



# Seleção de Hiperparâmetros no Treinamento de Redes Neurais

## Hyperparameter Selection for Training Neural Networks

Roberto Marafon Leandro<sup>1</sup>, Sarah Morgana Meurer<sup>2</sup>,  
Daniel Gomes de Pinho Zanco<sup>3</sup>, Ranniery Maia<sup>4</sup>, Eduardo Vinícius Kuhn<sup>5</sup>

### RESUMO

Neste trabalho de pesquisa, é apresentada uma abordagem sistemática para a seleção de hiperparâmetros no treinamento de duas arquiteturas relevantes de aprendizado profundo encontradas na literatura, empregadas na classificação de patologias cardíacas. Especificamente, foi criado um *script* em Python que utiliza a base de dados multirrotulo PTB-XL ECG e métricas adequadas para realizar uma busca exaustiva em uma grade de valores predefinidos para os hiperparâmetros. Resultados demonstraram um desempenho semelhante para ambas as arquiteturas quando se empregam valores de hiperparâmetros apropriados.

**PALAVRAS-CHAVE:** Busca exaustiva; classificação multirrotulo; eletrocardiograma.

### ABSTRACT

In this research work, a systematic approach is presented for hyperparameter selection in the training of two relevant deep-learning architectures from the literature, which are used in the classification of cardiac pathologies. Specifically, a Python script was developed using the PTB-XL ECG multi-label database along with appropriate metrics to perform an exhaustive grid search over a predefined range of hyperparameter values. Results demonstrated similar performance for both architectures when appropriate hyperparameter values were employed.

**KEYWORDS:** Exhaustive grid search, multilabel classification, electrocardiogram.

### INTRODUÇÃO

Doenças cardiovasculares representam aproximadamente 32% das causas de morte no mundo registradas em 2019 (WORLD HEALTH ORGANIZATION (WHO), 2021). Dessas mortes, uma importante parte poderia ser evitada através do diagnóstico precoce da doença por meio de um exame cardíaco não invasivo, denominado eletrocardiograma (ECG), o qual fornece uma representação da atividade elétrica do coração. A partir do ECG, características cardíacas, como o ritmo e a morfologia dos batimentos, podem ser identificadas e avaliadas (SANNINO; PIETRO, 2018). Contudo,

<sup>1</sup> Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Discente do Curso de Engenharia Eletrônica. Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil. E-mail: robertomarafonleandro@gmail.com. ID Lattes: 0895287239273706.

<sup>2</sup> Discente do Programa de Pós-Graduação em Tecnologias em Biociências (PPGBio). Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil. E-mail: sarah\_morgana18@hotmail.com. ID Lattes: 8816747216670793.

<sup>3</sup> Egresso do Programa de Pós-graduação em Engenharia Elétrica. Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil. E-mail: dangpzanco@gmail.com. ID Lattes: 7771398727588537.

<sup>4</sup> Docente do Departamento de Informática e Matemática Aplicada. Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte, Brasil. E-mail: rmaia@dimap.ufrn.br. ID Lattes: 7607725638613830.

<sup>5</sup> Docente do Curso de Engenharia Eletrônica. Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil. E-mail: kuhn@utfpr.edu.br. ID Lattes: 2456654064380180.



a interpretação do ECG é uma tarefa complexa, requerendo assim perícia e atenção do profissional de saúde na detecção de alterações morfológicas. Destaca-se que até 33% das interpretações de ECG por médicos não cardiologistas apresentam erros de diagnóstico significativos (MELE, 2008).

Visando aprimorar o processo de diagnóstico, ferramentas computacionais voltadas para a área de eletrofisiologia clínica vêm sendo desenvolvidas, usando arquiteturas de aprendizado profundo (SANNINO; PIETRO, 2018). Dentre essas arquiteturas, merecem destaque as introduzidas por Rajpurkar *et al.* (RAJPURKAR *et al.*, 2017) e Ribeiro *et al.* (RIBEIRO *et al.*, 2020), as quais foram posteriormente adaptadas por Sarah *et al.* (MEURER, 2022) para classificar cinco patologias cardíacas em uma base de dados pública (WAGNER *et al.*, 2020). Embora desempenho satisfatório foi observado por Sarah *et al.* (MEURER, 2022), o ajuste dos hiperparâmetros seguiu as diretrizes fornecidas nos trabalhos originais, quando disponíveis. Entretanto, em virtude de alterações realizadas quanto a base de dados e adaptações necessárias nas arquiteturas, não se pode assegurar que os hiperparâmetros utilizados resultam no desempenho ótimo; especialmente, diante da natureza empírica do processo de seleção de hiperparâmetros (GOODFELLOW; BENGIO; COURVILLE, 2016), (RASCHKA; LIU; MIRJALILI, 2022).

Neste contexto, levando em consideração as arquiteturas adaptadas de Sarah *et al.* (MEURER, 2022) e a importância da seleção de hiperparâmetros, este trabalho de pesquisa tem os seguintes objetivos:

- Desenvolver um *script* em Python que auxilie na escolha dos valores dos hiperparâmetros utilizados no processo de treinamento;
- Realizar experimentos usando um conjunto de dados e métricas padronizadas em um espaço de busca pré-definido; e
- Apresentar e discutir os resultados obtidos, fornecendo diretrizes consistentes para a comparação de desempenho.

## FORMULAÇÃO DO PROBLEMA

A base de dados utilizada, conhecido como PTB-XL e disponível em (WAGNER *et al.*, 2020), compreende 21.837 registros de sinais de ECG de 12 derivações, cada um com 10 segundos de duração e uma frequência de amostragem de 100 Hz. Este conjunto inclui sinais de ECG que foram anotados por cardiologistas e, também, de metadados que contêm declarações de diagnóstico, seguindo o padrão descrito na Norma ISO 11073-91064:2009 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO), 2009). Tais declarações de diagnóstico podem ser agrupadas em 5 superclasses não mutuamente exclusivas, o que resulta em um problema de classificação multirrotulo. Essa base de dados é dividida em conjuntos de treinamento, validação e teste, conforme descrito em (WAGNER *et al.*, 2020), sendo apenas os conjuntos de treinamento e validação são utilizados aqui (RASCHKA; LIU; MIRJALILI, 2022, Cap. 6).

Com respeito às arquiteturas, são consideradas duas arquiteturas adaptadas em (MEURER, 2022), a saber:



- A primeira, proposta por Rajpurkar *et al.* (RAJPURKAR *et al.*, 2017), composta por um total de 34 camadas. Essas camadas estão organizadas em 1 bloco de entrada, seguido por 16 blocos residuais, cada um com 2 camadas convolucionais e conexões de atalho, e finalizando com um bloco de saída que utiliza a função de ativação *sigmoid*.
- A segunda arquitetura, desenvolvida por Ribeiro *et al.* (RIBEIRO *et al.*, 2020), é composta por um total de 10 camadas. Essas camadas são organizadas em 1 bloco de entrada, seguido por 4 blocos residuais (com modificações em relação à arquitetura original), cada um contendo 2 camadas convolucionais e conexões de atalho. O bloco de saída utiliza a função de ativação *sigmoid*.

## IMPLEMENTAÇÃO PROPOSTA

A implementação proposta utiliza as bibliotecas Tensorflow/Keras e se baseia em uma busca exaustiva em grade no espaço pré-definido de hiperparâmetros. Esse espaço inclui variações de tamanho de lote (*batch size*) de 16, 32, 64, 128 e 256, otimizadores como SGD, RMSProp e Adam, e taxas de aprendizado (*learning rate*) de 0,001, 0,01 e 0,1. Essa abordagem resulta 90 rodadas distintas de treinamento. Durante o treinamento, são aplicados *callbacks* da biblioteca Keras, como *ReduceLROnPlateau* e *EarlyStopping*, para ajustar a taxa de aprendizado e interromper o treinamento antecipadamente (com um limite máximo de 100 épocas). Ao término de cada rodada de treinamento, o modelo obtido é avaliado utilizando a matriz de confusão multirrotulo (MLCM), cuja metodologia foi apresentada em (HEYDARIAN; DOYLE; SAMAVI, 2022). A partir da MLCM, diversas métricas são calculadas, incluindo *precision*, *recall* e *F1-score*, bem como suas médias *micro*, *macro* e *weighted*. Os resultados obtidos em cada rodada são registrados em um arquivo *.csv* para análises posteriores.

## RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os resultados da métrica (*macro avg*) *F1-score* no conjunto de validação para as arquiteturas de Rajpurkar *et al.* (RAJPURKAR *et al.*, 2017) e Ribeiro *et al.* (RIBEIRO *et al.*, 2020), considerando o espaço de busca de hiperparâmetros pré-definido. Resultados próximos de zero indicam que a combinação de hiperparâmetros em questão não leva à convergência da arquitetura durante o treinamento. De acordo com os resultados, a arquitetura de Rajpurkar *et al.* (RAJPURKAR *et al.*, 2017), utilizando um *batch size* de 128, otimizador Adam e *learning rate* de 0,01, alcança um valor máximo de (*macro avg*) *F1-score* de 0,71. Por outro lado, a arquitetura de Ribeiro *et al.* (RIBEIRO *et al.*, 2020), com um *batch size* de 64, otimizador RMSProp e *learning rate* de 0,01, atinge um (*macro avg*) *F1-score* máximo de 0,70. Portanto, verifica-se que ambas as arquiteturas demonstram desempenho semelhante, quando valores adequados de hiperparâmetros são empregados. Isso permite a realização de comparações justas de desempenho entre os modelos.



# XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR

Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão  
20 a 23 de novembro de 2023 - Campus Ponta Grossa, PR



SEI-SICITE  
2023

Tabela 1 – Resultados obtidos considerando a métrica (*macro avg*) *F1-score* para as arquiteturas de Rajpurkar *et al.* (RAJPURKAR *et al.*, 2017) e de Ribeiro *et al.* (RIBEIRO *et al.*, 2020).

		Adam			RMSProp			SGD		
		0,001	0,01	0,1	0,001	0,01	0,1	0,001	0,01	0,1
Rajpurkar	16	0,69	0,69	0,47	0,67	0,67	0	0,49	0,62	0,63
	32	0,70	0,70	0,50	0,68	0,66	0,67	0,57	0,59	0,64
	64	0,70	0,69	0,66	0,69	0,68	0,67	0,33	0,58	0,66
	128	0,67	0,71	0,68	0,65	0,68	0,66	0,24	0,53	0,52
	256	0,68	0,68	0,65	0,65	0,69	0,60	0,39	0,50	0,61
Ribeiro	16	0,65	0,69	0,17	0,67	0,67	0,66	0,61	0,64	0,67
	32	0,68	0,69	0,61	0,67	0,66	0,68	0,62	0,66	0,66
	64	0,68	0,69	0,53	0,63	0,70	0,62	0,59	0,66	0,67
	128	0,67	0,69	0,67	0,64	0,66	0,68	0,56	0,61	0,64
	256	0,64	0,66	0,68	0,65	0,68	0,69	0,45	0,63	0,66

Fonte: Aatoria Própria.

## CONCLUSÕES

Neste estudo, um *script* em Python foi desenvolvido para facilitar a seleção de valores de hiperparâmetros relevantes no treinamento das arquiteturas descritas em (MEURER, 2022), destinadas à classificação de patologias cardíacas em uma base de dados multirrótulo. Com base nos resultados obtidos, foi constatado que ambas as arquiteturas apresentam desempenho semelhante quando combinações apropriadas de hiperparâmetros são empregadas, sendo menos significativo o impacto dos hiperparâmetros no desempenho da arquitetura de Ribeiro *et al.* (RIBEIRO *et al.*, 2020). Visando dar continuidade à pesquisa, pretende-se agora ampliar a busca para incluir hiperparâmetros internos das arquiteturas consideradas.

## Agradecimentos

Agradecimentos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Universidade Tecnológica Federal do Paraná pelo financiamento parcial do projeto.

## Disponibilidade de código

O código pode ser encontrado em [https://github.com/lablapse/ecg\\_signal\\_processing](https://github.com/lablapse/ecg_signal_processing).

## Conflito de interesse

Não há conflito de interesse.

## REFERÊNCIAS

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. [S.l.]: MIT Press, 2016.



# XIII Seminário de Extensão e Inovação XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR

Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão  
20 a 23 de novembro de 2023 - Campus Ponta Grossa, PR

SEI-SICITE  
2023



HEYDARIAN, Mohammadreza; DOYLE, Thomas E.; SAMAVI, Reza. MLCM: Multi-label confusion matrix. **IEEE Access**, v. 10, p. 19083–19095, 2022.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **Health informatics – Standard communication protocol – Part 91064: Computer-assisted electrocardiography**. Geneva, CH, 2009.

MELE, Paul. Improving electrocardiogram interpretation in the clinical setting. **J. Electrocardiology**, v. 41, p. 438–439, 2008.

MEURER, Sarah Morgana. **Detecção e classificação de patologias cardíacas em eletrocardiogramas utilizando redes neurais profundas**. TCC do Curso de Engenharia Eletrônica, Universidade Tecnológica Federal do Paraná, Toledo, PR: [s.n.], 2022.

RAJPURKAR, Pranav et al. Cardiologist-level arrhythmia detection with convolutional neural networks. **ArXiv**, abs/1707.01836, 2017.

RASCHKA, Sebastian; LIU, Yuxi (Hayden); MIRJALILI, Vahid. **Machine Learning with Pytorch and Scikit-Learn**. 1. ed. [S.l.]: Packt, 2022. ISBN 978-1-80181-931-2.

RIBEIRO, Antônio H. et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. **Nature Commun.**, v. 11, 2020.

SANNINO, G.; PIETRO, G. De. A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. **Future Generation Comput. Syst.**, v. 86, p. 446–455, 2018.

WAGNER, Patrick et al. PTB-XL, a large publicly available electrocardiography dataset. **Sci Data**, v. 7, 2020.

WORLD HEALTH ORGANIZATION (WHO). **Cardiovascular diseases (CVDs)**. 2021. Disponível em: [🔗](#).