# As instruções destacadas em amarelo devem ser deletadas e substituídas pelo texto do manuscrito. O artigo não deve ter qualquer identificação dos autores, estando essas restritas a capa.

# Título em Português

Machine Learning em Saúde: Estudo com SVM no Reconhecimento de Doenças Cardíacas.

## Resumo

Estruturado em um único parágrafo, em até 250 palavras, considerando Objetivo, Método, Resultados e Conclusão. Deve-se utilizar fonte Times New Roman 12, com espaçamento de 1,5 cm entre as linhas. **Objetivo:** Avaliar o desempenho de diferentes configurações do classificador Support Vector Machine (SVM) no diagnóstico de doenças cardíacas foi o objetivo deste trabalho, desenvolvido com o conjunto de dados *Heart Disease Cleveland*. A proposta integra o 3º UMC SUMMIT de Pesquisa, Inovação e Extensão, com foco na aplicação de técnicas de aprendizado de máquina em saúde.. **Métodos:** O conjunto de dados contém atributos clínicos e laboratoriais de pacientes, categorizados em presença ou ausência de doença. As variáveis contínuas foram padronizadas e os dados divididos em treino e teste (80/20). Um *pipeline* com normalização e SVM foi implementado, iniciando com modelo linear. Em seguida, realizou-se busca em grade (*GridSearchCV*) dos hiperparâmetros C, *kernel* e *gamma*, utilizando validação cruzada KFold (10 divisões). A avaliação considerou a acurácia média e métricas da matriz de confusão.. **Resultados:** Os resultados mostraram que o modelo linear obteve acurácia próxima a 85%. Entretanto, configurações não lineares, especialmente o *kernel* radial, alcançaram melhor equilíbrio entre sensibilidade e especificidade. O modelo otimizado apresentou acurácia média acima de 88% e reduziu falsos negativos, fator crítico no contexto clínico. A variação entre os *folds* manteve-se controlada, reforçando a estabilidade do método, embora a limitação do tamanho da base de dados exija cautela em generalizações.. **Conclusão:** Conclui-se que o SVM é uma ferramenta eficaz no apoio ao diagnóstico de doenças cardíacas, com destaque para o uso do *kernel* radial, evidenciando o potencial do aprendizado de máquina em diagnósticos mais precisos e precoces..

## Palavras-chave

aprendizado de máquina; validação cruzada; SVM; diagnóstico médico; doenças cardíacas.

# Title in English

Machine Learning in Healthcare: A Study Using SVM for the Recognition of Heart Diseases.

## Abstract

Structured in a single paragraph, with up to 250 words, considering Objective, Method, Results, and Conclusion. The text should use Times New Roman font, size 12, with 1.5 cm line spacing. **Objective:** Evaluate the performance of different configurations of the Support Vector Machine (SVM) classifier in the diagnosis of heart diseases was the objective of this study, conducted using the *Heart Disease Cleveland* dataset. The study is part of the 3rd UMC SUMMIT on Research, Innovation, and Extension, focusing on the application of machine learning techniques in healthcare. **Methods:** The dataset contains clinical and laboratory attributes of patients, categorized as presence or absence of disease. Continuous variables were standardized, and the data were split into training and testing sets (80/20). A pipeline integrating normalization and SVM was implemented, starting with a linear model. Subsequently, a grid search (*GridSearchCV*) was performed on the hyperparameters C, kernel, and gamma, using 10-fold cross-validation (KFold). Performance was evaluated using mean accuracy and confusion matrix metrics. **Results:** Results showed that the linear model achieved approximately 85% accuracy; however, non-linear configurations, particularly the radial kernel, achieved a better balance between sensitivity and specificity. The optimized model reached mean validation accuracy above 88% and reduced false negatives, a critical factor in clinical contexts. Variation across folds remained controlled, demonstrating method stability, although the dataset size limits caution in generalization.. **Conclusion:** It is concluded that SVM is an effective tool for supporting heart disease diagnosis, highlighting the radial kernel and demonstrating the potential of machine learning for more precise and early diagnoses..

## Keywords

machine learning; cross-validation; SVM; medical diagnosis; heart diseases.

# Título en Español

Aprendizaje Automático en Salud: Estudio con SVM para el Reconocimiento de Enfermedades Cardíacas.

## Resumen

Estructurado en un solo párrafo, con un máximo de 250 palabras, considerando Objetivo, Método, Resultados y Conclusión. El texto debe utilizar la fuente Times New Roman, tamaño 12, con un espaciado de 1,5 cm entre líneas. **Objetivo:** Evaluar el desempeño de diferentes configuraciones del clasificador Support Vector Machine (SVM) en el diagnóstico de enfermedades cardíacas fue el objetivo de este estudio, realizado con el conjunto de datos *Heart Disease Cleveland*. El estudio forma parte del 3º UMC SUMMIT de Investigación, Innovación y Extensión, centrado en la aplicación de técnicas de aprendizaje automático en salud.. **Métodos:** El conjunto de datos contiene atributos clínicos y de laboratorio de pacientes, categorizados como presencia o ausencia de enfermedad. Las variables continuas fueron estandarizadas y los datos se dividieron en conjuntos de entrenamiento y prueba (80/20). Se implementó un *pipeline* que integra normalización y SVM, iniciando con un modelo lineal. Posteriormente, se realizó una búsqueda en cuadrícula (*GridSearchCV*) de los hiperparámetros C, kernel y gamma, utilizando validación cruzada KFold de 10 divisiones. La evaluación consideró la precisión media y métricas de la matriz de confusión.. **Resultados:** Los resultados mostraron que el modelo lineal alcanzó aproximadamente un 85% de precisión; sin embargo, las configuraciones no lineales, especialmente el kernel radial, lograron un mejor equilibrio entre sensibilidad y especificidad. El modelo optimizado alcanzó una precisión media de validación superior al 88% y redujo los falsos negativos, un factor crítico en contextos clínicos. La variación entre los *folds* se mantuvo controlada, demostrando la estabilidad del método, aunque el tamaño limitado de la base de datos requiere precaución en la generalización.. **Conclusión:** Se concluye que el SVM es una herramienta eficaz para apoyar el diagnóstico de enfermedades cardíacas, destacando el kernel radial y evidenciando el potencial del aprendizaje automático para diagnósticos más precisos y tempranos.

## Palabras clave

aprendizaje automático; validación cruzada; SVM; diagnóstico médico; enfermedades cardíacas.