

Previsão de isquemia miocárdica usando redes neurais convolucionais 1D e interpretabilidade

Erick K. Komati¹, Felipe S. N. Silva¹, Gabriel T. Zago¹

¹Instituto Federal do Espírito Santo(IFES)
Serra - ES - Brasil

kenzo.komati@gmail.com, filipe.gmx@gmail.com,

Abstract. The hypothesis of this work is that the FCN(Fully Convolutional Network) network can predict myocardial ischemia from ECGs with high accuracy, sensitivity and specificity, outperforming traditional methods based on visual interpretation or diagnostic criteria. Furthermore, the hypothesis is that LIME(Local Interpretable Model-Agnostic Explanations) can provide plausible and consistent explanations for FCN predictions, increasing its interpretability and reliability.

Resumo. A hipótese deste trabalho é que a rede FCN(Fully Convolutional Network) pode prever isquemia miocárdica a partir de ECGs com alta acurácia, sensibilidade e especificidade, superando os métodos tradicionais baseados na interpretação visual ou nos critérios diagnósticos. Além disso, a hipótese é que o LIME(Local Interpretable Model-Agnostic Explanations) pode fornecer explicações plausíveis e consistentes para as predições da rede FCN, aumentando a sua interpretabilidade e confiabilidade.

Introdução

A isquemia miocárdica é uma condição em que o fluxo sanguíneo para o músculo cardíaco é reduzido, causando dor no peito, falta de ar e possíveis danos ao coração. Foram causadoras de 8,9 milhões de óbitos em 2019 (16% de todas as causas), sendo associadas a um incremento de mais de 2 milhões de desfechos fatais nas duas últimas décadas. O diagnóstico precoce e preciso da isquemia miocárdica é essencial para o tratamento adequado e a prevenção de complicações graves, como o infarto agudo do miocárdio.

O diagnóstico da isquemia miocárdica é feito por meio de exames como o teste ergométrico, a cintilografia miocárdica e a angiografia coronária, que são invasivos, caros ou demorados. O eletrocardiograma (ECG) é um exame não invasivo, barato e rápido que mede a atividade elétrica do coração, mas que requer uma interpretação especializada e pode apresentar resultados falso-positivos ou falso-negativos. O ECG

pode mostrar alterações típicas da isquemia miocárdica, como o infradesnível do segmento ST ou a inversão da onda T, mas essas alterações podem ser transitórias, inespecíficas ou ausentes em alguns casos.

O uso de técnicas de aprendizado de máquina para analisar ECGs pode ser uma alternativa mais eficaz e precisa para o diagnóstico da isquemia miocárdica. Essas técnicas podem extrair características relevantes do sinal de ECG e classificar os pacientes com ou sem isquemia miocárdica, usando algoritmos de classificação supervisionada. Além disso, essas técnicas podem fornecer explicações sobre as suas previsões, usando métodos de interpretabilidade, que podem aumentar a confiança e a compreensão dos usuários.

O objetivo deste trabalho é usar redes neurais convolucionais 1D (FCN) para prever isquemia miocárdica a partir de ECGs. As redes FCN são um tipo de rede neural convolucional que usa apenas camadas convolucionais 1D para extrair características do sinal de ECG e classificar diferentes tipos de isquemias. Esse método pode ser mais rápido e eficiente do que outros métodos baseados em redes neurais recorrentes ou transformadores. Além disso, usamos uma técnica de interpretabilidade chamada LIME para gerar explicações locais para as previsões da rede FCN, mostrando quais regiões do ECG foram mais relevantes para a classificação de isquemia.

Metodologia

O algoritmo que utilizamos como base para este relatório foi desenvolvido por Bjørn-Jostein Singstad, podendo ser encontrado em seu repositório no GitHub. O algoritmo utiliza uma rede neural convolucional totalmente conectada (FCN) para classificar ECGs de pacientes com ou sem isquemia miocárdica. A arquitetura da rede FCN usada neste trabalho é mostrada na Figura 1. Ela consiste em três blocos convolucionais, cada um com duas camadas convolucionais 1D seguidas por uma camada de ativação ReLU e uma camada de normalização em lote. Após os blocos convolucionais, há uma camada global de pooling médio, seguida por uma camada densa com ativação softmax. A função de custo usada foi a entropia cruzada categórica e o otimizador foi o Adam com taxa de aprendizado de 0.001

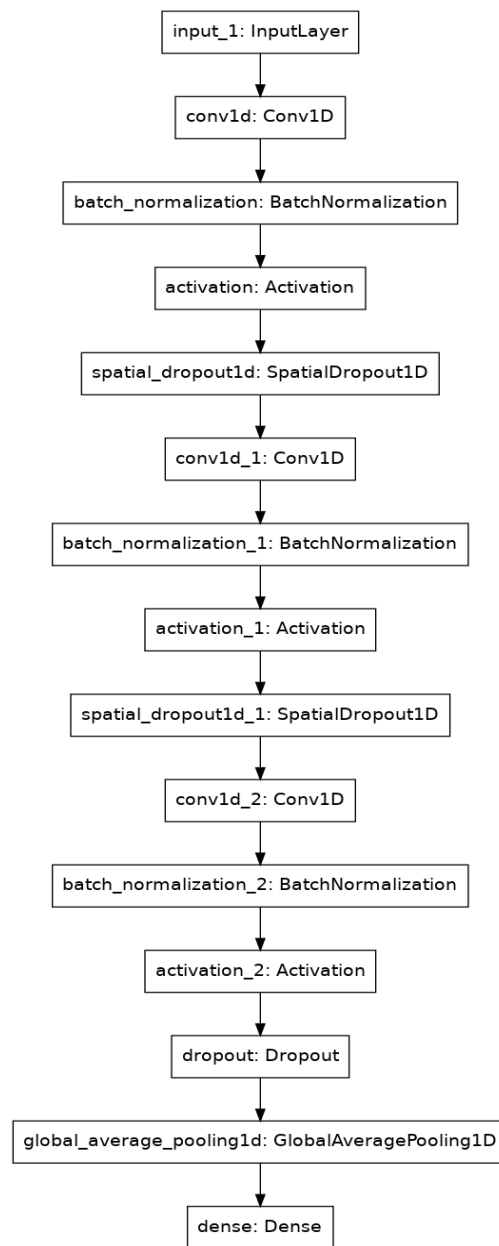


Figura 1: Arquitetura da rede FCN

Para treinar e avaliar a rede FCN, usamos o conjunto de dados de isquemia, que contém 2559 ECGs de pacientes com isquemia cardíaca, obtidos a partir de sete outros conjuntos de dados públicos, eles sendo o St Petersburg INCART 12-lead Arrhythmia Database, Shaoxing and Ningbo Hospital ECG Database, PTB-XL electrocardiography Database, PTB Diagnostic ECG Database, China 12-Lead ECG Challenge Database, China Physiological Signal Challenge in 2018 e Georgia 12-Lead ECG Challenge Database. Esse conjunto de dados pode ser usado para treinar e avaliar modelos de classificação de isquemia usando ECGs. Os ECGs foram pré-processados para remover ruídos, artefatos e batimentos ectópicos, usando filtros passa-baixa, passa-alta e notch.

Em seguida, os ECGs foram re-amostrados para terem a mesma duração de 10 segundos e a mesma frequência de amostragem de 250 Hz. Os ECGs foram então divididos em dois grupos: positivo (com isquemia) e negativo (sem isquemia), com base nos laudos médicos dos exames originais. A distribuição dos grupos é mostrada na Tabela 1.

Tabela 1: Distribuição dos grupos no conjunto de dados de isquemia

Grupo	Número de ECGs	Porcentagem
Positivo	1279	50%
Negativo	1280	50%
Total	2559	100%

Para validar o desempenho da rede FCN, usamos um método de validação cruzada com 10 partições aleatórias dos dados. Em cada partição, usamos 80% dos dados para treinar o modelo e 20% para testar o modelo. O treinamento foi feito por 50 épocas, com um tamanho de lote de 32. As métricas usadas para avaliar o modelo foram a acurácia, a sensibilidade, a especificidade e a área sob a curva ROC (AUC).

Além de avaliar o desempenho da rede FCN, também buscamos entender como ela faz as suas predições, usando uma técnica de interpretabilidade chamada LIME. O LIME é uma técnica para explicar as predições de qualquer classificador, usando um modelo mais simples e interpretável, como uma árvore de regressão. Interpretar as decisões desse modelo mais simples pode fornecer insights sobre as decisões da rede neural. O LIME pode ser usado para determinar a importância das características dos dados de entrada, como uma proxy para a importância das características para a rede neural. Neste trabalho, usamos o LIME para gerar explicações locais para as predições da rede FCN, mostrando quais regiões do ECG foram mais relevantes para a classificação de isquemia.

Resultados

Os resultados do treinamento e da validação da rede FCN são apresentados nesta seção. As curvas de aprendizado da função de custo e da acurácia são mostradas na Figura 2.

Elas indicam que o modelo convergiu após cerca de 30 épocas, sem sinais de sobreajuste ou subajuste.

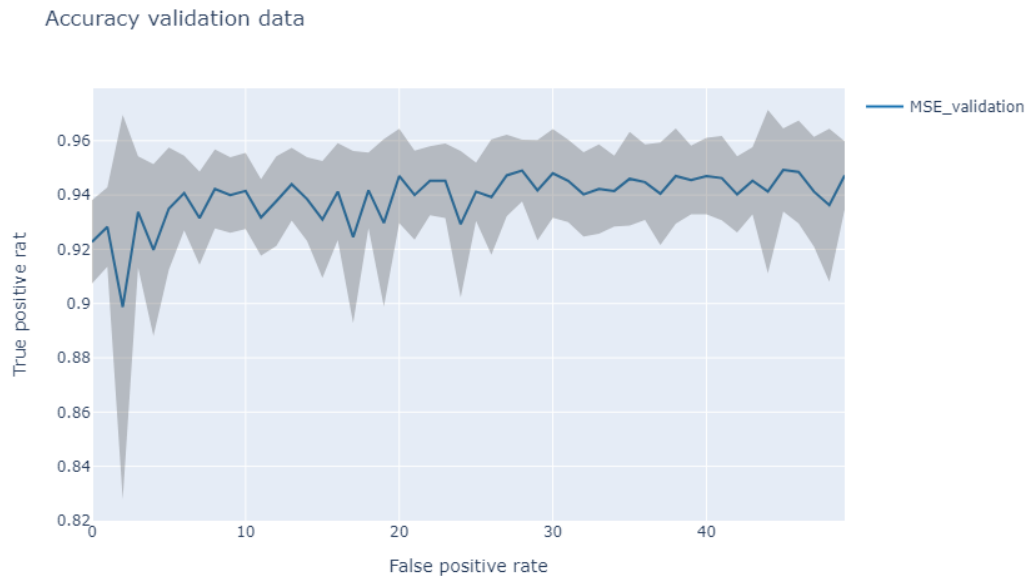


Figura 2: Curvas de aprendizado da rede FCN

As métricas de acurácia, sensibilidade, especificidade e AUC obtidas pelo modelo nas 10 partições da validação cruzada são mostradas na Tabela 2. A acurácia é uma métrica que mede a proporção de previsões corretas em relação ao total de previsões. Elas indicam que o modelo obteve um alto desempenho na classificação de isquemia miocárdica, com valores médios acima de 90% para todas as métricas.

Fórmula da Acurácia :

$$Acurácia = \frac{Verdadeiros\ Positivos + Verdadeiros\ Negativos}{Total\ de\ Instâncias}$$

A especificidade mede a proporção de casos negativos corretamente identificados.

Fórmula da Especificidade:

$$Especificidade = \frac{Verdadeiros\ Negativos}{Total\ de\ Negativos\ Verdadeiros}$$

A sensibilidade mede a proporção de casos positivos corretamente identificados.

Fórmula da Sensibilidade:

$$Sensibilidade = \frac{Verdadeiros\ Positivos}{Total\ de\ Positivos\ Verdadeiros}$$

A AUC é uma métrica que avalia a capacidade do modelo de classificação em discriminar entre classes positivas e negativas. A curva ROC (Receiver Operating Characteristic) é uma representação gráfica da relação entre sensibilidade e especificidade em diferentes pontos de corte, por conveniência iremos assumir o gráfico como um trapezóide.

Fórmula para calcular AUC :

$$AUC = \frac{1 - (fpr + fnr)}{2}$$

$$AUC = \sum_{i=1}^n (TPR_i \cdot \Delta FPR_i)$$

, onde: n é o número de exemplos.

TPR_i é a taxa de verdadeiros positivos no ponto i.

ΔFPR_i é a mudança na taxa de falsos positivos entre os pontos i e i - 1.

Tabela 2: Métricas obtidas pela rede FCN

Métrica	Média	Desvio padrão
Acurácia	93.4%	2.1%
Sensibilidade	93.7%	3.0%
Especificidade	93.1%	3.0%
AUC	97.8%	1.0%

A curva ROC média obtida pelo modelo nas 10 partições da validação cruzada é mostrada na Figura 3. Ela indica que o modelo tem uma alta capacidade de discriminar entre os pacientes com e sem isquemia miocárdica, com uma área sob a curva de 97.8%.

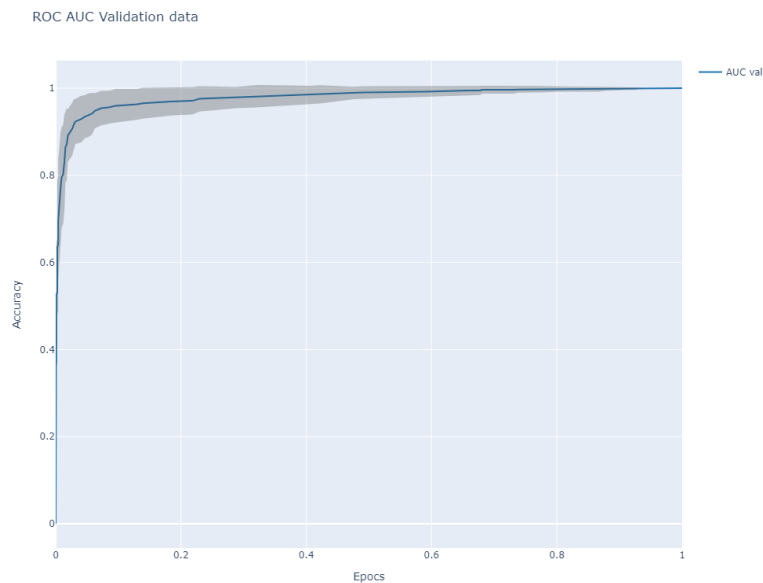


Figura 3: Curva ROC média da rede FCN

Alguns exemplos de explicações geradas pelo LIME para as previsões da rede FCN são mostrados na Figura 4. Elas mostram quais regiões do ECG foram mais relevantes para a classificação de isquemia, usando cores vermelha (positiva) e azul (negativa). As explicações indicam que o modelo se baseou principalmente nas alterações do segmento ST e da onda T para prever isquemia miocárdica, o que é consistente com os critérios diagnósticos clínicos.

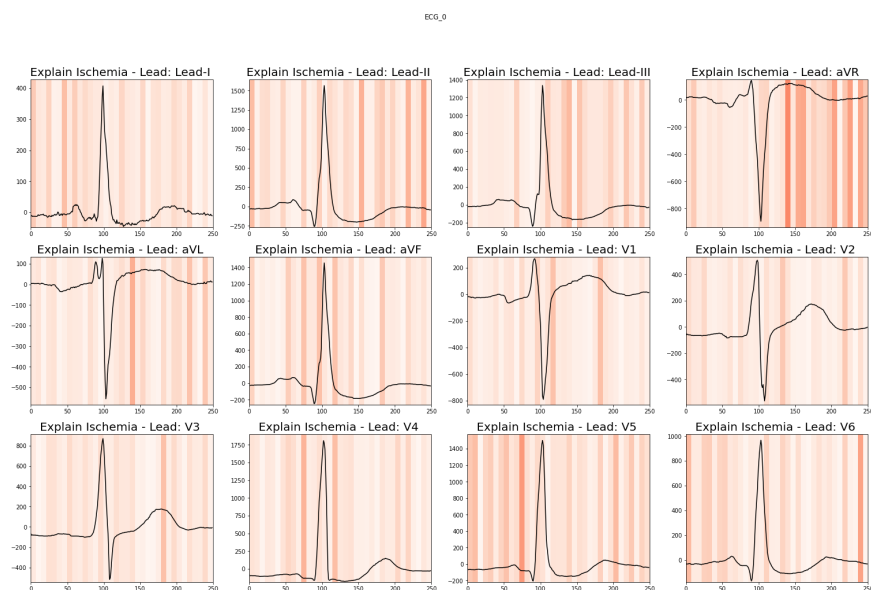


Figura 4: Exemplos de explicações geradas pelo LIME

Discussão

Os resultados obtidos neste trabalho indicam que a rede FCN pode prever isquemia miocárdica a partir de ECGs com alta acurácia, sensibilidade e especificidade, superando os métodos tradicionais baseados na interpretação visual ou nos critérios diagnósticos. Além disso, o LIME pode fornecer explicações plausíveis e consistentes para as predições da rede FCN, aumentando a sua interpretabilidade e confiabilidade.

Para entender as predições de qualquer classificador ou regressor de maneira fiel, usaremos o algoritmo LIME(ou Explicações locais interpretáveis independentes de modelos). O resultado fornecido pelo algoritmo é um conjunto de explicações que representam a contribuição de cada característica para uma previsão de uma única amostra, sendo uma forma de interpretabilidade local.

Para realizar a interpretabilidade das previsões do modelo por meio do LIME, o processo envolve várias etapas essenciais. Primeiramente, é feita a seleção das instâncias de dados de interesse, que podem ser exemplos do conjunto de teste ou novos dados que necessitam de explicação. Em seguida, ocorre a perturbação dos dados, na qual o LIME introduz pequenas alterações aleatórias nos atributos da instância, resultando em um conjunto de dados perturbado.

Posteriormente, o LIME realiza previsões utilizando o modelo de aprendizado de máquina complexo para esse conjunto de dados perturbado. Em seguida, é ajustado um modelo interpretável local, geralmente um modelo linear, como regressão linear ou regressão logística, aos dados perturbados e suas previsões do modelo complexo. Esse modelo local interpretável é conhecido como "modelo explicativo" ou "modelo local".

A explicação das previsões é gerada utilizando o modelo local, revelando como as mudanças introduzidas pela perturbação afetaram as previsões do modelo complexo. Os coeficientes do modelo local são examinados para compreender a importância relativa de cada atributo na previsão. Para avaliar a qualidade da explicação proporcionada pelo LIME, métricas como o coeficiente de determinação (R^2) ou métricas específicas do problema podem ser empregadas, permitindo uma análise objetiva da interpretabilidade das previsões do modelo.

Um dos pontos fortes do método proposto é a sua simplicidade e eficiência, pois usa apenas camadas convolucionais 1D para extrair características do sinal de ECG e classificar diferentes tipos de arritmias. Esse método pode ser mais rápido e eficiente do que outros métodos baseados em redes neurais recorrentes ou transformadores, que podem exigir mais recursos computacionais e tempo de treinamento. Além disso, o método proposto pode ser facilmente adaptado para outros tipos de classificação de ECG, como a detecção de fibrilação atrial, bloqueio cardíaco ou síndrome do QT longo.

Outro ponto forte do método proposto é a sua interpretabilidade, pois usa o LIME para gerar explicações locais para as previsões da rede FCN, mostrando quais regiões do ECG foram mais relevantes para a classificação de isquemia. Essas explicações podem ajudar os usuários a entender como o modelo faz as suas decisões, quais são as evidências que suportam ou contradizem as suas previsões e quais são os possíveis fatores de confusão ou incerteza. Além disso, essas explicações podem aumentar a confiança e a aceitação dos usuários no modelo, bem como facilitar a sua validação e verificação.

No entanto, o método proposto também tem algumas limitações que devem ser consideradas. Uma delas é o tamanho e a qualidade dos dados usados para treinar e avaliar o modelo. O conjunto de dados de isquemia é relativamente pequeno (2559 ECGs) e pode não representar toda a variabilidade e complexidade dos casos reais de isquemia miocárdica. Além disso, os dados podem conter ruídos, artefatos ou erros de rotulagem que podem afetar o desempenho e a generalização do modelo. Portanto, seria necessário usar outros conjuntos de dados maiores e mais diversificados para confirmar os resultados obtidos neste trabalho.

Outra limitação é a generalização do modelo para outros cenários clínicos. O modelo foi treinado e testado usando ECGs de pacientes com isquemia cardíaca, mas não foi avaliado em pacientes com outras condições cardíacas ou não cardíacas que podem interferir na análise do ECG. Por exemplo, pacientes com hipertrofia ventricular esquerda, bloqueio de ramo esquerdo ou marcapasso podem apresentar alterações no ECG que podem ser confundidas com isquemia miocárdica. Além disso, pacientes com outras doenças sistêmicas, como diabetes, hipertensão ou insuficiência renal, podem ter um maior risco de isquemia miocárdica e podem requerer uma avaliação mais cuidadosa. Portanto, seria necessário testar o modelo em diferentes populações e contextos para verificar a sua robustez e aplicabilidade.

Uma terceira limitação é a possibilidade de viés ou erro nas explicações geradas pelo LIME. O LIME é uma técnica que usa um modelo mais simples e interpretável, como uma árvore de regressão, para explicar as previsões de qualquer classificador. No entanto, esse modelo mais simples pode não capturar toda a complexidade e a não linearidade da rede neural, podendo gerar explicações que não refletem fielmente as suas decisões. Além disso, o LIME pode ser sensível aos parâmetros usados para gerar as explicações, como o número de amostras, o tipo de perturbação e a medida de proximidade. Portanto, seria necessário avaliar a qualidade e a consistência das explicações geradas pelo LIME, usando outros métodos de interpretabilidade ou validação humana.

Conclusão

Neste trabalho, usamos redes neurais convolucionais 1D (FCN) para prever isquemia miocárdica a partir de ECGs. As redes FCN são um tipo de rede neural convolucional que usa apenas camadas convolucionais 1D para extrair características do sinal de ECG e classificar diferentes tipos de arritmias. Além disso, usamos uma técnica de interpretabilidade chamada LIME para gerar explicações locais para as predições da rede FCN, mostrando quais regiões do ECG foram mais relevantes para a classificação de isquemia.

Os resultados obtidos neste trabalho indicam que a rede FCN pode prever isquemia miocárdica a partir de ECGs com alta acurácia, sensibilidade e especificidade, superando os métodos tradicionais baseados na interpretação visual ou nos critérios diagnósticos. Além disso, o LIME pode fornecer explicações plausíveis e consistentes para as predições da rede FCN, aumentando a sua interpretabilidade e confiabilidade.

Portanto, concluímos que o método proposto é viável e eficaz para prever isquemia miocárdica a partir de ECGs. Esse método pode ser uma alternativa mais rápida e precisa para o diagnóstico da isquemia miocárdica, podendo auxiliar os profissionais de saúde na tomada de decisão clínica. Além disso, esse método pode aumentar a compreensão e a confiança dos usuários no modelo, fornecendo explicações sobre as suas predições.

As contribuições deste trabalho são:

- O uso de redes neurais convolucionais 1D (FCN) para prever isquemia miocárdica a partir de ECGs, com alto desempenho e eficiência.
- O uso de uma técnica de interpretabilidade (LIME) para gerar explicações locais para as predições da rede FCN, mostrando quais regiões do ECG foram mais relevantes para a classificação de isquemia.
- A avaliação do método proposto usando um conjunto de dados público e um método de validação cruzada com 10 partições aleatórias dos dados.

As implicações deste trabalho são:

- A possibilidade de melhorar o diagnóstico da isquemia miocárdica, usando um exame não invasivo, barato e rápido como o ECG.
- A possibilidade de aumentar a interpretabilidade e a confiabilidade das redes neurais para análise de ECGs, usando técnicas de interpretabilidade como o LIME.
- A possibilidade de estimular novas pesquisas na área de aprendizado de máquina aplicado à cardiologia, usando outros conjuntos de dados, outros modelos de redes neurais ou outras técnicas de interpretabilidade.

Referências

SINGSTAD, Bjørn-Jostein. Median Heartbeat Classification: Ischemia. Kaggle, 2023.

Disponível em:

<https://www.kaggle.com/code/bjoernjostein/median-heartbeat-classification-ischemia/notebook>

SINGSTAD, Bjørn-Jostein. Ischemia Dataset. Kaggle, 2023. Disponível em:

<https://www.kaggle.com/datasets/bjoernjostein/ischemia-dataset>

Fully Convolutional Network (Semantic Segmentation). Great Learning, 2022.

Disponível em:

<https://www.mygreatlearning.com/blog/fcn-fully-convolutional-network-semantic-segmentation/>

What is a fully convolution network?. Stack Exchange, 2020. Disponível em:

<https://ai.stackexchange.com/questions/21810/what-is-a-fully-convolution-network>

VON WANGENHEIM, Aldo. Deep Learning::Glossário. LAPiX, 2023. Disponível em:

[https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningglossario/#Redes totalmente convolucionais Fully Convolutional Network FCN](https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningglossario/#Redes%20totalmente%20convolucionais%20Fully%20Convolutional%20Network%20FCN)

Sobre A Interpretabilidade Do Modelo Baseado Em Aprendizado De Máquina Para

Prever Hipertensão. Biomedical house, 2023. Disponível em:

<https://pt.biomedicalhouse.com/3392353-on-the-interpretability-of-machine-learning-based-model-for-predicting-hypertension>

Mortalidade por Doença Isquêmica do Coração no Brasil – Disparidades no Nordeste.

Scielo, 2021. Disponível em:

<https://www.scielo.br/j/abc/a/NDLScWpgmyKGZB7byCdTYKL/>

Local Interpretable Model-Agnostic Explanations. Papers With Code, 2016. Disponível

em: <https://paperswithcode.com/method/lime>

EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC,

INFORMEDNESS, MARKEDNESS & CORRELATION. Arxiv, 2011. Disponível em:

<https://arxiv.org/ftp/arxiv/papers/2010/2010.16061.pdf>