



Centro de Ciências Biológicas e da Saúde

Programa de Doutorado em Biotecnologia – Rede Nordeste de Biotecnologia

(RENORBIO)

**Classificação da Doença Arterial Coronariana usando a Variabilidade da  
Frequência Cardíaca: Uma abordagem de Rede Neural Profunda com Explicação  
Agnóstica de Modelo Interpretável Localmente (LIME)**

*São Luís*

2023

*Antonio Carlos Pereira Silva Filho*

**Classificação da Doença Arterial Coronariana usando a Variabilidade da  
Frequência Cardíaca: Uma abordagem de Rede Neural Profunda com Explicação  
Agnóstica de Modelo Interpretável Localmente (LIME)**

Tese apresentada ao Programa de Pós-Graduação  
Rede Nordeste de Biotecnologia da Universidade  
Federal do Maranhão, com o objetivo de qualificação  
para adquirir título de Doutor em Biotecnologia.

**Área de Concentração:** Biotecnologia em Saúde

**Orientador:** Profa. Dra. Maria do Socorro Cartágenes

**Co-orientador:** Prof. Dr. Marcelo Sousa Andrade

**Coordenador:** Prof. Dr. Antonio Marcus Paes

São Luís

2023

Silva Filho, Antonio Carlos.

Classificação das Doença Arterial Coronariana usando a Variabilidade da Frequência Cardíaca : Uma abordagem de Rede Neural Profunda com Explicação Agnóstica de Modelo Interpretável Localmente LIME / Antonio Carlos Silva Filho. - 2023.

38 p.

Coorientador(a): Marcelo Sousa de Andrade.

Orientador(a): Maria do Socorro de Sousa Cartágenes.

Tese (Doutorado) - Programa de Pós-graduação em Biotecnologia - Renorbio/ccbs, Universidade Federal do Maranhão, São Luís, 2023.

1. Aprendizado de Máquina. 2. LIME. 3. Rede Neural Profunda. 4. Variabilidade da Frequência Cardíaca. I. de Sousa Cartágenes, Maria do Socorro. II. Sousa de Andrade, Marcelo. III. Título.

UNIVERSIDADE FEDERAL DO MARANHÃO  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO REDE NORDESTE DE BIOTECNOLOGIA

ANTONIO CARLOS PEREIRA SILVA FILHO

**Classificação das Doença Arterial Coronariana usando a Variabilidade da  
Frequência Cardíaca: Uma abordagem de Rede Neural Profunda com Explicação  
Agnóstica de Modelo Interpretável Localmente (LIME)**

Aprovado em \_\_\_\_\_ de \_\_\_\_\_ de 2023

---

Profa. Dra. Maria do Socorro Cartágenes  
Orientadora

---

Profa. Dra. Nilviane Soares Pires  
Examinador I

---

Profa. Dr. Giordano Bruno Soares Souza  
Examinador II

---

Profa. Dra. Aldeídia Pereira de Oliveira  
Examinador III

---

Profa. Dra. Maria do Desterro Soares Brandão  
Suplente

---

Profa. Dr. Giovanni Lucca França da Silva  
Suplente

## **Agradecimentos**

Antes de mais nada a Deus, autor da criação, e que tudo sabe e tudo vê. Aos meus pais, Antônia Silva Oliveira, que sempre teve o sonho de ser professora, mas que devido às dificuldades da vida, não conseguiu. Ao meu pai, meu chará (na verdade, eu xará dele), Antonio Carlos, motorista de caminhão que terá a honra de chamar seu filho de Doutor.

À minha querida irmã desconhecida para mim, Caroline, que faleceu aos dois anos de idade. Quisera eu ter a chance de ter te conhecido e compartilhado essa alegria, quem sabe seria mais inteligente e dedicada do que eu jamais poderia ser.

À minha princesa herdeira, Liz, que veio em uma hora tão turbulenta, mas que é a luz que ilumina minha vida, e que me dá forças pra lutar e continuar todo dia.

À minha rainha e companheira, Sara. Passamos tantas coisas juntos, tantas dificuldades, quero muito ter a chance de te chamar de Doutora também. Você merece o mundo meu amor, e eu estou aqui pra te dar. Te amo.

*“Savor what you feel and what you see  
Things that may not seem important now  
But may be tomorrow”*

*Chuck Schuldiner na música Symbolic, Death*

## **Classificação de Doença Arterial Coronariana usando a Variabilidade da Frequência Cardíaca: Uma abordagem de Rede Neural Profunda com Explicação Agnóstica de Modelo Interpretável Localmente (LIME)**

### **Resumo**

**Introdução:** A Doença Arterial Coronariana é umas das doenças cardiovasculares que mais mata no mundo hoje. O Aprendizado de Máquina tem sido amplamente aplicado para detectar precocemente indicadores de risco de doença cardiovascular em pacientes no mundo inteiro. Entretanto, modelos complexos e sem clareza para os profissionais da saúde impedem e dificultam a implementação dos modelos de Aprendizado de Máquina em escala. **Objetivo:** Visamos aplicar um Rede Neural Profunda capaz de detectar sujeitos com Doença Arterial Coronariana de sujeitos saudáveis, usando índices de Variabilidade da Frequência Cardíaca como parâmetro, com o menor número possível de variáveis, aumentando a confiança nos resultados do modelo e fazendo a coleta de dados para a implantação do modelo em um ambiente clínico. **Métodos:** Utilizamos dados coletados de eletrocardiogramas 24h (holter) do banco de dados *Telemetric and Holter Electrocardiogram Warehouse* (THEW) de 354 pacientes. Os dados de variabilidade da frequência cardíaca foram extraídos dos holters e funcionaram como entrada de uma rede neural profunda, e as variáveis mais explicativas dos modelos foram encontradas usando o algoritmo de explicação local agnóstica de modelo (LIME). **Resultados:** O domínio do tempo e da frequência mostrou maior precisão e menor perda. O LIME foi usado para identificar as cinco variáveis mais explicativas, que mais tarde foram reintroduzidas no modelo separadamente. A precisão de previsão e a perda foram mantidas, indicando que as variáveis destacadas pela LIME foram

as mais importantes e mais explicativas. **Conclusão:** Os dados de variabilidade da frequência cardíaca podem ser usados para classificar a doença arterial coronariana de indivíduos saudáveis usando redes neurais profundas, e a LIME pode simplificar o modelo, aumentando sua confiabilidade.

**Palavras-chave:** variabilidade da frequência cardíaca, rede neural profunda, aprendizado de máquina, LIME



## **Classification of Coronary Artery Disease using Heart Rate Variability: A Deep Neural Network approach using Locally Interpretable Model-agnostic Explanation (LIME)**

### **Abstract**

**Aim:** We aimed to create a DNN capable of detecting subjects with coronary artery disease from healthy subjects, using heart rate variability indexes as a parameter, with as few variables as possible, increasing trust in model output and making the data collection for model deployment in a clinical setting. **Methods:** We used data collected from 24h holter data from Telemetric and Holter Electrocardiogram Warehouse (THEW) database of 354 patients. Heart rate variability data was extracted from the holters and worked as input of a deep neural network, and the most explanatory models' variables were found using the LIME. **Results:** Time and frequency domain showed higher accuracy and lower loss. LIME was used to identify the five most explanatory variables, that were later reintroduced to the model separately. The accuracy and loss were maintained, indicating that the variables highlighted by LIME were the most important and most explanatory. **Conclusion:** Heart rate variability data can be used to evaluate coronary artery disease from healthy subjects using deep neural networks, and the LIME can simplify the model, increasing its trustworthiness.

**Keywords:** heart rate variability; deep neural network; machine learning; LIME

## Sumário

Lista de siglas.....	11
Introdução.....	12
Objetivos.....	14
<b>Capítulo I.....</b>	<b>15</b>
<b>Revisão de Literatura.....</b>	<b>15</b>
Referencial teórico.....	16
A Doença Cardiovascular.....	16
O Sistema Nervoso Autônomo.....	18
Variabilidade da Frequência Cardíaca.....	21
Referências.....	25
<b>Capítulo II.....</b>	<b>26</b>
<b>Artigo submetido.....</b>	<b>26</b>
<b>Anexo I.....</b>	<b>36</b>
<b>Material Suplementar.....</b>	<b>36</b>
<b>Anexo II - Certificado de Registro de Software.....</b>	<b>42</b>
<b>Anexo III - Comprovante de Submissão.....</b>	<b>44</b>
<b>Anexo IV - Artigo Publicado.....</b>	<b>46</b>

## Lista de siglas

AVE: Acidente Vascular Encefálico

DAC: Doença Arterial Coronariana

DCV: Doença Cardiovascular

DNN: *Deep Neural Network*

FC: Frequência Cardíaca

LIME: *Locally Interpretable Model-agnostic Explanation*

ML: *Machine Learning*

LF: *low frequency* (baixa frequência)

HF: *high frequency* (alta frequência)

IM: Infarto do Miocárdio

SDNN: *Standard Deviation of Normal-to-Normal intervals*

HR: *Heart Rate* (frequência cardíaca)

SDANN: *Standard Deviation of All Normal-to-Normal intervals*

VLF: *Very Low Frequency* (muito baixa frequência)

SD1: *Standard Deviation 1* (desvio padrão 1)

SD2: *Standard Deviation 2* (desvio padrão 2)

VFC: Variabilidade da Frequência da Cardíaca

## Introdução

A doença cardiovascular (DCV) tem sido a principal causa de morte no mundo desde 1990, com aproximadamente 240 mortes a cada 100.000 casos no mundo inteiro (1), sendo a maioria dos casos Infarto Agudo do Miocárdio e Acidente Vascular Encefálico (2). A variabilidade da frequência cardíaca (VFC) tem sido reconhecida como um sinal robusto de saúde cardiovascular e um marcador não invasivo de doenças cardiovasculares La Rovere et al. (2003 e Wulsin et al. (2015). A VFC pode ser facilmente extraída do monitoramento de rotina com eletrocardiograma, tornando a coleta e análise dos dados uma tarefa fácil para clínicos e pesquisadores.

As tentativas bem sucedidas em uma grande variedade de campos têm aumentado o interesse no estudo do Aprendizado de Máquina (*Machine Learning*, ML) na área da saúde pública, de tal forma que, os esforços recentes têm utilizado dados relacionados à saúde como parâmetros para modelos de ML, na tentativa da identificação precoce da DCV, como o Acidente Vascular Encefálico Dritsas; Trigka (2023 e Livne et al. (2018 e Nawabi et al. (2021) e o IM Dritsas; Trigka (2023 e Mandair et al. (2020). Um dos modelos ML mais utilizados é conhecido como Aprendizagem Profunda, que é composto de uma Rede Neural Profunda (*Deep Neural Network*, DNN), capaz de encontrar padrões em dados não estruturados, melhorando em relação aos modelos ML tradicionais Camburu (2020).

Entretanto, estudos recentes apontaram que o diagnóstico ML "out-of-the-box"<sup>1</sup> não é confiável para os clínicos e, ainda mais, pode não levar a

---

<sup>1</sup> Modelos de ML que não são individualizados para cada caso e que não permitem (ou são muito complexos) entendimento da razão ou do padrão encontrado para justificar as decisões tomadas.

aumentos gerais na qualidade da decisão tomada pelos profissionais Ribeiro et al. (2016).

Com base nesta questão de confiança, um grupo de cientistas da Universidade de Washington desenvolveu um modelo explicativo, chamado Explicação Agnóstica de Modelo Interpretável Localmente (LIME) Ribeiro et al. (2016). O LIME é capaz de explicar qualquer modelo de ML, retornando as variáveis mais explicativas para previsões individuais, escolhendo se a previsão é precisa e dando a oportunidade para a otimização do modelo.

## **Objetivos**

### **Objetivo Geral**

Implementar uma DNN capaz de detectar sujeitos com doença arterial coronária de sujeitos saudáveis, usando índices de VFC como parâmetro, com o mínimo possível de variáveis, aumentando a confiança nos resultados do modelo, e facilitando a coleta de dados para a implantação do modelo em um ambiente clínico.

### **Objetivos Específicos**

- Determinar a acurácia do modelo na classificação de pacientes com DAC de pacientes saudáveis utilizando a VFC;
- Determinar quais as variáveis mais explicativas para a classificação de pacientes com DAC de pacientes saudáveis utilizando a VFC;
- Reduzir o número de variáveis necessárias para a utilização do modelo;

# *Capítulo I*

## *Revisão de Literatura*

## Referencial teórico

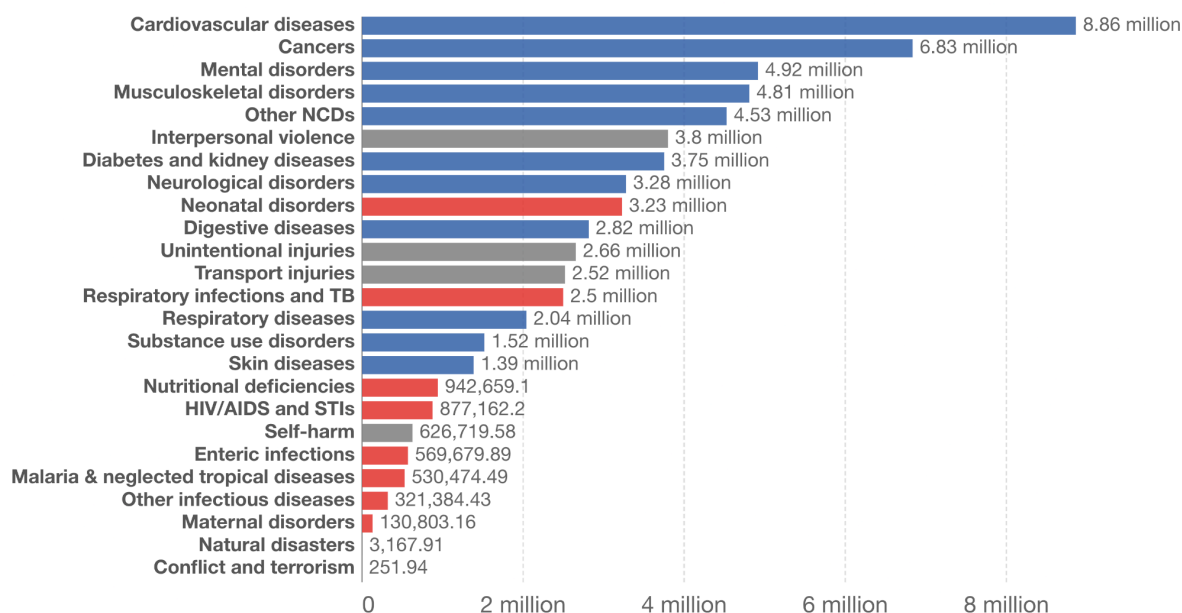
### A Doença Cardiovascular

A doença cardiovascular é a doença que mais mata no mundo moderno. Ao analisar o quadro da distribuição de causa de morte no Brasil, 15,52% das mortes foram causadas por DCV's (Figura 1). Apenas como exercício de comparação, desde Março de 2020, houve um acumulado de aproximadamente 692.000 mortes causadas por COVID-19, o que representa apenas 7% do total de mortes causadas por DCV's apenas em 2019.

### Burden of disease by cause, Brazil, 2019



Total disease burden, measured in Disability-Adjusted Life Years (DALYs) by sub-category of disease or injury. DALYs measure the total burden of disease – both from years of life lost due to premature death and years lived with a disability. One DALY equals one lost year of healthy life.



Source: IHME, Global Burden of Disease (2019)

OurWorldInData.org/burden-of-disease • CC BY

Note: Non-communicable diseases are shown in blue; communicable, maternal, neonatal and nutritional diseases in red; injuries in grey.

**Figura 1.** Total de mortes e suas causas no Brasil em 2019. Fonte: *OurWorldInData*

A fisiopatologia da doença cardiovascular é diversa e extremamente complexa, incluindo múltiplos fatores, desde fatores genéticos, ambientais, comportamentais, alimentares e emocionais. Além destes, a relação direta e indireta entre outras



doenças crônicas, como a Diabetes, Hipertensão Arterial, Obesidade e Sedentarismo, acentua de forma vertiginosa o risco de desenvolvimento de DCV's ou, de eventos cardiovasculares agudos, como O Infarto do Miocárdio e o Acidente Vascular Encefálico.

Ao longo dos anos, vários estudos acompanharam e determinaram indicadores que pudessem antever ou apontar risco de evento cardiovascular. Um dos mais famosos e conhecidos é o Framingham Heart Study Tsuji et al. (1996), que acompanha moradores da comunidade americana de Framingham. Este estudo teve descobertas extremamente importantes para o entendimento dos fatores de risco para as DCV's, como por exemplo, a hipercolesterolemia, hipertensão, obesidade, tabagismo, sedentarismo e diabetes. (Nayor et al. (2016)

Embora a doença cardiovascular seja uma causa comum de morte, é possível preveni-la ou retardar seu desenvolvimento através de estilos de vida saudáveis Dimovski et al. (2019). A prevenção primária inclui o controle de fatores de risco modificáveis, tais como dieta saudável, exercício regular, não fumar, controle de pressão arterial e diabetes, e tratamento precoce de fatores de risco não modificáveis, como história familiar de doença cardiovascular. (Dimovski et al. (2019 e Johansson et al. (2021)

Além disso, a monitoração de indicadores de risco cardiovascular é fundamental, pois permite identificar precocemente indivíduos com maior risco de desenvolver doença cardiovascular. A avaliação da variabilidade da frequência cardíaca (VFC) é um dos indicadores utilizados para isso, pois é um reflexo do equilíbrio entre o sistema nervoso simpático e parassimpático, e pode indicar o risco de eventos cardiovasculares futuros. (Thayer et al. (2010)

## *O Sistema Nervoso Autônomo*

O Sistema Nervoso Autônomo é a parte do sistema nervoso responsável por regular as funções involuntárias. O Sistema Nervoso Autônomo controla a maior parte das funções viscerais do organismo, e uma característica importante deste sistema é a velocidade e intensidade em que ele modifica funções, por exemplo a frequência cardíaca (FC) pelas vias simpática e parassimpática, estimulando o nó sinoatrial, fazendo-o despolarizar mais, e desta forma aumentar a FC, ou despolarizar menos, diminuindo a FC, tendo seus valores normais duplicados num curto espaço de tempo de 3 segundos; ou até mesmo em 10 a 15 segundos ele é capaz de aumentar a pressão arterial duas vezes em relação a pressão basal, ou nesse mesmo tempo pode diminuir a pressão arterial a níveis tão baixos ao ponto de causar desmaios (GUYTON, HALL, 2017).

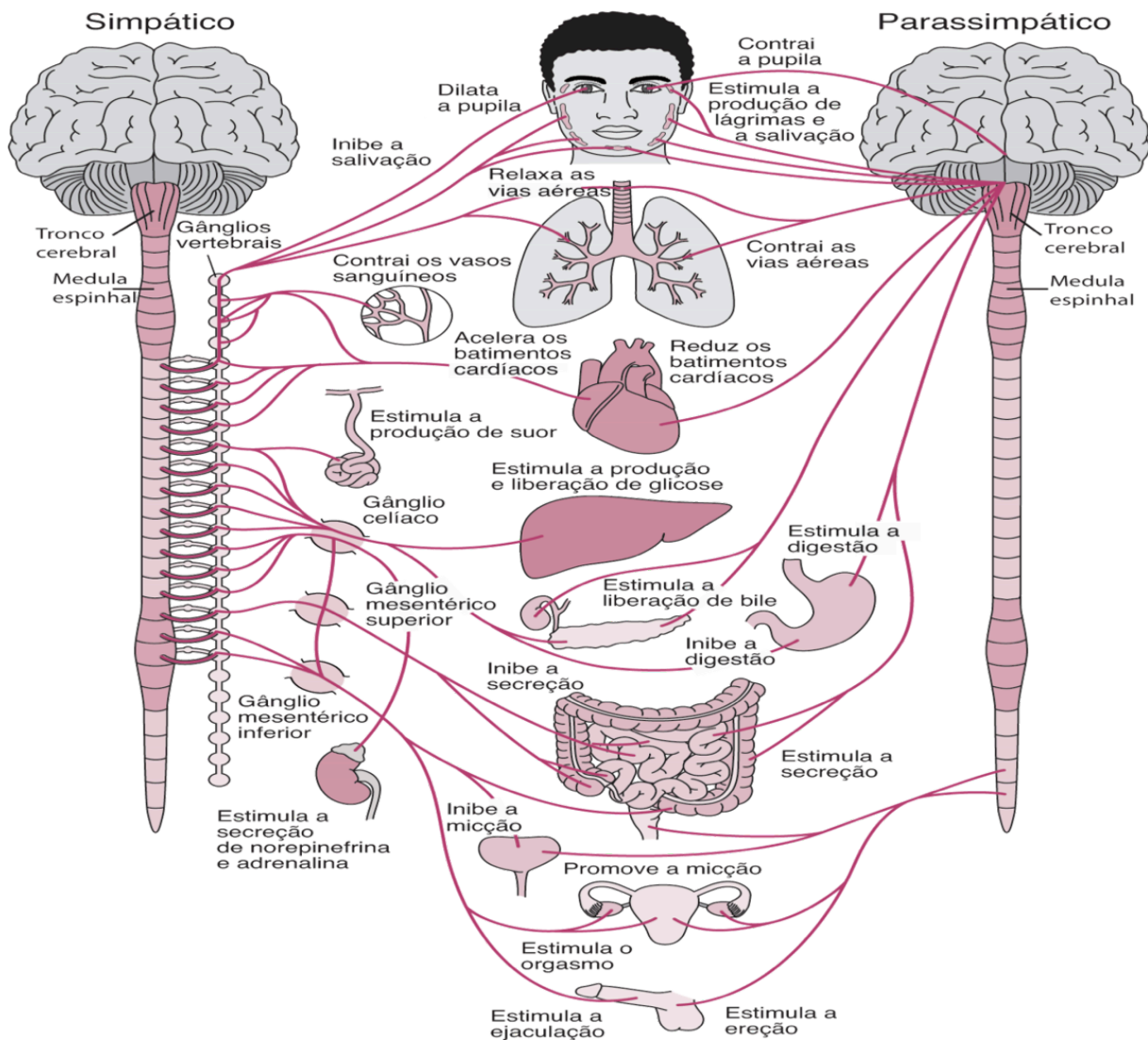
O SNS tem em sua organização geral duas cadeias de gânglios simpáticos paravertebrais, dois pré-vertebrais (o celíaco e o hipogástrico), e nervos que vão dos gânglios para vários órgãos internos. A formação do componente simpático é iniciada através dos nervos simpáticos que são originados nos segmentos T1 a L2 da medula espinhal, em seguida se distribuem para os órgãos sendo controlados por estímulos (BORTOLOTTO et al., 2013).

Tais estímulos são coordenados por componentes neurais com estruturas colinérgicas, chamadas neurônios pré-ganglionares (o corpo celular fica na ponta intermédio-lateral da medula espinhal e a fibra é direcionada para o nervo espinhal correspondente por uma raiz anterior da medula), e a estrutura adrenérgica, que são os neurônios pós-ganglionares, tem sua origem como maioria em um dos gânglios de cadeia simpática ou em gânglios pré-vertebrais, estimuladas pelo simpático. Esta medula secretará catecolaminas, principalmente adrenalina e noradrenalina, que

atuarão nos órgãos através da circulação sanguínea, onde seguem em direção aos órgãos efetores (BORTOLOTTI et al., 2013; GUYTON, HALL, 2017).

O SNP deixa o SNC pelos nervos cranianos (III [oculomotor], VII [facial], IX [glosssofaríngeo], e X [vago]), e pelo segundo e terceiro nervos espinhais, porém 75% das fibras nervosas parassimpáticas estão contidas no nervo vago, distribuídas para o coração, fígado e pulmão, principalmente, entre outros órgãos (GUYTON, HALL, 2017).

Através dos nervos sacrais 2 e 3 as fibras parassimpáticas sacrais saem do plexo sacral de cada lado da medula para serem distribuídas para o cólon descendente. Os neurônios pré-ganglionares e pós-ganglionares do SNP são ambos formados por fibras colinérgicas na sua maioria. As fibras pré-ganglionares passam por todo um trajeto até chegar ao órgão que vai estimular, havendo casos que as fibras pós-ganglionares estejam localizadas na própria parede dos órgãos que vai inervar. Contudo, existem algumas exceções de nervos que não seguem essa rota, como alguns nervos cranianos parassimpáticos (GUYTON, HALL, 2017).



**Figura 1.** Sistema nervoso autônomo. Fonte: Low (2023)

Os SNS e SNP se diferenciam também pelos neurotransmissores que liberam em suas terminações sinápticas, na porção pós-ganglionar. Entretanto, na sinapse entre neurônios pré e pós-ganglionares ocorre a liberação de acetilcolina para ambos. Assim, pode-se dividir os dois sistemas em: Colinérgico (SNP), onde há liberação de acetilcolina, e Adrenérgico (SNS), que secretam, principalmente, noradrenalina, exceto em órgãos onde o SNS passa a ser colinérgico (DOUGLAS, 2000).

Sabe-se da necessidade de ocorrer uma neuromodulação no corpo humano denominada de homeostase, é com a modulação neural que o sistema cardiovascular através da influência tônica e reflexa sofre alterações de forma aguda e crônica, podendo sofrer alterações também a nível fisiológico e patológico (MOSTARDA et al., 2009).

A inervação de nervos aferentes e eferentes para o coração é feita pelas terminações simpáticas que cobrem todo miocárdio, e os nervos parassimpáticos sendo encontrados no nodo sinoatrial e nodo atrioventricular, onde o controle neural depende intimamente da FC, sendo dependente de um bom funcionamento barorreflexo (MOSTARDA et al., 2009).

#### *Variabilidade da Frequência Cardíaca*

A Variabilidade da Frequência Cardíaca (VFC) é uma análise indireta da modulação Sistema Nervoso Autônomo, que é um sistema neural importante para controlar e regular vários processos fisiologicamente importantes para o corpo Strüven et al. (2021). A VFC é uma avaliação da variação da frequência cardíaca em resposta a estímulos externos e internos, como a respiração, mudanças na posição do corpo, atividades físicas e estados emocionais. Ela pode ser usada para avaliar a saúde cardiovascular, as respostas ao estresse e até mesmo a resposta a medicamentos. (Kim et al. (2018 e Pumprla et al. (2002)

A VFC é composta de vários domínios, como o domínio do tempo, que analisa a variabilidade no espaço de tempo entre os batimentos cardíacos; o domínio da frequência, que analisa o espectro de frequência dos batimentos cardíacos. Cygankiewicz; Zareba (2013 e Malik (1996).

Vários estudos ao longo dos anos têm demonstrado a correlação da baixa VFC e do desbalanço autonômico (quando há predominância da modulação simpática sobre a parassimpática) com a doença cardiovascular e aumento do risco de morte súbita. Montano et al. (2009) e Rovere et al. (1998) e Thayer et al. (2010).

Denomina-se VFC as variações do ciclo cardíaco no tempo transcorrido entre duas ondas R consecutivas (R-R), que são intervalo entre os batimentos consecutivos do eletrocardiograma. A FC é regulada de forma instantânea por vários mecanismos fisiológicos, tanto pela atividade simpática quanto pela vagal. Sendo assim, a FC tem seu ritmo sinusal normal devido a influência dinâmica de vários mecanismos fisiológicos que o regulam instantaneamente, como a atividade simpática e parassimpática que são um dos mecanismos responsáveis por esse controle (NGOMANE et al., 2022).

Dessa forma, quanto maior a variabilidade de tempo dos R-R, maior será a atividade parassimpática, quando o inverso acontece existe maior atividade simpática. A praticidade na coleta da FC tem sido ponto chave para o crescente número de estudos em diferentes condições associadas ao repouso e ao exercício. O equilíbrio entre as atividades simpática e parassimpática influencia de forma significativa diversas condições funcionais e clínicas, o que torna imprescindível avaliar esta modulação (MOSTARDA et al., 2009).

No componente fisiológico de batimentos cardíacos, as terminações parassimpáticas liberam o neurotransmissor acetilcolina na fenda sináptica que causa a despolarização do nodo sinoatrial e, por apresentar uma velocidade de remoção muito rápida, provoca oscilações na duração dos intervalos R-R, acarretando em variações rítmicas na FC. Já a noradrenalina, liberada pelos terminais simpáticos, possui uma velocidade de remoção lenta, ocasionando uma

variação rítmica na FC, que pode ser observada somente em registros de longo prazo, definindo a VFC como ajuste entre a modulação rápida e a lenta. Contudo, essas variações na FC são atribuídas, principalmente, pelas oscilações da atividade parassimpática e, portanto, a amplitude da VFC reflete a atividade vagal sobre o coração (NGOMANE et al., 2022).

O método da VFC permite analisar o controle neural cardíaco em diversos períodos tanto em curtos quanto em longos, além das condições fisiológicas (durante o sono, monitoramento de 24 horas, repouso, exercício físico e bloqueio farmacológico), e patológicas (NGOMANE et al., 2022).

Os métodos para análise da VFC podem ser feitos por cálculo de índices baseado em operações estatísticas dos intervalos R-R (domínio do tempo), e por uma análise espectral de intervalos R-R ordenados (domínio da frequência). Essas análises são realizadas em segmentos curtos - 0,5 a 5 minutos - ou em gravações de eletrocardiograma em períodos longos de 24 horas. (MOSTARDA et al., 2009).

Os métodos geométricos transformam uma série de intervalos R-R em um gráfico para posterior cálculo da distribuição da densidade de sinais. As medidas do domínio da frequência são derivadas da análise do espectro de potência que apresenta a distribuição da densidade em função da frequência. Esta análise decompõe a FC em seus componentes causadores, os apresentando de acordo com a frequência com que alteram a FC (MOSTARDA et al., 2010).

Entre os componentes de alta e baixa frequência, que são assim chamados devido ao fato do nervo vago e o sistema simpático enviarem, respectivamente, uma maior ou menor frequência de impulsos sobre o nodo sinusal, são mensurados em unidades absolutas de potência ( $\text{ms}^2$ ), ou podem ser expressos em valores relativos à potência total subtraindo-se o componente de muito baixa frequência (unidades

normalizadas, alta frequência %, e baixa frequência %). A relação entre a potência de baixa para a de alta frequência (baixa frequência/alta frequência) pode ser considerada uma medida do balanço simpátovagal (MOSTARDA et al., 2009).

### *Aprendizado Supervisionado e Não Supervisionado*

O aprendizado supervisionado e não supervisionado são duas abordagens fundamentais no campo do aprendizado de máquina Schmidt et al. (2024). No aprendizado supervisionado, os algoritmos são treinados usando dados rotulados. Isto é, cada exemplo de dados no conjunto de treinamento vem com um rótulo correspondente ou resultado. O algoritmo aprende a prever o rótulo a partir das características dos dados. Por outro lado, no aprendizado não supervisionado, os algoritmos são treinados usando dados não rotulados. Isso significa que o algoritmo não recebe informações sobre os resultados esperados, e sim sobre a estrutura dos dados. O objetivo do aprendizado não supervisionado é encontrar padrões, agrupamentos ou relações entre os dados, sem a necessidade de rótulos prévios.

Na área de aprendizado de máquina, os algoritmos de aprendizado supervisionado e não supervisionado desempenham funções cruciais na análise e previsão de dados. Para o aprendizado supervisionado, os três algoritmos mais comumente usados são Regressão Linear, Árvores de Decisão e Máquinas de Vetores de Suporte. Esses algoritmos aprendem com dados rotulados para prever resultados de dados novos e não vistos Rashidi et al. (2021). Por outro lado, os algoritmos de aprendizado não supervisionado, como K-means Clustering, Hierarchical Clustering e Principal Component Analysis, são projetados para identificar padrões e estruturas em dados não rotulados Yan; Wang (2022).



## *O Aprendizado de Máquina e a Doença Cardiovascular*

A utilização de aprendizado de máquina tem sido amplamente explorada no contexto da doença cardiovascular. Através da análise avançada de dados, algoritmos de *machine learning* podem identificar padrões e prever eventos cardiovasculares com alta precisão. Essa abordagem promissora pode facilitar o diagnóstico precoce, prevenir complicações e melhorar os resultados clínicos para pacientes com doenças cardíacas Dritsas; Trigka (2023 e Terzi; Arikan (2023).

O conceito central do aprendizado de máquina está em sua capacidade de aprender com os dados e fazer previsões ou tomar decisões sem ser explicitamente programado. Isso é obtido por meio do uso de algoritmos que aprendem iterativamente com os dados, ajustando seus parâmetros para otimizar o desempenho. No contexto das doenças cardiovasculares, os algoritmos de aprendizado de máquina podem ser treinados em grandes conjuntos de dados que contêm uma ampla gama de informações sobre os pacientes, incluindo dados demográficos, histórico médico, testes de diagnóstico e resultados de tratamentos Chen et al. (2023 e Qayyum (2023).

Entre muitos algoritmos, as redes neurais profundas, especificamente, revolucionaram o campo da inteligência artificial (IA) e demonstraram desempenho excepcional em vários domínios, como visão computacional, processamento de linguagem natural e reconhecimento de fala. As redes neurais profundas (DNNs) são uma classe de algoritmos de aprendizado de máquina inspirados na estrutura e no funcionamento do cérebro humano Camburu (2020 e Sugomori et al. (2017).

As DNN's consistem em várias camadas de neurônios artificiais interconectados, também conhecidos como nós. Cada nó recebe entradas da camada anterior e executa uma série de operações matemáticas nessas entradas para produzir uma saída. Essas saídas são então passadas para a próxima camada e esse processo continua até a camada final, que produz a saída desejada Camburu (2020).

A principal característica das DNN's é sua capacidade de aprender e extrair automaticamente recursos significativos de dados brutos. Isso é obtido por meio de um processo chamado treinamento, em que a rede é exposta a um grande conjunto de dados e ajusta seus parâmetros internos para minimizar a diferença entre a saída prevista e a saída real. Esse processo envolve o uso de uma função de perda, que quantifica a diferença entre os resultados previstos e os reais, e um algoritmo de otimização, que atualiza os parâmetros da rede para minimizar essa diferença Kingma; Ba (2014).

A utilização de algoritmos de Machine Learning (ML) na área da saúde tem o potencial de revolucionar a forma como as doenças cardiovasculares são diagnosticadas. Estes algoritmos, quando treinados com grandes conjuntos de dados de saúde, podem aprender a identificar padrões e fazer previsões com uma precisão que muitas vezes supera a dos humanos.

## Referências

- ABADI, M.; AGARWAL, A.; BARHAM, P.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. **arXiv [cs.DC]**, 14. mar. 2016. Disponível em: <<http://arxiv.org/abs/1603.04467>>. .
- AGARAP, A. F. Deep Learning using Rectified Linear Units (ReLU). **arXiv [cs.NE]**, 22. mar. 2018. Disponível em: <<http://arxiv.org/abs/1803.08375>>. .
- CAMBURU, O.-M. **Explaining Deep Neural Networks**. 2020.
- CHEN, L.; JI, P.; MA, Y.; RONG, Y.; REN, J. Custom machine learning algorithm for large-scale disease screening - taking heart disease data as an example. **Artificial intelligence in medicine**, v. 146, p. 102688, 2023.
- CYGANKIEWICZ, I.; ZAREBA, W. Heart rate variability. **Handbook of clinical neurology**, v. 117, p. 379–393, 2013.
- DIMOVSKI, K.; ORHO-MELANDER, M.; DRAKE, I. A favorable lifestyle lowers the risk of coronary artery disease consistently across strata of non-modifiable risk factors in a population-based cohort. **BMC public health**, v. 19, n. 1, p. 1575, 2019.
- DRITSAS, E.; TRIGKA, M. Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction. **Sensors**, v. 23, n. 3, 2023. Disponível em: <<http://dx.doi.org/10.3390/s23031161>>. .
- GBD 2019 RISK FACTORS COLLABORATORS. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. **The Lancet**, v. 396, n. 10258, p. 1223–1249, 2020.
- GBD 2019 VIEWPOINT COLLABORATORS. Five insights from the Global Burden of Disease Study 2019. **The Lancet**, v. 396, n. 10258, p. 1135–1159, 2020.
- JACOBS, M.; PRADIER, M. F.; MCCOY, T. H.; et al. How machine-learning recommendations influence clinician treatment selections: the example of antidepressant selection. **Translational psychiatry**, v. 11, n. 1, p. 1–9, 2021. Nature Publishing Group. Acesso em: 9/10/2022.
- JAYALAKSHMI, T.; SANTHAKUMARAN, A. Statistical normalization and back propagation for classification. **International Journal of Computer Theory and Engineering**, v. 3, n. 1, p. 1793–8201, 2011.
- JOHANSSON, A.; DRAKE, I.; ENGSTRÖM, G.; ACOSTA, S. Modifiable and Non-Modifiable Risk Factors for Atherothrombotic Ischemic Stroke among Subjects in the Malmö Diet and Cancer Study. **Nutrients**, v. 13, n. 6, 2021. Disponível em: <<http://dx.doi.org/10.3390/nu13061952>>. .
- KIM, H.-G.; CHEON, E.-J.; BAI, D.-S.; LEE, Y. H.; KOO, B.-H. Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. **Psychiatry investigation**, v. 15, n. 3, p. 235–245, 2018.
- KINGMA, D. P.; BA, J. Adam: A Method for Stochastic Optimization. **arXiv [cs.LG]**, 22. dez. 2014. Disponível em: <<http://arxiv.org/abs/1412.6980>>. .
- LA ROVERE, M. T.; PINNA, G. D.; MAESTRI, R.; et al. Short-term heart rate variability strongly predicts sudden cardiac death in chronic heart failure patients. **Circulation**, v. 107, n. 4, p. 565–570, 2003.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.
- LIVNE, M.; BOLDSSEN, J. K.; MIKKELSEN, I. K.; et al. Boosted Tree Model Reforms Multimodal Magnetic Resonance Imaging Infarct Prediction in Acute Stroke. **Stroke; a journal of cerebral circulation**, v. 49, n. 4, p. 912–918, 2018.
- MALIK, M. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European Society of Cardiology and the North .... **Annals of noninvasive electrocardiology: the official journal of the International Society for Holter and Noninvasive Electrocardiology, Inc**, 1996. Wiley Online Library. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1542-474X.1996.tb00275.x>>. .
- MANDAIR, D.; TIWARI, P.; SIMON, S.; COLBORN, K. L.; ROSENBERG, M. A. Prediction of incident myocardial infarction using machine learning applied to harmonized electronic health

record data. **BMC medical informatics and decision making**, v. 20, n. 1, p. 252, 2020.

MONTANO, N.; PORTA, A.; COGLIATI, C.; et al. Heart rate variability explored in the frequency domain: a tool to investigate the link between heart and behavior. **Neuroscience and biobehavioral reviews**, v. 33, n. 2, p. 71–80, 2009.

NAWABI, J.; KNIEP, H.; ELSAYED, S.; et al. Imaging-Based Outcome Prediction of Acute Intracerebral Hemorrhage. **Translational stroke research**, v. 12, n. 6, p. 958–967, 2021.

NAYOR, M.; ENSERRO, D. M.; VASAN, R. S.; XANTHAKIS, V. Cardiovascular Health Status and Incidence of Heart Failure in the Framingham Offspring Study. **Circulation. Heart failure**, v. 9, n. 1, p. e002416, 2016.

ORR, G. B.; MÜLLER, K.-R. **Neural Networks: Tricks of the Trade**. Springer, 2003.

PICHOT, V.; ROCHE, F.; CELLE, S.; BARTHÉLÉMY, J.-C.; CHOUCOU, F. HRVanalysis: A Free Software for Analyzing Cardiac Autonomic Activity. **Frontiers in physiology**, v. 7, p. 557, 2016.

PUMPRLA, J.; HOWORKA, K.; GROVES, D.; CHESTER, M.; NOLAN, J. Functional assessment of heart rate variability: physiological basis and practical applications. **International journal of cardiology**, v. 84, n. 1, p. 1–14, 2002.

QAYYUM, S. N. A comprehensive review of applications of Artificial Intelligence in Echocardiography. **Current problems in cardiology**, p. 102250, 2023.

RAMASUBRAMANIAN, K.; SINGH, A. Deep Learning Using Keras and TensorFlow. In: K. Ramasubramanian; A. Singh (Orgs.); **Machine Learning Using R: With Time Series and Industry-Based Use Cases in R**. p.667–688, 2019. Berkeley, CA: Apress.

RASHIDI, H. H.; TRAN, N.; ALBAHRA, S.; DANG, L. T. Machine learning in health care and laboratory medicine: General overview of supervised learning and Auto-ML. **International journal of laboratory hematology**, v. 43 Suppl 1, n. S1, p. 15–22, 2021. Wiley.

RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. “why should I trust you?”: Explaining the predictions of any classifier. **arXiv [cs.LG]**, 16. fev. 2016. Disponível em: <<https://dl.acm.org/doi/abs/10.1145/2939672.2939778>> .

ROVERE, M. T. L.; BIGGER, J. T.; MARCUS, F. I.; MORTARA, A.; SCHWARTZ, P. J. Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. **The Lancet**, v. 351, n. 9101, p. 478–484, 1998.

SCHMIDT, R. A.; SEAH, J. C. Y.; CAO, K.; et al. Generative Large Language Models for Detection of Speech Recognition Errors in Radiology Reports. **Radiology. Artificial intelligence**, v. 6, n. 2, p. e230205, 2024.

SONG, T.; QU, X. F.; ZHANG, Y. T.; et al. Usefulness of the heart-rate variability complex for predicting cardiac mortality after acute myocardial infarction. **BMC cardiovascular disorders**, v. 14, p. 59, 2014.

STRÜVEN, A.; HOLZAPFEL, C.; STREMMEL, C.; BRUNNER, S. Obesity, Nutrition and Heart Rate Variability. **International journal of molecular sciences**, v. 22, n. 8, 2021. Disponível em: <<http://dx.doi.org/10.3390/ijms22084215>> .

SUGOMORI, Y.; KALUZA, B.; SOARES, F. M.; SOUZA, A. M. F. **Deep Learning: Practical Neural Networks with Java**. Packt Publishing Ltd, 2017.

TERZI, M. B.; ARIKAN, O. Machine learning based hybrid anomaly detection technique for automatic diagnosis of cardiovascular diseases using cardiac sympathetic nerve activity and electrocardiogram. **Biomedizinische Technik. Biomedical engineering**, 2023. Disponível em: <<http://dx.doi.org/10.1515/bmt-2022-0406>> .

THAYER, J. F.; YAMAMOTO, S. S.; BROSSCHOT, J. F. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. **International journal of cardiology**, v. 141, n. 2, p. 122–131, 2010.

TSUJI, H.; LARSON, M. G.; VENDITTI, F. J., Jr; et al. Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study. **Circulation**, v. 94, n. 11, p. 2850–2855, 1996.

WULSIN, L. R.; HORN, P. S.; PERRY, J. L.; MASSARO, J. M.; D’AGOSTINO, R. B. Autonomic Imbalance as a Predictor of Metabolic Risks, Cardiovascular Disease, Diabetes, and Mortality. **The Journal of clinical endocrinology and metabolism**, v. 100, n. 6, p. 2443–2448, 2015.

YAN, J.; WANG, X. Unsupervised and semi-supervised learning: the next frontier in machine learning for plant systems biology. **The Plant journal: for cell and molecular biology**, v. 111, n. 6, p. 1527–1538, 2022.

## *Capítulo II*

### *Artigo submetido*

# CLASSIFICAÇÃO DA DOENÇA ARTERIAL CORONARIANA USANDO A VARIABILIDADE DA FREQUÊNCIA CARDÍACA: UMA ABORDAGEM DE REDE NEURAL PROFUNDA COM EXPLICAÇÃO AGNÓSTICA DE MODELO INTERPRETÁVEL LOCALMENTE (LIME)

## CLASSIFYING CORONARY ARTERY DISEASE USING HEART RATE VARIABILITY AND A DEEP NEURAL NETWORK WITH LIME EXPLANATIONS

**Resumo:** A Doença Arterial Coronariana é uma das doenças cardiovasculares que mais mata no mundo hoje. O Aprendizado de Máquina tem sido amplamente aplicado para detectar precocemente indicadores de risco de doença cardiovascular em pacientes no mundo inteiro. Entretanto, modelos complexos e sem clareza para os profissionais da saúde impedem e dificultam a implementação dos modelos de Aprendizado de Máquina em escala. Visamos aplicar um Rede Neural Profunda capaz de detectar sujeitos com Doença Arterial Coronariana de sujeitos saudáveis, usando índices de Variabilidade da Frequência Cardíaca como parâmetro, com o menor número possível de variáveis para a implantação do modelo em um ambiente clínico. Métodos: Utilizamos dados coletados de eletrocardiogramas 24h (holter) do banco de dados Telemetric and Holter Electrocardiogram Warehouse (THEW) de 354 pacientes. Os dados de variabilidade da frequência cardíaca foram extraídos dos *holters* e funcionaram como entrada de uma rede neural profunda, e as variáveis mais explicativas dos modelos foram encontradas usando o algoritmo de explicação local agnóstica de modelo (LIME). O domínio do tempo e da frequência mostraram maior precisão e menor perda. O LIME identificou as cinco variáveis mais explicativas, que mais tarde foram reintroduzidas no modelo separadamente. A precisão e a perda foram mantidas, indicando que as variáveis destacadas pela LIME foram as mais importantes e mais explicativas. Os dados de variabilidade da frequência cardíaca podem ser usados para avaliar a doença arterial coronariana de indivíduos saudáveis usando redes neurais profundas, e a LIME pode simplificar o modelo, aumentando sua confiabilidade.

**Palavras-chave:** variabilidade da frequência cardíaca, rede neural profunda, aprendizado de máquina, LIME

**Abstract:** We aimed to create a DNN capable of distinguishing subjects with coronary artery disease from healthy subjects, using heart rate variability indexes as a parameter, with as few variables as possible, to increase trust in the model's output and make the data collection for model deployment in a clinical setting easier. We used data collected from 24h holter data from Telemetric and Holter Electrocardiogram Warehouse (THEW) database of 354 patients. Heart rate variability data was extracted from the Holter and used as input for a deep neural network, and the most explanatory model variables were found using LIME. Time and frequency domain showed higher accuracy and lower loss. LIME was used to identify the five most explanatory variables, which were then reintroduced to the model separately. The accuracy and loss were maintained, indicating that the variables highlighted by LIME were the most important and explanatory. Heart rate variability data can be used to differentiate between coronary artery disease and healthy subjects using deep neural networks, and LIME can simplify the model, increasing its trustworthiness.

**Keywords:** heart rate variability; deep neural network; machine learning; LIME

### Introduction

Cardiovascular disease (CVD) has been the leading cause of death worldwide since 1990, with approximately 240 deaths per 100.000 cases worldwide GBD 2019 Viewpoint Collaborators (2020), most of the cases being myocardial infarction (MI) and stroke GBD 2019 Risk Factors Collaborators (2020). Heart rate variability (HRV) has been recognized as a robust indicator of cardiovascular health and a non-invasive marker for cardiovascular disease Malik (1996 e Thayer et al. (2010 e Wulsin et al. (2015). HRV can be easily extracted from routine ECG monitoring, making the collection and analysis of the data an easy task for physicians and researchers.

Successful attempts in a wide variety of fields have increased interest in the study of machine learning (ML) in public health, such that recent efforts have used health-related data as input for ML models, in pursuit of early identification of cardiovascular disease (CVD) risk factors, such as stroke La Rovere et al. (2003 e Livne et al. (2018 e Nawabi et al. (2021) and MI Mandair et al. (2020 e Song et al. (2014). One of the most used ML models is known as Deep Learning, which is composed of a Deep Neural Network (DNN), capable of finding patterns in unstructured data, improving upon traditional ML models LeCun et al. (2015).

However, recent studies have pointed out that ML "out-of-the-box" diagnoses are not trusted by clinicians, and even more, may not lead to overall increases in the quality of decisions taken by practitioners Jacobs et al. (2021 e Ribeiro et al. (2016).

Based on this trust issue, a group of scientists at the University of Washington developed an explanatory model, called Locally Interpretable Model-Agnostic Explanation (LIME) Jacobs et al. (2021). The LIME model is capable of explaining any model by giving the most explanatory variables for individual predictions, choosing if the prediction is accurate, and giving the opportunity for model optimization.

We aimed to create a DNN capable of detecting subjects with CAD from healthy subjects, using heart rate variability indexes as a parameter, with as few variables as possible, increasing trust in the model's output and making data collection easier for model deployment in a clinical setting. The importance of detecting the most explanatory variables is that we can simplify the analysis and interpretation of the DNN model, aiming for future deployment of a system that detects CAD from normal subjects at the clinical stage.

## **Materials and methods**

### ***Database***

The sample consisted of 354 patients divided into three groups (Healthy, n=128; CAD, n=226). The sample dataset is part of the Telemetric and Holter Electrocardiogram Warehouse (THEW) database, from the University of Rochester Medical Center. After an official contact with the THEW coordinators, we downloaded and analyzed the data using the HRVanalysis software Pichot et al. (2016). The database E-HOL-03-0271-002 comprised 226 24-hour Holter recordings from subjects with CAD. For the Healthy group, we used the E-HOL-03-0202-003 database of 128 24-hour Holter recordings of healthy subjects.

### ***Preprocessing***

#### ***Normalization***

Data normalization is an important step in ML preprocessing because it reduces computing time and improves the performance of the algorithm Jayalakshmi; Santhakumaran (2011). The normalization was made according to the minimum/maximum procedure, according to the following formula:



$$(1) \quad x = \frac{x - x_{min}}{x_{max} - x_{min}}$$

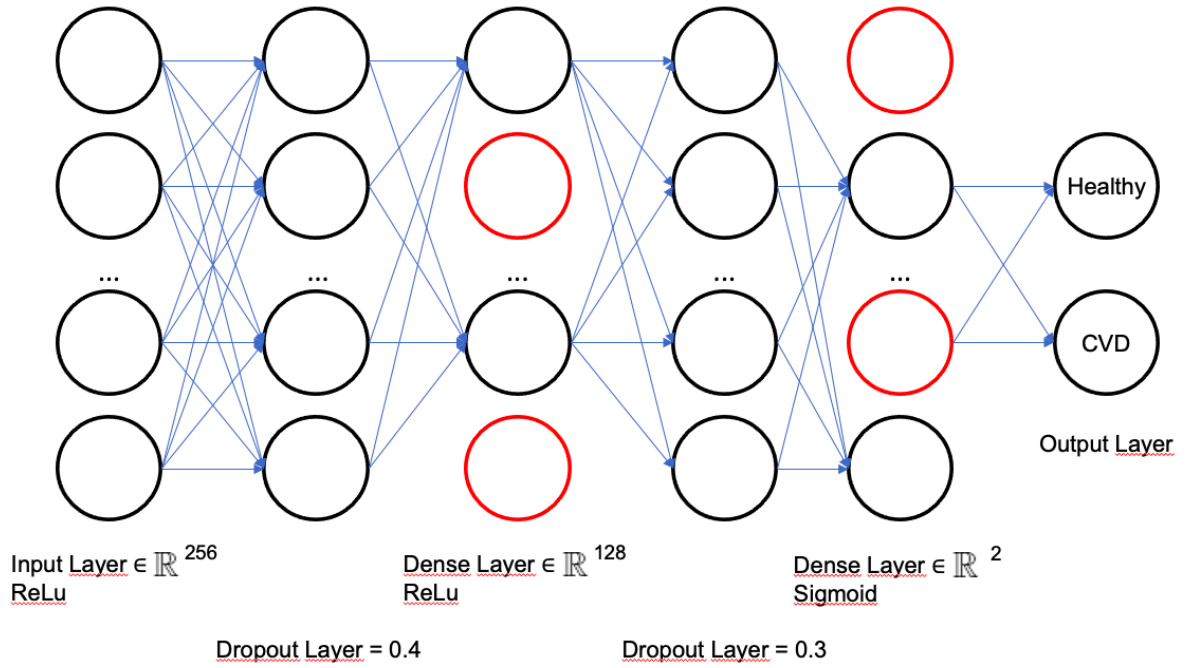
### *Heart rate variability - feature extraction*

We used the following HRV indexes: Total variance ( $\text{ms}^2$ ), mean of all RR intervals (ms), SDNN (ms, standard deviation of all normal to normal intervals), RMSSD (ms, root mean square of the successive differences of the normal to normal intervals), Time domain (Mean RR [ms], Mean HR [bpm], NN20 [ms], pcNN20 [%], NN30 [ms], pcNN30 [%], NN50 [ms], pcNN50 [%], SDNN [ms], RMSSD [ms], SDANN [ms], SDNNIDX [ms]), Frequency domain (Ptot [ $\text{ms}^2$ ], ULF [ $\text{ms}^2$ ], VLF [ $\text{ms}^2$ ], LF [ $\text{ms}^2$ ], HF [ $\text{ms}^2$ ]), Poincaré plot (Centroid, SD1 [ms], SD2 [ms], SD1/SD2 [ms], SD1nu [%], SD2nu [%]), Geometrical indexes (Tri ind, TINN, X, Y, N, M), Symbolic analysis (0V [absolute], 0V% [%], 1V [absolute], 1V% [%], 2LV [absolute], 2LV% [%], 2UV [absolute], 2UV% [%], MP [absolute], MP% [%]). LF ( $\text{ms}^2$  and, normalized units, low-frequency spectral density), HF ( $\text{ms}^2$  and normalized units, high-frequency spectral density), LF/HF (ratio of low and high frequency domains). We used all indexes in 24 hours, day and night timeframes.

### *Deep Learning model - pattern classification*

We used a deep neural network model to analyze the data. The model was composed of one input layer followed by five other layers, as represented in Figure 1. Our model was implemented using the Keras package and the TensorFlow backend Abadi et al. (2016 e Ramasubramanian; Singh (2019) . The model was trained using the R statistical language on a conventional computer with a GPU. The numerical values of HRV were used as input data and then split into training and test sets (75% for training and 25% for validation). We used a maximum epoch selection according to early stopping, which is a standard practice for neural networks as it ensures the maximum accuracy and minimum loss using the lowest number of iterations (epochs) possible Orr; Müller (2003). We also used different unit counts in the dense layers (256 in the first layer, 128 in the second layer, and two in the output layer), employing the Rectified Linear Units (ReLU) activation function with Adam optimization Agarap (2018 e Kingma; Ba (2014) (The Sigmoid activation function was used in the output layer), and a Binary Cross-entropy loss function. The Locally Interpretable Model-Agnostic Explanation (LIME) was applied using the lime package for R (<https://CRAN.R-project.org/package=lime>). The most explanatory variables were grouped and listed in Table 1, and then re-analyzed separately.

**Figure 1.** Graphical representation of the DNN used for CAD detection



ReLU, rectified linear units; CVD, cardiovascular disease

*Model metrics - results evaluation*

To evaluate the model’s ability to predict desired outcomes, we used common model metrics such as accuracy,

$$accuracy = \frac{True\ positive}{(True\ positive + True\ negative)}$$

loss function,

$$H_p(q) = -\frac{1}{N} \sum_{i=0}^n y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

where p is probability, y is the data point label, and N is the sample

sensitivity,

$$Sensitivity = \frac{True\ positive}{True\ positive + False\ negative}$$

and specificity

$$Specificity = \frac{True\ negative}{False\ positive + True\ negative}$$

## Results and Discussion

Table 1 shows the results related to individuals with CAD. The results show that the time domain variables with 25 epochs showed accuracy of 0.89, loss of 0.37, sensitivity of 0.94 and specificity of 0.78. The frequency domain variables with 30 epochs showed accuracy of 0.84, loss of 0.53, sensitivity of 0.95 and specificity of 0.62. Using Poincaré plot variables with 30 epochs, the accuracy was 0.89, loss 0.57, sensitivity 0.90 and specificity 0.85. The geometrical index variables with 30 epochs showed accuracy of 0.75, loss of 0.79, sensitivity of 0.92 and specificity of 0.50. With the symbolic analysis, using 30 epochs, the results obtained were the accuracy of 0.75, loss of 0.81, sensitivity of 0.96 and specificity of 0.05.

**Table 1.** Most explanatory features according to LIME in CAD

Domains	Features
<b>CAD and Healthy</b>	
Time Domain	Mean HR (bpm/24h), Mean HR (bpm/Day), SDNN (ms/24h), SDANN (ms/24h), SDANN (ms/Day)
Frequency Domain	VLF (ms <sup>2</sup> /24h), VLF (ms <sup>2</sup> /Day), LF (ms <sup>2</sup> /24h), LF (ms <sup>2</sup> /Day), LF/HF (Day)
Poincaré Plot	SD2 (Night), SD1 (nu/24h), SD2 (nu/24h), SD1/SD2 (24h), Centroid (Day)
Geometrical Indexes	X (24h), Y (24h), N (Day), M (Day), X (Night)
Symbolic Analysis	2UV (24h), 2UV (Day), 0V (Day), MP (Night)

**Table 2.** Performance of different attributes according to the DNN for Coronary Artery Disease classification

All features	Accuracy (%)	Loss (%)	Sensitivity (%)	Specificity (%)	Epochs
Time domain	<b>0.89</b>	<b>0.37</b>	0.94	0.78	<b>25</b>
Frequency domain	0.84	0.53	0.95	0.62	30
Poincaré plot	<b>0.89</b>	0.57	0.90	<b>0.85</b>	30
Geometrical indexes	0.75	0.79	0.92	0.50	30
Symbolical analysis	0.75	0.81	<b>0.96</b>	0.05	30
<b>Top 5 LIME</b>					
Time domain	<b>0.85</b>	<b>0.32</b>	0.91	<b>0.74</b>	<b>25</b>
Frequency domain	<b>0.85</b>	0.47	<b>0.97</b>	0.58	30
Poincaré plot	0.81	0.62	0.91	0.61	30
Geometrical indexes	0.83	0.45	0.95	0.54	30
Symbolical analysis	0.73	0.94	0.93	0.08	30



Using LIME, five variables were chosen with the greatest predictive power. Using five time domain variables with 25 epochs, the accuracy was 0.85, the loss was 0.32, the sensitivity was 0.91 and the specificity was 0.74. With five variables in the frequency domain and using 30 epochs, the accuracy was 0.85, the loss was 0.47, the sensitivity was 0.97 and the specificity was 0.58. With five Poincaré plot variables and using 30 epochs, the accuracy was 0.81, the loss was 0.62, the sensitivity was 0.91 and the specificity was 0.61. With the five variables of the geometrical indexes and using 30 epochs, the accuracy was 0.83, the loss was 0.45, the sensitivity was 0.95 and the specificity was 0.54. With five variables from the symbolic analysis and using 30 epochs, the accuracy was 0.73, the loss was 0.94, the sensitivity was 0.93 and the specificity was 0.08.

The results show that the time domain variables with 20 epochs showed accuracy of 0.92, loss of 0.32, sensitivity of 0.88 and specificity of 0.96. With the frequency domain variables using 15 epochs, the accuracy was 0.89, loss was 0.40, sensitivity was 0.89 and specificity was 0.88. Using Poincaré plot variables with 40 epochs the accuracy was 0.96, loss 0.28, sensitivity 0.94 and specificity 1.00. Using geometrical indexes with 100 epochs, the accuracy was 0.73, loss was 2.30, sensitivity was 0.92 and specificity was 0.50. In the symbolic analysis with 100 epochs, the accuracy obtained was 0.52, loss of 14.54, sensitivity of 0.92 and specificity of 0.05.

Using LIME, five variables were chosen with the greatest predictive power. Using five time-domain variables with 20 epochs, the accuracy was 0.95, loss was 0.04, sensitivity was 0.96 and specificity was 1.00. Using five frequency-domain variables with 15 epochs, the accuracy was 0.83, loss was 0.37, sensitivity was 0.86 and specificity was 0.81. With five Poincaré plot variables and using 40 epochs, the accuracy was 0.94, loss was 0.30, sensitivity was 0.93 and specificity was 0.95. With five variables of the geometric indexes and using 100 epochs, the accuracy was 0.77, loss was 0.58, sensitivity was 0.89 and specificity was 0.52. Using five variables of the symbolic analysis and with 100 epochs, the accuracy was 0.63, loss was 9.75, sensitivity was 0.94 and specificity was 0.11.

The aim of this study was to propose a deep neural network (DNN) model for CAD disease prediction through HRV and to understand the most explanatory HRV variables in the model according to LIME, aiming to simplify the understanding of the model and to increase trustiness. To the best of our knowledge, this is the first study to use DNNs to predict cardiovascular outcomes through HRV and to explain the model with LIME.

We used the LIME approach in our DNN model aiming to understand the complexity of the model and to extract the most explanatory variables for future model replication and clinical deployment. The LIME explanatory model was developed by a group of computer scientists at the University of Washington with the objective of ranking the most important variables on a complex model, such as a DNN Ribeiro et al. (2016). The LIME allows us to interpret and explain the DNN model such as the variables used for the prediction. The explanation makes it possible to understand the relationship between a variable and the model applied.

Recent studies have shown that the use of ML to support clinicians in medical decisions may not increase the quality of the treatment, which can be explained by the lack of reasoning in the ML models Jacobs et al. (2021) e Ribeiro et al. (2016). Thus, to explain what variable is most important for the prediction makes the process more direct and increases the human's trust in the DNN model and in machine learning in general Jacobs et al. (2021).

The data in Table 2 shows that the Time domain and the Poincaré plot are the best predictors of CAD for the overall DNN model. After the LIME extraction, the Time and Frequency domains were the best predictors for CAD. However, the data in Table 3 shows

that the Poincaré plot variables are the best predictors of MI for the overall DNN model. After the LIME extraction, the Time domain variables were the best predictors for CAD.

The clinical importance of this study lies in the ease of using HRV data, which is extracted from the electrocardiogram, a routine exam in clinical practice, making it possible to use ML algorithms, such as DNNs, to support human decisions based on data and Artificial Intelligence.

Most of the studies involving HRV and ML use short-term HRV instead of longer periods, as recommended by MalikMalik (1996). Therefore, we used 24-hour holter ECG data, which can improve data interpretation and lead to better conclusions.

Comparing before and after LIME, CAD datasets showed a similar accuracy, but with fewer variables, showing that LIME highlighted the key variables for the model decision. This finding is important because simplifies the model while retaining a high accuracy (for example, from 16 variables to five in the Time Domain), making it easier for clinicians and practitioners to understand the reasoning behind the model and giving physiological plausibility to the decision making, increasing the trustiness in the DNN model.

We also found that after the LIME explanation, CAD datasets were explained by both the Time and Frequency domains, which are by far the most used and understood domains for autonomic modulation. This increases the level of interpretation and familiarity with the model and with autonomic modulation in general, thus facilitating understanding and clinical decisions.

## **Final points**

In conclusion, this study demonstrates the potential clinical importance of using HRV data and ML algorithms, such as DNNs, to support human decisions based on data and Artificial Intelligence. By utilizing 24-hour holter ECG data, this study improves data interpretation and leads to better conclusions compared to previous studies that used shorter-term HRV.

The use of LIME, an explanation technique, showed that CAD datasets maintained a similar accuracy while highlighting key variables for the model decision. This simplification of the model, from a complex DNN to a more interpretable model, allows for better understanding of the underlying mechanisms and variables that contribute to the decision-making process.

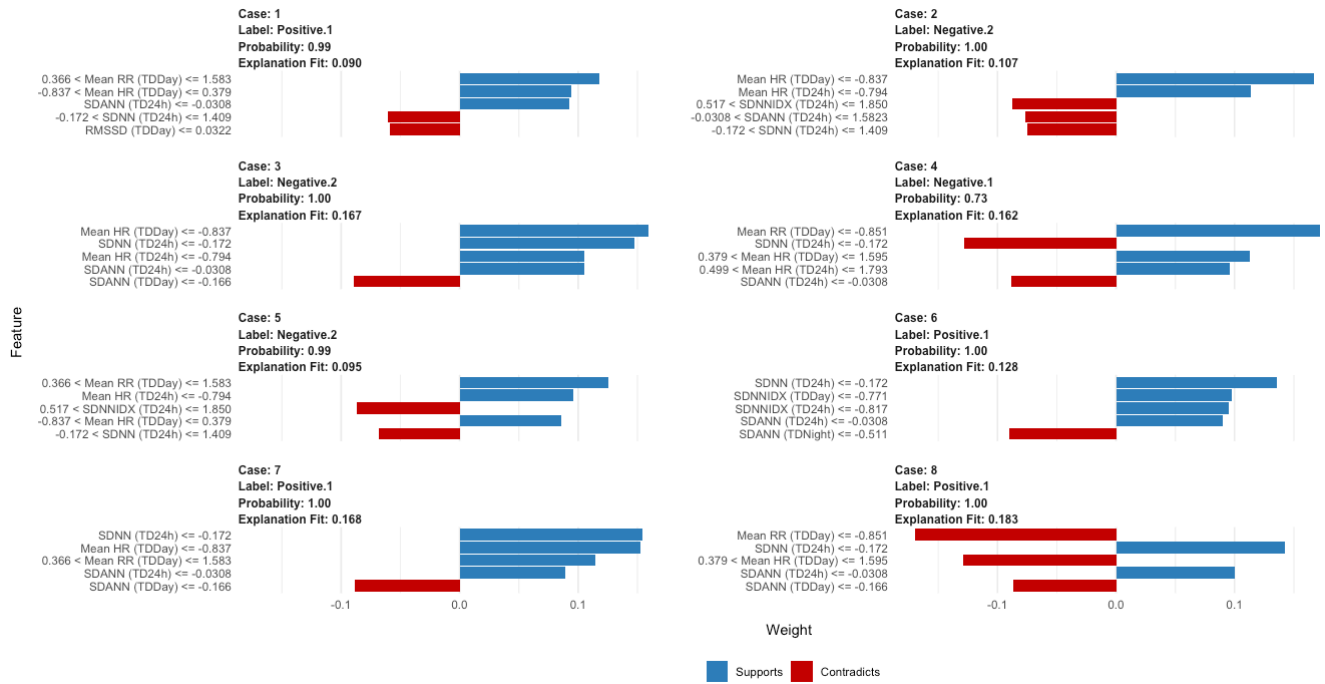
## References

*Anexo I*

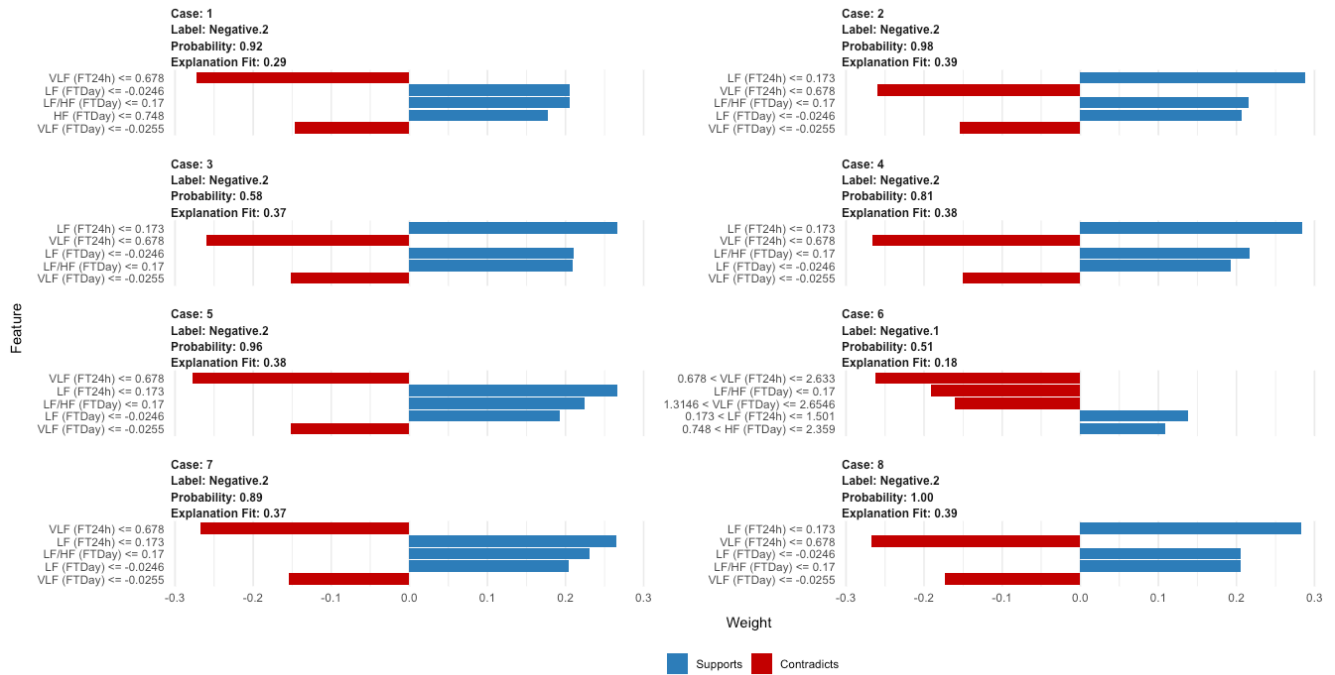
*Material Suplementar*



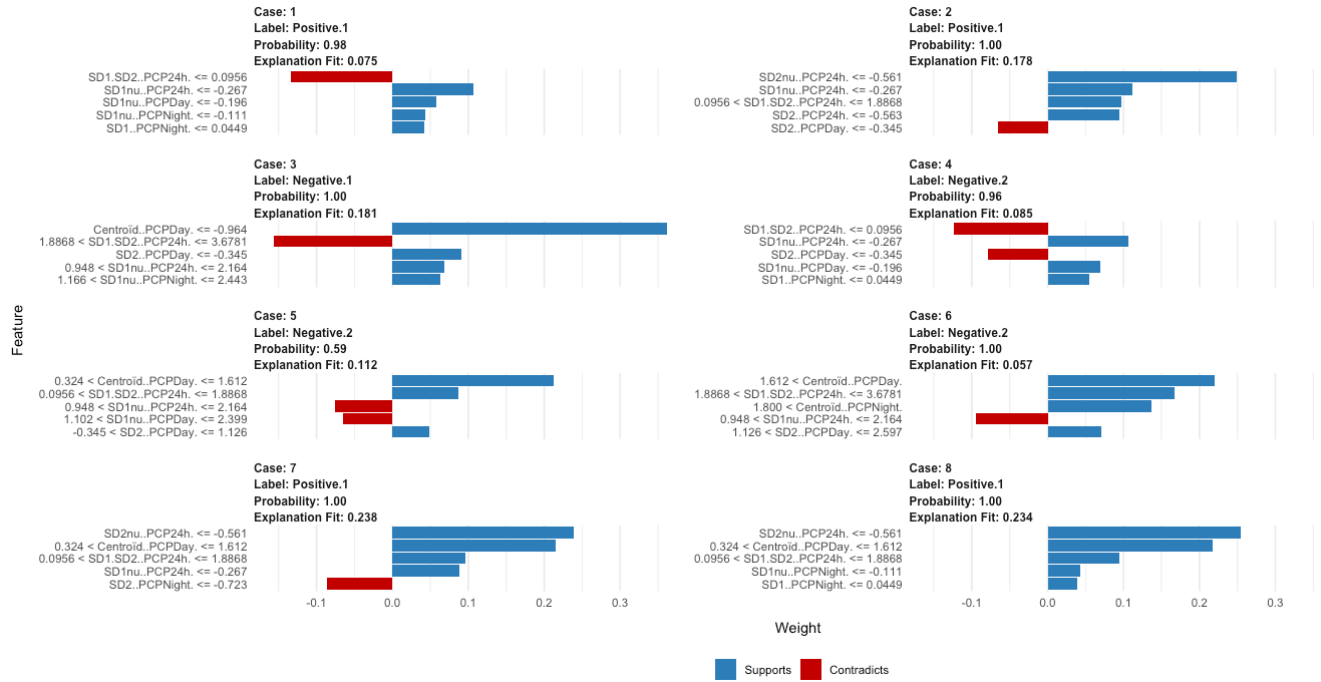
**Figure 2.** LIME's five most explanatory variables form the Time Domain of the CAD dataset



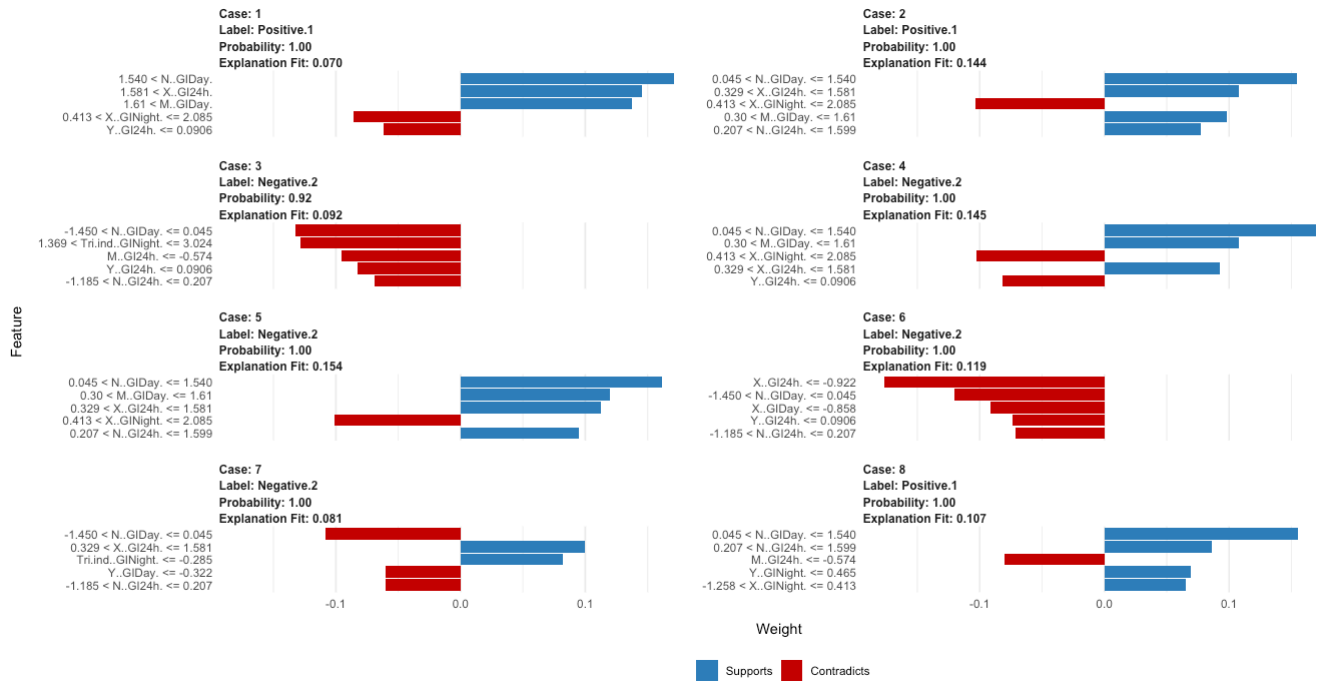
**Figure 3.** LIME's five most explanatory variables form the Frequency Domain of the CAD dataset



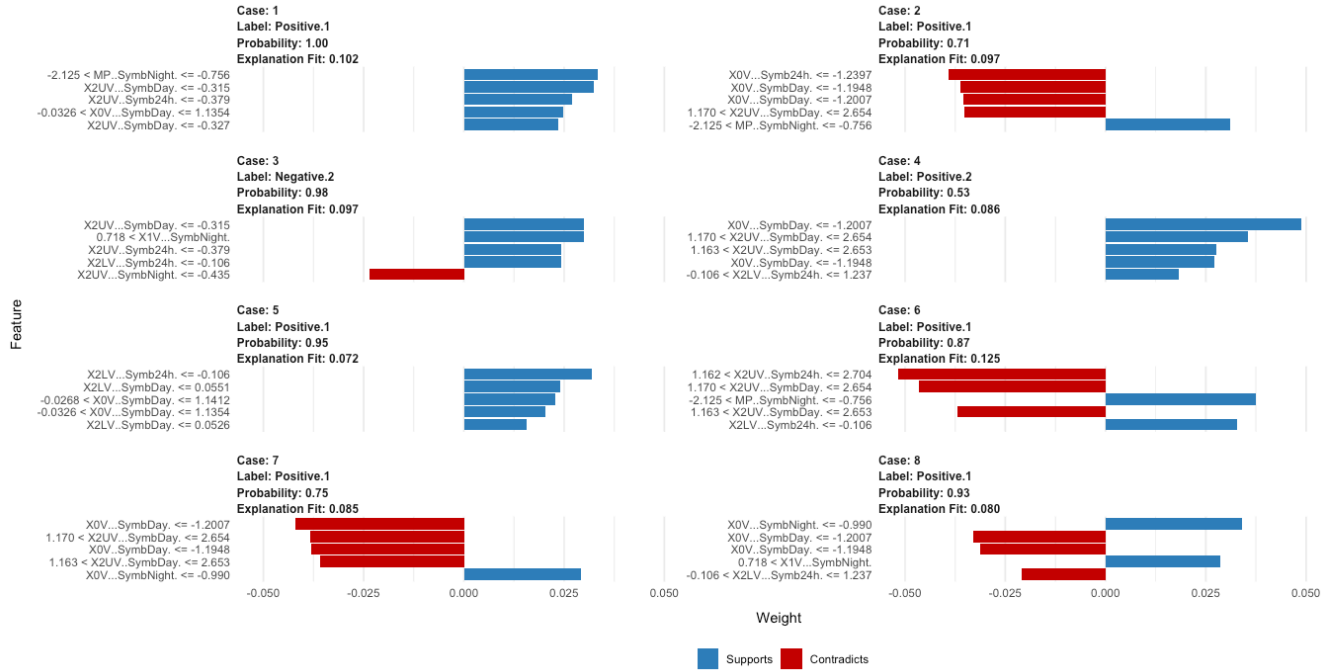
**Figure 4.** LIME's five most explanatory variables form the Poincaré plot of the CAD dataset



**Figure 5.** LIME's five most explanatory variables form the Geometric indexes of the CAD dataset



**Figure 6.** LIME's five most explanatory variables form the Symbolic analysis of the CAD dataset



*Anexo II - Certificado de  
Registro de Software*



REPÚBLICA FEDERATIVA DO BRASIL  
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS  
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL  
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

## Certificado de Registro de Programa de Computador

Processo Nº: **BR512023003689-7**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 01/06/2020, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

**Título:** Rede Neural Profunda com Explicação Agnóstica de Modelo Interpretável Localmente para Classificação da Doença Arterial Coronariana usando a Variabilidade da Frequência Cardíaca

**Data de criação:** 01/06/2020

**Titular(es):** UNIVERSIDADE FEDERAL DO MARANHÃO

**Autor(es):** ANTONIO CARLOS PEREIRA SILVA FILHO; SARA RAQUEL DUTRA MACEDO; MARCELO DE SOUZA ANDRADE; MARIA DO SOCORRO SOUSA CARTAGENES

**Linguagem:** R

**Campo de aplicação:** MT-06

**Tipo de programa:** IA-01

**Algoritmo hash:** SHA-512

**Resumo digital hash:**

a117e489c7cfbeb1bd1aaa14cef5da8d9013621cf8ced4cc31688ccb21828ea452cbac580f40e02e4f128922cf38a5526626101174db0c5101f6069107bc5579

**Expedido em:** 05/12/2023

**Aprovado por:**

Carlos Alexandre Fernandes Silva  
Chefe da DIPTO

*Anexo III - Comprovante  
de Submissão*

*Revista Vivências - Qualis A4*





Antonio Carlos Filho <antoniocarlos.j77@gmail.com>

---

**[Vivências] Agradecimento pela submissão**

1 mensagem

**Revista Vivências** <vivencias@reitoria.uri.br>  
Para: Antonio Carlos Filho <antoniocarlos.j77@gmail.com>

5 de dezembro de 2023 às 14:15

Antonio Carlos Filho,

Agradecemos a submissão do trabalho "CLASSIFYING CORONARY ARTERY DISEASE USING HEART RATE VARIABILITY AND A DEEP NEURAL NETWORK WITH LIME EXPLANATIONS" para a revista Vivências.  
Acompanhe o progresso da sua submissão por meio da interface de administração do sistema, disponível em:

URL da submissão: <http://revistas.uri.br/index.php/vivencias/authorDashboard/submission/1249>  
Login: acfilho

Em caso de dúvidas, entre em contato via e-mail.

Agradecemos mais uma vez considerar nossa revista como meio de compartilhar seu trabalho.

Revista Vivências

Revista vivencias <http://revistas.uri.br/index.php/vivencias>

*Anexo IV - Normas da  
Submissão*

*Revista Vivências - Qualis A4*