

Aplicação de Modelos de Classificação para Detecção do Risco de Óbito em Pacientes com Insuficiência Cardíaca

Ericksulino Manoel de Araújo Moura Vitor José Ferreira dos Santos de Santana

Insuficiência Cardíaca

- Incapacidade do coração suprir demandas metabólicas;
- · Causada por doença arterial coronariana, hipertensão e outras;
- Retenção hídrica, intolerância ao esforço;
- · Alta mortalidade e impacto na qualidade de vida.





Classificadores

- Um classificador é um algoritmo de aprendizado que categoriza novos dados usando exemplos rotulados.
- Random Forest: várias árvores que votam para a decisão final;
- Gradient Boosting: aprendizado sequencial que corrige erros anteriores;
- Naive Bayes: modelo probabilístico que assume independência entre atributos;
- · KNN: Classifica com base nos vizinhos mais próximos.

Redes Neurais Artificiais

- Inspiração em redes neurais biológicas;
- Neurônios artificiais com entradas e saídas;
- Capacidade de modelar padrões complexos;
- · Aplicada em classificação, reconhecimento de padrões e outros.



MLP (Multilayer Perceptron)

- Arquitetura com múltiplas camadas ocultas;
- Funções de ativação não-lineares (ReLU, sigmoid);
- · Aprendizado supervisionado com retropropagação.

Objetivo

O trabalho tem como objetivo aplicar técnicas de aprendizado de máquina para prever a ocorrência de óbito durante o tratamento de pacientes diagnosticados com insuficiência cardíaca.







Trabalhos Relacionados

Trabalho	Métricas	Classificadores	Usa RNA
		Random Forest	
		KNN Decision Tree	
		Decision Tree Naive Bayes	
[Gürfidan and Ersoy 2021]	Acurácia	SVM	Não
[Sarijaloo et al. 2021]	AUC e Recall	Regressão Logística	Não
[Kyodo et al. 2023]	Acurácia	Bayesian-Gaussian	Não
[Smole et al. 2021]	AUC	Boosted Trees	Não
	F-measure,	Random Forest	
	AUC-ROC e	KNN	
	curva de Precisão	Naive Bayes	
[Oladimeji and Oladimeji 2020]	-Revocação.	SVM	Não
[McGilvray et al. 2022]	C-statistic	LSTM	Sim
		KNN	
		Naive Bayes	
	Acurácia	Random Forest	
	Precisão	Gradient Boosting	
Este artigo	Recall	MLP	Sim

Base de Dados

- Heart Failure Clinical Records (UCI);
- 299 pacientes (105 mulheres, 194 homens, 40–95 anos)
- Coletados em 2015 no Paquistão (2 hospitais)
- 13 atributos (binários e numéricos)
- Variável alvo: evento de morte (0 = sobreviveu, 1 = óbito)
 - 0 age
 1 anaemia
 2 creatinine_phosphokinase
 3 diabetes
 4 ejection_fraction
 5 high_blood_pressure
 6 platelets
 7 serum_creatinine
 8 serum_sodium
 9 sex
 10 smoking
 11 time
 12 DEATH_EVENT

Metodologia

- Divisão em 70% Treinamento e 30% Teste
- Random Forest;
- Gradient Boosting;
- Naive Bayes;
- · KNN;
- MLP (scikit-learn):
 - 2 camadas ocultas (128, 64 neurônios);
 - ReLU, normalização, até 1000 iterações.

Metodologia

- RNA (Keras):
 - 3 camadas ocultas (128, 64, 32), ReLU;
 - Dropout: 40%, 30%, 0%;
 - Saída com sigmoide (binária);
 - Otimizador: Adam + Binary Cross-Entropy;
 - 30 épocas, batch size 16, validação 30%.
- Métricas: Acurácia, Precisão e Recall.

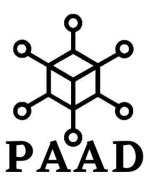
Resultados

Métrica Classificador	Acurácia	Precisão	Recall
Random Forest	$\textbf{0.836} \pm \textbf{0.036}$	$\textbf{0.783} \pm \textbf{0.098}$	0.697 ± 0.026
RNA	0.800 ± 0.038	0.689 ± 0.082	0.610 ± 0.120
Gradient Boosting	0.798 ± 0.011	0.710 ± 0.040	0.641 ± 0.047
Naive Bayes	0.791 ± 0.013	0.781 ± 0.048	0.497 ± 0.056
MLP	0.762 ± 0.035	0.646 ± 0.063	0.600 ± 0.068
KNN	0.613 ± 0.029	0.293 ± 0.112	0.172 ± 0.102

Conclusão

- Random Forest obteve a melhor acurácia na predição de óbito;
- Alta precisão e recall indicam potencial clínico para apoio à decisão;
- RNA teve bom desempenho, mas sensível ao número de amostras;
- Modelos mostram-se úteis para triagem e intervenções precoces;
- Limitações: base de dados pequena e falta de validação externa;
- Futuro: ampliar dados, otimizar modelos e explorar interpretabilidade.





Aplicação de Modelos de Classificação para Detecção do Risco de Óbito em Pacientes com Insuficiência Cardíaca

Ericksulino Manoel de Araújo Moura Vitor José Ferreira dos Santos de Santana