSTM. Além disso, a base de dados utilizada neste trabalho se difere das que foram utilizadas por Goodfellow *et al.* (2018) e por Ribeiro *et al.* (2023), visto que, Goodfellow *et al.* (2018) usa a base de dados do Desafio *Physionet* de 2017 e Ribeiro *et al.* (2023) utiliza a base de dados privados coletados do PACS.

A relação entre este trabalho e Filho e Coutinho (2022) acontece pelo fato de que será utilizado a mesma base de dados, ou seja, *PTB-XL* para realizar as predições de diagnósticos de doenças cardiovasculares. Entretanto, Filho e Coutinho (2022) faz uso das três abordagens da literatura descritas para classificação de eventos com AM, enquanto este trabalho só utiliza AP.

Na Tabela 1, é apresentado um resumo das características encontradas nos trabalhos relacionados descritos, anteriormente, e a comparação com o projeto apresentado neste trabalho.

Tabela 1 – Análise comparativa entre os trabalhos relacionados e este trabalho

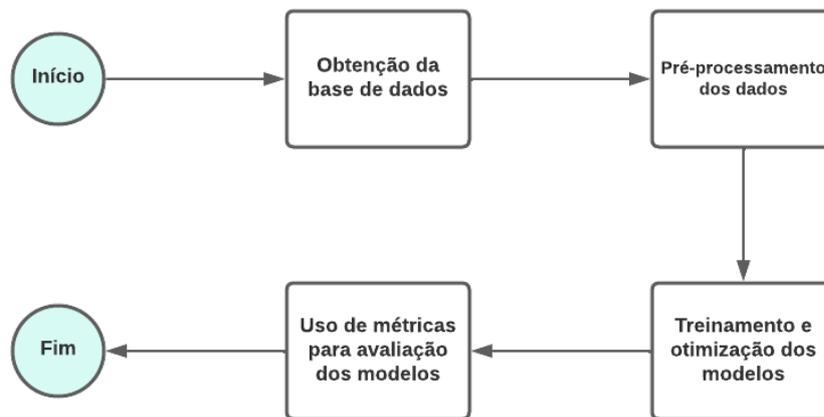
<b>Trabalho</b>	<b>Base de dados</b>	<b>Técnicas</b>	<b>Métricas</b>
Pyakillya <i>et al.</i> (2017)	Não informa	CNN	Acurácia
Goodfellow <i>et al.</i> (2018)	Desafio Physionet de 2017	CNN	Precisão, revocação, F1, acurácia
Filho e Coutinho (2022)	PTB-XL	LR, DT, RF, XGBoost e LSTM	Acurácia, precisão, revocação, f1-score
Ribeiro <i>et al.</i> (2023)	Dados privados coletados do PACS	CNN	Sensibilidade, Especificidade, f1-score, AUROC, acurácia
Este trabalho	PTB-XL	LSTM e Transformer	precisão, revocação, f1-score

Fonte: Elaborado pela autora (2023)

## 4 PROCEDIMENTOS METODOLÓGICOS

Para alcançar os objetivos propostos neste trabalho, se faz necessário um conjunto de etapas a serem seguidas, conforme pode ser observado no fluxograma mostrado na Figura 7.

Figura 7 – Fluxograma de procedimentos metodológicos



### 4.1 Obtenção da base de dados

A primeira etapa para realizar o desenvolvimento deste trabalho consiste em realizar a pesquisa e a escolha de uma base de dados no domínio de doenças cardiovasculares. A base de dados que será utilizada é a PTB-XL<sup>1</sup>. Essa base de dados consiste em um conjunto de dados de ECG clínico de tamanho sem precedentes com modificações propostas para avaliar algoritmos de AM (WAGNER *et al.*, 2020). Além disso, o conjunto de dados possui diversidade, tanto em termos de qualidade de sinal (com 77,01% de alta qualidade de sinal), quanto em uma ampla cobertura de patologias, incluindo muitas ocorrências de diferentes doenças, entretanto, também inclui uma grande proporção de amostras de controle saudáveis, uma característica raramente encontrada em conjuntos de dados clínicos (WAGNER *et al.*, 2020). A diversidade presente no PTB-XL o destaca como uma fonte valiosa para treino e avaliação de algoritmos em cenários do mundo real, no qual os algoritmos de AM tem de funcionar de forma confiável, independente das condições de gravação ou da possível presença de dados de má qualidade. Para a classificação há 44 diagnósticos, 24 subclasses e 5 superclasses (WAGNER *et al.*, 2020). Nesta base de dados, será realizada uma classificação multiclasse das 5 superclasses utilizando aprendizado

<sup>1</sup> PTB-XL, a large publicly available electrocardiography dataset - <https://physionet.org/content/ptb-xl/1.0.1/>

supervisionado.

## 4.2 Pré-processamento dos dados

Como segunda etapa do processo, tem-se o pré-processamento dos dados. Esta etapa é fundamental, tendo em vista que afeta diretamente o desempenho e a eficácia dos modelos. De uma forma geral, no pré-processamento pode ser realizada a *limpeza dos dados*, com o intuito de eliminar informações ausentes, inconsistentes ou incorretas. Neste trabalho, é realizada uma limpeza nos dados com o objetivo de eliminar valores ausentes do conjunto de dados. Além disso, se faz necessário realizar a codificação de variáveis categóricas nas superclasses de diagnósticos, onde estas variáveis são convertidas em um formato numérico adequado ao esperado por cada modelo.

## 4.3 Treinamento e otimização dos modelos

Na terceira etapa, levando em consideração o que foi apresentado no Capítulo 2, o presente trabalho utilizará modelos de AP por meio das arquiteturas LSTM e *Transformer*. Na construção dos modelos foi utilizado o *Keras*<sup>2</sup>, que é uma API de alto nível construída sobre a biblioteca *TensorFlow*<sup>3</sup>. Essa biblioteca apresenta bons padrões de redes neurais e o funcionamento é eficiente sem que se tenha que gastar muito tempo com ajustes dos parâmetros de forma manual (CHOLLET, 2017).

Para a etapa de treinamento, os dados do conjunto de treinamento serão separados em dois conjuntos, um contendo os dados de treino propriamente ditos e o outro será um conjunto de validação. Além disso, será escolhida uma métrica para avaliar durante o treinamento e obter o melhor modelo levando em consideração esta métrica. Dessa forma, os melhores modelos serão utilizados para realizar as predições no conjunto de teste.

## 4.4 Avaliação dos modelos

Por fim, os dados de teste serão submetidos ao modelo que obteve a melhor performance para a predição. Após isso, o modelo será avaliado por meio das métricas de avaliação para classificação apresentadas na seção 2.2.3 deste trabalho, ou seja, precisão, revocação e

---

<sup>2</sup> Keras - <https://keras.io/api/>

<sup>3</sup> TensorFlow - <https://www.tensorflow.org/>

*f1-score*. O desempenho avaliado de cada modelo por meio das métricas será apresentado através de tabelas. Outro ponto que contribuirá para a avaliação dos modelos é um valor de limiar denominado, neste trabalho, como *threshold*, que tem como objetivo observar os resultados obtidos pelas métricas em diferentes pontos de decisão, funcionando como um valor crítico que determina a classe final. Os valores de *threshold* escolhidos foram 0.3, 0.5 e 0.75, tendo em vista que as previsões dos modelos que serão avaliadas pelas métricas são vetores de probabilidades, diferentes valores de *threshold* oferecem impactos que permitem diferentes observações sobre cada métrica e sobre o desempenho dos modelos.

## 5 RESULTADOS E DISCUSSÕES

Neste capítulo, serão apresentados os experimentos realizados e seus resultados, assim como as discussões a cerca desses resultados.

### 5.1 Obtenção da base de dados

#### 5.1.1 Base de dados PTB-XL

As informações fornecidas nesta seção são derivadas do trabalho apresentado em Wagner *et al.* (2020) e do trabalho apresentado em Filho (2022).

Neste trabalho, serão utilizados os dados da base de dados PTB-XL. Esta base de dados é disponibilizada por um projeto do instituto nacional de metrologia da Alemanha, o *Physikalisch-Technische Bundesanstalt* (PTB). O conjunto de dados inclui registros clínicos de ECG de 12 derivações, cada um com uma duração de 10 segundos, provenientes de 18.885 pacientes. A distribuição no conjunto de dados é equilibrada em termos de sexo (52% masculino e 48% feminino). Além disso, abrange uma ampla faixa etária, variando de 0 a 95 anos, com uma mediana de 62 anos e um intervalo interquartil de 22 anos. Os registros de ECG foram anotados por até dois cardiologistas, potencialmente com múltiplas declarações de ECG provenientes de um conjunto de 71 declarações distintas, em conformidade com o padrão SCP-ECG. Estas declarações englobam declarações de forma, ritmo e diagnóstico em um formato unificado e legível por máquina.

Além disso, o conjunto de dados apresenta uma organização hierárquica composta por 5 superclasses e 24 subclasses em relação aos rótulos de diagnósticos. As superclasses são Normal (NORM), Infarto do Miocárdio (MI, do inglês *Myocardial Infarction*), Distúrbio de Condução (CD, do inglês *Conduction Disturbance*), Alteração ST/T (STTC, do inglês *ST/T-Change*) e Hipertrofia (HYP, do inglês *Hypertrophy*). A Tabela 2 mostra informações gerais da base de dados PTB-XL.

Tabela 2 – Informações sobre a base de dados PTB-XL

<b>Formato</b>	Séries temporais
<b>Tamanho</b>	21.837 ECGs clínicos de 12 derivações
<b>Superclasses</b>	NORM, MI, CD, STTC, HYP
<b>Frequência de amostragem</b>	500 Hz e 100 Hz

Fonte: Adaptado de Filho (2022)

Além de anotações na forma de declarações de ECG, acompanhadas por informações de probabilidade para declarações de diagnóstico, há também metadados adicionais, por exemplo, na forma de declarações de qualidade de sinal anotadas de forma manual. Os metadados correspondentes foram registrados em um banco de dados por uma enfermeira. Na Tabela 3 é ilustrado o resumo dos metadados do conjunto de dados e na Tabela 4 é ilustrada a organização hierárquica dos diagnósticos e a descrição de cada um.

Tabela 3 – Descrição dos atributos da base de dados

Seção	Variável	Tipo de dado	Descrição
Identificadores	ecg_id	Inteiro	Identificador único do Eletrocardiograma (ECG)
	patient_id	Inteiro	Identificador único do paciente
	filename_lr	String	Caminho para os dados em forma de onda (100 Hz)
	filename_hr	String	Caminho para os dados em forma de onda (500 Hz)
Metadados gerais	age	Inteiro	Idade de gravação em anos
	sex	Categórico	Sexo (0 para Masculino, 1 para Feminino)
	height	Inteiro	Altura em centímetros
	weight	Inteiro	Peso em quilogramas
	nurse	Categórico	Enfermeira envolvida (pseudônimo)
	site	Categórico	Local da gravação (pseudônimo)
	device	Categórico	Dispositivo de gravação
	recording_date	Tipo data	Data e hora de gravação de ECG
Diagnósticos de ECG	report	String	Relatório de ECG do cardiologista de diagnóstico
	scp_codes	Dicionário	Declarações SCP-ECG
	heart_axis	Categórico	Eixo elétrico do coração
	infarction_stadium1	Categórico	Estágio de infarto
	infarction_stadium2	Categórico	Segundo estágio de infarto
	validated_by	Categórico	Cardiologista de validação (pseudônimo)
	second_opinion	Booleano	Sinal para segunda opinião (divergente)
	initial_autogenerated_report	Booleano	Relatório inicial gerado automaticamente pelo dispositivo de ECG
	validated_by_human	Booleano	Validado por humano
Metadados de sinal	baseline_drift	String	Desvio da base
	static_noise	String	Zumbido elétrico / Ruído estático presente
	burst_noise	String	Ruído de ruptura
	electrodes_problems	String	Problemas de eletrodos
	extra_beats	String	Batidas extras
	pacemaker	String	Marca-passo
Parte da validação cruzada	start_fold	Inteiro	Parte estratificada sugerida

Fonte: Adaptado de Wagner *et al.* (2020)

Tabela 4 – Visão geral da declaração de diagnóstico

Diagnóstico	Descrição	Superclasse	Subclasse
LAFB	Bloqueio fascicular anterior esquerdo	CD	LAFB/LPFB
IRBBB	Bloqueio incompleto de ramo direito	CD	IRBBB
AVB	Bloqueio AV de primeiro grau	CD	_AVB
IVCD	Distúrbio de condução intraventricular não específico (bloqueio)	CD	IVCD
CRBBB	Bloqueio completo do ramo direito	CD	CRBBB
CLBBB	Bloqueio completo do ramo esquerdo	CD	CLBBB
LPFB	Bloqueio fascicular posterior esquerdo	CD	LAFB/LPFB
WPW	Síndrome de Wolff-Parkinson-White	CD	WPW
ILBBB	Bloqueio de ramo esquerdo incompleto	CD	ILBBB
3AVB	Bloqueio AV de terceiro grau	CD	_AVB
2AVB	Bloqueio AV de segundo grau	CD	_AVB
LVH	Hipertrofia ventricular esquerda	HYP	LVH
LAO/LAE	Sobrecarga/aumento do átrio esquerdo	HYP	LAO/LAE
RVH	Hipertrofia ventricular direita	HYP	RVH
RAO/RAE	Sobrecarga/aumento do átrio direito	HYP	RAO/RAE
SEHYP	Hipertrofia septal	HYP	SEHYP
IMI	Infarto do miocárdio inferior	MI	IMI
ASMI	Infarto do miocárdio ântero-septal	MI	AMI
ILMI	Infarto do miocárdio inferolateral	MI	IMI
AMI	Infarto do miocárdio anterior	MI	AMI
ALMI	Infarto do miocárdio anterolateral	MI	AMI
INJAS	Lesão subendocárdica em derivações ântero-septais	MI	AMI
LMI	Infarto lateral do miocárdio	MI	LMI
INJAL	Lesão subendocárdica em derivações anterolaterais	MI	AMI
IPLMI	Infarto do Miocárdio Inferoposterolateral	MI	IMI
IPMI	Infarto do miocárdio ínfero-posterior	MI	IMI
INJIN	Lesão subendocárdica em derivações inferiores	MI	IMI
PMI	Infarto do miocárdio posterior	MI	PMI
INJLA	Lesão subendocárdica em derivações laterais	MI	AMI
INJIL	Lesão subendocárdica em derivações inferolaterais	MI	IMI
NORM	ECG normal	NORM	NORM
NDT	Anormalidades T não diagnósticas	STTC	STTC
NST_	Alterações ST não específicas	STTC	NST_
DIG	Efeitos digitálicos	STTC	STTC
LNGQT	Intervalo QT longo	STTC	STTC
ISC_	Isquêmico não específico	STTC	ISC_
ISCAL	Isquêmico em derivações anterolaterais	STTC	ISCA
ISCIN	Isquêmico em derivações inferiores	STTC	ISCI
ISCIL	Isquêmico em derivações inferolaterais	STTC	ISCI
ISCAS	Isquêmico em derivações ântero-septais	STTC	ISCA
ISCLA	Isquêmico nas derivações laterais	STTC	ISCA
ANEUR	Alterações ST-T compatíveis com aneurisma ventricular	STTC	STTC
EL	Distúrbio eletrolítico ou droga (antigo EDIS)	STTC	STTC
ISCAN	Isquêmico em derivações anteriores	STTC	ISCA

Fonte: Adaptado de Wagner *et al.* (2020)

### 5.1.2 Pré-processamento dos dados

Para realizar o pré-processamento dos dados, é previsto na seção 4.2 a realização de procedimentos para o processamento de dados, como eliminação de amostras com dados ausentes e codificação de variáveis categóricas, com o intuito de melhorar a eficiência e eficácia

dos modelos, além de transformar os dados em um formato esperado pelos algoritmos.

Primeiramente, para o pré-processamento dos dados foi utilizado o código exemplo oferecido pela *PhysioNet*<sup>1</sup> que pode ser encontrado no Anexo A. Neste código, é implementado um exemplo de como carregar os dados em forma de onda e nele já é proposto uma divisão dos dados em treino e teste.

Os arquivos processados são compostos por dois grupos de ECG caracterizados por suas taxas de amostragem, os ECG com taxa de amostragem de 100Hz fazem parte de um grupo e os ECG com taxa de amostragem de 500Hz fazem parte de outro grupo. Além disso, há dois arquivos com dados dos pacientes e os diagnósticos. O pré-processamento, primeiramente, é feito em dois arquivos do tipo csv<sup>2</sup>.

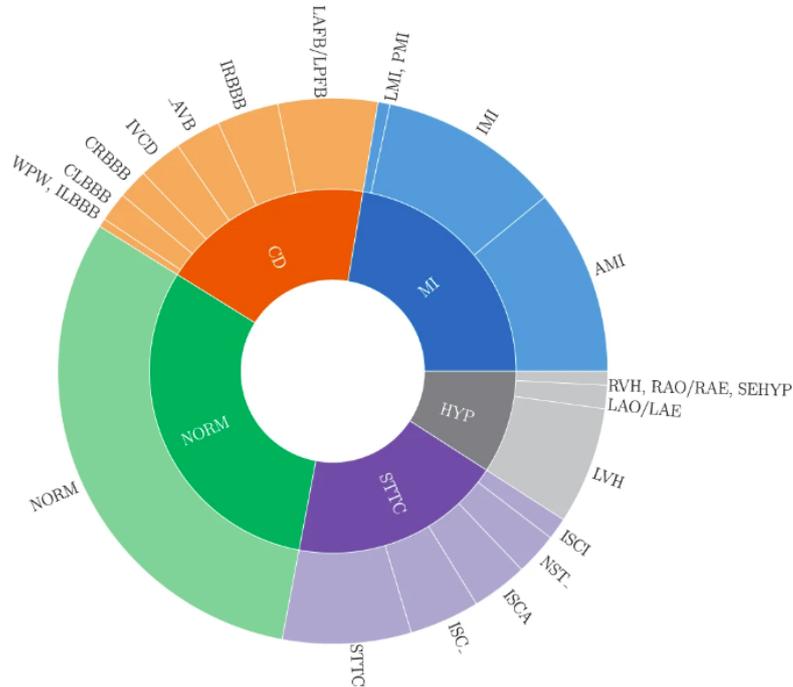
Nesse contexto, o arquivo que possui informações médicas como o índice de todos os exames realizados, informações demográficas, anotação de diagnóstico, forma e ritmo, entre outras informações, denominado *ptbxl\_database.csv* é carregado. Na coluna *scp\_codes*, são apresentadas as declarações que foram informadas pelos médicos e suas probabilidades na forma de dicionário. O outro arquivo chama-se *scp\_statements.csv*, ele possui a descrição de todas as anotações de diagnóstico usadas pelos médicos e anotações de forma e ritmo, além de outros dados referentes às anotações. Estas anotações estão organizadas em uma superclasse e uma subclasse, respeitando a hierarquia das anotações de diagnósticos que foi mencionada anteriormente. Neste trabalho, são observados os dados do tipo diagnóstico em termos de superclasses informadas e para realizar a classificação há 5 superclasses. Na Figura 8 é mostrado uma visão geral gráfica de todo o conjunto de dados em termos de superclasses e subclasses.

---

<sup>1</sup> PhysioNet - <https://physionet.org/>

<sup>2</sup> csv - *comma-separated-values* (valores separados por vírgulas)

Figura 8 – Conjunto de dados PTB-XL em termos de superclasses e subclasses de diagnóstico representado graficamente



Fonte: Wagner *et al.* (2020)

#### 5.1.2.1 Remoção de dados ausentes

Com o objetivo de verificar se há no conjunto de dados valores ausentes, foi implementada uma função chamada *remove*. Ela é responsável por verificar se este fato acontece e retorna os valores que não são vazios no conjunto de dados. Após utilizar esta função, foram encontrados 405 dados ausentes.

#### 5.1.2.2 Codificação de Variáveis Categóricas

Para realizar a tarefa de predição dos diagnósticos se fez necessário utilizar a técnica de codificação *Multi-hot encoding*, visto que, as amostras podem ter mais de um diagnóstico anotado, a técnica *One-hot encoding* não representa a melhor opção para realizar a codificação (FILHO; COUTINHO, 2022).

#### 5.1.3 Treinamento dos Modelos

Os algoritmos abordados implementam as arquiteturas de redes neurais LSTM e *Transformer*. A seguir serão apresentadas e discutidas as arquitetura de redes neurais propostas para então ser feito o treinamento, validação e teste dos modelos.

### 5.1.3.1 Arquitetura da Rede Neural LSTM

A arquitetura é composta pelas seguintes camadas:

- i. Uma camada de entrada com 12 neurônios.
- ii. Uma camada *Dropout* para regularização.
- iii. Uma camada *Batch Normalization* para normalizar as saídas da camada anterior.
- iv. Uma camada oculta com a quantidade de blocos ou neurônios LSTM em conformidade com a quantidade de valores a serem previstos, que são 5.
- v. Uma camada de *Dropout*.
- vi. Uma camada *Batch Normalization*.
- vii. Uma camada densa de saída para a classificação multiclasse.

A função de ativação utilizada é a sigmóide padrão, devido à técnica *Multi-hot encoding*, visto que, a função de ativação sigmóide tem o objetivo de prever a probabilidade de que um determinado exemplo pertença a uma classe particular (RASCHKA; MIRJALILI, 2019).

### 5.1.3.2 Arquitetura da Rede Neural Transformer

Para a construção de um bloco da arquitetura, a rede neural possui as seguintes camadas:

- i. Uma camada de atenção de entrada.
- ii. Uma camada densa com função de ativação ReLu.
- iii. Uma camada para normalização.
- iv. Uma camada de *Dropout*.
- v. Uma camada densa residual para a conexão residual.

Na construção do modelo *Transformer*, tem-se as seguintes camadas:

- i. Uma camada *Flatten* para realizar "achatamento" da entrada.
- ii. Uma camada de saída densa com função de ativação sigmóide que faz a predição dos valores multiclasse.

### 5.1.3.3 Hiperparâmetros

Para a rede neural LSTM foram utilizados os hiperparâmetros número de épocas igual a 30 e tamanho do lote igual a 32. Os mesmo hiperparâmetros foram utilizados para a rede neural *Transformer*. O tamanho do lote determina a quantidade de amostras que serão

propagadas pela rede e o número de épocas especifica a quantidade de vezes que o algoritmo vê o conjunto de dados (FILHO; COUTINHO, 2022).

Para a rede neural *Transformer*, foram treinados três modelos. Os demais hiperparâmetros desses modelos estão apresentados na Tabela 5, onde *n\_blocks* é o número total de blocos em um modelo *Transformer*, *n\_heads* é o número de cabeças que controla quantas cabeças independentes existem em cada camada de atenção e *ff\_dim* é a dimensão da camada de *feed forward*.

Tabela 5 – Hiperparâmetros da rede neural *Transformer*

	<i>n_blocks</i>	<i>n_heads</i>	<i>ff_dim</i>
modelo 1	6	4	512
modelo 2	6	8	512
modelo 3	4	10	512

Fonte: Elaborado pela autora (2023)

#### 5.1.4 Avaliação dos Modelos

Nessa etapa, os modelos são avaliados pelas métricas precisão, revocação e *f1-score* após serem expostos ao conjunto de teste. Nessas métricas, foi observada a média ponderada pelo número de instâncias verdadeiras para cada rótulo, devido ao objetivo ser a classificação das 5 superclasses (NORM, CD, STTC, MI e HYP). Além disso, serão utilizados *thresholds* para observar os resultados em diferentes pontos de decisão, funcionando como um valor crítico que determina a classe final.

Os modelos *Transformer* não apresentaram um desempenho tão satisfatório na classificação das 5 superclasses. O modelo 1 obteve um desempenho inferior em relação aos outros dois modelos, como pode ser visto na Tabela 6. Os modelos 2 e 3 obtiveram uma melhora nos resultados dependendo do valor de *threshold*, como pode ser visto nas Tabelas 7 e 8.

Tabela 6 – Desempenho do modelo 1 da rede neural *Transformer*

<i>Threshold</i>	Precisão	Revocação	<i>F1-score</i>
0.3	0.31	0.32	0.24
0.5	0.28	0.03	0.05
0.75	0.28	0.02	0.02

Fonte: Elaborado pela autora (2023)

Tabela 7 – Desempenho do modelo 2 da rede neural *Transformer*

<i>Threshold</i>	Precisão	Revocação	<i>F1-score</i>
0.3	0.36	0.50	0.36
0.5	0.48	0.36	0.25
0.75	0.37	0.26	0.22

Fonte: Elaborado pela autora (2023)

Tabela 8 – Desempenho do modelo 3 da rede neural *Transformer*

<i>Threshold</i>	Precisão	Revocação	<i>F1-score</i>
0.3	0.30	0.62	0.32
0.5	0.32	0.58	0.32
0.75	0.35	0.17	0.20

Fonte: Elaborado pela autora (2023)

O modelo que apresentou o melhor desempenho na classificação foi a rede neural LSTM. Esta rede neural obteve resultados superiores aos modelos *Transformer*, exceto no caso em que o *threshold* é igual a 0.5 para o modelo 3 do *Transformer* e para o LSTM na métrica de revocação, como pode ser observado nas Tabela 8 e 9.

Tabela 9 – Desempenho da rede neural LSTM

<i>Threshold</i>	Precisão	Revocação	<i>F1-score</i>
0.3	0.69	0.77	0.72
0.5	0.84	0.57	0.65
0.75	0.83	0.33	0.41

Fonte: Elaborado pela autora (2023)

Por meio das Tabelas 6 e 9, é possível observar que o modelo 1 do *Transformer* mostrou-se inferior em todas as métricas, para os três valores de *threshold*, em comparação a rede neural LSTM. Ao analisar as Tabelas 7 e 9, embora o desempenho do *Transformer* tenha melhorado em relação ao modelo anterior, também nota-se a superioridade de desempenho da rede neural LSTM quando comparado ao segundo modelo.

Ao analisar a Tabela 8, é possível notar que o terceiro modelo *Transformer* apresentou um resultado superior na métrica de revocação para *threshold* igual a 0.3 e 0.5 em relação aos outros modelos desta arquitetura. Cabe ressaltar, que um *threshold* de 0.3 significa que o modelo está sendo menos rigoroso ao classificar exemplos como positivos, aumentando a taxa de verdadeiros positivos e, por consequência, diminuindo a taxa de falsos negativos. Isso, no contexto de diagnósticos de DCV, é algo relevante, uma vez que a consequência de um falso negativo é significativamente alta.

Já um *threshold* de 0.5 exige uma probabilidade um pouco mais elevada para classi-

ficar um exemplo como positivo. Entretanto, isso pode resultar ainda em um aumento na taxa de verdadeiros positivos, tendo em vista que exemplos com uma probabilidade relativamente baixa, mas maiores que 0.5, são classificados como positivos. Também é importante ressaltar que essa abordagem pode resultar em uma alta taxa de falsos positivos, pois exemplos com probabilidade próxima de 0.5 podem ser classificados como positivos, embora sejam negativos. Contudo, levando em consideração a primeira observação sobre esse valor de *threshold*, ele também apresenta relevância no contexto de diagnósticos de DCV, pois apresenta a possibilidade de aumento na taxa de verdadeiros positivos.

Ao observar os resultados de precisão, revocação e *f1-score*, embora tenham sido observados resultados que podem apresentar relevância para o contexto desse trabalho na métrica de revocação, a rede neural *Transformer* apresentou dificuldade para prever corretamente as 5 classes, o que pode configurar um problema, levando em consideração que o domínio dos dados é da área da saúde e os modelos apresentaram uma margem de erro significativa ao classificar os diagnósticos corretamente. A rede neural LSTM obteve melhores resultados e apresentou menos dificuldade na tarefa de classificação, entretanto, ainda há espaço para melhoria considerando as consequências que um diagnóstico errado pode ocasionar.

## 6 CONCLUSÃO E TRABALHOS FUTUROS

Neste capítulo, serão apresentadas as considerações finais e as oportunidades para trabalhos futuros.

### 6.1 Considerações Finais

As doenças cardiovasculares são um conjunto de doenças do coração e dos vasos sanguíneos. A detecção dessas doenças o mais rápido possível é de extrema importância para que o tratamento com aconselhamento e medicamentos possa ser começado. Neste trabalho, foi proposto o uso de aprendizado profundo com o intuito de realizar a predição de diagnósticos de doenças cardiovasculares de forma automática, podendo, assim, contribuir para a área da saúde. Este trabalho apresentou um estudo comparativo entre a rede neural *Transformer* e a rede neural recorrente LSTM.

A base de dados utilizada foi a PTB-XL, a qual apresenta uma organização hierárquica de 5 superclasses e 24 subclasses em relação a rótulos de diagnósticos. Neste trabalho, foi realizada uma classificação multiclasse das 5 superclasses NORM, MI, CD, STTC e HYP. Os resultados dos experimentos mostraram um melhor desempenho da rede neural LSTM ao realizar a tarefa de classificação das 5 classes em comparação a rede neural *Transformer*. A LSTM apresentou para a métrica de precisão resultados entre, aproximadamente, 70% a 84% para classificação das 5 superclasses de diagnósticos e o *Transformer* apresentou precisão igual ou inferior a 48%. Para a métrica de revocação, a LSTM obteve resultados de até 77% e o *Transformer* obteve valores que alcançaram até 62%. Por fim, para a métrica *f1-score*, a LSTM apresentou valores entre 41% a 72% e o *Transformer* obteve valores menores, de até 36%.

O presente trabalho reforça a ideia de como diferentes métodos apresentados a contextos diferentes podem ter seus desempenhos afetados quando comparados. Isso pode ser observado ao passo que, embora, segundo Academy (2022), o *Transformer* tenha se tornado de forma rápida um modelo de escolha para PLN, substituindo redes neurais recorrentes, inclusive a LSTM, neste trabalho, ele não obteve melhores resultados que esta rede neural recorrente.

### 6.2 Trabalhos futuros

Como trabalhos futuros, apresenta-se como oportunidade realizar mais estudos para a otimização de hiperparâmetros ou para realizar formas diferentes de pré-processamento dos dados

para os modelos e, também, ampliar a tarefa de classificação para as 24 subclasses de diagnósticos e os 44 diagnósticos. Além disso, outro ponto a ser estudado, seria buscar compreender o uso de diferentes arquiteturas de redes neurais como as redes neurais convolucionais que podem ser uma boa alternativa para trabalhar com dados com 3 dimensões, tendo em vista, seu bom desempenho em trabalhar com imagens.

## REFERÊNCIAS

- ACADEMY, D. S. Transformadores – o estado da arte em processamento de linguagem natural. In: **Deep Learning Book**. [S.l.: s.n.], 2022. cap. 85. Disponível em: <https://www.deeplearningbook.com.br/transformadores-o-estado-da-arte-em-processamento-de-linguagem-natural/>. Acesso em: 29 nov. 2023.
- BARBOSA, N. K. G. da S.; CORREIA, M. A.; SILVA, J. S. da; SILVA, R. P. L. Perfil de pacientes adultos com doenças cardiovasculares no brasil: uma revisão integrativa. **Revista Eletrônica da Estácio Recife**, v. 6, n. 2, 2020.
- CHOLLET, F. **Deep learning with Python**. [S.l.]: Manning, 2017.
- CHOLLET, F. **Deep learning with Python**. [S.l.]: Manning, 2018.
- EHLERS, R. S. Análise de séries temporais. **Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná**, v. 1, p. 1–118, 2007.
- FERREIRA, M. T. **Métodos lineares e não lineares de análise de séries temporais e sua aplicação no estudo da variabilidade da frequência cardíaca de jovens saudáveis**. Universidade Estadual Paulista (UNESP), 2010.
- FILHO, F. R. da S. **Aprendizado de Máquina para Predição de Doenças Cardiovasculares**. TCC — Universidade Federal do Ceará, 2022.
- FILHO, F. S.; COUTINHO, E. Aprendizado de máquina para predição de diagnósticos de doenças cardiovasculares. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE, 22., 2022, Teresina. **Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde**. Porto Alegre, RS, Brasil: SBC, 2022. p. 358–369.
- GÉRON, A. **Mãos à Obra: Aprendizado de máquina com scikit-learn & tensorflow**. [S.l.]: Starlin Alta Editora e Consultoria Eireli, 2019.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. Disponível em: <http://www.deeplearningbook.org>.
- GOODFELLOW, S. D.; GOODWIN, A.; GREER, R.; LAUSSEN, P. C.; MAZWI, M.; EYTAN, D. Towards understanding ecg rhythm classification using convolutional neural networks and attention mappings. In: DOSHI-VELEZ, F.; FACKLER, J.; JUNG, K.; KALE, D.; RANGANATH, R.; WALLACE, B.; WIENS, J. (Ed.). **Proceedings of the 3rd Machine Learning for Healthcare Conference**. [S.l.]: PMLR, 2018. (Proceedings of Machine Learning Research, v. 85), p. 83–101.
- MACENO, L. K.; GARCIA, M. d. S. Fatores de risco para o desenvolvimento de doenças cardiovasculares em jovens adultos/risk factors for the development of cardiovascular diseases in young adults. **Brazilian Journal of Health Review**, v. 5, n. 1, p. 2820–2842, 2022.
- MOHAN, S.; THIRUMALAI, C.; SRIVASTAVA, G. Effective heart disease prediction using hybrid machine learning techniques. **IEEE access**, IEEE, v. 7, p. 81542–81554, 2019.
- MS. **Prevenção é o melhor remédio contra doenças cardiovasculares**. 2021. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/noticias/2021-1/setembro/prevencao-e-o-melhor-remedio-contra-doencas-cardiovasculares>. Acesso em: 24 mai. 2022.

- PYAKILLYA, B.; KAZACHENKO, N.; MIKHAILOVSKY, N. Deep learning for ecg classification. **Journal of Physics: Conference Series**, IOP Publishing, v. 913, n. 1, p. 012004, out 2017.
- RAMOS, Â. P.; SOUSA, B. S. Eletrocardiograma: princípios, conceitos e aplicações. **Centro de Estudos de Fisiologia do Exercício**, v. 9, 2007.
- RASCHKA, S.; MIRJALILI, V. **Python machine learning**: Machine learning and deep learning with python, scikit-learn, and tensorflow 2. [S.l.]: Packt Publishing Ltd, 2019.
- RIBEIRO, E.; DIAS, F.; SOARES, Q.; KRIEGER, J.; GUTIERREZ, M. Deep learning approach for detection of atrial fibrillation and atrial flutter based on ecg images. In: SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE, 23., 2023, São Paulo/SP. **Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde**. Porto Alegre, RS, Brasil: SBC, 2023. p. 509–514.
- RUSSELL, P. N. S. **Inteligência Artificial**. [S.l.]: Elsevier Editora Ltda, 2013.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017.
- WAGNER, P.; STRODTHOFF, N.; BOUSSELJOT, R.-D.; KREISELER, D.; LUNZE, F. I.; SAMEK, W.; SCHAEFFTER, T. Ptb-xl, a large publicly available electrocardiography dataset. **Scientific data**, Nature Publishing Group UK London, v. 7, n. 1, p. 154, 2020.
- WEN, Q.; ZHOU, T.; ZHANG, C.; CHEN, W.; MA, Z.; YAN, J.; SUN, L. Transformers in time series: A survey. **arXiv preprint arXiv:2202.07125**, 2022.
- WHO. **Cardiovascular diseases (CVDs)**. [S.l.]: Disponível em: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Acesso em: 23 nov. 2023, 2023.

**ANEXO A – PRÉ-PROCESSAMENTO DOS DADOS FORNECIDO PELA PHYSIONET**

## Código-fonte 1 – Exemplo de código em Python para pré-processamento dos dados

```
1 import pandas as pd
2 import numpy as np
3 import wfdb
4 import ast
5
6 def load_raw_data(df, sampling_rate, path):
7     if sampling_rate == 100:
8         data = [wfdb.rdsamp(path+f) for f in df.filename_lr
9                 ]
10    else:
11        data = [wfdb.rdsamp(path+f) for f in df.filename_hr
12              ]
13        data = np.array([signal for signal, meta in data])
14    return data
15
16 path = 'path/to/ptbxl/'
17 sampling_rate=100
18
19 Y = pd.read_csv(path+'ptbxl_database.csv', index_col='
20     ecg_id')
21 Y.scp_codes = Y.scp_codes.apply(lambda x: ast.literal_eval(
22     x))
23
24 X = load_raw_data(Y, sampling_rate, path)
25
26 agg_df = pd.read_csv(path+'scp_statements.csv', index_col
27     =0)
28 agg_df = agg_df[agg_df.diagnostic == 1]
```

```
25 def aggregate_diagnostic(y_dic):
26     tmp = []
27     for key in y_dic.keys():
28         if key in agg_df.index:
29             tmp.append(agg_df.loc[key].diagnostic_class)
30     return list(set(tmp))
31
32 Y['diagnostic_superclass'] = Y.scp_codes.apply(
33     aggregate_diagnostic)
34
35 test_fold = 10
36
37 X_train = X[np.where(Y.strat_fold != test_fold)]
38 y_train = Y[(Y.strat_fold != test_fold)].
39     diagnostic_superclass
40
41 X_test = X[np.where(Y.strat_fold == test_fold)]
42 y_test = Y[Y.strat_fold == test_fold].diagnostic_superclass
```