

João Victor Sousa

Título do seu TCC

Ilhéus - Bahia

2025

João Victor Sousa

Título do seu TCC

Trabalho apresentado ao Curso de Y da Universidade X como requisito parcial para obtenção do título de Bacharel.

Universidade Estadual de Santa Cruz
Curso de Graduação em Ciência da Computação

Orientador: Otacílio José Pereira

Ilhéus - Bahia

2025

Resumo

Escreva aqui o resumo do seu trabalho. Inclua objetivos, metodologia, resultados e conclusões.

Palavras-chave: palavra1. palavra2. palavra3.

Sumário

Sumário	3	
1	INTRODUÇÃO	5
2	METODOLOGIA	6
2.1	Particionamento dos dados e classes	6
2.2	Métricas	7
2.3	Modelos utilizados	8
2.4	Organização da entrada	8
3	RESULTADOS	9
4	CONCLUSÃO	10
	REFERÊNCIAS	11

Lista de Siglas

AUC	Area Under Curve
RNN	Redes Neurais Recorrentes
CNN	Redes Neurais Convolucionais
LSTM	Long Short-Term Memory
AAMI	Association for the Advancement of Medical Instrumentation
GRU	Gated Unit Recurrent
ECG	Eletrocardiograma
TP	Verdadeiro Positivo
FP	Falso Positivo
TN	Verdadeiro Negativo
FN	Falso Negativo

1 Introdução

Texto da introdução.

2 Metodologia

Primeiro, foi necessário definir qual banco de dados seria usado para treino e validação. Foi escolhido o banco MIT-BIH Arrhythmia Database (Moody e Mark 2005). O banco contém 48 horas de registro de ECGs compostos da seguinte forma: 23 registros escolhidos aleatoriamente de um conjunto de 4000 gravações de 24 horas de um ambulatório e 25 registros escolhidos de modo a incluir arritmias raras, mas clinicamente significantes. Por ser um banco recomendado pela AAMI, optou-se por utilizá-lo.

Em seguida, foi escolhido os tipos de arritmia, definido a estratégia de validação, as métricas o particionamento dos dados e os modelos

2.1 Particionamento dos dados e classes

Os dados foram particionados seguindo o particionamento inter-paciente, proposto por Chazel et al. (apud Silva et al. 2025) no qual cada batimento de um determinado paciente não pode aparecer ao mesmo tempo no conjunto de treino e validação. O objetivo é garantir a capacidade de generalização do modelo para qualquer paciente. Além disso, como recomendado pela AAMI, pacientes com marcapasso tiveram seus registros excluídos. Os registros 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, e 230 são utilizados para treino. Esse conjunto é chamado de DS1 e os demais (100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234) formam o conjunto de teste, chamados de DS2.

A AAMI define cinco classes de arritmia: N, SVEB, VEB, F, P e Q:

Tabela 1 – Particionamento inter-paciente proposto por Chazel et al.

Conjunto	N	SVEB	VEB	F	Q	T
DS1	45.866	944	3.788	415	8	51.021
DS2	44.259	1.837	3.221	388	7	49.712
Total	90.125	2.781	7.009	803	15	100.733

Fonte: Adaptado de Silva et al. (2025).

O conjunto DS1 foi então dividido em treino e validação usando validação cruzada de 2 partições, primeiramente, e 5 nos modelos finais.

Neste trabalho, foi adotado a classificação binária onde a classe positiva indica arritmia e a negativa um batimento normal.

2.2 Métricas

As métricas escolhidas foram: sensibilidade, precisão, acurácia, *F1 score* e AUC.

A sensibilidade pode ser interpretada como a capacidade do modelo em achar as classes positivas, no contexto desse trabalho, os batimentos arritmicos. Ela é calculada pela seguinte equação:

$$\text{Sensibilidade} = \frac{TP}{TP + FN} \quad (2.1)$$

Onde TP são os verdadeiros positivos e FN os falso negativos. A precisão calcula dos batimentos classificados como arritmicos, quais realmente são e é calculada como:

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (2.2)$$

Onde FP indica os falsos positivos. Já o *F1 score* é a média harmônica entre a precisão e a sensibilidade:

$$F1 \text{ score} = \frac{2 \cdot \text{Precisão} \cdot \text{Sensibilidade}}{\text{Sensibilidade} + \text{Precisão}} \quad (2.3)$$

Já a acurácia mede o certo geral do modelo, tanto para as classes positivas quanto para as negativas:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.4)$$

A AUC mede a capacidade do modelo em separar as classes positivas das negativas e vai de 0 a 1. Sendo 1 uma separação perfeita e 0,5 um modelo equivalente a uma classificação aleatória.

A matriz de confusão é uma forma gráfica de visualizar os acertos do modelo (positivos verdadeiros e verdadeiros negativos) e seus erros (falsos positivos e falsos negativos)

Tabela 2 – Exemplo de matriz de confusão binária

Classe Verdadeira	Classe Predita	
	Positiva	Negativa
Positiva	TP	FN
Negativa	FP	TN

Fonte: Elaborado pelo autor.

2.3 Modelos utilizados

Como o diagnóstico de arritmias depende tanto do aspecto morfológico dos ECGs quanto do aspecto sequencial, isto é, o ritmo, foram escolhidos modelos RNNs que possuem bom desempenho em dados sequenciais como séries temporais e linguagem natural (James et al. 2023). Já as redes CNN possuem grande capacidade de extração de características e podem aprender os aspectos morfológicos dos ECGs.

2.4 Organização da entrada

Foi escolhido de cada ECG a lead II, recomendada por facilitar a detecção das arritmias. Cada batimento foi segmentado utilizando a biblioteca neuroki2 (Makowski et al. 2021). Com os batimentos segmentados foi montado uma matriz tridimensional no formato: (*batch size*, tamanho da sequência, características). As características inicialmente correspondem as amostras do ECGs e demais características como intervalo RR, média do intervalo RR para o paciente e a média do 5 intervalos RR anteriores e posteriores. Em cada sequência de batimentos, só há batimentos exclusivos de um único paciente.

3 Resultados

Texto dos resultados.

4 Conclusão

Texto da conclusão.

Referências

James et al. 2023 JAMES, G. et al. *An Introduction to Statistical Learning: With Applications in Python*. Cham: Springer, 2023. E-book. Disponível em: <<https://www.statlearning.com/>>. Acesso em: 28 ago. 2025.

Makowski et al. 2021 MAKOWSKI, D. et al. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, Springer Science and Business Media LLC, v. 53, n. 4, p. 1689–1696, feb 2021. Disponível em: <<https://doi.org/10.3758%2Fs13428-020-01516-y>>.

Moody e Mark 2005 MOODY, G. B.; MARK, R. G. *MIT-BIH Arrhythmia Database*. 2005. Disponível em: <<https://www.physionet.org/physiobank/database/mitdb/>>. Acesso em: 28 set. 2025.

Silva et al. 2025 SILVA, G. A. L. et al. *A Systematic Review of ECG Arrhythmia Classification: Adherence to Standards, Fair Evaluation, and Embedded Feasibility*. 2025. ArXiv:2503.07276v1. Acesso em: 28 set. 2025.