

**CENTRO UNIVERSITÁRIO SENAC SANTO AMARO**

Diego de Moraes Aguiar

Gustavo Donnangelo Cassettari

Leandro Sartini de Campos

**TCC I - RECONHECIMENTO E MONITORAÇÃO DE PADRÕES ASSOCIADOS A MORTE SÚBITA CARDÍACA (MSC) DE ELETROCARDIOGRAMA (ECG) POR MEIO DE REDES NEURAIS ARTIFICIAS (RNA)**

**SÃO PAULO**

**2021**

Diego de Moraes Aguiar

Gustavo Donnangelo Cassettari

Leandro Sartini de Campos

**TCC I - RECONHECIMENTO E MONITORAÇÃO DE PADRÕES ASSOCIADOS A MORTE SÚBITA CARDÍACA (MSC) DE ELETROCARDIOGRAMA (ECG) POR MEIO DE REDES NEURAIS ARTIFICIAS (RNA)**

Trabalho de conclusão de curso, apresentado como exigência parcial para aprovação do curso de Engenharia da Computação, Centro Universitário SENAC sob a orientação dos professores Ms. Sergio Tavares, Dr. Jorge de Oliveira Echeimberg.

**SÃO PAULO**

**2021**

Diego de Moraes Aguiar

Gustavo Donnangelo Cassettari

Leandro Sartini de Campos

**TCC I - RECONHECIMENTO E MONITORAÇÃO DE PADRÕES ASSOCIADOS A MORTE SÚBITA CARDÍACA (MSC) DE ELETROCARDIOGRAMA (ECG) POR MEIO DE REDES NEURAIS ARTIFICIAS (RNA)**

Trabalho de conclusão de curso, apresentado como exigência parcial para obtenção de grau de Bacharel em Engenharia da Computação, pelo Centro Universitário SENAC.

Data de aprovação: XX / XX / XXXX

**BANCA EXAMINADORA**

Prof. Ms. Sergio Tavares  
Centro Universitário SENAC

Prof. (Nome do professor avaliador)   
Instituição do avaliador

Prof. (Nome do professor avaliador)   
Instituição do avaliador

*“A dúvida é o princípio da sabedoria”*

-Aristóteles

**AGRADECIMENTOS**

Gostaríamos de agradecer e dedicar esse trabalho às seguintes pessoas:

Ao Prof. Dr. Jorge de Oliveira Echeimberg pelo apoio em todo o percurso do TCC, por ter nos orientado e nos provido de toda informação necessária, por estar sempre disposto a nos ajudar, fornecendo seu tempo para nos ouvir e auxiliar nos problemas encontrados.

Ao Prof. Dr. Fábio L. S. Cuppo, por todo incentivo, apoio e as valiosas contribuições dadas durante todo processo.

Ao Prof. Ms. Sérgio Tavares, por assumir um projeto já estava em andamento e ainda assim entender o que estava ocorrendo e nos auxiliar com suas indicações que fizeram a diferença.

Ao Prof. Dr. Daminello Raimundo da Faculdade de Medicina do ABC, Membro pesquisador do Laboratório de Delineamento de estudos e Escrita Científica da FMABC, por ter nos ajudado com revisões relacionadas aos dados que usamos.

Por fim, à nossa família e amigos por sempre ter nos acompanhado e ajudado durante esse período, oferecendo apoio de diferentes maneiras.

**RESUMO**

A Organização Mundial da Saúde (OMS) estima que 30% das mortes no mundo são decorrentes de doenças cardiovasculares (DCV) e a previsão é que até 2030 aproximadamente 23,4 milhões pessoas morrerão destas causas. Esse evento, provocará um aumento na demanda por sistemas de processamento de sinais cardíacos para prevenção e tratamento de saúde. Atualmente o eletrocardiograma (ECG) é o procedimento médico mais utilizado à análise dos sinais cardíacos, porém ainda não existe uma solução de ECG capaz de prever o risco de morte súbita cardíaca (MSC), entretanto, a evolução tecnológica tem se mostrado grande aliada da medicina trazendo soluções inovadoras para previsão, prevenção e tratamento de doenças através, por exemplo, da Inteligência Artificial que atualmente é capaz de reconhecer padrões por meio de dados e inferir o risco de doenças ao paciente. Desta forma, este projeto consiste no desenvolvimento de um protótipo de ECG que será conectado de forma não invasiva aos usuários coletando os sinais elétricos do coração, que serão analisados e processados por Redes Neurais Artificiais (RNA) com o objetivo de prever a MSC através dos padrões cardíacos do usuário e enviar alertas contendo localização e tempo estimado para MSC às pessoas e profissionais de saúde responsáveis pelo usuário e indicar o endereço do hospital mais próximo. Desse modo, o sistema poderá colaborar para reduzir o risco de MSC.

**Palavras chaves:** 1. Morte súbita cardíaca 2. ECG 3. Inteligência Artificial 4. Redes Neurais Artificiais 5. Prevenção.

**ABSTRACT**

The World Health Organization estimates that 30% of the deaths in the world are caused by cardiovascular diseases and it is predicted that in 2030 approximately 23,4 million people will die by that cause. This event will cause an increase on the demand for system that process cardiac signals to prevent diseases and treat the health of patients. Actually the electrocardiogram (ECG) is the most used medical procedure to analyse cardiac signals, however, the digital transformation has been showing as a great ally of the medicine bringing innovative solutions to predictions, prevention and treatment of diseases through, for example, Artificial Intelligence which is capable of recognizing patterns with data to infer the risk of diseases in patients. Thus, this project consists on the development of a ECG prototype where it will be connected in non-invasive ways to the patients collecting electrical heart signals, that will be analysed and processed by Artificial Neural Network with the objective of predicting Sudden Cardiac Death from cardiac patterns of the patient and sending a warning with location and times to the Sudden Cardiac Death, and indicating the closest hospital. This way, the system can collaborate to reduce the risk of Sudden Cardiac Death

**Keywords:**  1. Sudden Cardiac Death; 2. ECG; 3. Artificial Intelligence; 4. Artificial Neural Network. 5. Prevention.

**SUMÁRIO**

1. INTRODUÇÃO 12

1.1 PROBLEMA DE PESQUISA 13

1.2 JUSTIFICATIVA 13

1.3 OBJETIVOS 13

1.3.1 Objetivo Geral 13

1.3.2 Objetivos Específicos 13

2. REFERENCIAL TEÓRICO 15

2.1 TECNOLOGIA NA MEDICINA 15

2.2 CORAÇÃO HUMANO 16

2.2.1 Eletrofisiologia 17

2.2.2 Atividade Elétrica no Coração 18

2.2.3 Nódulo Sinoatrial (SA) 19

2.2.4 Nódulo Atrioventricular (AV) 20

2.2.5 Excitação Natural do Coração 21

2.2.6 Ritmo Normal e Arritmia 22

2.2.7 Morte Súbita Cardíaca 23

2.3 ELETROCARDIOGRAMA (ECG) 23

2.3.1 Protótipo ECG 29

2.3.1.1 Aquisição de Sinal 30

2.3.1.2 Amplificador de Instrumentação 30

2.3.1.3 Filtro Passa Alta 32

2.3.1.4 Filtro Passa Baixa 34

2.3.1.5 Passa Notch ou Rejeita-Faixa 35

2.3.2 Taxa de Amostragem 38

2.3.3 Software de Simulação 38

2.4 INTELIGÊNCIA ARTIFICIAL 38

2.4.1 Análise dos Dados 40

2.4.1.1 Testes de Hipóteses 40

2.4.1.2 Teste de Kolmogorov-Smirnov 40

2.4.1.3 Teste de Qui-Quadrado de Aderência 41

2.4.1.4 Correlações 41

2.4.2 Machine Learning 42

2.4.2.1 Métricas de Erro 42

2.4.2.1.1 MAE ou Média Absoluta do Erro 43

2.4.2.1.2 MSE ou Média do Erro Quadrático 43

2.4.2.1.3 Matriz de Confusão 43

2.4.2.1.4 Precisão 44

2.4.2.1.5 Especificidade 45

2.4.3 Redes Neurais 45

2.4.3.1 O Neurônio 45

2.4.3.2 Tipos de Redes Neurais 49

3. MATERIAIS E MÉTODO 51

3.1 METÓDO 51

3.2 DESENVOLVIMENTO DO PROTÓTIPO 51

3.2.1 Hardware 51

3.2.2 Software 52

3.2.2.1 Cloud: Computação em Nuvem 52

3.2.2.2 Serviço de Nuvem 53

3.2.1 A Infraestrutura 53

3.2.2 Marcadores de Morte Súbita 54

3.3 SUJEITOS DA PESQUISA E AMOSTRA 55

3.4 CRONOGRAMA DAS ATIVIDADES 56

3.5 PLANEJAMENTO DO DESENVOLVIMENTO DO PROTÓTIPO 57

3.5.1 Scrum 57

3.5.2 Organização Durante o Desenvolvimento de Software 58

4. DESENVOLVIMENTO 60

5. CONSIDERAÇÕES FINAIS 61

REFÊRENCIAS 62

APÊNDICE 67

**LISTA DE QUADROS**

[**Quadro 1** - Principais marcadores de MSC 55](#_Toc42712299)

**LISTA DE TABELAS**

[**Tabela 1** - Morfologia padrão das ondas do ECG 28](#_Toc42712316)

[**Tabela 2 -** Cronograma de desenvolvimento do projeto 56](#_Toc42712317)

**LISTA DE FIGURAS**

[**Figura 1** - Estrutura do coração e fluxo sanguíneo 19](#_Toc64744136)

[**Figura 2** - Sistema de condução elétrica cardíaca 22](#_Toc64744137)

[**Figura 3** - Potenciais de ação em cada parte do coração 24](#_Toc64744138)

[**Figura 4** – Potenciais de ação coletados pelo ECG 25](#_Toc64744139)

[**Figura 5** - Disposição convencional dos eletrodos para o registro das derivações eletrocardiográficas padrão 28](#_Toc64744140)

[**Figura 6** - Diagrama de blocos do ECG 28](#_Toc64744141)

[**Figura 7** - Configuração da onda cardíaca 29](#_Toc64744142)

[**Figura 8** - Espectro de potência relativo do complexo QRS, ondas P e T, atrito do eletrodo e ruído muscular 32](#_Toc64744143)

[**Figura 9** - Diagrama dos módulos conectados para o funcionamento do ECG 33](#_Toc64744144)

[**Figura 10** - Diagrama de um amplificador de instrumentação e sua expressão de ganho 34](#_Toc64744145)

[**Figura 11** - Amplificador operacional utilizando 3 amplificadores utilizado no protótipo 35](#_Toc64744146)

[**Figura 12** - Esquematização do filtro passa alta 36](#_Toc64744147)

[**Figura 13** - Filtro passa alta com amplificador utilizado no protótipo 36](#_Toc64744148)

[**Figura 14** - Esquematização do filtro passa baixa 37](#_Toc64744149)

[**Figura 15** - Filtro passa baixa utilizado no protótipo 37](#_Toc64744150)

[**Figura 16** - Esquematização do filtro passivo rejeita-faixa 38](#_Toc64744151)

[**Figura 17** - Filtro rejeita faixa utilizado no protótipo 39](#_Toc64744152)

[**Figura 18** - Esquematização do protótipo de ECG completo 40](#_Toc64744153)

[**Figura 19** - Evolução da necessidade de processamento para algoritmos mais pesados de Inteligência Artificial 42](#_Toc64744154)

[**Figura 20** - Matriz de confusão 47](#_Toc64744155)

[**Figura 21** - Curva AUC-ROC 49](#_Toc64744156)

[**Figura 22** - Estrutura básica simplificada de um neurônio 50](#_Toc64744157)

[**Figura 23** - Primeira imagem do potencial de ação medido na tela de um osciloscópio por Hodgkin e Huxley 51](#_Toc64744158)

[**Figura 24** - Modelo matemático simples de um neurônio conhecido como Perceptron 52](#_Toc64744159)

[**Figura 25** - A esquemática de uma RNA compostas por duas camadas e apenas um neurônio de saída 52](#_Toc64744160)

[**Figura 26** - Topologia das mais variadas arquiteturas de uma RNA 54](#_Toc64744161)

[**Figura 27** - Arquitetura utilizada no projeto 58](#_Toc64744162)

[**Figura 28 -** Kanban organizado na ferramenta Trello 63](#_Toc64744163)

[**Figura 29** - GitHub do TCC com o CookieCutter 64](#_Toc64744164)

[**Figura 30** – Circuito montado do Multisim 65](#_Toc64744165)

[**Figura 31** – Teste no Multisim com um sinal de onda cardíaca 66](#_Toc64744166)

[**Figura 32** – Circuito montado na protoboard 66](#_Toc64744167)

[**Figura 33** – Sinal cardíaco registrado pelo osciloscópio 67](#_Toc64744168)

[**Figura 34** – Sinal cardíaco registrado pelo osciloscópio 67](#_Toc64744169)

[**Figura 35** – Teste de sinal no filtro Notch 60 Hz - Ativo 68](#_Toc64744170)

[**Figura 36** – Teste de sinal no filtro Notch 60 Hz - Passivo 69](#_Toc64744171)

[**Figura 37** – Circuito final com filtro Notch ativo 69](#_Toc64744172)

[**Figura 38** – Simulação do circuito impresso no software KICAD 70](#_Toc64744173)

[**Figura 39** – Separação dos dados eletrocardiográficos entre treino e teste 71](#_Toc64744174)

[**Figura 40** – Curva de aprendizagem do modelo em treino e validação 72](#_Toc64744175)

[**Figura 41** – Gráfico de comparação entre o sinal real e o previsto 1 73](#_Toc64744176)

[**Figura 42** – Gráfico de comparação entre o sinal real e o previsto 2 74](#_Toc64744177)

[**Figura 43** – Dados coletados pelo ECG com ruídos 75](#_Toc64744178)

[**Figura 44** – Dados coletados pelo ECG após o tratamento 76](#_Toc64744179)

[**Figura 45** – Parábola dividida em treinamento e teste para validação do modelo LSTM 77](#_Toc64744180)

[**Figura 46** – Curva de aprendizagem do modelo em treino e validação para a parábola 78](#_Toc64744181)

[**Figura 47** – Gráfico de comparação entre os dados reais e a previsão do modelo LSMT 78](#_Toc64744182)

[**Figura 48** – Gráfico da parábola com o modelo prevendo o restante da curva 79](#_Toc64744183)

[**Figura 49** – Gráfico da base de um paciente com fibrilação 80](#_Toc64744184)

[**Figura 50** – Curva de aprendizagem do modelo em treino e validação com a base de fibrilação 81](#_Toc64744185)

[**Figura 51** – Gráfico de comparação entre o sinal real e o previsto com a base de fibrilação 81](#_Toc64744186)

[**Figura 52** – Resultado do modelo com apenas 04 épocas de treinamento 84](#_Toc64744187)

[**Figura 53** – Gráfico de poder de separação do modelo 86](#_Toc64744188)

[**Figura 54** – Curva ROC da RCN 86](#_Toc64744189)

[**Figura 55** – Teste de especificidade vs sensibilidade do modelo 87](#_Toc64744190)

[**Figura 56** – Percentual de predições corretas do modelo pacientes com MSC 89](#_Toc64744191)

[**Figura 57** – Percentual de predições corretas do modelo pacientes Saudáveis 89](#_Toc64744192)

[**Figura 58** – Paciente 38 de MSC com quantidade de predições do modelo no tempo 90](#_Toc64744193)

[**Figura 59** – Paciente 43 de MSC com quantidade de predições do modelo no tempo 90](#_Toc64744194)

[**Figura 60** – Paciente 51 de MSC com quantidade de predições do modelo no tempo 91](#_Toc64744195)

[**Figura 61** – Resultado do modelo para pacientes com FCC 92](#_Toc64744196)

[**Figura 62** – Resultado do modelo rebalanceado com 05 épocas de treinamento 93](#_Toc64744197)

[**Figura 63** – Curva ROC após o rebalanceamento 94](#_Toc64744198)

[**Figura 64** – Teste de especificidade vs sensibilidade do modelo rebalanceado 95](#_Toc64744199)

[**Figura 65** - Porcentagem de acertos e erros para pacientes de MSC 95](#_Toc64744200)

[**Figura 66** - Porcentagem de acertos e erros para pacientes saudáveis 96](#_Toc64744201)

[**Figura 67** - Porcentagem de acertos com dados SINUS para pacientes com FCC 97](#_Toc64744202)

[**Figura 66** – Dashboard com os dados do paciente 98](#_Toc64744203)

# INTRODUÇÃO

Morte súbita cardíaca (MSC) é um evento inesperado, de causa cardíaca. A Organização Mundial da Saúde (OMS) estima que a morte de pessoas por causa de doenças cardiovasculares crescerá mundialmente atingindo 23,4 milhões de casos em 2030. E prevê que muitas destas mortes, ocorrerão sem qualquer sintoma prévio de doença cardiovascular (OMS, 2017). Este aspecto potencialmente provocará uma maior demanda por sistemas de processamento de sinais cardíacos para prevenção e tratamento de saúde.

Segundo estudos da Cardiac Electrophysiology Clinics (2017), a principal causa de MSC está relacionada a doença estrutural do coração, podendo ocorrer por doenças genéticas, problemas nas vias arteriais, síndromes e casos relacionados a polarização e despolarização precoce.

De acordo com o Instituto do Coração (InCor - HC/FMUSP, 2002), mais de 95% das MSC ocorrem fora do ambiente hospitalar. Sendo o principal mecanismo de morte a fibrilação ventricular. A rápida desfibrilação e o suporte básico de vida podem aumentar a taxa de sobrevida em longo prazo. Casos em que o atendimento ocorre no período de cinco a sete minutos as chances de sobrevida após uma parada cardíaca é maior do que 49%.

Atualmente não existe solução de ECG que possa inferir adequadamente o risco de MSC, porém existem parâmetros[[1]](#footnote-1) que podem ser úteis para a análise e previsão do risco de MSC (RIERA; BARROS; SHENASA, 2017).

Desta forma, este estudo consiste no desenvolvimento de um protótipo de ECG que será conectado aos usuários coletando os sinais elétricos do coração, que serão analisados e processados por Redes Neurais Artificiais (RNA) com o objetivo de prever a MSC através dos padrões cardíacos do usuário e enviar um aviso para a hospital mais próximo para acelerar o atendimento e consequentemente diminuir o risco de morte do usuário.

* 1. PROBLEMA DE PESQUISA

Mesmo após mais de 100 anos da sua descoberta o ECG continua sendo o método mais comum de diagnóstico clínico. Porém atualmente não existe um dispositivo ou meio que possibilite o monitoramento e a previsão da morte súbita cardíaca.

* 1. JUSTIFICATIVA

Conforme dados da Organização Pan-Americana de Saúde (2017), estima-se que 31% de todas as mortes no mundo tem como causa Doenças Cardiovasculares (DCV). A identificação de padrões e a previsão através de Inteligência Artificial podem auxiliar o médico na redução do risco de morte súbita, particularmente se for possível prever eventos indesejáveis para o paciente.

* 1. OBJETIVOS
     1. Objetivo Geral

Desenvolver um protótipo, para reconhecimento e monitoração de padrões de sinais associados a MSC.

* + 1. Objetivos Específicos

1 – desenvolver um protótipo de ECG não invasivo e portátil para a coleta de sinais elétricos cardíacos;

2 – desenvolver um aplicativo para receber os dados do ECG e enviar para a infraestrutura Web;

3 – configurar uma infraestrutura na Amazon Web Service (AWS) para armazenamento e tratamento dos dados coletados;

4 – criar filtros e realizar normalizações nos dados para posterior treinamento da RNA;

5 – aplicar um método de comunicação para informar médicos e hospitais da possível MSC.

# REFERENCIAL TEÓRICO

Esta seção tem como objetivo realizar a fundamentação teórica que envolve o tema abordado, através de pesquisas bibliográficas de autores que já discorreram sobre o assunto e com isso, o embasamento para o desenvolvimento do nosso protótipo.

* 1. TECNOLOGIA NA MEDICINA

Atualmente, a tecnologia vem se mostrando indispensável para a rotina das pessoas em todos os sentidos. É quase impossível pensar em qualidade de vida, por exemplo, sem levar em consideração a grande contribuição da tecnologia na saúde.

Os avanços são inúmeros — sistemas operacionais eficientes e seguros, equipamentos modernos, adoção de prontuários eletrônicos, avanços em engenharia e tecnologia biomédica, entre outros. Neste contexto, as áreas da medicina têm recebido suporte tecnológico, o que permite maior eficácia nos tratamentos e, principalmente, diagnósticos mais rápidos e efetivos.

Por meio dos *Wearables*, pessoas diabéticas ou com doenças que necessitam de um controle mais rigoroso das condições físicas podem contar com as possibilidades do monitoramento em tempo real. *Wearables* é o nome dado às tecnologias vestíveis, como os relógios inteligentes que possibilitam controlar o batimento cardíaco de quem os utiliza, por exemplo (OPTVISION, 2018).

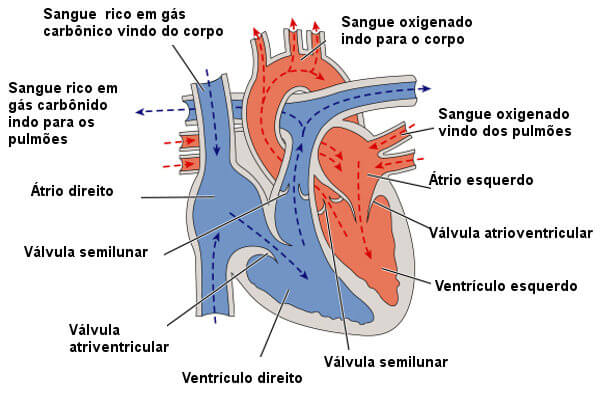
A evolução da tecnologia na medicina ao longo dos anos é marcante, tanto para profissionais quanto para pacientes. Basta um pouco de atenção para percebermos que, a todo momento, surgem novidades na saúde que qualificam as ações de prevenção, diagnóstico e tratamento de doenças e outras condições médicas (MORSCHI, 2018).

A era da informação não deixou à margem a área da saúde. De fato, a tecnologia ultrapassou o processamento padrão de dados para funções administrativas comuns em todas as organizações, tais como recursos humanos, folhas de pagamento, sistemas de contabilidade, entre outros, e agora desempenha um papel fundamental tanto no cuidado ao paciente, na **interpretação do eletrocardiograma (nosso grifo)**, como em escalas de trabalho, prescrição, relatório de resultados e sistemas de prevenção (PINOCHET; LOPES; SILVA, 2014).

* 1. CORAÇÃO HUMANO

O coração humano, ilustrado na Figura 1, é responsável pelo bombeamento do sangue através do organismo. É formado por duas bombas distintas: o coração direito e o coração esquerdo. O coração direito é responsável pela circulação do sangue nos pulmões, e o coração esquerdo é responsável pela circulação do sangue nos órgãos periféricos (GUYTON; HALL, 1996). Cada um desses corações é uma bomba pulsátil com duas câmaras, denominadas de átrio e ventrículo. A principal função do átrio é mover o sangue para o ventrículo, atuando como uma bomba auxiliar. O ventrículo, por sua vez, fornece a força principal para a circulação do sangue para todo o organismo (GUYTON; HALL, 1996) (FROWNFELTER; DEAN, 2004).

**Figura 1** - Estrutura do coração e fluxo sanguíneo



Fonte: Brasil Escola (2020)

Os átrios e os ventrículos contraem-se em tempos diferentes. Quando os ventrículos se encontram contraídos (sístole ventricular), as válvulas atrioventriculares (tricúspide e mitral) estão fechadas e as válvulas das artérias aorta e pulmonar estão abertas. A partir deste momento, os ventrículos começam a relaxar-se (diástole ventricular), as válvulas atrioventriculares se abrem, as válvulas das artérias aorta e pulmonar se fecham e os átrios se contraem (sístole atrial) provocando a passagem do sangue dos átrios para os ventrículos. Após o enchimento dos ventrículos, ocorre a contração destes (sístole ventricular) e fecham-se as válvulas atrioventriculares, abrem-se as válvulas das artérias aorta e pulmonar e o sangue é impulsionado para os pulmões através das artérias pulmonares e para o organismo, através da artéria aorta. A sístole ventricular, por sua vez, ocorre ao mesmo tempo em que ocorre a diástole atrial quando o sangue das veias cavas e das veias pulmonares flui para os átrios direito e esquerdo. As propriedades de automaticidade (capacidade de iniciar seu próprio batimento) e ritmicidade (regularidade na atividade de marcapasso) permitem a um coração bater e produzir o ciclo cardíaco (AZEVEDO, 1999 apud BERNE et al., 2004).

O sistema cardiovascular ou circulatório agrupa os órgãos destinados à circulação do sangue e da linfa. É constituído pelo coração e por uma rede fechada de canais distribuídos por todo o organismo humano que são: os vasos arteriais, os vasos venosos e os vasos linfáticos (AZEVEDO, 1999 apud BERNE et al., 2004).

Assim sendo, o coração também funciona como uma bomba eletromecânica que produz sinais eletromagnéticos pela operação do sistema cardiovascular, gerando pulsos elétricos que alteram o campo elétrico em torno do coração e que pode ser medido e amplificado através de dispositivos como o eletrocardiógrafo.

* + 1. Eletrofisiologia

O corpo humano é constituído de aproximadamente 60% de água distribuída pelos vasos sanguíneos, pelas células e por entre as células. Esta água é composta de substâncias químicas dissolvidas como: proteínas, glicose, gordura, ureia, sais, etc. Alguns desses sais como cloreto de sódio, cloreto de potássio e cloreto de cálcio, quando dissolvidos na água, têm seus elementos constituintes dissociados e carregados de energia eletroquímica. Esses sais constituem os íons, alguns carregados positivamente e outros negativamente. Os principais íons do corpo humano são: o sódio, o cloro, o potássio e o cálcio (AZEVEDO, 1999 apud BERNE et al., 2004).

A água e os íons se deslocam entre os compartimentos do corpo humano, atravessando as membranas de suas células. O movimento da água e dos íons entre esses compartimentos é determinado por diferenças elétricas, osmóticas, oncóticas e mecanismos biológicos. O deslocamento dos íons, através das membranas das células cardíacas, produz corrente elétrica de natureza eletroquímica (AZEVEDO, 1999 apud BERNE et al., 2004).

Mudanças na permeabilidade da membrana celular alteram o deslocamento desses íons através da membrana. A permeabilidade da membrana a um determinado íon, a diferença de concentração dele e a diferença de potencial transmembrana definem a quantidade resultante deste íon a ser deslocada através dela (AZEVEDO, 1999 apud BERNE et al., 2004).

* + 1. Atividade Elétrica no Coração

O coração tem uma capacidade de auto excitação, ou seja, o miocárdio possui determinado tipo de célula capaz de gerar e propagar, de forma organizada, um estímulo elétrico por todas as câmaras do coração, que acarreta a contração organizada dos átrios e ventrículos e, como consequência, o batimento cardíaco, conforme ilustra a Figura 2 a seguir:

**Figura 2** - Sistema de condução elétrica cardíaca



Fonte: Shutterstock (2020)

As estruturas envolvidas neste processo são:

* nódulo sinoatrial ou sinusal, no qual o impulso auto excitador rítmico é gerado;
* vias internodais, que conduzem o impulso do nódulo sinoatrial para o nódulo atrioventricular;
* nódulo atrioventricular, no qual o impulso proveniente dos átrios é retardado antes de passar para os ventrículos;
* feixe de His, que conduz o impulso dos átrios aos ventrículos;
* feixes esquerdo e direito das fibras de Purkinje, que conduzem o impulso a todas as partes dos ventrículos (MOHRMAN; HELLER, 2006).
  + 1. Nódulo Sinoatrial (SA)

As ondas de despolarização, em corações normais, são iniciadas em fibras especializadas chamadas nódulo sinoatrial (nódulo SA). O nódulo SA localiza-se na parede superior do átrio direito, onde há a junção do átrio com a veia cava. Ele tem o tamanho aproximado de uma ponta de lápis e atua como marca-passo primário do coração. O nódulo SA não tem os canais de sódio usuais, nele a despolarização é realizada por íons Ca2+. O nódulo SA atua como oscilador biológico e dispara entre 60 e 100 vezes por minuto; esta taxa é aumentada pelos estímulos oriundos dos nervos simpáticos do coração e diminuída pelos estímulos dos nervos parassimpáticos (HOBBIE; ROTH, 2007); (BERBARI, 2000); (NAZERAN, 2006).

* + 1. Nódulo Atrioventricular (AV)

Como átrios e ventrículos são separados por tecido fibroso que não transmite o impulso, a condução direta entre átrios e ventrículos não ocorre. A única conexão entre os átrios e os ventrículos é um tecido condutivo chamado nódulo atrioventricular (nódulo AV) (HOBBIE; ROTH, 2007); (NAZERAN, 2006).

A onda de ativação segue um caminho que passa pelo nódulo AV e prossegue através do feixe de His nos seus galhos (ramos) esquerdo e direito e por fim, às fibras de Purkinje que são espalhadas e infiltradas no tecido ventricular endocárdico. O impulso cardíaco é atrasado no nódulo AV por cerca de 100 ms. O atraso do nódulo AV é benéfico visto que a ativação do músculo cardíaco inicia suas sucessivas contrações mecânicas. Este atraso dá tempo suficiente para o término da contração atrial e o bombeamento de sangue dentro dos ventrículos. Quando o impulso cardíaco atinge o feixe de His a condução é muito rápida, resultando no início da ativação ventricular sobre uma larga faixa. As subsequentes propagações elétricas de célula a célula são altamente sequenciais e coordenadas resultando em uma ação de bombeamento dos ventrículos altamente síncrona e eficiente. O nódulo AV despolariza espontaneamente a uma taxa média de 50 batimentos por minuto, mas como a despolarização do nódulo SA é mais rápida é essa última que dita o ritmo cardíaco em condições normais (HOBBIE; ROTH, 2007); (NAZERAN, 2006).

Diversas cardiopatias podem ser detectadas a partir do estudo gráfico da despolarização e repolarização das células cardíacas. O médico especialista consegue identificar a condutibilidade das paredes e com isso, se há alguma região necrosada; consegue identificar também qual marca-passo natural está funcionando, se o nódulo SA ou o nódulo AV; consegue identificar ainda, se um nódulo está interferindo no outro; entre outros problemas (HOBBIE; ROTH, 2007).

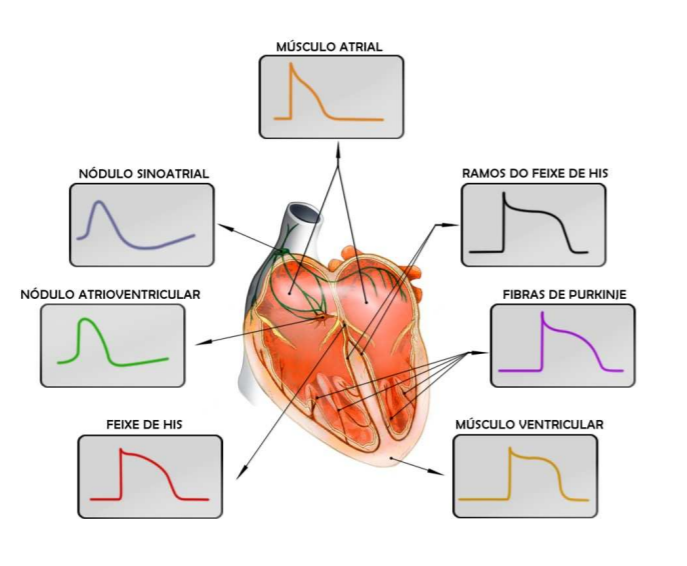
* + 1. Excitação Natural do Coração

O estímulo elétrico originado no nódulo sinoatrial se propaga pelos dois átrios causando a contração dos mesmos e, consequentemente, o bombeamento do sangue para os ventrículos. Entre os átrios e os ventrículos, existe uma estrutura fibrosa que apresenta baixa condutividade elétrica, impedindo a passagem dos impulsos elétricos dos átrios para os ventrículos, a não ser através de um grupo de células conhecido como nódulo atrioventricular (OLIVEIRA, 2010).

As células do nódulo atrioventricular apresentam uma baixa velocidade de condução, permitindo que a contração dos ventrículos ocorra apenas após a ejeção do sangue contido nos átrios. A extremidade inferior do nódulo atrioventricular se prolonga pelo feixe de His[[2]](#footnote-2) até a extremidade inferior dos ventrículos por meio das fibras de Purkinje. Logo após, o potencial elétrico é conduzido para o miocárdio ventricular causando a contração dos ventrículos (OLIVEIRA, 2010).

Com isso, cada estrutura do coração relacionada a condução gera potenciais de ação e consequentemente as formadas de onda de cada ciclo cardíaco, conforme ilustrado na Figura 3.

**Figura 3** - Potenciais de ação em cada parte do coração

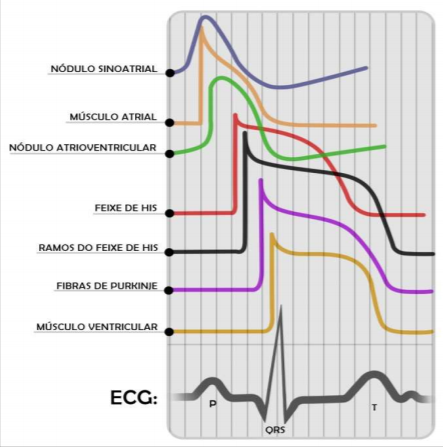


Fonte: Tomazati (2010)

Cada potencial de ação contribui para a formação do sinal de ECG, sendo que a maior contribuição se dá pelos músculos atriais e ventriculares por representar a maior parte do tecido do coração (OLIVEIRA, 2010).

A Figura 4, ilustra o sinal de eletrocardiograma destacando as cinco ondas características provenientes dos diversos potenciais de ação.

**Figura 4** – Potenciais de ação coletados pelo ECG



Fonte: Tomazati (2010)

* + 1. Ritmo Normal e Arritmia

O ritmo cardíaco comandado pelo nódulo SA, onde cada ciclo cardíaco tem a mesma duração e todas as ondas do registro sucedem-se de forma regular é classificado como ritmo normal (AZEVEDO, 1999).

Arritmia cardíaca é um distúrbio caracterizado por batimentos cardíacos em ritmo irregular que engloba ritmos desordenados ou a interrupção de um ritmo normal (AZEVEDO, 1999).

* + 1. Morte Súbita Cardíaca

A morte súbita cardíaca descreve uma morte natural inesperada por uma causa cardíaca dentro de um curto espaço de tempo, geralmente menor que uma hora do início dos sintomas, em uma pessoa sem qualquer condição que parecia fatal. Essa morte rápida é muito atribuída a arritmia cardíaca, porém com o avanço das capacidades de monitoramento a partir de desfibriladores cardíacos implementados, agora é bem reconhecido que classificações baseadas em circunstâncias clínicas podem ser enganosas e muitas vezes impossíveis, pois 40% das mortes súbitas não são testemunhadas (ZIPES; WELLENS, 1998).

A maior causa da morte súbita cardíaca é quando existe um problema estrutural no coração, na sua maior parte doenças ateroscleróticas, que representam aproximadamente 85% de todos os casos. Os outros 15% são causados por doenças cardiopáticas sem problemas estruturais aparentes (RIERA; BARROS; SHENASA, 2017).

Pelo fato de a morte súbita cardíaca ser uma síndrome multifatorial complexa, sua patofisiologia e iniciadores são muito mal entendidos. Devido a morte súbita cardíaca ter um perfil de risco multifatorial, existe uma grande razão para se utilizar múltiplos marcadores de risco, ou marcadores de morte súbita, que refletem diferentes faces da atividade elétrica de um coração (RIERA; BARROS; SHENASA, 2017).

* 1. ELETROCARDIOGRAMA (ECG)

Em 1903, Einthoven registrou várias anomalias cardíacas em humanos usando um galvanômetro com fios. Ele introduziu conceitos na eletrocardiografia que ainda hoje são utilizados como: os nomes dados as ondas de P a U, a disposição dos eletrodos para os canais bipolares, bem como a definição do primeiro modelo teórico para o coração, o dipolo simples variante no tempo (FOTIADIS et al., 2006).

O eletrocardiógrafo é um galvanômetro que amplifica, mede e registra a intensidade da corrente elétrica gerada pela diferença de potencial elétrico entre os eletrodos. Os sistemas de derivações usados para registro do ECG de rotina são orientados em certos pontos do corpo. O ECG detecta alterações temporais do potencial elétrico, entre um ponto na superfície da pele e um ponto indiferente, ou entre pares de pontos na superfície da pele (AZEVEDO, 1999 apud BERNE et al., 2004).

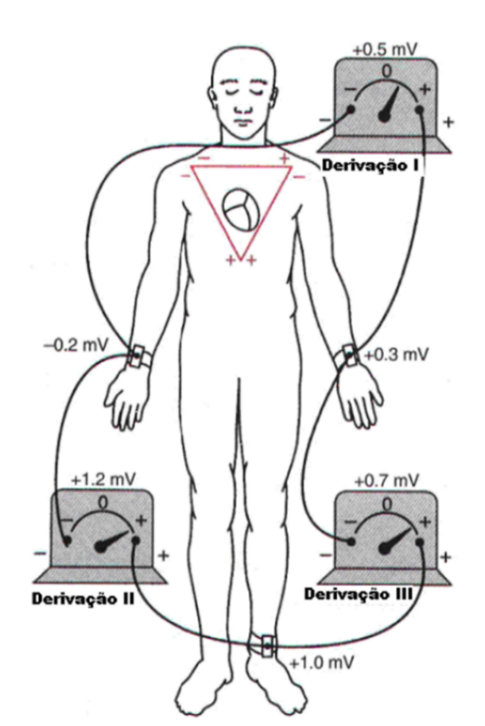
O ECG é um procedimento não invasivo, informativo, reproduzível e de baixo custo. Por essa razão é a ferramenta mais importante e mais utilizada para o diagnóstico da função cardiovascular (AZEVEDO, 1999 apud SANDHAM, 2007).

Com essas alterações temporais do potencial elétrico, cria-se um sistema de derivações. Segundo Azevedo (1999), o sistema de derivações original do ECG foi concebido por Dutchman Willem Einthoven. Neste sistema de derivações, a soma dos vetores de toda a atividade elétrica cardíaca, a qualquer momento, é chamada vetor cardíaco resultante.

De acordo com Azevedo (1999), considera-se que esta força elétrica está no centro de um triângulo equilátero, cujos vértices estão localizados no ombro esquerdo, no ombro direito e na região púbica. Este triângulo, chamado de Triângulo de Einthoven, está orientado no plano frontal do corpo. Os eletrodos são, por conveniência, conectados aos antebraços direito e esquerdo. De modo similar, a perna representa uma extensão do sistema de derivações do púbis, e desta forma, o terceiro eletrodo é usualmente conectado ao tornozelo esquerdo.

A Figura 5, ilustra a disposição convencional dos eletrodos para o registro das derivações eletrocardiográficas padrão. O triângulo de Einthoven está sobreposto ao tórax.

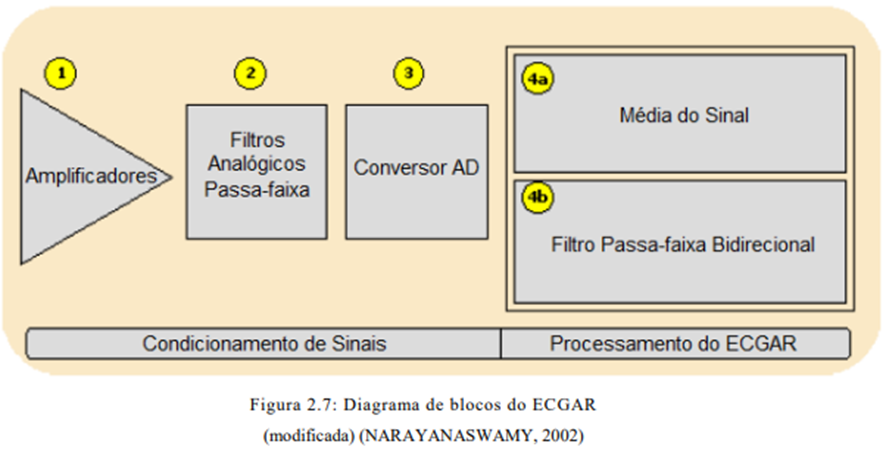
**Figura 5** - Disposição convencional dos eletrodos para o registro das derivações eletrocardiográficas padrão



Fonte: Guyton; Hall (1996)

Os eletrocardiógrafos digitais de alta resolução têm a habilidade de registrar sinais de baixíssima magnitude (~1 µV[[3]](#footnote-3)) que ocorrem após o complexo QRS e são denominados potenciais tardios (NARAYANASWAMY, 2002). A Figura 6, ilustra o diagrama de blocos deste tipo de ECG.

**Figura 6** - Diagrama de blocos do ECG

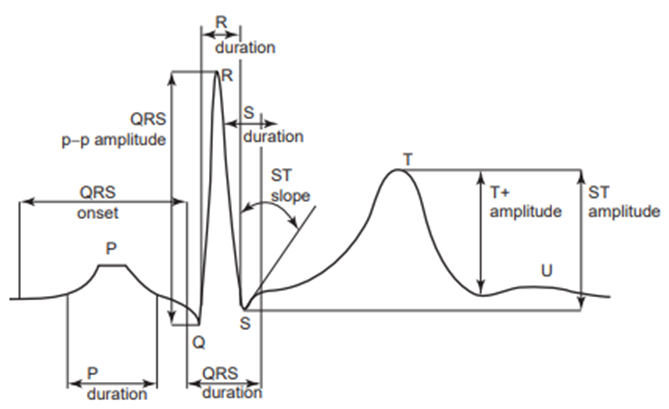


Fonte: Narayanaswamy (2002)

Dutchman Willem Einthoven designou as deflexões P, Q, R, S e T, padronizou a velocidade de registro, a técnica de padronização e a tomada de derivações, empregando, nesse primeiro período, as derivações: mão direita e mão esquerda (AZEVEDO, 1999 apud BERNE et al., 2004).

A Figura 7, ilustra a representação de uma onda cardíaca referenciando os seus pontos.

**Figura 7** - Configuração da onda cardíaca



Fonte: Fotiadis et al. (2006)

A onda P representa a ativação dos átrios. Quando a frente de onda de ativação do nódulo SA encontra a massa aumentada no músculo atrial, o início da atividade elétrica de um ciclo cardíaco é observado na superfície da pele e a primeira onda é encontrada no ECG (BERBARI, 2000).

A onda Q é a deflexão inicial para baixo, a onda R é a deflexão para cima e a onda S é a deflexão final para baixo. A polaridade e a presença dessas três ondas dependem da localização dos eletrodos no corpo e bem como de diversas anomalias que podem existir. O complexo QRS, é mostrado quando o sistema His-Purkinje ativa o músculo ventricular e a grande massa dos ventrículos é excitada. Essa excitação gera a contração que provê a principal força para o sangue circular nos órgãos e representa a sístole do ciclo cardíaco (BERBARI, 2000).

Em alguns indivíduos, um pequeno pico pode ocorrer no final, ou um pouco depois, da onda T e é conhecido como onda U. Sua origem nunca foi completamente estabelecida, mas é considerado um sinal de repolarização (BERBARI, 2000).

A Tabela 1, representa as deflexões P, Q, R, S e T mencionadas anteriormente que são captadas pelo ECG e suas principais características.

**Tabela 1** - Morfologia padrão das ondas do ECG

|  |  |
| --- | --- |
| **Característica do ECG** | **Descrição** |
| **Onda P** | * Deflexão obtusa simétrica da linha isoelétrica * Precede o complexo QRS * Normalmente é rítmica com pequenas variações e sem uniformidades * Duração < 0,12 s * Amplitude < 0,3 mV * Mesma frequência do complexo QRS |
| **Complexo QRS** | * Todas as três ondas que constituem são deflexões abruptas da linha isoelétrica * Duração < 0,12 s * Amplitude < 1,0 mV (0,6 mV – 1,0 mV) * A onda Q tem duração < 0,03 s e amplitude de cerca de 1/4 da onda R * A onda S tem duração < 0,04 s e amplitude de cerca de 2/5 * da onda R |
| **Onda T** | * Deflexão obtusa da linha isoelétrica * Segue o complexo QRS * Duração < 0,2 s * Amplitude < 0,4 mV (0,2 mV – 0,4 mV) * Assimétrico com o pico perto do fim |
| **Onda U** | * Pequena deflexão obtusa (às vezes inexistente) da linha isoelétrica * É localizada perto do fim da onda T * Duração < 0,04 s * Amplitude < 0,5 mm |
| **Intervalo PR** | * É medido do começo da onda P até o começo do complexo QRS * Duração normal de 0,12 s – 0,22 s |

Fonte: Adaptada FOTIADIS et al. (2006)

A Figura 8, ilustra o espectro de potência relativo do sinal de ECG, complexo QRS, ondas P e T, atrito do eletrodo com a pele e ruídos musculares.

**Figura 8** - Espectro de potência relativo do complexo QRS, ondas P e T, atrito do eletrodo e ruído muscular



Fonte: Tompkins (1993)

* + 1. Protótipo ECG

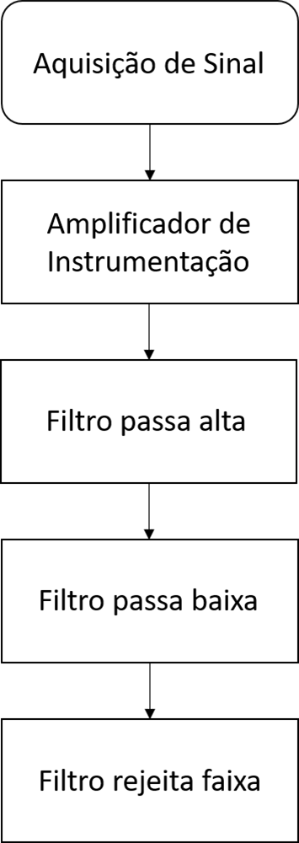
O desenvolvimento do protótipo de ECG tem como base o projeto de um circuito eletrocardiógrafo produzido por Thiago Moreira, engenheiro eletrônico da Força Aérea Brasileira. Especialista em eletrônica, telecomunicações e sistemas embarcados.

O protótipo possui os seguintes módulos:

* aquisição de sinal;
* amplificador de instrumentação;
* filtro passa alta;
* filtro passa baixa;
* filtro Notch.

Esses módulos são conectados em serie conforme ilustrado na Figura 9:

**Figura 9** - Diagrama dos módulos conectados para o funcionamento do ECG



Fonte: Elaborada pelos autores (2020)

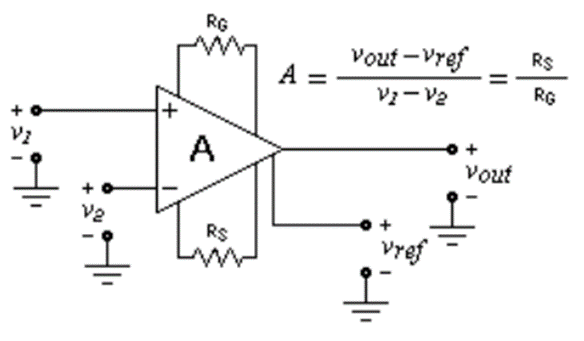
* + - 1. Aquisição de Sinal

A coleta dos sinais eletrocardiográficos é realizada por meio de eletrodos cardíacos [[4]](#footnote-4) descartáveis, no protótipo são utilizados 3 eletrodos, um em cada braço e um na perna direita ou no abdômen mais próximo a perna direita, realizando a aquisição da atividade elétrica no paciente.

* + - 1. Amplificador de Instrumentação

Segundo Bowers (1990), o termo parece ter surgido no fim da década de 60 denotando geralmente um amplificador que executava uma conversão de um sinal diferencial na entrada para um sinal de saída simples com um ganho precisamente definido. Desta maneira, um amplificador de instrumentação pode ser ilustrado como mostra o diagrama da Figura 10, que representa um circuito na qual as entradas se aplicam dois sinais de tensão ( e ) e que disponibiliza em sua saída a diferença entre estes dois sinais ampliada por um ganho .

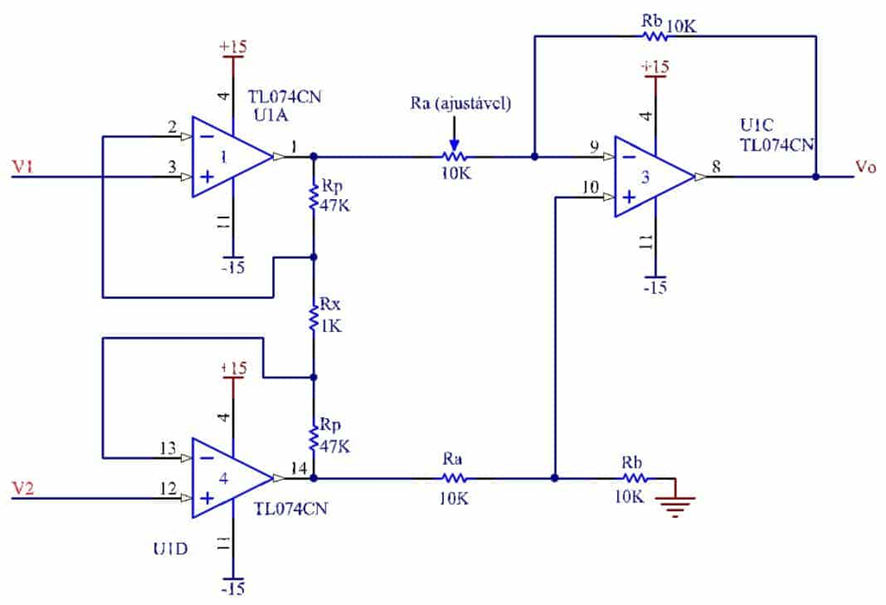
**Figura 10** - Diagrama de um amplificador de instrumentação e sua expressão de ganho



Fonte: Grebene (1984)

O ganho precisamente definido a que se referiu Bowers é geralmente dado pela razão entre duas resistências (RS e RG). Outra característica do amplificador de instrumentação é que ele opera sempre em malha fechada e, portanto, não requer realimentação externa.

**Figura 11** - Amplificador operacional utilizando 3 amplificadores utilizado no protótipo



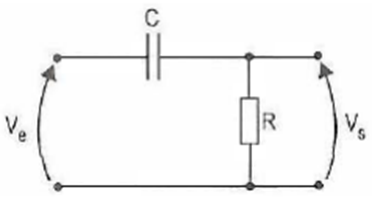
Fonte: Moreira (2016)

Segundo Moreira (2016), o amplificador ilustrado na Figura 11 tem ganho de 95 com o resistor Ra ajustado em 10k. Devido a imprecisão dos componentes o resistor Ra deve ser ajustado buscando o melhor desempenho.

* + - 1. Filtro Passa Alta

Para ondas senoidais de frequências altas, a reatância capacitiva[[5]](#footnote-5) assume valores baixos em comparação ao valor da resistência. Dessa maneira, a tensão de saída será praticamente igual à tensão de entrada. Para frequências baixas, a reatância capacitiva assume valores altos em comparação com o valor da resistência, atenuando a tensão de saída para um valor praticamente nulo. Deste modo, o filtro permite a passagem de sinais de frequências altas, sendo por isto denominado filtro passa alta (CAPUAN; MARINO, 2005).

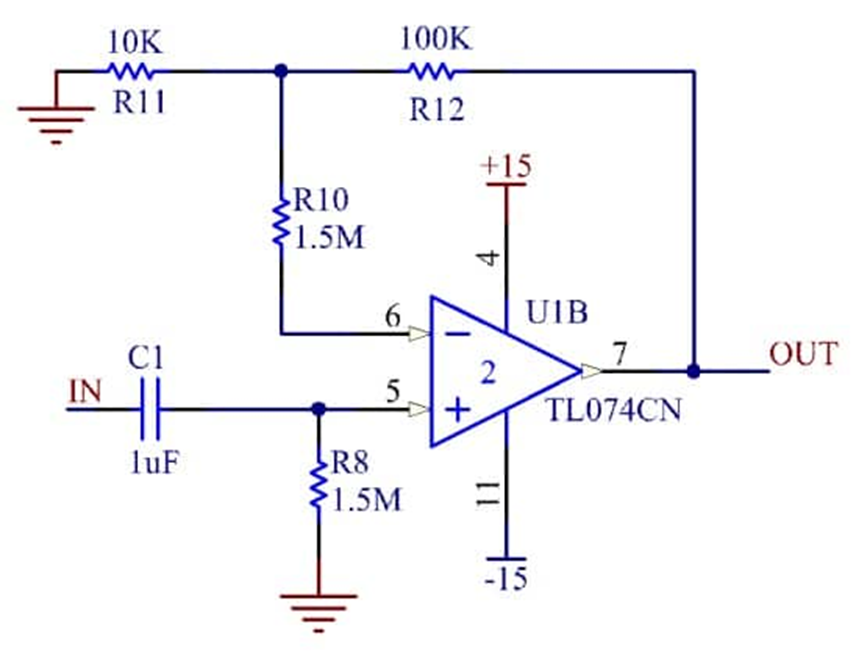
**Figura 12** - Esquematização do filtro passa alta



Fonte: Capuan; Marino (2005)

No protótipo desenvolvido por Moreira (2016) é utilizado um sinal de corte próximo a de 0,1 Hz e adicionado um amplificador de ganho 10. A Figura 13, ilustra o circuito do filtro passa alta com amplificador utilizado no protótipo.

**Figura 13** - Filtro passa alta com amplificador utilizado no protótipo

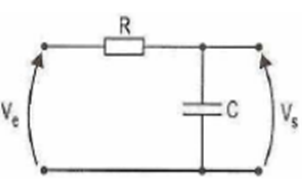


Fonte: Moreira (2016)

* + - 1. Filtro Passa Baixa

Para ondas senoidais de frequências baixas, a reatância capacitiva assume valores altos em comparação ao valor da resistência. Dessa maneira, a tensão de saída será praticamente igual à tensão de entrada. Para frequências altas, a reatância capacitiva assume valores baixos em comparação com o valor da resistência, atenuando a tensão de saída para um valor praticamente nulo. Desse modo, o filtro permite a passagem de sinais de frequências baixas, sendo por isso denominado filtro passa baixa (CAPUAN; MARINO, 2005).

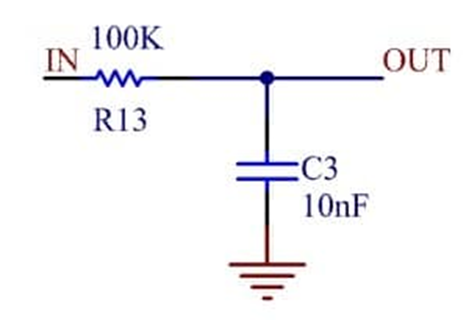
**Figura 14** - Esquematização do filtro passa baixa



Fonte: Capuan; Marino (2005)

No protótipo é utilizado um sinal de corte aproximado de 150 Hz. A Figura 15, ilustra o circuito do filtro passa baixa.

**Figura 15** - Filtro passa baixa utilizado no protótipo

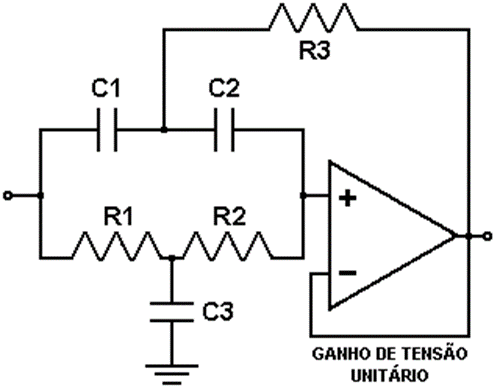


Fonte: Moreira (2016)

* + - 1. Passa Notch ou Rejeita-Faixa

Um Filtro Passivo Rejeita-Faixa é um circuito que atenua, “impede” a passagem de sinais de tensão e corrente com frequências situadas numa faixa intermediária, “permitindo” a passagem de sinais com frequências acima ou abaixo dessa faixa. Essa faixa intermediária é delimitada por uma Frequência de Corte Inferior e uma Frequência de Corte Superior (MUSSOI, 2004). A Figura 16, ilustra o circuito do filtro rejeita-faixa.

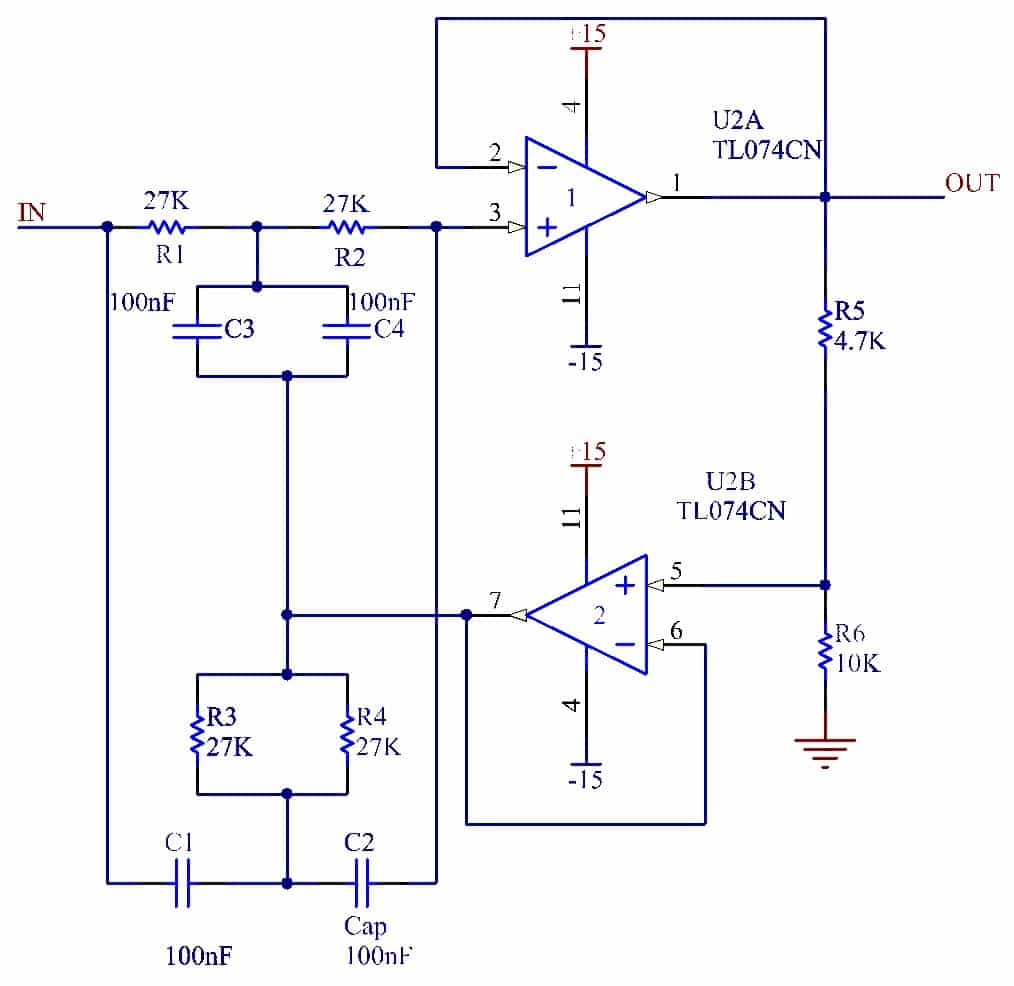
**Figura 16** - Esquematização do filtro passivo rejeita-faixa



Fonte: Physika (2012)

O filtro rejeita-faixa é utilizado para eliminar a interferência de sinal da energia elétrica que é de 60 Hz. O filtro utilizado tem como topologia conhecida como duplo T, onde os amplificadores funcionam como buffer. A Figura 17, ilustra o circuito do filtro rejeita faixa utilizado em nosso protótipo.

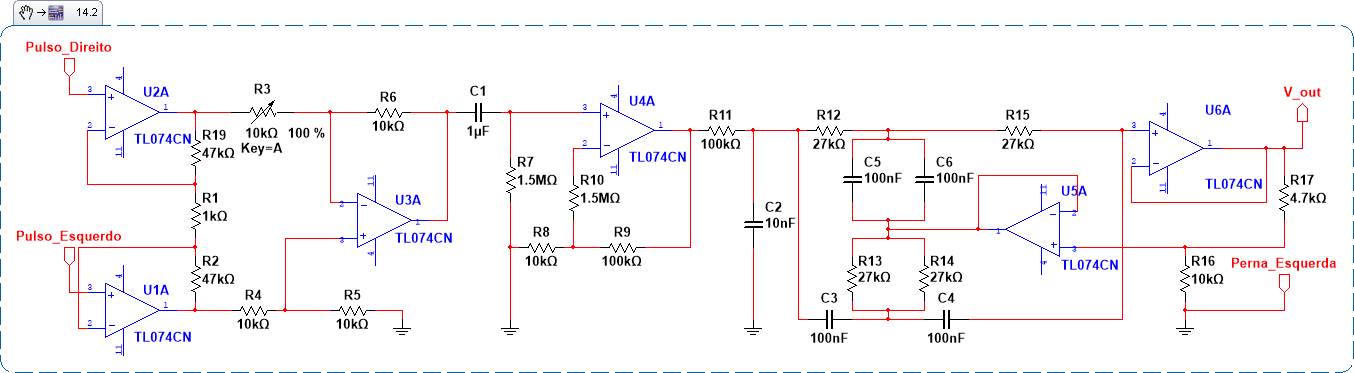
**Figura 17** - Filtro rejeita faixa utilizado no protótipo



Fonte: Moreira (2016)

Após a construção de todos os equipamentos, o circuito deve ser como ilustrado na Figura 18.

**Figura 18** - Esquematização do protótipo de ECG completo



Fonte: Adaptado Moreira (2016)

* + 1. Taxa de Amostragem

Para a captação de sinais de ECG por qualquer equipamento eletrocardiográfico é necessário que seja possível fazer a conversão dos sinais analógicos para sinais digitais que possam ser processados em um microcomputador. Idealmente, a taxa de amostragem deve ser 250 Hz ou superior, com resolução de digitalização igual ou superior a 12 bits (BARBOSA; MONTEIRO; WOLF, 2004).

* + 1. Software de Simulação

Uma parte muito importante no desenvolvimento de protótipos é a utilização de softwares de simulação, nesse caso especificamente, na simulação de circuitos elétricos/eletrônicos para elaboração do hardware. Através da simulação é possível projetar um modelo computacional com o objetivo de experimentar e entender o comportamento de sistemas reais.

O software Multisim é um ambiente de simulação padrão da indústria de software de projeto de circuitos para o ensino e pesquisa de eletrônica digital, analógica e de potência (NATIONAL INSTRUMENTS CORPORATION, 2020).

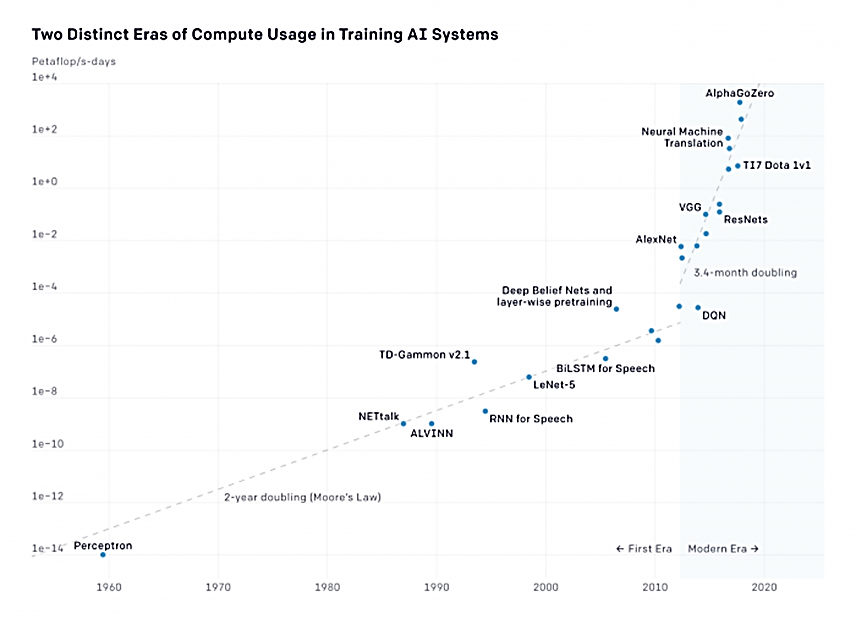
Segundo a National Instruments Corporation (2020), o Multisim possibilita criar esquemáticos interativos, permitindo analisar instantaneamente o comportamento de circuitos eletrônicos, o que faz com que pesquisadores, projetistas e estudantes consigam reduzir a interação dos protótipos físicos com as placas de circuitos impresso e assim reduzir o custo de desenvolvimento.

* 1. INTELIGÊNCIA ARTIFICIAL

A Inteligência Artificial (IA) vem ganhando cada vez mais espaço em nossa sociedade (MARGETTS; DOROBANTU, 2019). Hoje na medicina, já existem múltiplas empresas oferecendo serviços por meio da IA. Algumas empresas que oferecem este tipo de serviço, por exemplo, a Eye Diagnosis - IDx, que desenvolve algoritmos de IA para detectar problemas oculares, a HEARTFLOW que utiliza RNA para criar modelos 3D personalizados do coração de um paciente específico, a HEARTVISTA que faz diagnóstico por meio de imagens de ressonância nuclear magnética, entre muitas outras.

A primeira publicação sobre o assunto, foi em 1943, no artigo intitulado “*a logical calculus of the ideas immanent in nervous activity*” (MCCULLOCH; PITTS, 1943), quando ainda não existia processamento o suficiente. A partir de 1959 até 2012, seguindo a Lei de Moore o poder de processamento dobrava a cada 2 anos acompanhando a evolução do processamento da época (HAO, 2019) e também segundo a autora, de 2012 em diante, os modelos de IA vem dobrando a necessidade de processamento a cada 3.4 meses, ou seja, sete vezes mais rápido do que os anos anteriores.

**Figura 19** - Evolução da necessidade de processamento para algoritmos mais pesados de Inteligência Artificial



Fonte: OpenAI (2015)

* + 1. Análise dos Dados

Antes de iniciar qualquer modelo de Machine Learning, primeiro vem toda a análise dos dados, e a mesma não existe sem a estatística.

Estatística é muito importante quando falamos de dados, porém dados sozinhos não são interessantes. É a interpretação dos dados que estamos interessados (MAHESH, 2019).

* + - 1. Testes de Hipóteses

Um teste de hipótese avalia duas declarações mutuamente exclusivas sobre uma população[[6]](#footnote-6) para determinar qual a declaração é mais bem suportada pelos dados da amostra (MAHESH, 2019).

De acordo com LEE (2019), um teste de hipótese nos retorna duas saídas, a primeira sendo a estatística que é um valor atribuído para a hipótese e dependendo daquele valor a hipótese pode ter diferentes interpretações, a segunda sendo o p-valor que é uma probabilidade, e esta serve para definir se a hipótese é aceita ou rejeitada.

A princípio será utilizado os testes de Kolmogorov-Smirnov e o Qui-Quadrado de aderência.

* + - 1. Teste de Kolmogorov-Smirnov

De acordo com Arnold e Emerson (2011), o teste de Kolmogorov-Smirnov (EQ. 1) consiste em analisar duas amostras e verificar se suas populações possuem uma distribuição parecida.

Na (EQ. 1) é o supremo de x, ou seja, o menor número do limite superior, representa a função de distribuição acumulada assumida para os dados, ou seja, se ela for uma distribuição normal e a função de distribuição acumulada empírica dos dados.

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 1) |

O teste é utilizado para verificar se o sinal real e o sinal previsto pertencem a uma distribuição normal, com isso pode-se fazer uma verificação de correlação que vai depender se a distribuição é normal ou tende ao caos.

* + - 1. Teste Qui-Quadrado de Aderência

O teste de Qui-Quadrado (EQ. 2) é um procedimento estatístico utilizado por pesquisadores para examinar as diferenças entre duas variáveis categóricas de mesma população (FOLEY, 2019).

Na (EQ. 2) é o número de observações do tipo i, é a frequência esperada do tipo i e é o número de células na tabela.

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 2) |

De acordo com Foley (2019), é bom quando é necessário estimar o quanto a distribuição observada se parece com uma distribuição esperada. Com isso, é possível determinar a distância e o quanto as curvas estarão parecidas.

* + - 1. Correlações

Além dos testes de hipóteses, também é utilizado correlações para determinar se existe relação estatística entre duas variáveis.

De acordo com Hauke e Kossowski (2011), para identificar estas relações podemos utilizar dois métodos, o método de Charles Spearman ou o método de Karl Pearson, a principal diferença entre os dois métodos é que o método de Pearson trabalha melhor com distribuições normais, onde é identificado por meio do teste de Kolmogorov-Smirnov citado anteriormente, enquanto a correlação de Spearman trabalha melhor com outras distribuições.

* + 1. Machine Learning

Um modelo de Machine Learning supervisionado aprende com dados históricos, e estes dados são utilizados para o treinamento[[7]](#footnote-7) e validação do modelo.

Depois da validação e os testes serem feitos, e chegar na comprovação de que o modelo está eficaz, então este modelo deve ser produtizado[[8]](#footnote-8) dentro de uma infraestrutura em nuvem. Para realizar a validação é necessário determinar uma métrica de erro para o modelo.

* + - 1. Métricas de Erro

Um modelo de Machine Learning modula sua função a partir de uma métrica de erro, ele tenta minimizar o erro para poder performar melhor. Existem vários tipos de métricas de erro e cada uma tem sua particularidade.

Algumas das métricas de erro existentes para modelos lineares são a Média Absoluta do Erro ou Média do Erro Quadrático.

* + - * 1. *MAE ou Média Absoluta do Erro*

O MAE (EQ. 3) calcula o residual para cada ponto de dado, utilizando apenas o valor absoluto da diferença de cada ponto para que valores positivos e negativos não se cancelem, feito isso é tirada a média destes residuais (DATAQUEST, 2020).

Na (EQ. 3) o é total de pontos de dados, o valor atual e o valor previsto.

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 3) |

* + - * 1. *MSE ou Média do Erro Quadrático*

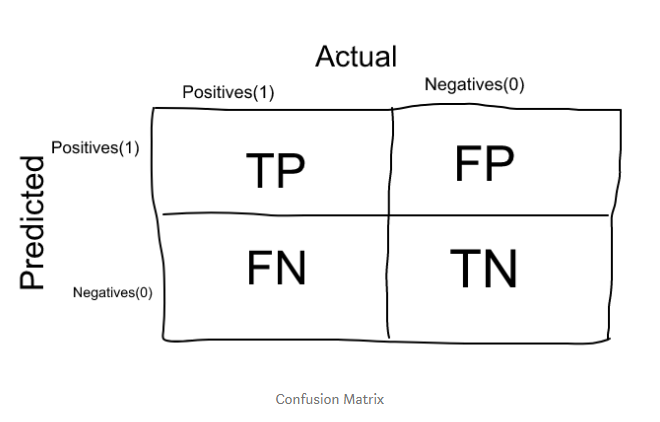
O MSE (EQ. 4) é bem parecido com o MAE, porém, ao invés de utilizar o valor absoluto ele eleva todas as diferenças ao quadrado antes de somar, assim então tirando a média (DATAQUEST, 2020).

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 4) |

* + - * 1. *Matriz de Confusão*

A matriz de confusão é uma das métricas mais fáceis e intuitivas usadas para encontrar a precisão, a especificidade do modelo, entre outros. Ela é utilizada para problemas de classificação onde a saída pode ser dois ou mais tipos de classes (SUNASRA, 2017).

**Figura 20** - Matriz de confusão



**Fonte: Sunasra (2017)**

* + - * 1. *Precisão*

A precisão é uma medida que nos conta se a proporção de verdadeiros positivos, ou seja, a proporção de que algo que a previsão disse que ia acontecer, realmente aconteceu (SUNASRA, 2017).

Na (EQ. 5) em que TP é a quantidade de verdadeiros positivos e FP é a quantidade de falsos positivos.

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 5) |

* + - * 1. *Especificidade*

A especificidade é uma medida que nos conta se a proporção de verdadeiros negativos, ou seja, a proporção de que algo que a previsão afirmou que não ia acontecer, realmente não aconteceu (SUNASRA, 2017).

Na (EQ. 6) temos que TN é a quantidade de verdadeiros negativos e FP é a quantidade de falsos positivos.

|  |  |  |
| --- | --- | --- |
|  |  | (EQ. 6) |

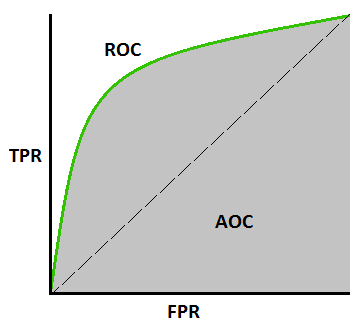
* + - * 1. *Curva AUC-ROC*

A curva AUC-ROC é uma análise utilizada como métrica de performance para problemas de classificação. ROC do inglês *Receiver Operating Characteristics* é um método gráfico para avaliação, organização e seleção de sistemas de diagnóstico e/ou predição (EGAN, 1975).

Enquanto AUC do inglês *Area Under the Curve* mede a separabilidade do modelo. Essa análise permite identificar o quanto o modelo é capaz de distinguir entre classes. Quanto maior a AUC, o modelo consegue prever melhor que 0 é 0 e que 1 é 1 (NARKHEDE, 2018).

Por exemplo, quanto maior a AUC mais o modelo consegue distinguir se um paciente com probabilidade de MSC, realmente terá a MSC. A Figura 21 ilustra uma curva AUC-ROC, com a taxa de verdadeiro positivo (*True Positive Rate*, TPR, em inglês) que é a probabilidade de detecção, enquanto a taxa de falso positivo (*False Positive Rate,* FPR, em inglês) é a probabilidade de alarme falso.

**Figura 21** - Curva AUC-ROC

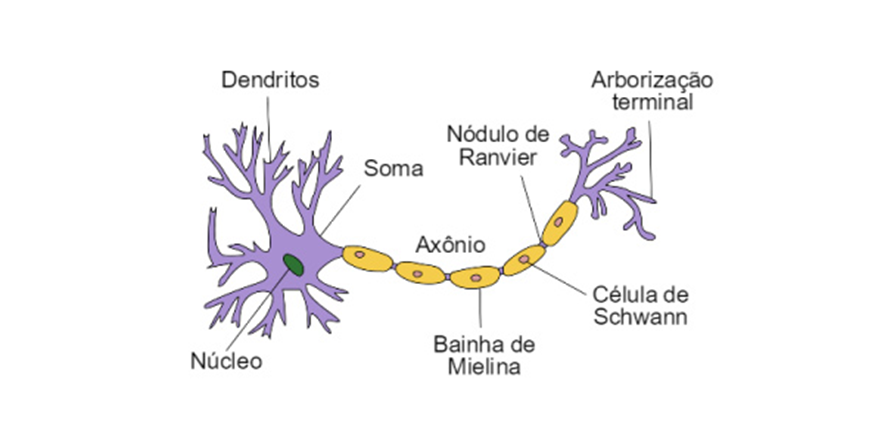


**Fonte: Narkhede (2018)**

* + 1. Redes Neurais
       1. O Neurônio

Para o entendimento de uma rede neural, primeiro precisamos entender o neurônio biológico. O neurônio é uma célula nervosa, estrutura básica do sistema nervoso, comum à maioria dos vertebrados. Os neurônios são células altamente estimuláveis, que processam e transmitem informação através de sinais eletroquímicos. Uma das suas caraterísticas é a capacidade das suas membranas plasmáticas gerarem impulsos nervosos. A maioria dos neurônios, tipicamente, possui o corpo celular e dois tipos de prolongamentos citoplasmáticos, as dendrites e os axônios (MOREIRA, 2013). A Figura 21, ilustra a estrutura simplificada de um neurônio biológico.

**Figura 22** - Estrutura básica simplificada de um neurônio



Fonte: Moreira (2013)

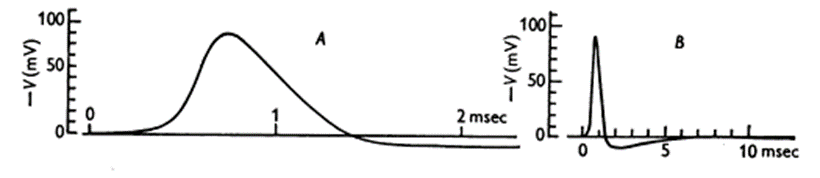
Segundo Moreira (2013), o corpo celular contém o núcleo e suas organelas, as dendrites são prolongamentos que recebem os estímulos provenientes de outros neurônios. O axônio é o prolongamento mais longe que transmite os impulsos nervosos provenientes do corpo celular e as terminações do axônio contém sinapses, que são estruturas especializadas onde são libertadas substâncias químicas.

Quando a terminação do axônio de um neurônio estabelece ligações com as dendrites ou corpo celular de um outro neurônio, as membranas modificam-se e formam uma sinapse, que permite que o impulso nervoso seja conduzido de um neurônio para o seguinte (MOREIRA, 2013).

Como em todas as células, o citoplasma do neurônio tem um excesso de carga negativa. A voltagem no interior do neurônio é geralmente 60-70 milivolts (mV) mais negativa que o exterior da célula. Esta diferença de carga entre o meio extracelular e o meio intracelular gera uma diferença de potencial elétrico entre as duas faces da membrana chamado de potencial de membrana, que quando a célula não está transmitindo impulsos nervosos é da ordem dos -70 mV – potencial de repouso (MOREIRA, 2013).

O sinal negativo indica, como referido anteriormente, que o interior das células tem maior carga negativa do que o exterior. O neurônio é sensível a qualquer químico ou físico que provoque uma alteração no potencial de repouso da membrana. A alteração mais extrema que pode ocorrer no potencial de membrana é o impulso nervoso (ou potencial de ação), que é uma rápida alteração do potencial elétrico, em que por breves instantes (1 ou 2 milissegundos) o interior da célula torna-se mais positivo que o exterior (MOREIRA, 2013). A Figura 22, ilustra o potencial de ação de neurônio biológico.

**Figura 23** - Primeira imagem do potencial de ação medido na tela de um osciloscópio por Hodgkin e Huxley



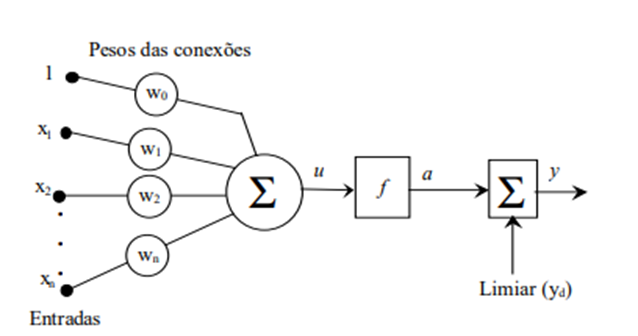
Fonte: Adaptado Hodgkin e Huxley (1952)

A transmissão de um impulso nervoso é um exemplo de uma resposta do tipo “tudo-ou-nada”, isto é, o estímulo tem que ter uma determinada intensidade para gerar um potencial de ação. O estímulo mínimo necessário para desencadear um potencial de ação é o estímulo limiar, e uma vez atingido este limiar, o aumento de intensidade não produz um potencial de ação mais forte, mas sim um maior número de impulsos por segundo. O potencial de ação gerado na membrana estimulada propaga-se à área vizinha, conduzindo à sua despolarização e assim por diante (MOREIRA, 2013).

Com o conhecimento básico do funcionamento de um neurônio foi possível a criação de um modelo matemático simplificado e generalista que pode representá-lo (VIJAY; et al, 2019).

A Figura 23, é um modelo matemático simplificado de um neurônio, que recebe sinais de outro neurônio. Cada sinal de entrada que vem do outro neurônio recebe um peso gerado pelo neurônio que vemos.

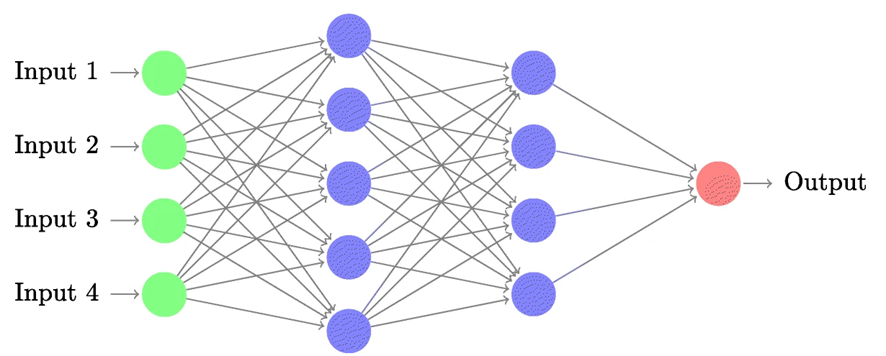
**Figura 24** - Modelo matemático simples de um neurônio conhecido como Perceptron



Fonte: Zuben (1993)

Portanto, uma RNA é uma coleção de neurônios artificiais, representados pelo modelo matemático descrito por McCulloch e Pitts. Cada conexão, representando a synapse do cérebro podendo transmitir sinal de um neurônio a outro (HOLMGREN, 2019). A Figura 24, ilustra a esquemática de uma RNA com 4 entradas, duas camadas internas e uma única saída.

**Figura 25** - A esquemática de uma RNA compostas por duas camadas e apenas um neurônio de saída



Fonte: Holmgren (2019)

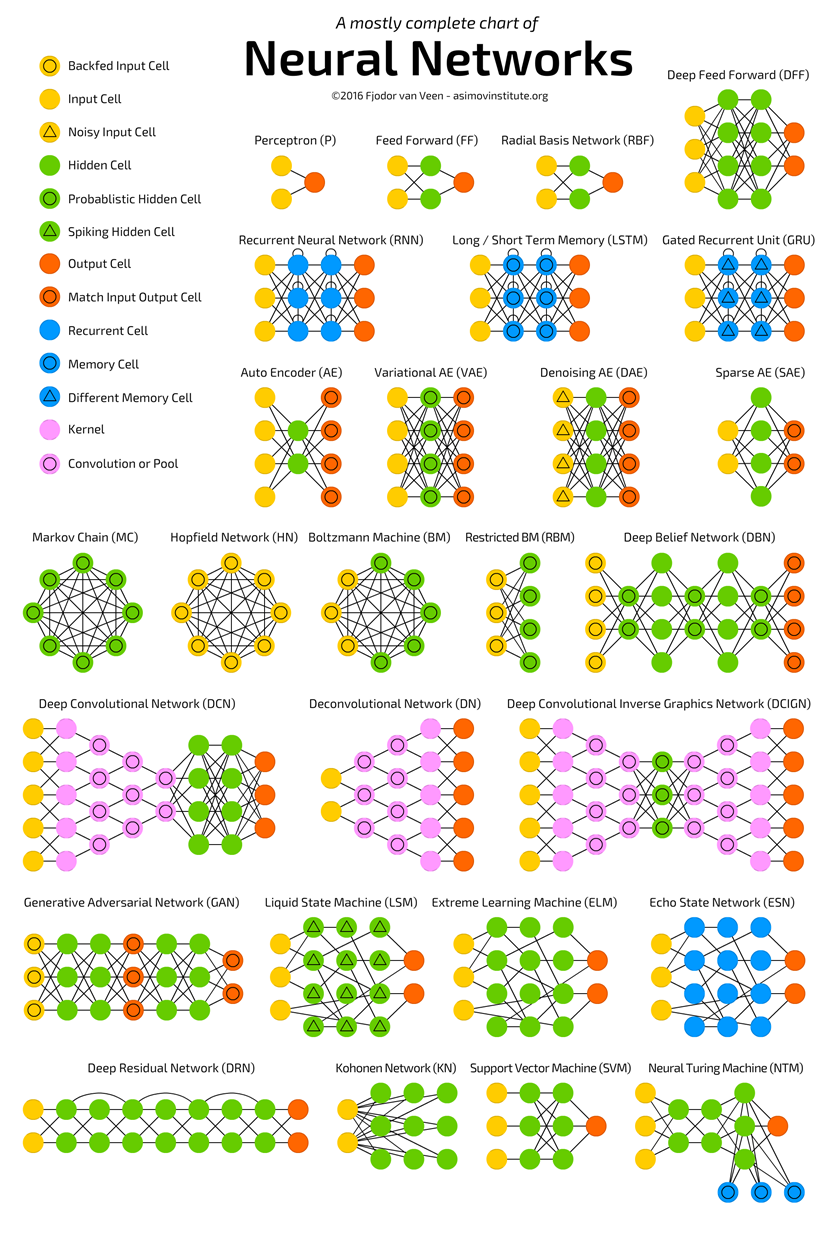
Segundo (GONÇAÇVES, 2018) a topologia de uma rede neural diz respeito à disposição dos neurônios na rede, como são estruturados. A topologia da rede está diretamente ligada ao tipo de algoritmo de aprendizagem utilizado.

Pode-se observar na figura feita por Holmgren, uma camada verde que é chamada de camada de entrada, as camadas azuis são chamadas de camadas escondidas e é lá onde estão nossos neurônios, a camada vermelha é a camada de saída, que é o resultado de saída do nosso modelo.

* + - 1. Tipos de Redes Neurais

Existem muitos tipos de RNA, com possibilidades distintas, sendo que cada arquitetura auxilia um tipo específico de dado e/ou um problema a ser solucionado (VEEN, 2017). A Figura 25, ilustra as principais topologias de RNA atualmente.

**Figura 26** - Topologia das mais variadas arquiteturas de uma RNA



Fonte: Veen (2017)

Como ferramenta principal para desenvolvimento de uma RNA existe o TensorFlow desenvolvido pela Google e para sua utilização é necessário o conhecimento da linguagem de programação Python, lógica, estrutura de algoritmos e conhecimento de estatística.

# MATERIAIS E MÉTODO

Esta seção apresenta os métodos adotados para a o desenvolvimento do projeto, baseado em simulações e pesquisas bibliográficas. A escrita do documento segue as normas ABNT no Guia de Normalização de monografias, dissertações e teses, do Centro Universitário Senac (SENAC, n.d.).

Este projeto é fundamentado por meio de pesquisa quantitativa e o modelo para implementação do estudo é baseado no método experimental.

O método experimental consiste, especialmente, em submeter os objetos de estudo à influência de certas variáveis, em condições controladas e conhecidas pelo investigador, para observar os resultados que a variável produz no objeto (GIL, 2008).

* 1. METÓDO

O projeto é divido em duas partes principais, o desenvolvimento do hardware que engloba todo o circuito elétrico e eletrônico que é conectado aos pacientes para coletar os sinais elétricos cardíacos e a parte de software que recebe estes dados através de uma infraestrutura em nuvem para analisar e processa-los através de RNAs, e apresentar os *outputs* relacionados ao risco de morte súbita cardíaca do paciente em um determinado período.

* 1. PLANEJAMENTO DO PROTÓTIPO
     1. Hardware

A construção do protótipo de ECG é baseada no modelo referenciado por Moreira (2016), que será divido em módulos e durante a montagem de cada módulo, será realizado simulações com ondas semelhantes ao sinal cardíaco real e aplicação de ruídos, afim de testar o funcionamento dos amplificadores e dos filtros e identificar possíveis falhas e anormalidades. Após os testes dos módulos, serão realizados testes do circuito completo e analisado se o comportamento do circuito condiz com o esperado para o estudo.

Após a construção, novas simulações devem ser realizadas afim de verificar se cada componente está funcionando corretamente e analisar se a saída do sinal corresponde com a expectativa.

Feito essa aquisição do sinal, os dados são encaminhados para a nuvem onde são armazenados e analisados.

* + 1. Utilização de dados

Serão realizados treinamentos das RNAs utilizando bases de dados de diversos pacientes que forneceram seus sinais eletrocardiográficos para pesquisadores do MIT e que foram posteriormente disponibilizados para a comunidade cientifica através da plataforma PhysioNet. Portanto, são dados secundários e não há necessidade de autorização do comitê de ética da área da saúde para a sua utilização.

As bases utilizadas possuem dados de pacientes saudáveis e doentes que foram ou não acometidos por MSC, os links de acesso as bases estão disponíveis no apêndice.

* + 1. Software

Serão realizados treinamentos das RNAs utilizando bases de dados de diversos pacientes que forneceram seus sinais eletrocardiográficos para pesquisadores do MIT e posteriormente disponibilizados para a comunidade cientifica através da plataforma *PhysioNet.* Portanto, são dados secundários e não há necessidade de autorização do comitê de ética da área da saúde para a sua utilização. Através desses dados, com o intuito de testar a viabilidade dos modelos e escolher o melhor para ser utilizado, para validar o sinal previsto comparando-o com o sinal real serão utilizados testes de hipótese estatísticos, o já citado, teste de Kolmogorov-Smirnov para verificar se a distribuição tende a normalidade. Verificado o teste, então poderemos usar as correlações para verificar se os sinais estão parecidos, por fim, será utilizado o teste de Qui-Quadrado para verificara distância entre os sinais.

Após as validações, será construída a infraestrutura em nuvem, que será conectada ao hardware. Todas as partes da infraestrutura devem ser concebidas ao mesmo tempo, depois da infraestrutura elaborada ela deve se conectar no hardware e então ter o modelo do paciente treinado com seu output saindo para *dashboards* de visualização, de fácil acesso para médico e paciente.

* + - 1. Cloud: Computação em Nuvem

Seguindo as tendências atuais de mercado este TCC será elaborado dentro de uma infraestrutura em nuvem. O serviço em nuvem nos possibilita trabalhar com grandes massas de dados sem a preocupação com servidores físicos, pois oferece uma infraestrutura pronta e que se redimensiona de acordo com a necessidade e utilização.

* + - 1. Serviço de Nuvem

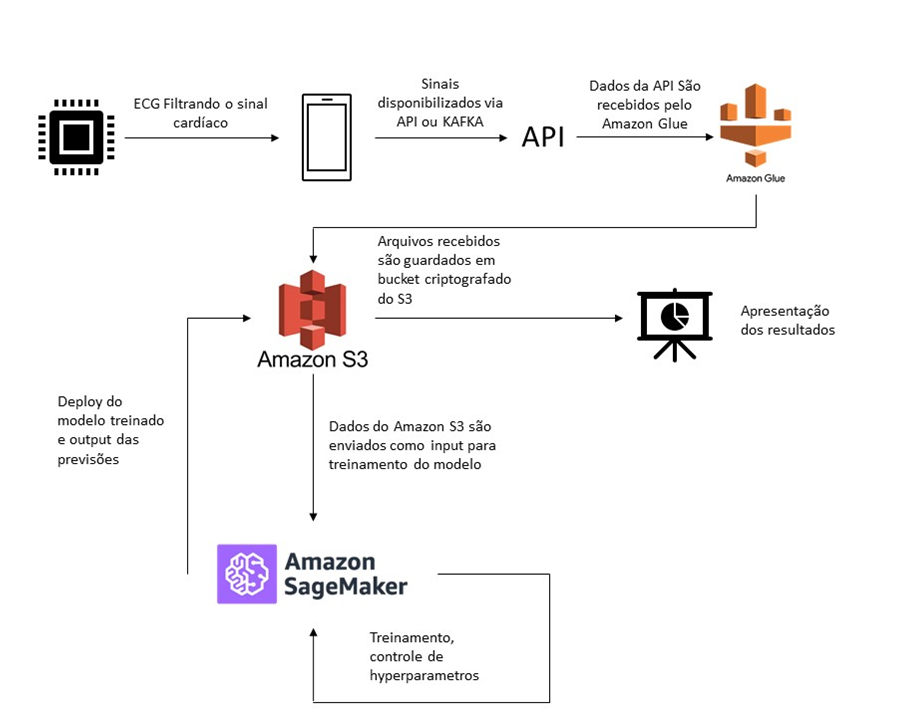
Atualmente, quando falamos de computação em nuvem existem alguns grandes nomes, que são: Google Cloud Platform, Amazon Web Services (AWS) , Microsoft Azure e IBM Cloud.

Dentre tantos recursos disponíveis a AWS se destaca no quesito análise e processamento de dados, apesar de ser um pouco mais complexa, com inúmeros serviços e meios de segurança para se gerenciar, ela atende a todos os requisitos da infraestrutura necessária e ainda oferta recursos para diminuir o custo da computação em nuvem.

* + 1. Arquitetura

Com isto, será criada a arquitetura representada na Figura 26, que ilustra toda a arquitetura do projeto e o fluxo de dados da solução, desde a coleta dos dados pelo ECG até o processamento e resposta.

**Figura 27** - Arquitetura utilizada no projeto



Fonte: Elaborada pelos autores (2020)

**Amazon S3**: É um serviço de armazenamento de objetos por meio de uma interface. Os dados podem ser divididos de maneira inteligente e a maioria dos serviços são dependentes dele para funcionar.

**Amazon Glue**: Serviço de ETL (Extract, Transform and Load) da AWS, ou seja, é um serviço de integração onde conseguimos capturar um dado, fazer as transformações necessárias utilizando Python ou Spark[[9]](#footnote-9) e depois carregar esses dados em algum sistema da AWS ou fora da mesma.

**Amazon SageMaker**: Serviço de Inteligência Artificial da AWS que visa fornecer um meio de trabalho mais fácil para cientistas de dados, enquanto oferece opções simples de automatização e Deploy[[10]](#footnote-10) de modelos de Machine Learning.

Portanto, foi desenvolvido um hardware que captura os sinais cardíacos e envia para o celular do paciente, estes dados são disponibilizados por API[[11]](#footnote-11) ou Streaming[[12]](#footnote-12) para que o Amazon Glue faça a extração e o carregamento dentro do S3, que é o meio de armazenamento principal, ou seja, um Object Schema[[13]](#footnote-13), o Amazon SageMaker por sua vez, treina o modelo utilizando o dado do S3 e toda a otimização de hiperparâmetros[[14]](#footnote-14) será feita por ele, e por fim, o modelo está pronto para uso fornecendo os dados para um outro objeto de *outputs* dentro do S3 que é disponibilizado ao médico e pacientes através de um método de visualização a ser definido.

* + 1. Marcadores de Morte Súbita

Através da análise do eletrocardiograma é possível extrair inúmeros marcadores de morte súbita como citado anteriormente, sendo que existem alguns marcadores conhecidos e principais na medicina, conforme apresentado no Quadro 1.

**Quadro 1** - Principais marcadores de MSC

|  |  |
| --- | --- |
| **Despolarização eletrocardiográfica, distúrbio de repolarização e marcadores de ECG da MSC:** | **Marcadores ECG de distúrbios de despolarização em SCD:** |
| O intervalo QT ou sístole elétrica:   * intervalo QT Prolongado; * intervalo QT-QTc encurtado. | Prolongamento da duração QRS (QRSd); |
| JT Prolongado; | Prolongação QRS em derivações precordiais direitas (de V1 até V3); |
| Dispersão QT Prolongada; | Uma onda S na liderança (0.1 mV e/ou 40 ms); |
| Síndrome da onda J; | Dispersão QRS; |
| Intervalo do pico até o fim da onda T; | Fragmentação QRS estreita ou larga (fQRS e fQRS larga); |
| Razão QT; | Ondas Epsilon; |
| Macro-ondas T; | Presença de potenciais ventriculares atrasados usando Eletrocardiógrafos de alta resolução. |
| Micro-ondas T; |  |

Fonte: RIERA; BARROS; SHENASA (2017)

Vamos elaborar um software para o monitoramento de possíveis marcadores de morte súbita que podem ser identificados através da utilização de RNA.

* + 1. Ferramenta de visualização

Como mencionado na arquitetura da solução os resultados precisão ser apresentados de alguma forma para os profissionais de saúde, portanto será utilizado a plataforma de análise visual Tableau para tornar mais visíveis e acessíveis as informações obtidas com o resultado do modelo de Inteligência Artificial.

O intuito é desenvolver um dashboard com as informações principais relacionadas a MSC como por exemplo, quantidade de casos analisados e previstos pelo modelo, dados dos sinais do paciente e data e hora da previsão.

* 1. SUJEITOS DA PESQUISA E AMOSTRA

Os sujeitos da pesquisa são pacientes indicados por médicos que suspeitam da necessidade de monitoramento cardíaco, por meio do protótipo fornecem amostras de sinais elétricos cardíacos servindo de base para o retreinamento do modelo, que através dos dados previamente disponibilizados é capaz de prever o risco da MSC para cada paciente separadamente.

Uma possível utilização se dá em Unidades de Tratamento Intensivo (UTIs), prevendo possíveis eventos inesperados e possibilitando a ação dos médicos que poderão salvar o paciente por questão de segundos.

Na área médica esse sistema é chamado de CAD, que vem do inglês Computer Aided Diagnostic encarregado pela detecção de anomalias gerando resultados que auxiliam o profissional de saúde à tomada de decisões rápidas e mais precisas. Porém, não exime o médico da responsabilidade do diagnóstico final.

* 1. CRONOGRAMA DAS ATIVIDADES

**Tabela 2 -** Cronograma de desenvolvimento do projeto

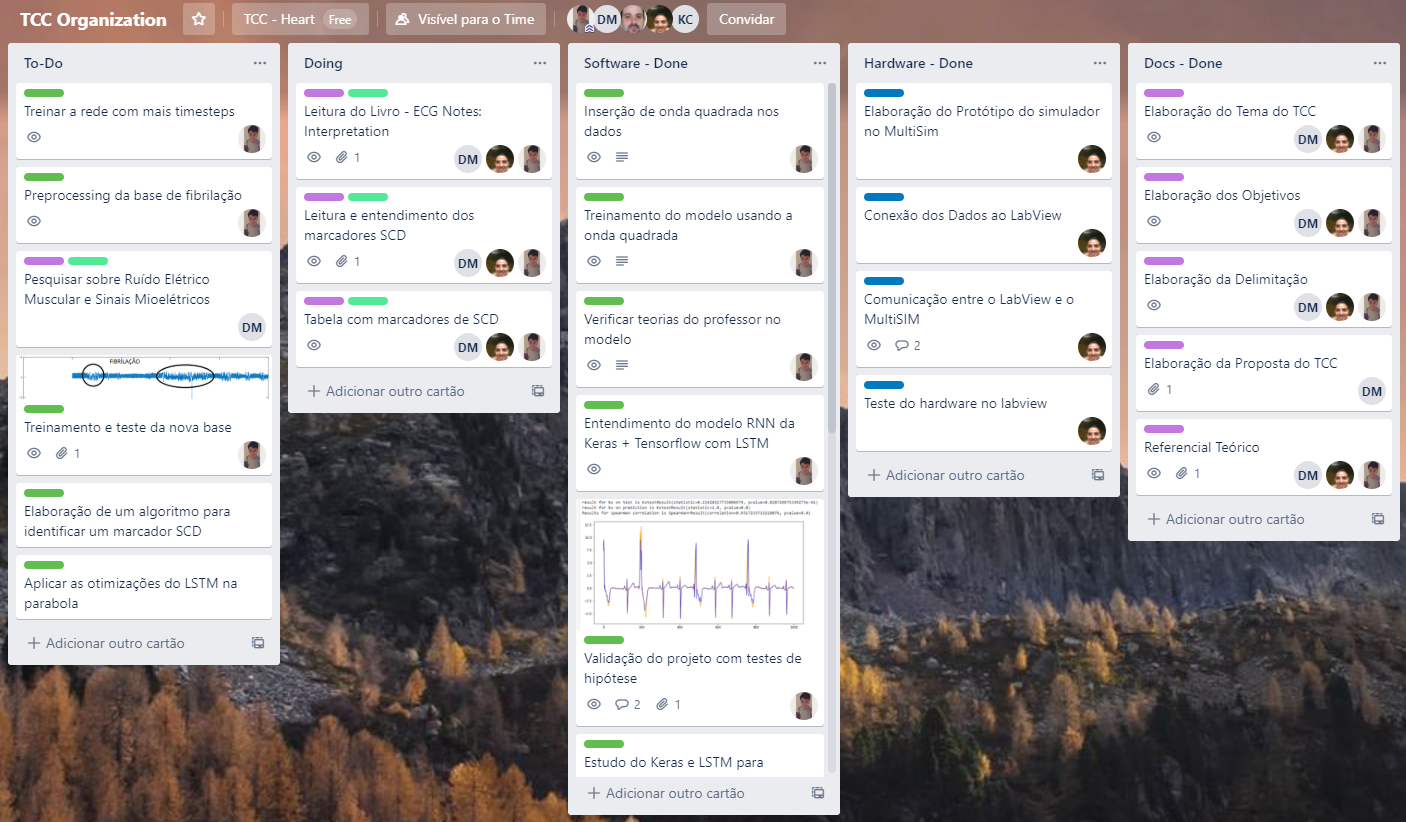
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Descrição da atividade** | **Atividade** | 2020 | | | | | | | | | | | | 2021 | | |
| JAN | FEV | MAR | ABR | MAI | JUN | JUL | AGO | SET | OUT | NOV | DEZ | JAN | FEV | MAR |
| **Desenvolvimento do Protótipo** | | | | | | | | | | | | | | | | |
| **Simulação por software** | **A1** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Donwnload e instalação dos softwares de simulação | A1.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Elaboração dos módulos na ferramenta | A1.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Simulação de cada módulo | A1.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Junção dos módulos para o protótipo completo | A1.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Simulação de toda estrutura | A1.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise dos sinais de saída | A1.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Elaboração do Hardware** | **A2** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Compra dos materiais necessários | A2.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Projeto e modelagem do hardware | A2.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Montagem do hardware | A2.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Modelo no Kicad | A2.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Testes de funcionamento | A2.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise dos sinais de saída | A2.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Elaboração do Software** | | | | | | | | | | | | | | | | |
| **Analise e entendimentos dos dados** | **A3** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Instalação e configuração das bibliotecas Python | A3.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Aquisição dos dados catalogados por bases médicas | A3.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise e entedimento inicial dos dados | A3.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Teste de viabilidade com modelos de RNN | A3.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise de sinais simples para o entendimento do modelo de RNN | A3.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Treinamento do modelo de RNN | A3.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Otimização de hiperparametros | A3.7 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Construção dã visualização no Tableau | A3.8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Construção da Infraestrutura na Nuvem** | **A4** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Configuração do S3 Amazon | A4.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Configuração do Glue Amazon | A4.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Verificação da consistencia dos dados no S3 | A4.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Deploy e acompanhamento do modelo** | **A5** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Configuração de hiperparametros no SageMaker | A5.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Treinamento do modelo no SageMaker | A5.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Deploy do modelo no SageMaker | A5.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Verificação da consistencia dos dados de saída do modelo | A5.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Testes de validação do modelo | A5.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Documentação e gerência** | | | | | | | | | | | | | | | | |
| **Elaboração do documento** | **A6** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pesquisa de referências bibliográficas | A6.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Escrita da monografia parcial (primeira etapa) | A6.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Escrita da monografia final | A6.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Elaboração da apresentação (primeira etapa) | A6.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Elaboração da apresentação final | A6.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Fonte: Elaborado pelos autores (2020)

* 1. PLANEJAMENTO DO DESENVOLVIMENTO DO PROTÓTIPO

Como forma de administrar o desenvolvimento do protótipo e os testes, foi utilizado como ferramenta o Kanban[[15]](#footnote-15), mais especificamente na plataforma do Trello, conforme ilustrado na Figura 27.

**Figura 28 -** Kanban organizado na ferramenta Trello



Fonte: Elaborada pelos autores (2020)

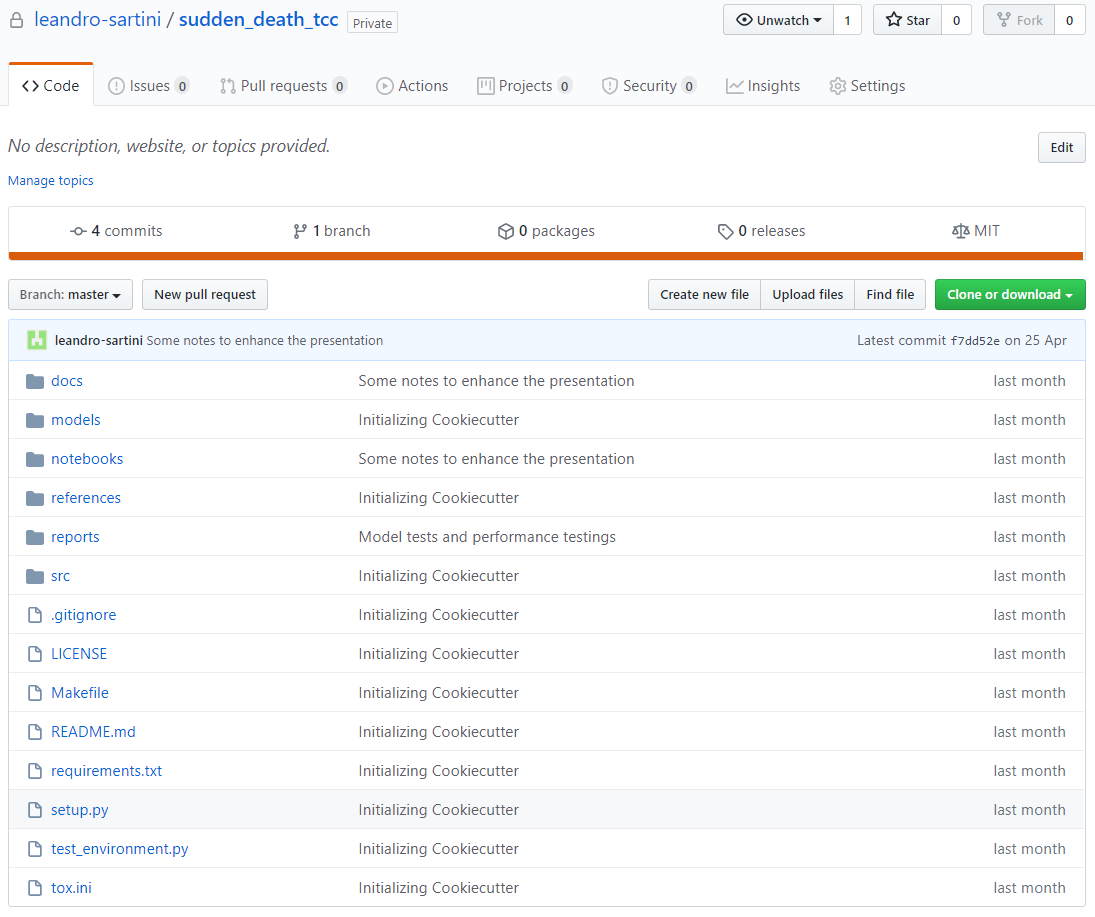
* + 1. Organização Durante o Desenvolvimento de Software

É comum em projetos de Ciência de Dados não haver muita organização, visto isso Audrey Roy criou uma biblioteca em Python chamada de cookiecutter em 2013. Ela cria projetos a partir de templates elaborados pela comunidade do Python e comunidade do cookiecutter.

Uma das comunidades do GIT[[16]](#footnote-16), conhecida como DrivenData elaborou uma estrutura de projetos para que Cientistas de Dados, se organizem melhor, de forma mais fácil não precisando existir o retrabalho toda vez que iniciar um projeto novo.

Além disso foi utilizado o GIT dentro da plataforma do GitHub para fazer o versionamento do projeto e deixar ele disponível para todos. A Figura 28, ilustra o GIT utilizado no projeto.

**Figura 29** - GitHub do TCC com o CookieCutter



Fonte: Elaborada pelos autores (2020)

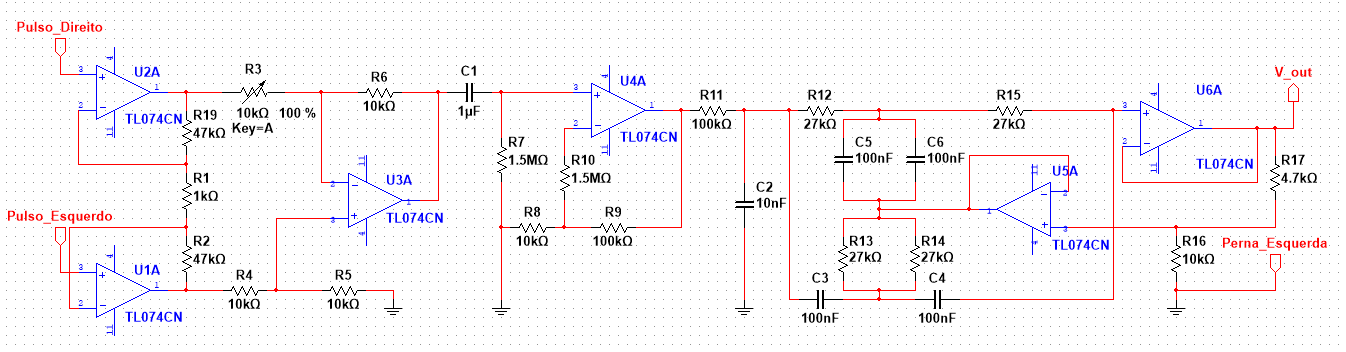
# DESENVOLVIMENTO

O desenvolvimento do nosso projeto, como mencionado anteriormente é divido em duas partes, o hardware elaborado para realizar a coleta dos sinais eletrocardiográficos do coração e o software que consiste no algoritmo de Inteligência Artificial responsável por reconhecer os sinais e prever a possível MSC do paciente.

* 1. Elaboração Hardware

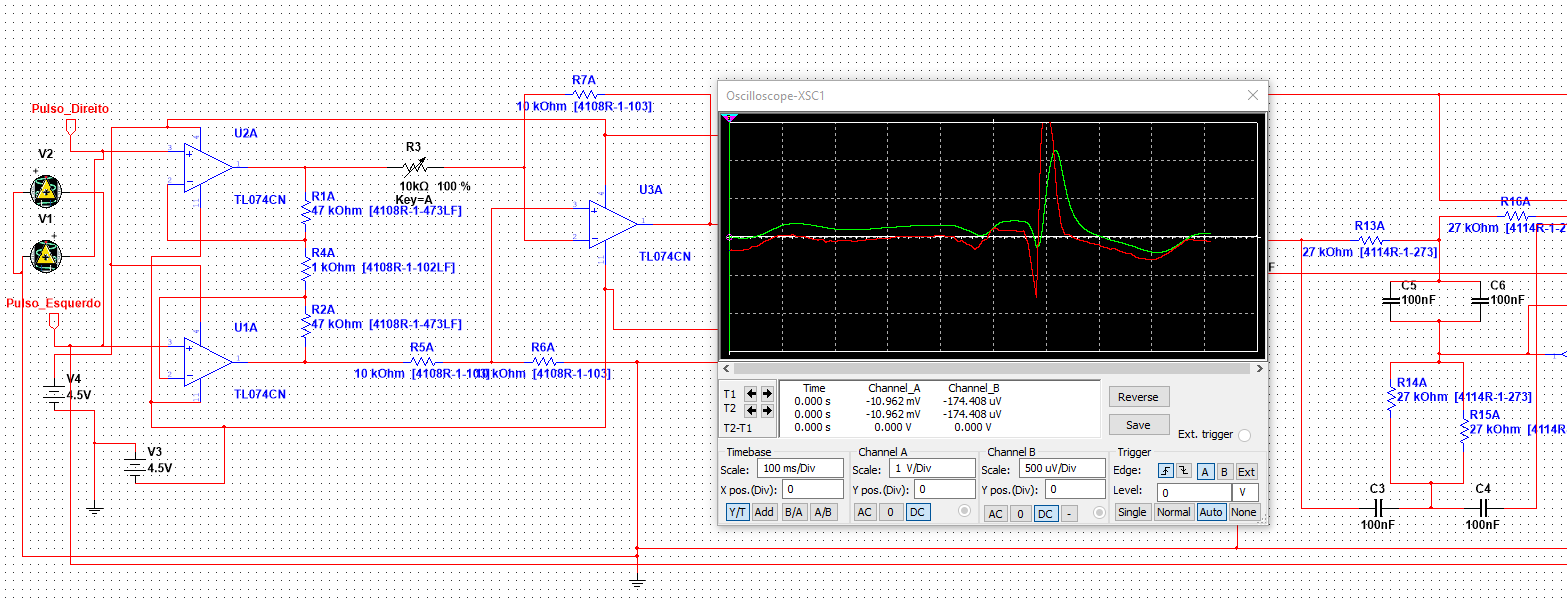
Como primeiro passo, foi realizado a uma simulação do circuito no software Multisim afim de testar e validar os filtros e o amplificador de instrumentação, aplicando uma onda senoidal, alterando a frequência da onda e verificando o comportamento dos filtros passa alta, passa baixa e rejeita faixa e após foi realizado outro teste com um sinal de onda cardíaca e o resultado estava dentro do esperado. As Figuradas 29 e 30 ilustram as simulações realizadas.

**Figura 30** – Circuito montado do Multisim



Fonte: Elaborada pelos autores (2020)

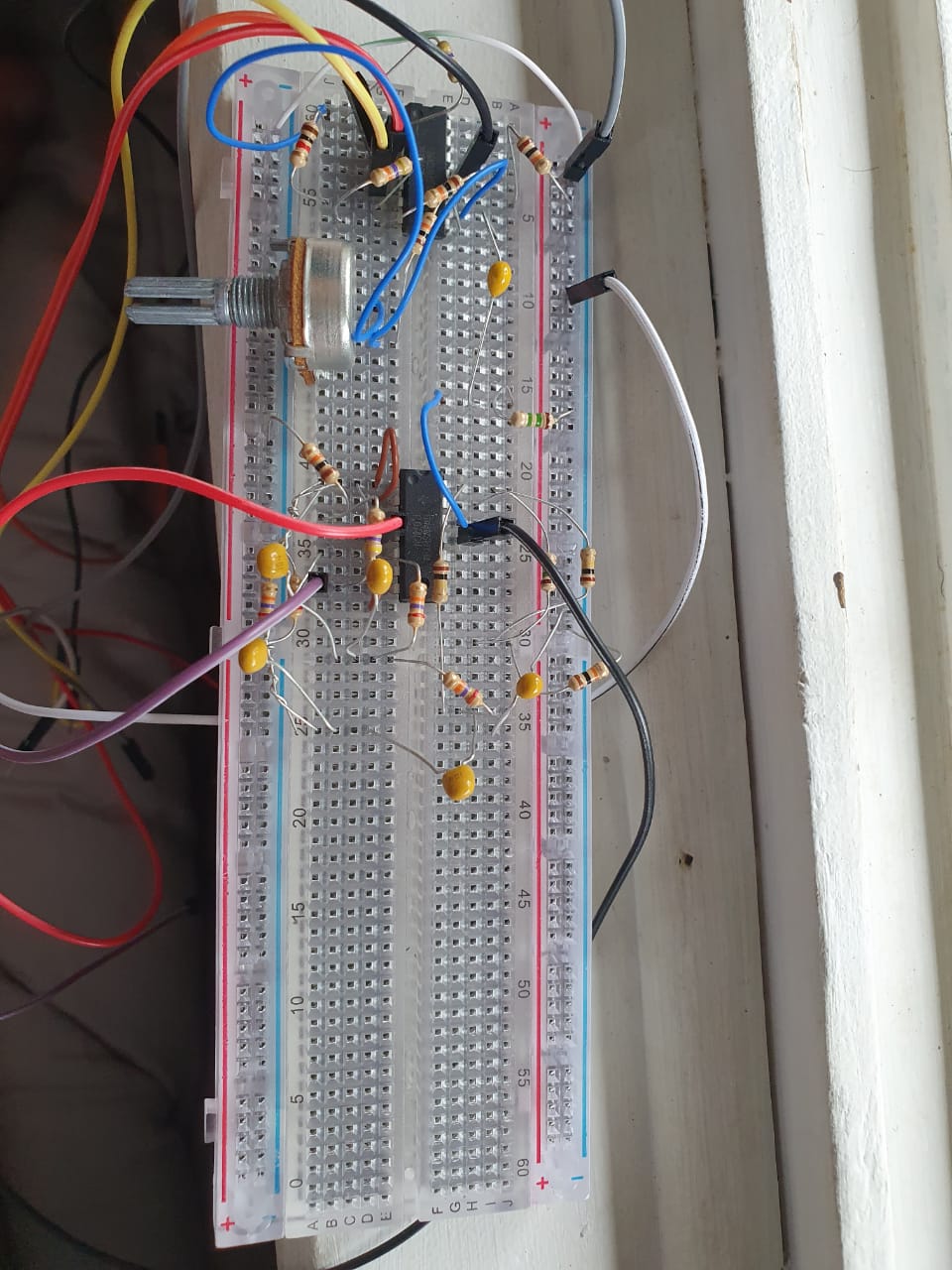
**Figura 31** – Teste no Multisim com um sinal de onda cardíaca



Fonte: Elaborada pelos autores (2020)

Após as simulações foi elaborado o circuito físico, conforme ilustrado na **Figura 31** e dado início aos testes utilizando equipamentos do laboratório de informática do Centro Universitário SENAC.

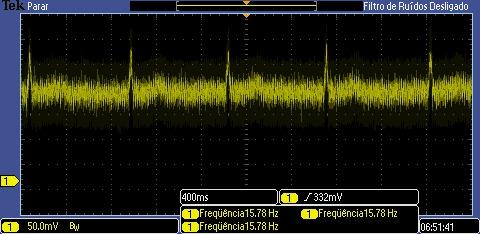
**Figura 32** – Circuito montado na protoboard



Fonte: Elaborada pelos autores (2020)

Realizado o primeiro teste dos componentes físicos utilizando um gerador de função, um osciloscópio e multímetros fornecidos no ambiente de testes do SENAC, porém foi identificado uma interferência de 60Hz no sinal possivelmente oriundos da rede de energia elétrica e dos equipamentos. Contudo, mesmo com a interferência mencionada foi possível registrar o sinal cardíaco do voluntário e integrante do grupo Gustavo Donnangelo. A Figura 32 ilustra o sinal obtido pelo osciloscópio.

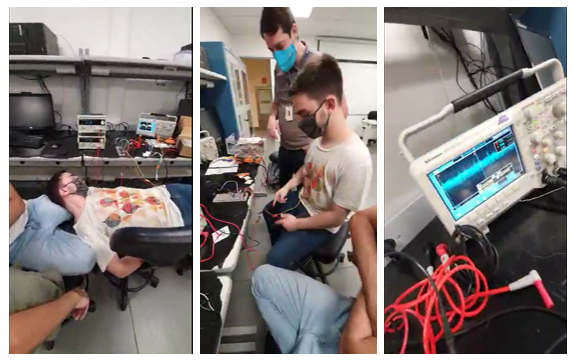
**Figura 33** – Sinal cardíaco registrado pelo osciloscópio



Fonte: Elaborada pelos autores (2020)

A seguir a Figura 33, ilustra os testes realizados com o circuito no laboratório do Centro Universitário SENAC, onde foi conectado os eletrodos no corpo do Gustavo Donnangelo e os dados eram captados pelo osciloscópio, com certo ruído, porém foi possível identificar as ondas R do sinal cardíaco.

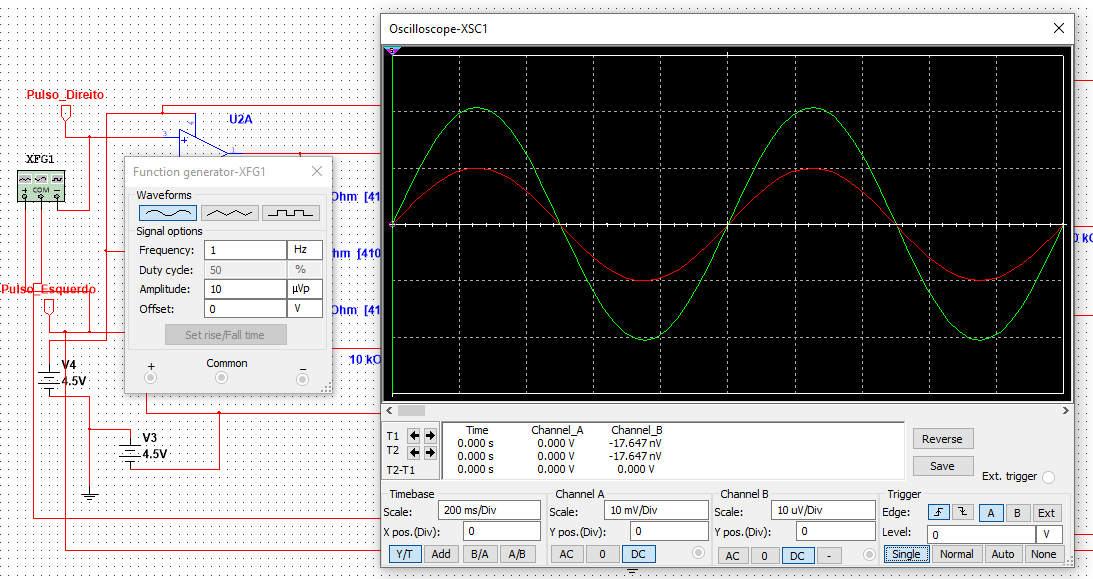
**Figura 34** – Sinal cardíaco registrado pelo osciloscópio



Fonte: Elaborada pelos autores (2020)

Após a identificação do ruído no sinal coletado, foi realizado testes em cada componente do circuito separadamente, com o objetivo de verificar falhas em componentes isolados ou nos filtros, com isso, foi identificado que o filtro Notch de 60Hz não estava bloqueando o sinal adequadamente e mesmo reconstruindo o filtro com outros componentes o problema persistiu, a Figura 32 ilustra o teste realizado no filtro Notch de 60 Hz com novos componentes.

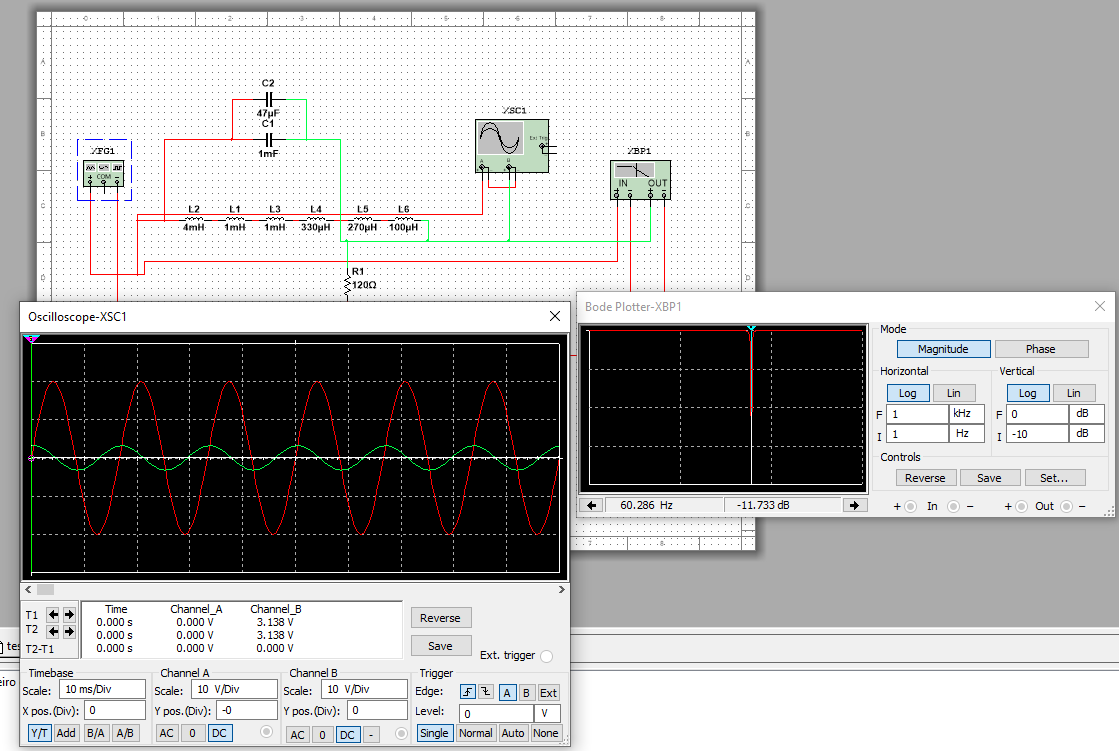
**Figura 35** – Teste de sinal no filtro Notch 60 Hz - Ativo



Fonte: Elaborada pelos autores (2020)

Como alternativa foi alterado o modelo do filtro Notch de um filtro ativo para um passivo, utilizando indutores, capacitores e resistores e não utilizando amplificadores operacionais. Após a alteração, foi realizado novamente o teste no filtro Notch e a frequência de 60Hz foi rejeitada com sucesso, conforme ilustrado na Figura 33 a seguir.

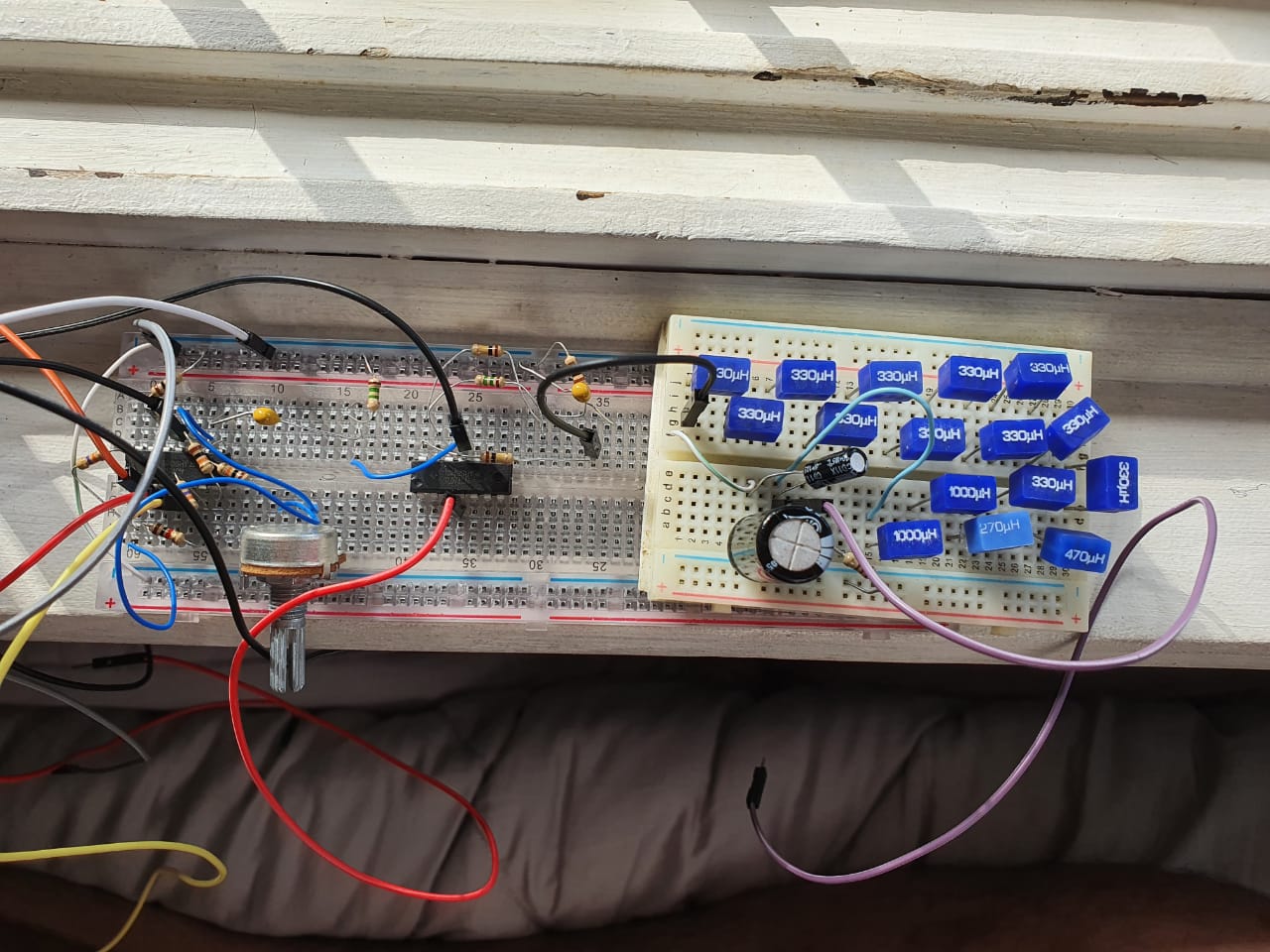
**Figura 36** – Teste de sinal no filtro Notch 60 Hz - Passivo



Fonte: Elaborada pelos autores (2020)

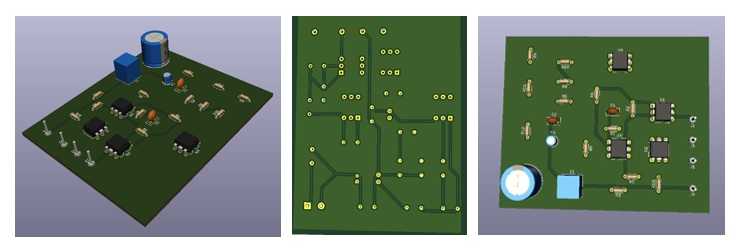
E por fim, após a validação de todos os componentes e filtros o circuito final foi o seguinte conforme ilustrado na Figura 36.

**Figura 37** – Circuito final com filtro Notch ativo

  
  
Fonte: Elaborada pelos autores (2020)

O circuito ilustrado na Figura 34 é a versão final, onde é possível coletar os sinais eletrocardiográficos de uma pessoa através de eletrodos posicionados em seu corpo. Os fios branco, laranja e amarelo são conectados nos eletrodos e conectados na perna esquerda e nos braços direito e esquerdo do indivíduo na qual pretende-se coletar o sinal. Uma observação é que devido a indisponibilidade de componentes os indutores foram colocados em série afim de obter o valor de 6,7mH utilizados na simulação.

**Figura 38** – Simulação do circuito impresso no software KICAD



Fonte: Elaborada pelos autores (2021)

* 1. Elaboração do Software

Para elaboração do algoritmo de inteligência artificial, no primeiro momento, foram coletadas bases de dados de sinais eletrocardiográficos do site physionet.org, dados estes que estão disponíveis publicamente para pesquisas cientificas, e com isso entender como estavam dispostos nas bases e como deveriam ser trabalhados.

Pelo fato de serem bases pesadas, foram necessários o acesso e o tratamento desses dados utilizando Python. A Tabela 3 ilustra como os dados são representados em pontos de mV a cada 0.008 segundos.

**Tabela 3 –** Demonstração dos valores na base Pyshionte.org

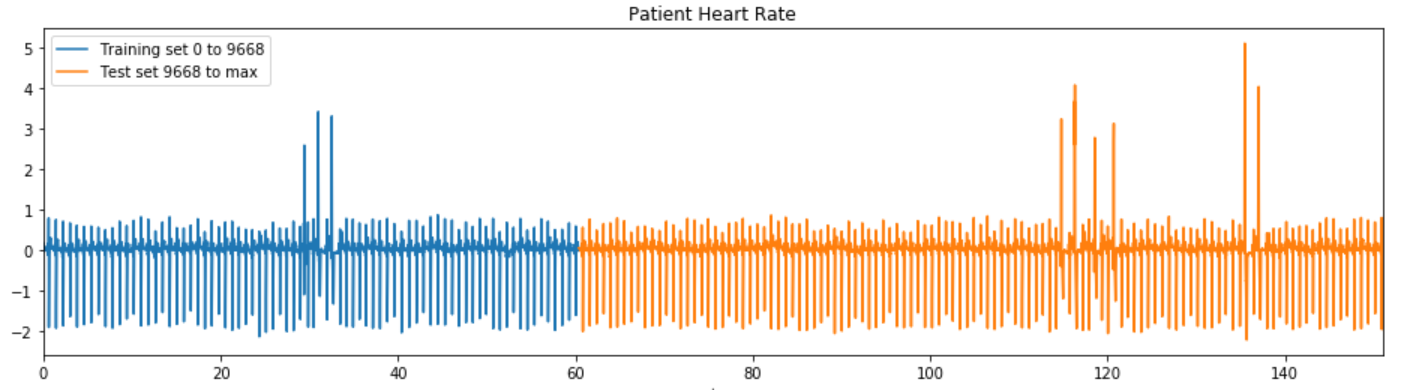
|  |  |
| --- | --- |
| **t** | **mV** |
| 0 | 0,16 |
| 0,008 | 0,18 |
| 0,016 | 0,20 |
| 0,023 | 0,22 |
| 0,031 | 0,24 |

Fonte: Elaborada pelos autores (2020)

Outro ponto importante foi entender como era este sinal que nós estávamos olhando, para ter um panorama geral e também para dividir nossa massa de dados entre treino e teste, onde o treino pegaria um pico dos dados enquanto o teste teria o restante, conforme ilustrado na Figura 38 que a parte azul representa os dados de treino e a laranja a massa de dados para teste.

O objetivo dessa separação tem o objetivo de treinar a rede neural com os dados antes da ocorrência de um pico para que o algoritmo possa captar um padrão e prever os próximos picos antes que eles ocorram, prevendo assim a MSC do paciente.

**Figura 39** – Separação dos dados eletrocardiográficos entre treino e teste

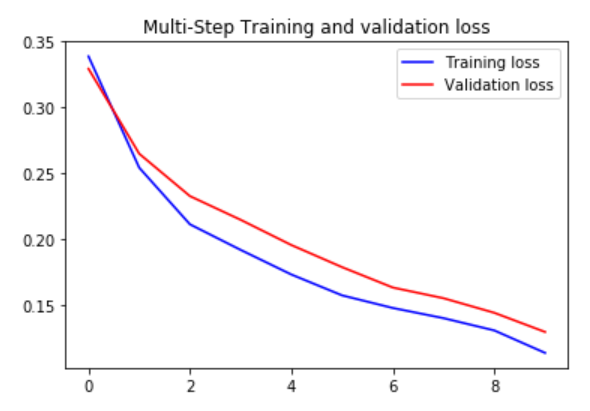


Fonte: Elaborada pelos autores (2020)

Após o entendimento das bases e como os pontos estavam distribuídos foi iniciado o desenvolvimento do modelo de Deep Learning para prever a MSC. A princípio foi decidido a utilização de um modelo de Deep Learning chamado de Rede Neural Recorrente com um suporte de LSTM (Long-Short-Term Memory).

Por indicação do professor orientador e também pelas literaturas, ela possui um poder maior de memória permitindo assim a identificação de padrões de dados de uma maneira mais precisa, a Figura 39 ilustra os primeiros resultados obtidos com o modelo LSTM.

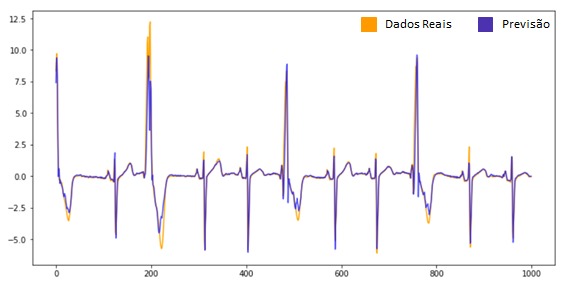
**Figura 40** – Curva de aprendizagem do modelo em treino e validação



Fonte: Elaborada pelos autores (2020)

O gráfico acima representa o erro do modelo e a cada passo de treinamento, a cada passo do treinamento este erro ia diminuindo. Com a utilização de 10 épocas foi possível prever um ponto no futuro, e as Figuras a seguir, representam alguns dos resultados obtidos pelo modelo:

**Figura 41** – Gráfico de comparação entre o sinal real e o previsto 1



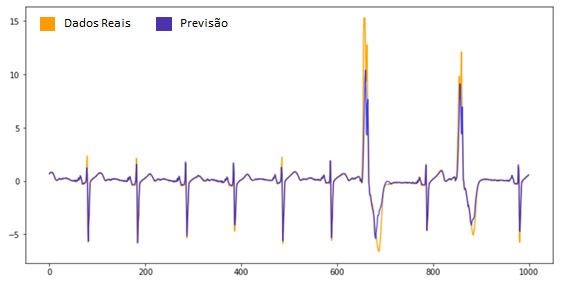
Fonte: Elaborada pelos autores (2020)

**Tabela 4 –** Valores do teste de comparação entre o sinal real e o previsto 1

|  |  |
| --- | --- |
| **Teste de Kolmogorov-Smirnov em Treino** | |
| Estatística | 0,214 |
| P-Valor | 1,00E-40 |
| **Teste de Kolmogorov-Smirnov em Teste** | |
| Estatística | 1 |
| P-Valor | 0 |
| **Correlação de Spearman** | |
| Estatística | 0,93 |
| P-Valor | 0 |

Fonte: Elaborada pelos autores (2020)

**Figura 42** – Gráfico de comparação entre o sinal real e o previsto 2



Fonte: Elaborada pelos autores (2020)

**Tabela 5 –** Valores do teste de comparação entre o sinal real e o previsto 2

|  |  |
| --- | --- |
| **Teste de Kolmogorov-Smirnov em Treino** | |
| Estatística | 0,305 |
| P-Valor | 1,00E-83 |
| **Teste de Kolmogorov-Smirnov em Teste** | |
| Estatística | 0,99 |
| P-Valor | 0 |
| **Correlação de Spearman** | |
| Estatística | 0,95 |
| P-Valor | 0 |

Fonte: Elaborada pelos autores (2020)

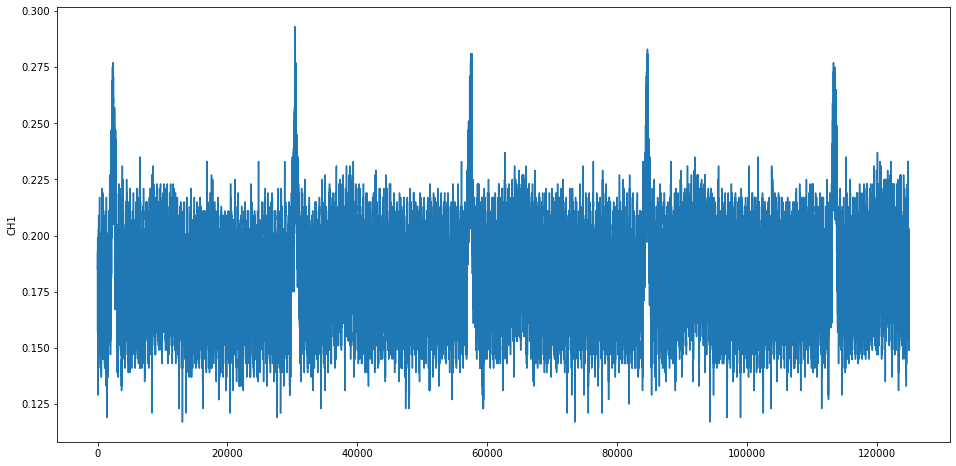
Os resultados obtidos foram muito bons com ondas previstas semelhantes as testadas, contudo esse modelo inicial só era capaz de prever um ponto no futuro.

* + 1. Tratamento dos dados por software

Com o período da pandemia e o laboratório fechado, não conseguimos utilizar o osciloscópio para tratar o ruído dos dados obtidos pelo ECG, portanto partimos para uma alternativa voltada para o software.

A Figura 43, ilustra uma amostra de dados de 5 segundos coletada pelo nosso ECG, contendo ruídos.

**Figura 43** – Dados coletados pelo ECG com ruídos

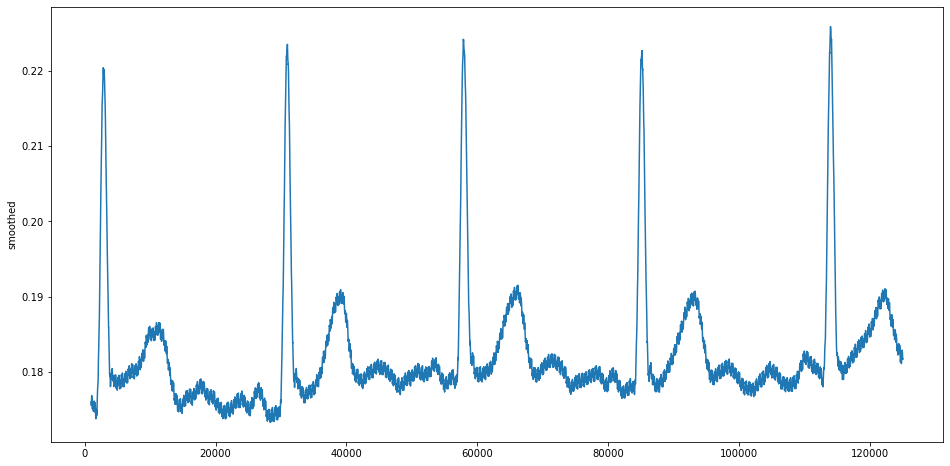


Fonte: Elaborada pelos autores (2021)

Para suavizar estes dados utilizamos um script em Python baseado em média móvel. Foram diversas tentativas para eliminar o ruído e chegar em algo mais próximo do sinal real para que o modelo de RNN possa prever a MSC.

Utilizamos 1000 passos para a suavização da média móvel, e com isso obtivemos o seguinte, conforme ilustra a Figura 44.

**Figura 44** – Dados coletados pelo ECG após o tratamento



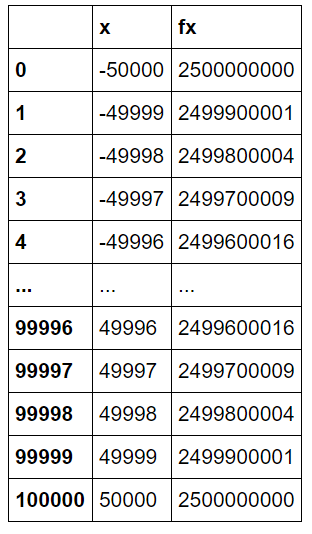
Fonte: Elaborada pelos autores (2021)

* + 1. Validação do modelo LSTM

O modelo apresentou resultados excelentes e com isso foi necessário a validação do mesmo, para tentar identificar falhas, viés, *overfitting* [[17]](#footnote-17) e, portanto, foi elaborado um script que gere uma parábola para validar os conceitos e testar novamente o modelo.

Foi criado uma parábola com 30 mil pontos e uma outra com 100, ambas para teste. A Tabela 6 ilustra os pontos em x e Fx que representam a parábola inserida no modelo.

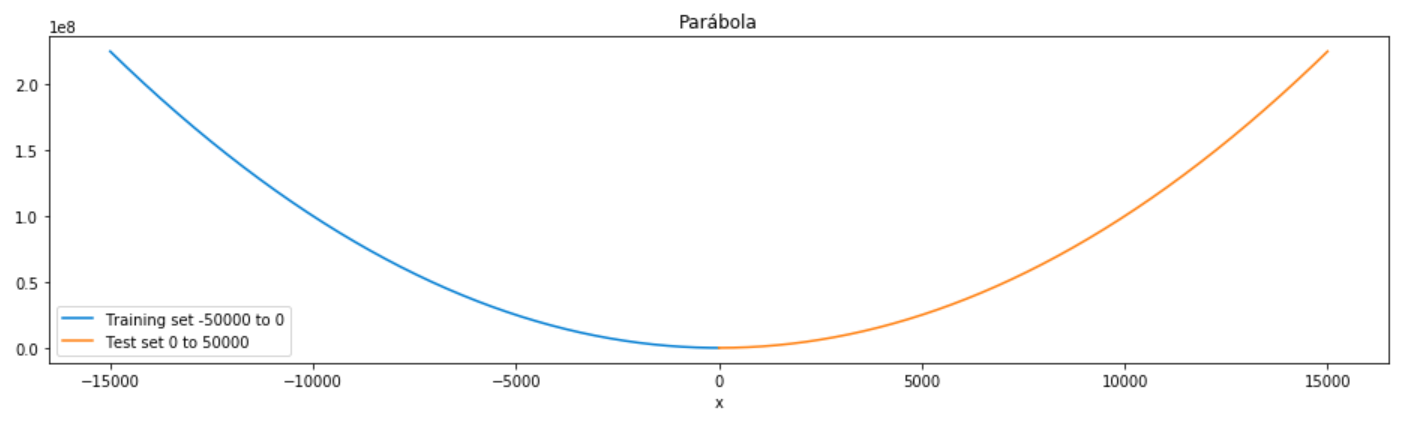
**Tabela 6 –** Valores dos pontos x e Fx da parábola



Fonte: Elaborada pelos autores (2020)

A Figura 42 ilustra a parábola utilizada para o treinamento do modelo e está dividida em duas partes, sendo a parte azul os dados de treinamento e amarelo o teste.

**Figura 45** – Parábola dividida em treinamento e teste para validação do modelo LSTM



Fonte: Elaborada pelos autores (2020)

Na curva de validação ilustrada na Figura 43 abaixo no instante 2.0 existe uma grande taxa de erro no treinamento, mas isso faz parte do gradiente descendente ou sua taxa de aprendizado, já que em algum ponto da curva o erro pode ser maior.

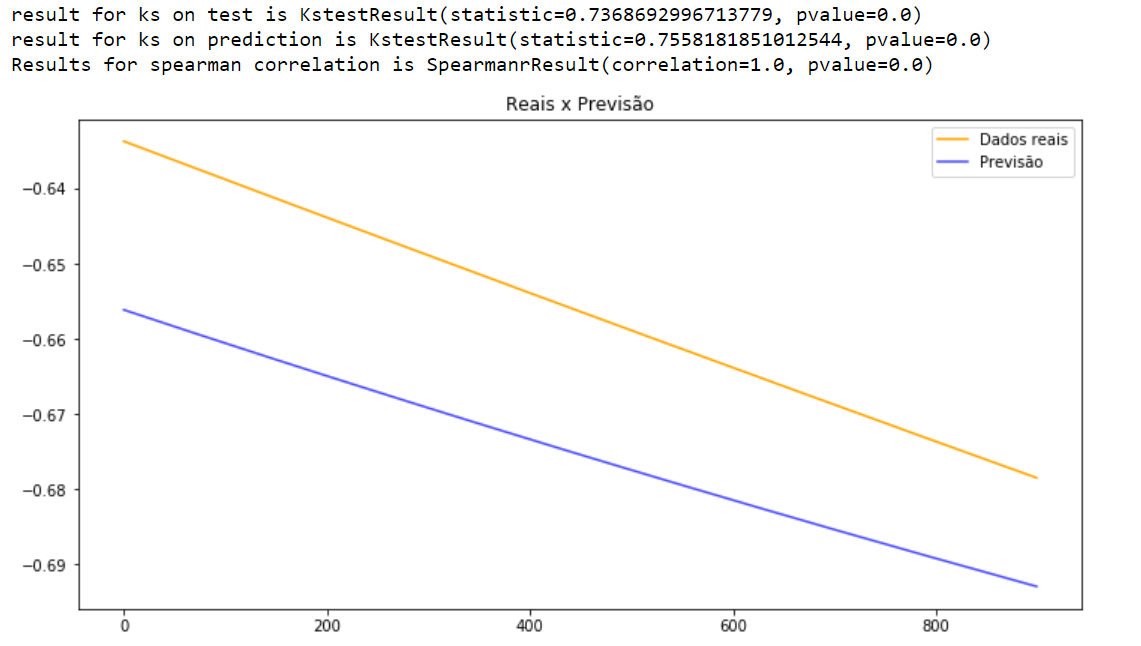
**Figura 46** – Curva de aprendizagem do modelo em treino e validação para a parábola



Fonte: Elaborada pelos autores (2020)

Outro gráfico interessante de ser avaliado é a comparação do resultado real com o resultado de previsão do modelo, conforme ilustrado na Figura 44.

**Figura 47** – Gráfico de comparação entre os dados reais e a previsão do modelo LSMT



Fonte: Elaborada pelos autores (2020)

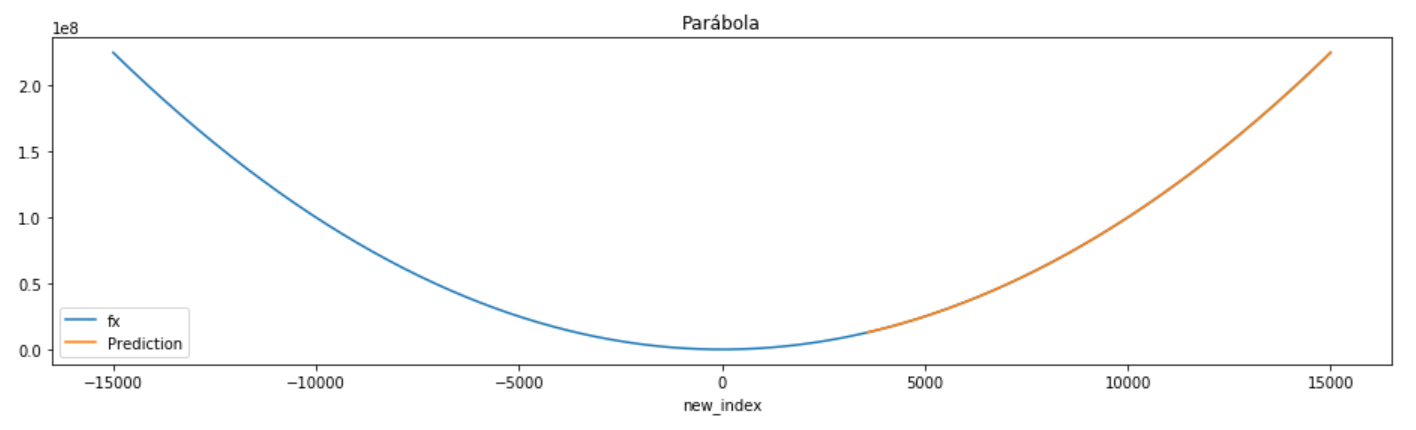
**Tabela 7 –** Valores do teste de comparação entre o sinal real e o previsto da parábola

|  |  |
| --- | --- |
| **Teste de Kolmogorov-Smirnov em Treino** | |
| Estatística | 0,736 |
| P-Valor | 0,00E+00 |
| **Teste de Kolmogorov-Smirnov em Teste** | |
| Estatística | 0,75 |
| P-Valor | 0 |
| **Correlação de Spearman** | |
| Estatística | 1 |
| P-Valor | 0 |

Fonte: Elaborada pelos autores (2020)

Pode-se observar que a correlação é perfeita e os testes de Kolmogorov Smirnov nos dizem para usar a correlação de Spearman já que nossa curva não segue a normalidade e sim o caos. A Figura 45 ilustra a parábola com a previsão completando sua curva, onde pode-se observar que a curva não destoa do padrão.

**Figura 48** – Gráfico da parábola com o modelo prevendo o restante da curva

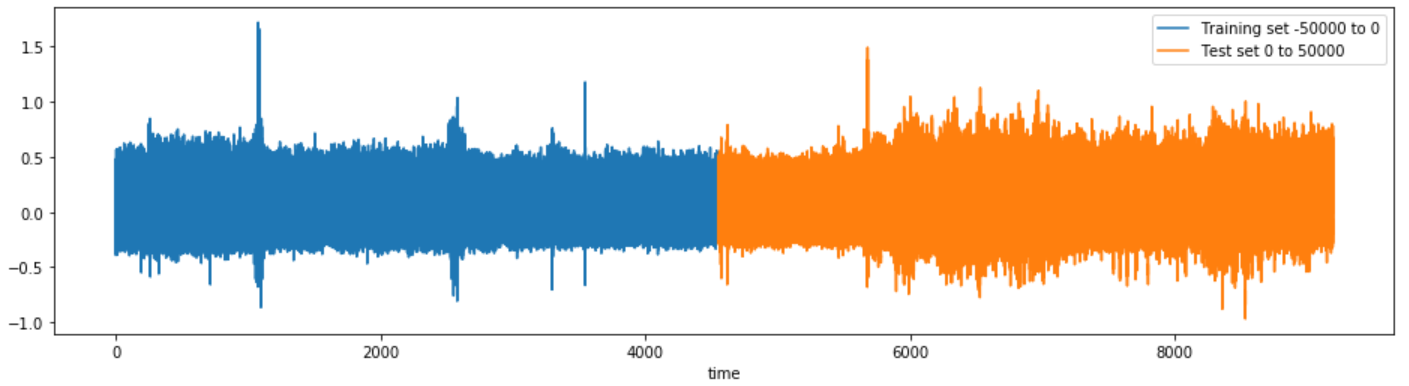


Fonte: Elaborada pelos autores (2020)

Após os testes anteriores com a Rede Neural Natural RNN, foram realizados alguns procedimentos com o intuito de aperfeiçoar o modelo com os parâmetros da parábola e consequentemente o modelo conseguir bons resultados com a base de fibrilação.

Com isso, o modelo foi retreinado a fim de aumentar os passos de previsão pois o ideal é prever o máximo de tempo antes da MSC. A Figura 46 ilustra uma base de fibrilação de um paciente que foi acometido por uma MSC e foi utilizada para treinamento e aprimoramento do modelo.

**Figura 49** – Gráfico da base de um paciente com fibrilação



Fonte: Elaborada pelos autores (2020)

Essa é uma base muito robusta onde foi utilizado 1000 pontos do passado para tentar prever um ponto a 500 passos do futuro. A Figura 47 ilustra a curva de aprendizagem do modelo para esses dados.

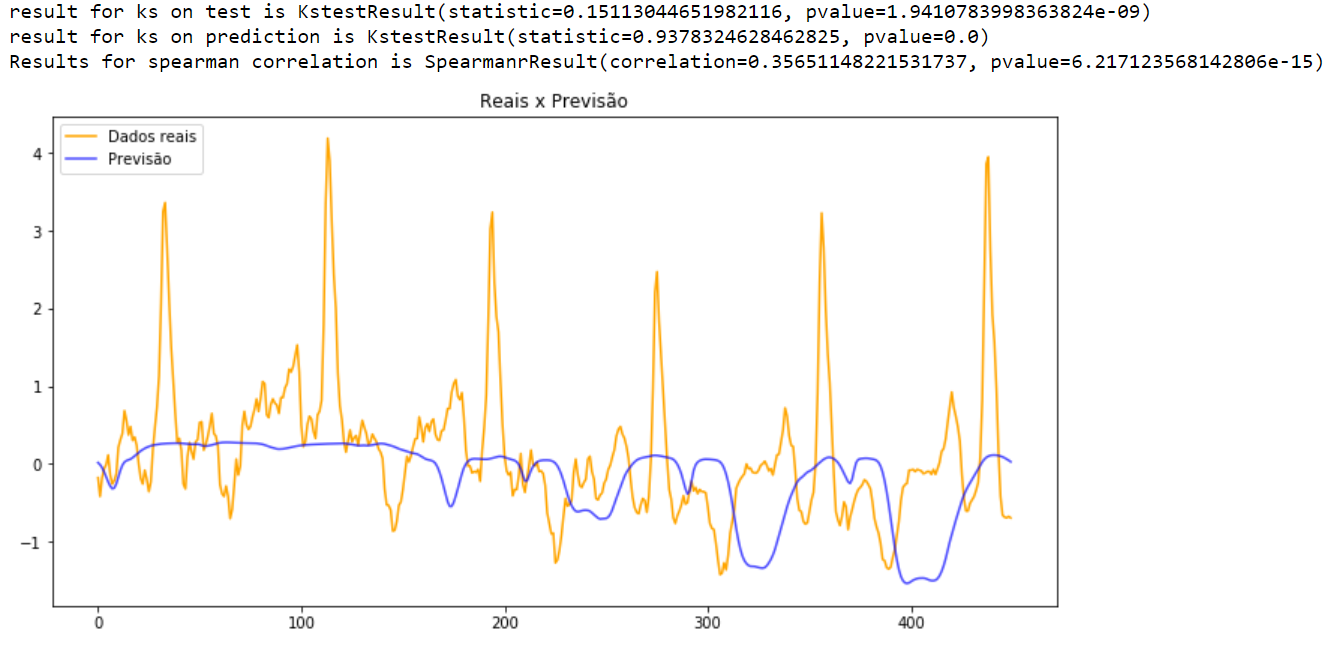
**Figura 50** – Curva de aprendizagem do modelo em treino e validação com a base de fibrilação



Fonte: Elaborada pelos autores (2020)

A curva de erro não era estável e o erro apresentado foi muito alto, as curvas não convergiram como deveria por mais que tenha sido utilizado o dobro de pontos para realizar a previsão, além de que o custo computacional foi muito alto, já que o tempo médio de treinamento foi de 30 minutos em uma máquina com 300GB de RAM. A Figura 48 ilustra o gráfico de comparação entre os sinais reais e previstos para a base de fibrilação completa.

**Figura 51** – Gráfico de comparação entre o sinal real e o previsto com a base de fibrilação



Fonte: Elaborada pelos autores (2020)

**Tabela 8 –** Valores do teste de comparação entre o sinal real e o previsto do modelo com a base de fibrilação

|  |  |
| --- | --- |
| **Teste de Kolmogorov-Smirnov em Treino** | |
| Estatística | 0,15 |
| P-Valor | 1,94E+00 |
| **Teste de Kolmogorov-Smirnov em Teste** | |
| Estatística | 0,93 |
| P-Valor | 0 |
| **Correlação de Spearman** | |
| Estatística | 0,35 |
| P-Valor | 6,21 |

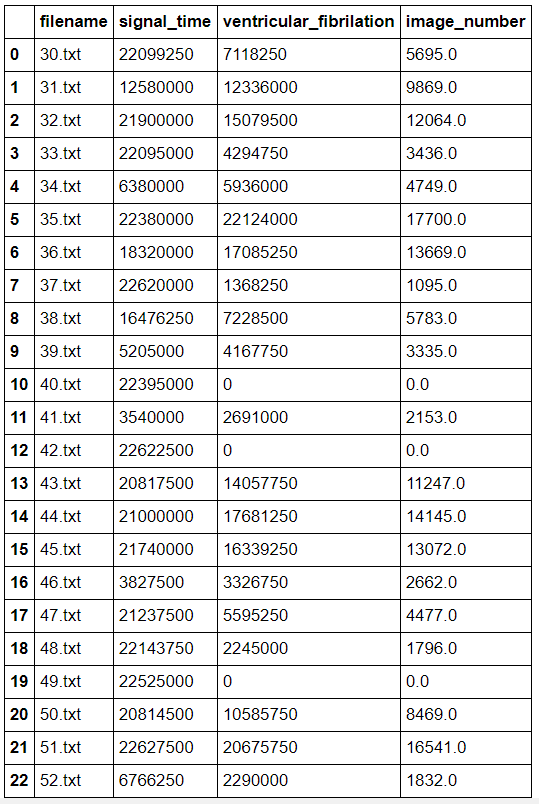
Fonte: Elaborada pelos autores (2020)

Por fim, após a atualização dos parâmetros com o treinamento realizado com a parábola e os novos testes de validação, os resultados foram insatisfatórios e foi decidido utilizarmos outra abordagem. Realizada com a transformação de todas as bases de pacientes com MSC em arquivos de imagens de 05 segundos e utilizar uma Rede Neural Convolucional (RNC) para prever se a pessoa terá uma MSC dentro do período de 24 horas.

* + 1. Elaboração do modelo RNC

Partindo dessa ideia de utilizar um modelo RNC ao invés do modelo anterior, como primeiro passo foi criado uma tabela de apoio com cada paciente que ocorreu a MSC e os respectivos instantes de tempo da MSC, conforme ilustrado na Tabela 9 a seguir:

**Tabela 9 –** Tabela auxiliar dos pacientes com MSC e as referências de imagem



Fonte: Elaborada pelos autores (2021)

Na hora de criar a divisão entre treino e validação foi utilizado o método de escolha aleatório e um balanceamento para determinar o número de imagens de pacientes de MSC, já que não era possível trabalharmos com uma quantidade imensa de imagens que poderiam enviesar nosso modelo, houve também uma separação de paciente para que os mesmos pacientes que não estivessem na massa de dados de treino, não aparecessem na massa de dados de teste, dessa maneira ficamos com:

* 5 pacientes de morte súbita no treino;
* 2 pacientes de morte súbita na validação;
* 88 pacientes saudáveis no treino;
* 27 pacientes de morte súbita na validação.

Após essa separação foi iniciado o processo de treinamento da RNC utilizando o TensorFlow e para criar a rede foi utilizado um padrão inicial que poderia ou não precisar de um *hyper parameter tuning[[18]](#footnote-18)* para melhorar o modelo e essa rede foi constituída de 03 blocos:

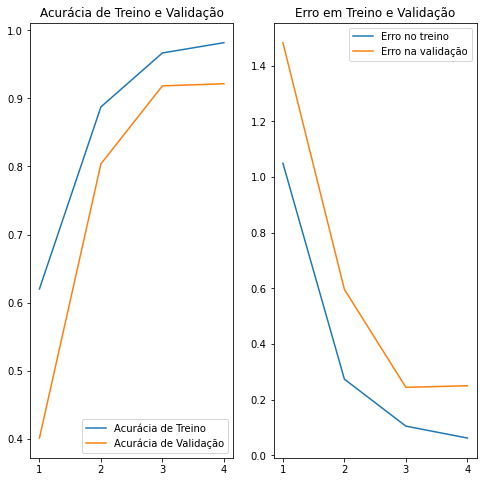
* O primeiro bloco contendo 16 neurônios;
* O segundo bloco contendo 32 neurônios;
* O terceiro bloco contendo 64 neurônios.

Além disso também possuí uma segunda camada com 128 neurônios que pode ser ativada por uma função específica do TensorFlow.

* + 1. Treinamento e validação do modelo RNC

Para o treinamento foi utilizado 15 épocas, ou seja, passos de treinamento, e uma função de early\_stopping caso a função comece a decair ou chegue em um valor muito bom: A Figura 49, ilustra os resultados de treinamento com apenas 04 épocas:

**Figura 52** – Resultado do modelo com apenas 04 épocas de treinamento



Fonte: Elaborada pelos autores (2021)

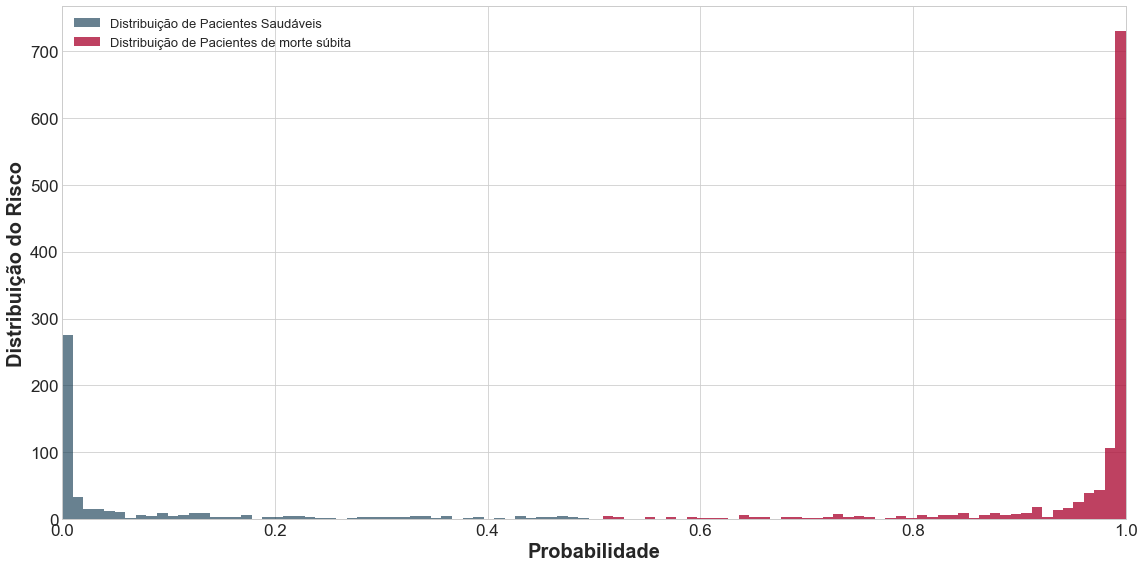
Em 4 épocas já um ótimo resultado da rede com uma acurácia de treino em 98% e uma acurácia de validação em 92%, lembrando que se obtivéssemos uma rede com uma acurácia de 100% em treino e validação isso seria sinal de *overfitting* ou possivelmente, imagens que não tem correlação alguma uma com a outra.

Expandindo um pouco mais o resultado da validação:

* Para os pacientes de MSC o resultado de acerto foi de 92,6% dos pontos analisados.
* Para os pacientes saudáveis o resultado de acerto foi de 88,6% dos pontos analisados.

Para ver isso um pouco melhor fora do texto pode-se analisar a Figura 50 que representa o gráfico de poder de separação do modelo:

**Figura 53** – Gráfico de poder de separação do modelo



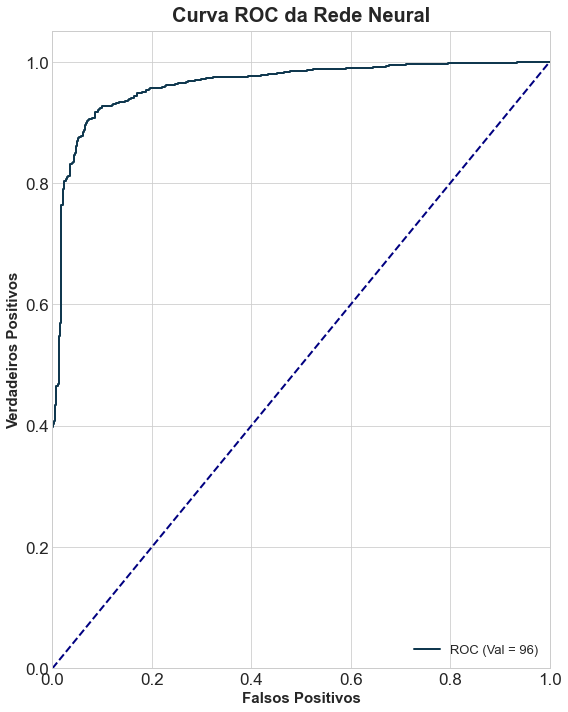
Fonte: Elaborada pelos autores (2021)

O Gráfico acima representa o poder de separação do modelo, com ele pode-se identificar o quão bem o modelo está separando suas previsões, neste gráfico o ideal é que seus picos ou caldas comumente chamadas em estatística estejam sempre nas pontas, e assim comprovar que o modelo está separando bem suas classificações.

Um modelo que tem caldas azuis e vermelhas misturadas nem sempre é um modelo ruim, tudo depende dos dados em que foi treinado, as vezes o melhor resultado possível a se obter é um resultado misturado.

Outra medida muito utilizada para se avaliar o modelo é a curva ROC que nos mostra o erro do modelo em relação a falso positivos e falso negativos, conforme ilustrado na Figura 51 a seguir.

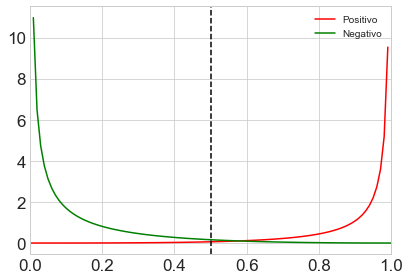
**Figura 54** – Curva ROC da RCN



Fonte: Elaborada pelos autores (2021)

E por fim o teste de especificidade x sensibilidade, ilustrado na Figura 52, que nos mostra o quanto mais próximo o modelo está da realidade. E todos esses testes e visualizações servem de embasamento do quanto o modelo está funcionando.

**Figura 55** – Teste de especificidade vs sensibilidade do modelo



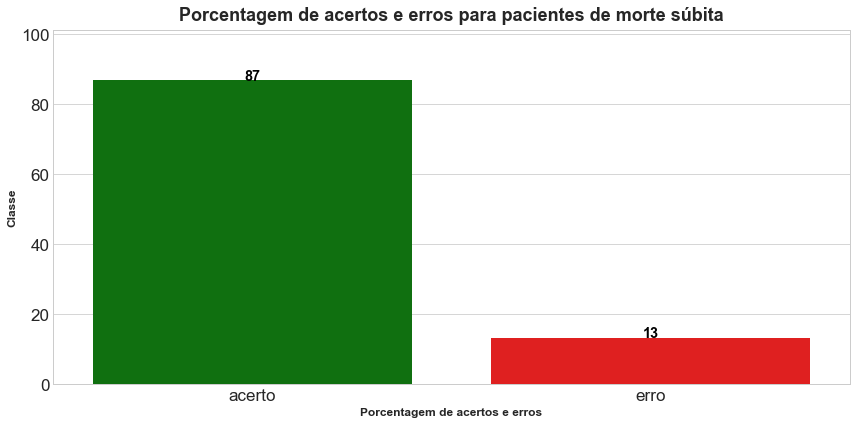
Fonte: Elaborada pelos autores (2021)

Com o teste de especificidade vs sensibilidade, da mesma forma que no gráfico de separação do modelo, ter caldas maiores nas pontas é sempre melhor, uma outra coisa que se pode observar é onde as caldas se cruzam. O resultado ideal ocorre quando elas se cruzam no eixo central!

Porém, isso só acontece com coisas que vemos na natureza, assim como funciona por exemplo uma curva normal. Outra coisa que é possível realizar com este teste é tentar mudar o limite de quando classificamos algo como 1 ou como 0, o padrão é sempre usar 50%, tudo que estiver acima de 50% é 1 enquanto tudo que estiver abaixo de 50% é 0, porém podemos mudar este limite de acordo com esse cruzamento, isso funciona muito bem para modelos de risco de crédito em banco por exemplo.

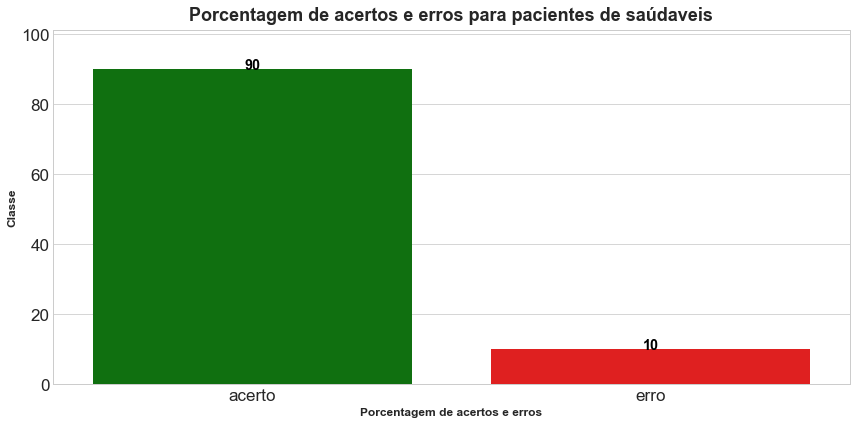
Após a realização dos testes no modelo, foi dado início na validação do modelo em novos pacientes com MSC pegando trechos antes de sua MSC e fazendo uma validação mais brusca por cima delas dessa forma conseguimos ver se existem comportamentos do coração que podem confundir a rede antes da morte súbita cardíaca, com isso conseguimos chegar numa métrica que nos mostrou que nosso modelo realmente é eficiente para cumprir sua função, conforme ilustrado nas Figuras 55 e 57.

**Figura 56** – Percentual de predições corretas do modelo pacientes com MSC



Fonte: Elaborada pelos autores (2021)

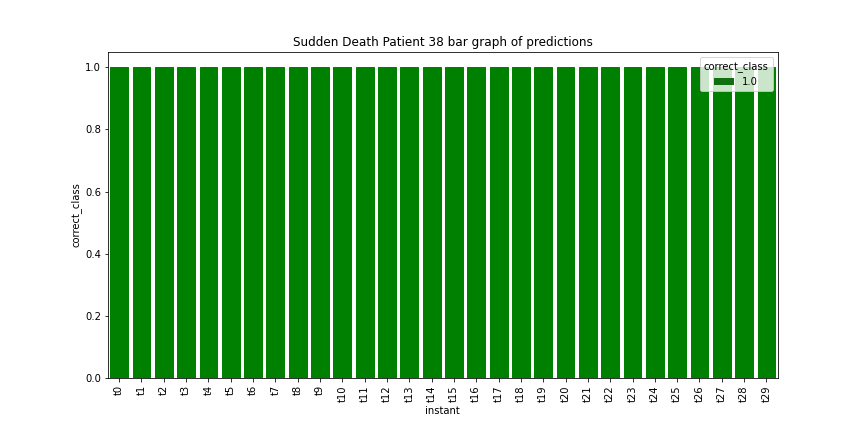
**Figura 57** – Percentual de predições corretas do modelo pacientes Saudáveis

****

Fonte: Elaborada pelos autores (2021)

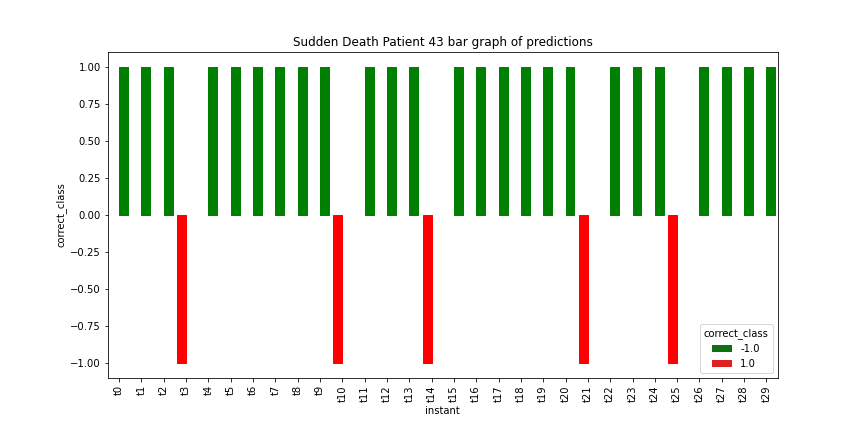
Como última etapa de teste, começamos a buscar trechos aleatórios de cada paciente e gerar imagens sequenciais de cada um deles para sustentar ainda mais o funcionamento do modelo. As figuras a seguir, ilustram alguns exemplos dos resultados, como auxílio no entendimento do poder de previsão do modelo em cima de pacientes de MSC.

**Figura 58** – Paciente 38 de MSC com quantidade de predições do modelo no tempo



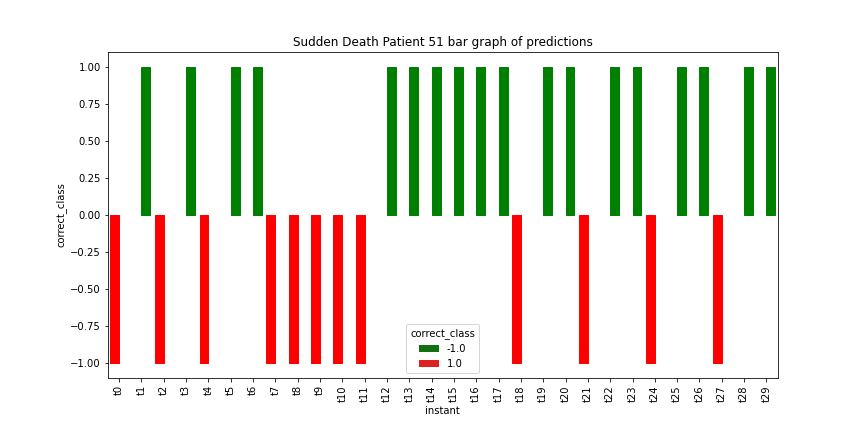
Fonte: Elaborada pelos autores (2021)

**Figura 59** – Paciente 43 de MSC com quantidade de predições do modelo no tempo



Fonte: Elaborada pelos autores (2021)

**Figura 60** – Paciente 51 de MSC com quantidade de predições do modelo no tempo

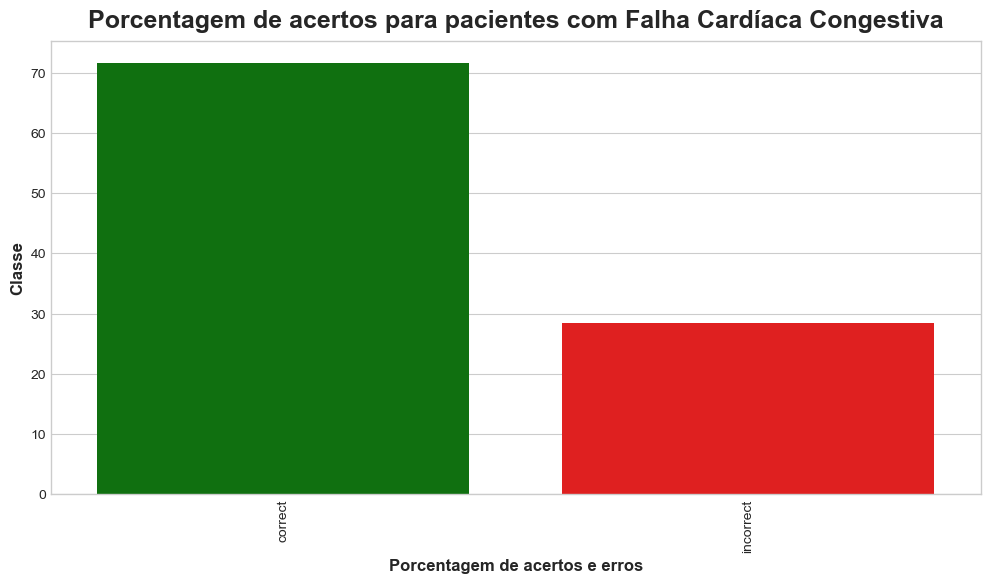


Fonte: Elaborada pelos autores (2021)

Nos gráficos acima pegamos 30 imagens de 5 segundos de cada pacientes em sequência, então tentamos prever se aquele paciente é um paciente onde ocorreu morte súbita cardíaca ou não. As barras verdes expressam um acerto da nossa rede neural, enquanto a barra vermelha expressa um erro da nossa rede neural.

Após realização dos testes foi sugerido que tentássemos utilizar o modelo em outras causas também, no caso apresentamos um sinal de pacientes que possuíam Falha Cardíaca Congestiva (FCC), nossa rede foi capaz de prever, porém com um erro um pouco maior, o que nos indicou a presença de algum padrão nas imagens.

**Figura 61** – Resultado do modelo para pacientes com FCC



Fonte: Elaborada pelos autores (2021)

* + 1. Rebalanceamento e ajustes finais do modelo

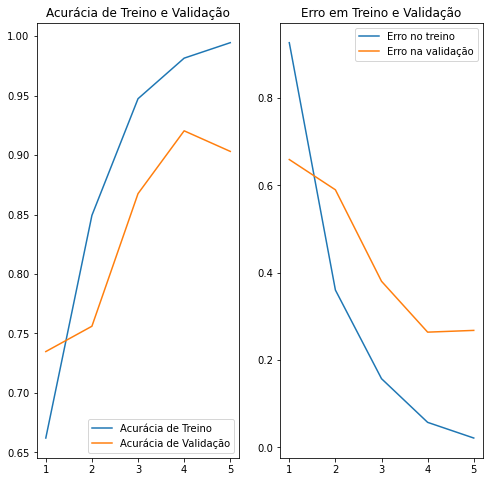
Após os testes com pacientes de FCC, decidimos retreinar o modelo usando um balanceamento um pouco melhor, haviam sido identificados pacientes onde a fibrilação ocorria desde o início do treinamento.

Desta forma ficaram balanceados 75% dos pacientes no treinamento e 25% na validação.

* 10 pacientes de morte súbita no treino;
* 3 pacientes de morte súbita na validação;
* 88 pacientes saudáveis no treino;
* 27 pacientes de morte súbita na validação.

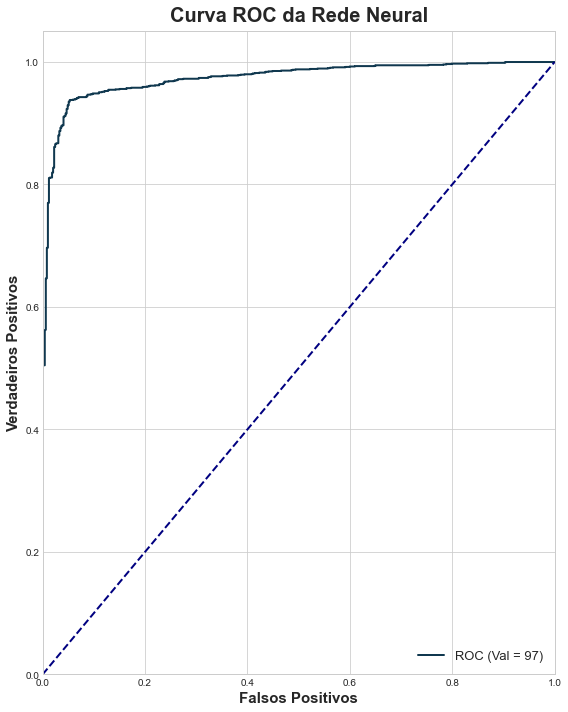
As figuras a seguir ilustram os resultados obtidos após o rebalanceamento do modelo:

**Figura 62** – Resultado do modelo rebalanceado com 05 épocas de treinamento



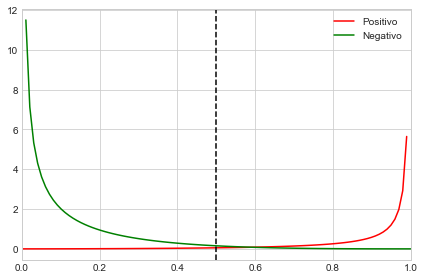
Fonte: Elaborada pelos autores (2021)

**Figura 63** – Curva ROC após o rebalanceamento



Fonte: Elaborada pelos autores (2021)

**Figura 64** – Teste de especificidade vs sensibilidade do modelo rebalanceado



Fonte: Elaborada pelos autores (2021)

Olhando para os gráficos de treinamento, a curva ROC e o gráfico de sensibilidade x especificidade não dá para notar muitas diferenças na rede, porém os próximos resultados nos mostram como isso impactou as previsões:

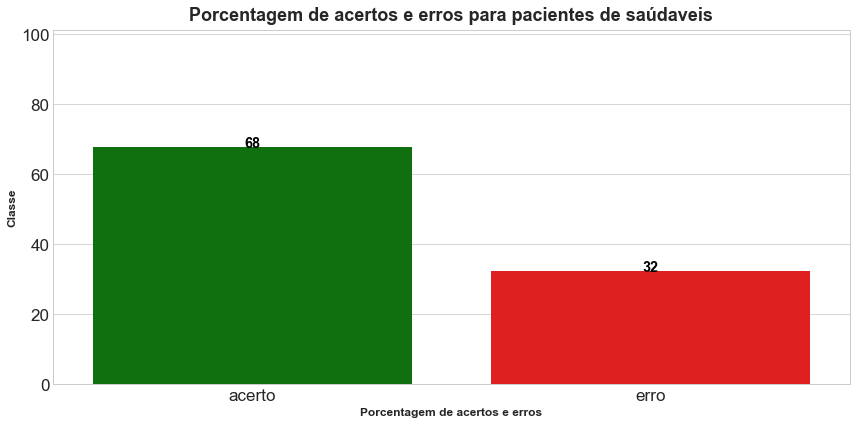
**Figura 65** - Porcentagem de acertos e erros para pacientes de MSC

****

Fonte: Elaborada pelos autores (2021)

Para os pacientes de morte súbita não tivemos quase nenhum erro, e isso é um resultado muito positivo, pois reflete onde estávamos querendo chegar, uma taxa de acerto muito maior para pacientes de morte súbita cardíaca.

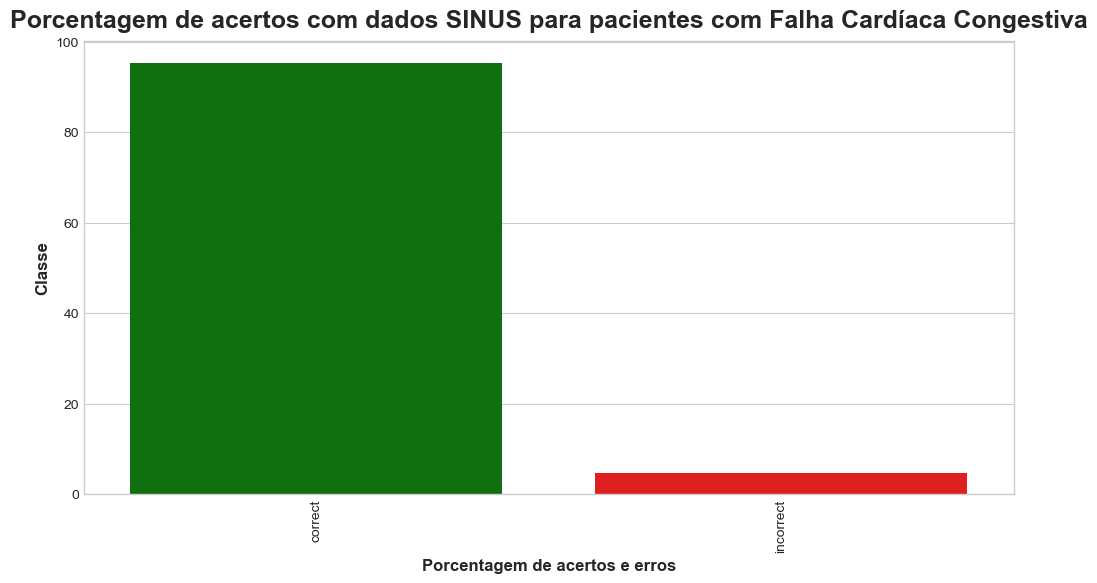
**Figura 66** - Porcentagem de acertos e erros para pacientes saudáveis

****

Fonte: Elaborada pelos autores (2021)

Já para os pacientes saudáveis tivemos um erro um pouco mais alto, todavia o balanceamento compensa pois estaríamos errando menos os pacientes que podem ter a MSC.

**Figura 67** - Porcentagem de acertos com dados SINUS para pacientes com FCC

****

Fonte: Elaborada pelos autores (2021)

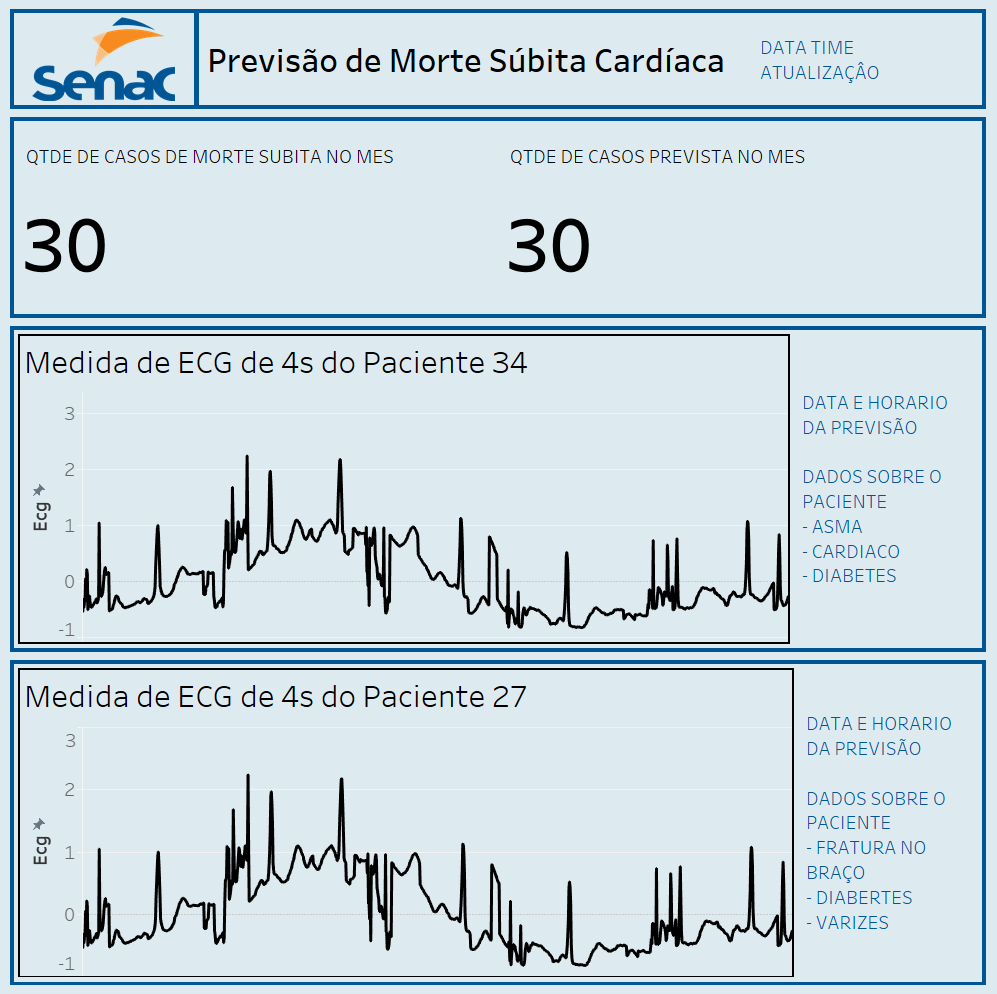
Já para os pacientes com Falha Cardíaca Congestiva, nós obtivemos um resultado ainda melhor, já que estamos acertando a maioria dos casos. No geral obtivemos resultados muito bons que podemos observar com as métricas de erros anteriores.

* + 1. Visualização dos dados

A visualização dos dados através de um dashboard é imprescindível para a análise dos sinais do paciente e para a visualização do resultado do modelo de RNN. Deste modo, elaboramos um painel na ferramenta Tableau onde são informados indicadores e dados de pacientes que geraram o alerta de morte súbita cardíaca, dentre eles:

Data e hora de atualização do painel, quantidade de casos de MSC no mês, quantidade de casos previstas pelo modelo e dados dos pacientes que geraram alerta facilitando as informações para uma possível emergência. A Figura 58, ilustra o dashboard desenvolvido.

**Figura 66** – Dashboard com os dados do paciente



Fonte: Elaborada pelos autores (2021)

# CONSIDERAÇÕES FINAIS

A realização desse projeto de conclusão de curso foi muito satisfatória, pois foi uma experiência nova para todo o grupo. Trabalhar em equipe com a orientação de nossos professores, ter de nos autogerenciar em todos os momentos para que o projeto não falhasse e concluir todas as etapas com êxito.

Durante o desenvolvimento do projeto surgiram alguns problemas, como por exemplo a pandemia do novo coronavírus Covid-19 que nos forçou a realizar algumas mudanças em nosso cronograma e na forma de trabalhar. Isso também dificultou o acesso aos equipamentos do laboratório do Centro Universitário SENAC que seriam essenciais para o desenvolvimento e validação do hardware, apesar disso conseguimos desenvolver um ECG capaz de coletar os sinais eletrocardiográficos de um paciente, com certo ruído.

No software obtivemos ótimos resultados, primeiro começamos com um estilo de rede neural que nos artigos lidos se provava competente para problemas de previsão de sinais, falhamos e com isso surgiu outra ideia, que era utilizar um “professor” para a rede, nosso professor seriam as imagens de 5 segundos dos sinais coletados. Os nossos primeiros resultados nos provaram que isso era um método muito forte, e já possuíamos um modelo dentro dos padrões aceitáveis, porém mais ideias foram surgindo e por fim conseguimos modelar uma rede capaz de acertar muitos pacientes de MSC sem conter viés, sem conter *overfitting* e ainda sim com uma boa taxa de acerto para pacientes saudáveis.

De fato, o modelo pode ser ajustado e aprimorado assim como o hardware que foi elaborado como prova de conceito para validar os procedimentos idealizados previamente. Mas consideramos que essa solução teve sucesso em todos os aspectos, tanto no hardware quanto no software, podendo assim ser de extrema importância para os profissionais da saúde. A junção do todo nos trouxe muito conhecimento e um resultado extremamente satisfatório.

**REFÊRENCIAS**

ARNOLD, T. B; EMERSON, J. W. **Nonparametric Goodness-of-fit Tests for Discrete Null Distributions**, 2011. Disponível em: <https://journal.r-project.org/archive/2011-2/RJournal\_2011-2\_Arnold+Emerson.pdf>. Acesso em: Maio de 2020.

AZEVEDO, D. F. **Iniciação à Eletrocardiografia**, Porto Alegre: Artmed,1999.

BARBOSA C. R. H; MONTEIRO E. C; WOLF, A. S. **Análise automática de sinais eletrocardiográficos por redes neurais artificiais**. Rio de Janeiro, 2004. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2004.

BERBARI, E. J. **Principles of Electrocardiography**. In: BRONZINO, J. D. The Biomedical Engineering Handbook: Second Edition. 2. Ed. Estados Unidos: CRC Press, 2000.

BERNE, R. M. et al. **Fisiologia**. 5 Ed. Rio de Janeiro: Elsevier, 2004.

BOWERS, D. F. **Applying Current Feedback to Voltage Amplifiers**. Peter Peregrinus Ltd. Londres, 1990.

CAPUANO, F. G; MARINO, M. A. M. **Laboratório de Eletricidade e Eletrônica - Teoria e Prática.** - 23ª edição - São Paulo: Editora Érica, 2005.

DATAQUEST. **Understanding Regression Error Metrics in Python**. Disponível em: <https://www.dataquest.io/blog/understanding-regression-error-metrics/>. Acesso em: Maio de 2020.

DOENÇAS CARDIOVASCULARES. **OPAS - organização pan-americana de saúde**, 2017. Disponível em: < https://www.paho.org/bra/index.php?option=com\_content&view=article&id=5253:doencas-cardiovasculares&Itemid=1096>. Acesso em: Abril de 2020.

EGAN, J. P. **Signal detection theory and ROC analysis**. New York, USA. Academic Press, 1975.

FOLEY, B. **An introduction to the Chi-Square Test & When to Use It**, 2018. Disponível em: <https://www.surveygizmo.com/resources/blog/introduction-to-Chi-Square-test-and-when-to-use-it/>. Acesso em: Maio de 2020.

FOTIADIS, D. I. et al. **Electrocardiogram (ECG): Automated Diagnosis**. In: AKAY, M. Wiley Encyclopedia of Biomedical Engineering. Estados Unidos: John Wiley & Sons, 2006.

FROWNFELTER, D.; DEAN, E. **Fisioterapia Cardiopulmonar – Princípios e Prática**. 3 Ed. Editora Revinter, 2004.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6 Ed. São Paulo: Atlas, 2008.

GREBENE, A. B. **Bipolar and Mos Analog Integrated Circuit Design**. New York: John Wiley & Sons, 1984.

GOLDBERGER, A., AMARAL, L., GLASS, L., HAUSDORFF, J., IVANOV, P. C., MARK, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: **Components of a new research resource for complex physiologic signals**. Circulation. 101 (23), pp. e215–e220.

GUYTON, A. C.; HALL, J. E. **Tratado de Fisiologia Médica**. Guanabara & Koogan, 1996.

HAO, Karen. **the computing power needed to train ai is now rising seven times faster than ever before**, MIT Technology Review, 2019.

HAUKE, J; KOSSOWSKI, T. **Comparison of values of pearson’s and spearman’s correlation coefficients on the same sets of data**. Institute of Socio-Economic Geography and Spatial Management. Poland, 2011.

HEARTFLOW, 2020. Transforming the diagnosis and management of coronary artery disease worldwide. Disponível em: <https://www.heartflow.com/>. Acesso em: Maio de 2020.

HEARTVISTA, 2019. Revolutionizing MRI through AI. Disponível em: <https://www.heartvista.ai/>. Acesso em: Maio de 2020.

HOBBIE, R. K.; ROTH, B. J. Intermediate Physics for Medicine and Biology. 4. Ed. Estados Unidos: **Springer Science Business Media**, 2007.

HODGKIN, A; HUXLEY, A**. A quantitative description of membrane current and its application to conduction and excitation in nerve**. The Physiological Society, 1952.

HOLMGREN, G; et al. **Artificial neural networks improve and simplify intensive care mortality prognostication: a national cohort study of 217, 289 first time intensive care unit admissions**. J intensive care, 2019.

IDx. Eye Diagnosis, 2018. Autonomous AI that instantly detects disease. Disponível em: < https://www.eyediagnosis.co/>. Acesso em: Maio de 2020.

LEE, C. H. **Agile in Data Science: Why My Scrum Doesn’t Work?.** 2019. Disponível em: <https://changhsinlee.com/agile-ds-scrum-kanban/>. Acesso em: Maio de 2020.

MAHESH. **Everything You Need To Know about Hypothesis Testing — Part I**. Towards Data Science, 2019. Disponível em: <https://towardsdatascience.com/everything-you-need-to-know-about-hypothesis-testing-part-i-4de9abebbc8a>. Acesso em: Maio de 2020.

MARGETTS, H; DOROBANTU C. **Rethink government with AI.** Nature, 2019.

MCCULLOCH W. S; PITTS W. **A logical calculus of the ideas immanent in nervous activity.** The bulletin of mathematical biophysics, 5(4):115-133, Dec 1943.

MOHRMAN, D.E.; HELLER, L.J. **Cardiovascular Physiology**. McGraw-Hill, 2006.

MOREIRA, Catarina. **Neurônio**. Revista de Ciência Elementar. Volume 1, 2013.

MOREIRA, Thiago. Projeto de circuito para Eletrocardiógrafo. **Embarcados**, 2016. Disponível em: <https://www.embarcados.com.br/projeto-eletrocardiografo/>. Acesso em: Maio de 2020.

MORSCH, José Aldair. **Tecnologia na medicina: avanços da medicina e o futuro da saúde**. Morsch, 2018. Disponível em: <https://telemedicinamorsch.com.br/blog/tecnologia-na-medicina>. Acesso em: Maio de 2020.

MORTE SÚBITA - O QUE PODEMOS FAZER ?. **Instituto do Coração (InCor - HC/FMUSP)**, 2002. Disponível em: <http://www.incor.usp.br/conteudo-medico/geral/morte%20subita.html>. Acesso em: Maio de 2020.

MUSSOI, F. L. R. **Respostas em Frequência - Filtros Passivos**. Centro Federal de Educação Tecnológica de Santa Catarina. Florianópolis, 2004.

NARAYANASWAMY, S. **High Resolution Electrocardiography**. Indian Pacing and Electrophysiology Journal, p. 50-56, 2002.

NARKHEDE, S. **Understanding AUC - ROC Curve**. Towards Data Science, 2018.

NAZERAN, H. **Computers in Electrocardiography**. In: WEBSTER, J. G. Encyclopedia Of Medical Devices And Instrumentation. 2. ed. Estados Unidos: John Wiley & Sons, v. 3, 2006. p. 34-53.

OLIVEIRA, A. T.; **Detecção do complexo QRS em sinais cardíacos utilizando FPGA**. UNICAMP, 2010.

OPEN AI, 2015. Discovering and enacting the path to safe artificial general intelligence. Disponível em: < https://openai.com/>. Acesso em: Maio de 2020.

PINOCHET, L.H.C; LOPES, A. S.L; SILVA, J. S. **Inovações e tendências aplicadas nas tecnologias de informação e comunicação na gestão da saúde.** RGSS, 2014.

PhysioNet. **The Research Resource for Complex Physiologic Signals**. MIT Laboratory for Computational Physiolog, 2020.

RIERA, A. R. P; BARROS, R. B; SHENASA, M. **Electrocardiographic Markers of Sudden Cardiac Death (Including Left Ventricular Hypertrophy)**. Card Electrophysiol Clin 9. Elsevier, 2017.

SANDHAM, W., HAMILTON, D. LAGUNA, P., COHEN, M. **Advances in Electrocardiogram Signal Processing and Analysis.** EURASIP Journal on Advances in Signal Processing, v. 2007, article ID 69169, 5 pages, 2007.

SENAC (n.d.). Guia de Normalização SENAC.

**SUDDEN CARDIAC DEATH**. Report of a WHO Scientific Group, World Health Organization. Geneva, 1985. Disponível em: <https://apps.who.int/iris/handle/10665/39554>. Acesso em: Maio de 2020.

SUNASRA, Mohammed. **Performance Metrics for Classification problems in Machine Learning**. 2017. Disponível em: <https://medium.com/@MohammedS/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b#:~:text=We%20can%20use%20classification%20performance,primarily%20used%20by%20search%20engines>. Acesso em: Junho de 2020.

TECNOLOGIA NA SAÚDE: A IMPORTÂNCIA DA TECNOLOGIA NO DIAGNÓSTICO. **Optivision Brasil**, 2018. Disponível em: < https://www.optivisionbrasil.com.br/blog/tecnologia-na-saude-a-importancia-da-tecnologia-no-diagnostico/>. Acesso em: Maio de 2020.

TOMPKINS J.W. **Biomedical Digital Signal Processing**. Prentice Hall, 1993.

TOMAZATI, A. O. **Detecção do complexo QRS em sinais cardíacos utilizando FPGA**. Dissertação de mestrado. Universidade Estadual de Campinas, Campinas – SP, Brasil. 2010.

VIJAY, C. J. et al. **A comprehensive review on mcculloch-pitts neuron model.** International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(6):1516-1522, Apr 2019.

VEEN, F. V. **Neural network zoo prequel: cells and layers.** The Asimov Institute, 2017**.** Disponível em: <https://www.asimovinstitute.org/author/fjodorvanveen/>. Acesso em: Maio de 2020.

ZHENG, Z.J. et al. **Sudden cardiac death in the United States, 1989 to 1998**. Circulation, 2001. Disponível em: <https://www.ahajournals.org/doi/epub/10.1161/hc4301.098254>. Acesso em: Maio de 2020.

ZIPES, D. P; WELLENS, H. J. J. **Sudden Cardiac Death**. Kluwe, Academic Publishers, 2000.

ZUBEN, F. J. V; **Redes Neurais Aplicadas ao Controle de Máquina de Indução**. Capítulo 2: Redes Neurais Artificiais. UNICAMP, 1993.

# APÊNDICE

PACIENTES SAUDÁVEIS:

Database ECG-IG: https://www.physionet.org/content/ecgiddb/1.0.0/

Database Combined measurement of ECG, Breathing and Seismocardiograms: https://www.physionet.org/content/cebsdb/1.0.0/

PACIENTES DOENTES:

Database Sudden Cardiac Death Holter : https://www.physionet.org/content/sddb/1.0.0/

Database BIDMC Congestive Heart Failure: https://physionet.org/content/chfdb/1.0.0/

1. Estes parâmetros citados pelos autores referem-se aos potenciais de ação analisados pelos ECGs. [↑](#footnote-ref-1)
2. O Feixe de His, ou fascículo atrioventricular, é um feixe de fibras localizado no interior do músculo cardíaco do septo interventricular. [↑](#footnote-ref-2)
3. Unidade de medida da diferença de potencial, equivalente a milionésimo de volt. Símb: µV. Etimologia (origem da palavra microvolt). [↑](#footnote-ref-3)
4. **Eletrodos cardíacos** são pequenos dispositivos que ficam em contato com pele do paciente durante o eletrocardiograma afim de coletar a atividade elétrica na pele do mesmo. [↑](#footnote-ref-4)
5. A oposição à passagem de uma corrente alternada oferecida por um capacitor é denominada "reatância capacitiva". [↑](#footnote-ref-5)
6. Em estatística entende-se população como uma amostra de dados ou eventos semelhantes. [↑](#footnote-ref-6)
7. Processo onde o modelo aprende com os dados modulando sua função de erro, validação = processo onde é testado a saída do modelo com base em dados reais para verificar a sua performance. [↑](#footnote-ref-7)
8. Em Data Science produtizar significa colocar o modelo em ambiente de produção. [↑](#footnote-ref-8)
9. Apache Spark é um framework de código fonte aberto para computação distribuída. [↑](#footnote-ref-9)
10. Implantação de Software. [↑](#footnote-ref-10)
11. API (Application Program Interface), é ponte de conexão entre sistemas para envio de dados. [↑](#footnote-ref-11)
12. Stream é fluxo de dados em tempo real de um sistema computacional. [↑](#footnote-ref-12)
13. Object Schema é um esquema de banco de dados movidos a arquivos. [↑](#footnote-ref-13)
14. Hiperparâmetros são parâmetros que devem ser definidos para a otimização de um modelo. [↑](#footnote-ref-14)
15. O Kanban é um método organizacional que visa aumentar a produtividade e otimizar a realização das tarefas e das entregas. [↑](#footnote-ref-15)
16. Git é um sistema de controle de versões distribuído, usado principalmente no desenvolvimento de software. [↑](#footnote-ref-16)
17. *overfitting* ou sobreajuste é um termo usado em estatística para descrever quando um modelo estatístico se ajusta muito bem ao conjunto de dados anteriormente observado, mas se mostra ineficaz para prever novos resultados. [↑](#footnote-ref-17)
18. No aprendizado de máquina, um hiperparâmetro é um parâmetro cujo valor é usado para controlar o processo de aprendizagem. [↑](#footnote-ref-18)