

# Implementación de Modelos de Aprendizaje Automático para la Predicción de Ataques Cardíacos: Un Enfoque Basado en Datos Reales

Laura Victoria Ramos Agudelo

**Resumen** - Este estudio se centró en la aplicación de técnicas de aprendizaje automático para la predicción de enfermedades cardíacas, utilizando un conjunto de datos clínicos y fisiológicos. Entre los modelos evaluados, destacó el Random Forest configurado con 30 árboles y una profundidad de 2, alcanzando un puntaje F1 de 0.8459, lo que demuestra su efectividad para predecir esta condición médica crítica. Además, se exploraron técnicas de extracción y selección de características, permitiendo comprender el impacto de la dimensionalidad en el rendimiento de los diferentes modelos. Esta optimización del conjunto de datos es crucial en aplicaciones prácticas donde el costo y tiempo de recolección de datos son factores determinantes, como en el campo médico. Los resultados obtenidos resaltan el potencial de los modelos de aprendizaje automático cuidadosamente ajustados, como el Random Forest, para proporcionar herramientas valiosas en la intervención temprana y la planificación del tratamiento en pacientes con riesgo de enfermedades cardíacas.

**Índice de Términos** - Aprendizaje Automático, Predicción de Enfermedades Cardíacas, Modelos de Clasificación, Redes Neuronales, Random Forest, Máquinas de Vectores de Soporte (SVM), Extracción de Características, Análisis de Componentes Principales (PCA), Base de Datos de Cleveland, Evaluación de Modelos.

## I. INTRODUCCIÓN

La predicción temprana y precisa de enfermedades cardíacas (HD por sus siglas en inglés 'Heart Disease') en el ámbito de la cardiología es crucial, ya que estas constituyen una de las principales causas de muerte a nivel mundial. Según la Organización Mundial de la Salud (OMS), la enfermedad isquémica del corazón y el accidente cerebrovascular son las principales causas de mortalidad global, responsables del 27% de todas las muertes, lo que equivale a aproximadamente 15 millones de personas en 2019 [1].

Además de estas dos condiciones, existen otras afecciones cardíacas significativas como el ataque cardíaco (Heart attack), la insuficiencia cardíaca (Heart failure), la cardiomiopatía

(Cardiomyopathy), la enfermedad cardíaca reumática (Rheumatic heart disease), la angina de pecho (Angina pectoris) y la arritmia (Arrhythmia). Todas estas condiciones están detalladas en la Figura 1[2] y contribuyen al amplio espectro de patologías que impactan la salud cardiovascular global.

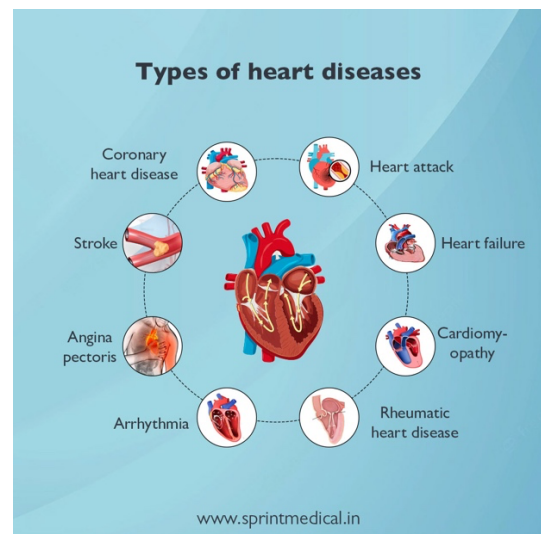


Fig. 1. Tipos de enfermedades cardíacas.

Dada relevancia de diagnósticos y tratamientos oportunos para combatir las enfermedades cardíacas, el problema de predicción que se plantea en este documento cobra una importancia considerable. Este problema se centra en determinar la presencia o ausencia de enfermedades cardíacas utilizando un conjunto de variables clínicas y fisiológicas.

Este caso de estudio se enfoca en una tarea de clasificación binaria destinada a determinar la presencia o ausencia de enfermedades cardíacas en individuos mediante el análisis de variables de entrada clínicas y fisiológicas que incluyen: la edad del paciente (age), el sexo (sex), el tipo de dolor torácico (cp), la presión arterial en reposo (restbps), el colesterol sérico (chol), el azúcar en sangre en ayunas (fbs), los resultados electrocardiográficos en reposo (restecg), la frecuencia cardíaca

máxima alcanzada (thalach), la presencia de angina inducida por el ejercicio (exang), la depresión del segmento ST inducida por ejercicio (oldpeak), la pendiente del segmento ST durante el ejercicio (slope), el número de vasos sanguíneos principales visibles en fluoroscopia (ca) y la presencia de anomalías tiroideas (thal).

La variable a predecir es el diagnóstico de enfermedad cardíaca (num), que refleja el estado angiográfico del paciente y toma el valor de 0 (sin presencia de estenosis significativa) o 1 (presencia significativa de estenosis que afecta más del 50% del diámetro arterial).

## II. TRABAJOS RELACIONADOS

Dentro del campo de la cardiología, la predicción de enfermedades cardíacas (HD) ha ganado una relevancia significativa, lo que ha fomentado la creación de diversos algoritmos dedicados a un diagnóstico preciso de estas condiciones. Es importante destacar que todos los estudios mencionados en esta sección utilizaron la misma base de datos de Cleveland, la cual también será utilizada en el presente análisis. Un ejemplo significativo es el trabajo de Fitriyani et al. [3], quienes mediante un meticuloso proceso de selección de características desarrollaron un modelo denominado HDPM. Este modelo se distingue por la integración de técnicas de aprendizaje automático de vanguardia, incluyendo el uso de DBSCAN para la detección y eliminación de datos atípicos, SMOTE-ENN para equilibrar el conjunto de entrenamiento desbalanceado, y la adopción de XGBoost como método de aprendizaje para generar el modelo de predicción.

La validación de este modelo se realizó mediante una validación cruzada de 10 pliegues, confirmando su robustez y efectividad con una precisión del 98.40% para el conjunto de datos de Cleveland, estableciendo así un precedente destacado para los sistemas de apoyo a las decisiones clínicas en el contexto de la alta precisión en la detección de enfermedades cardíacas.

En una línea similar, Al Reshan et al. [4] propusieron un modelo que es el resultado de una estrategia híbrida que integra diversas arquitecturas de redes neuronales profundas como las Redes Neuronales Artificiales (ANN), Convolucionales (CNN), y de Memoria a Corto y Largo Plazo (LSTM). Este enfoque multifacético no solo facilita el procesamiento y análisis de extensos conjuntos de datos complejos, sino que también refuerza la precisión de las predicciones. Mediante la implementación de una técnica de validación holdout 80:20, donde el 80% de los datos se empleó para el entrenamiento y el 20% para las pruebas, este sistema HDNN se evaluó rigurosamente, demostrando ser excepcionalmente eficaz con una precisión del 97.75% en el uso del conjunto de datos de Cleveland para enfermedades cardíacas.

No obstante, se reconoce que el estudio no está exento de

limitaciones. Las más notables incluyen la posible disminución de compatibilidad con varias técnicas de selección de características y una marcada vulnerabilidad ante bases de datos con numerosos valores ausentes, lo que sugiere la necesidad de refinamientos futuros para mejorar la robustez y generalización del modelo propuesto.

Otro enfoque destacado es el presentado por Mohan et al. [5], quienes aportaron una metodología híbrida para la predicción de enfermedades cardíacas, combinando la robustez del Bosque Aleatorio (RF) con la simplicidad del Modelo Lineal (LM), enfocado en el conjunto de datos de Cleveland. Esta aproximación, denominada Modelo Híbrido de Bosque Aleatorio con Modelo Lineal (HRFLM), resalta por alcanzar un nivel de precisión del 88.7. Este método resultó efectivo en reducir el conjunto de atributos críticos, proporcionando así un conjunto de datos optimizado para la entrada en Redes Neuronales Artificiales (ANN).

En su investigación, utilizaron 297 registros de pacientes del conjunto de datos, de los cuales 252 se usaron para entrenamiento y los restantes para pruebas. La selección y reducción estratégica de atributos no solo simplificó el modelo, sino que también permitió enfocar el proceso de aprendizaje en las características más impactantes, optimizando así la capacidad predictiva del modelo y estableciendo un precedente en la aplicación de técnicas de aprendizaje automático combinadas para la predicción de enfermedades cardíacas.

Además, Gupta et al. [6] exponen un enfoque computacional para el diagnóstico de afecciones cardíacas a través de técnicas de aprendizaje automático. Los autores introducen un sistema que aprovecha el Análisis Factorial de Datos Mixtos (FAMD) para extraer y derivar características del conjunto de datos de enfermedades cardíacas de Cleveland.

Para la validación del modelo, se implementa una partición con una proporción del 80:20, es decir, el 80% de los datos se utiliza para el entrenamiento del modelo, mientras que el 20% restante se emplea para probar su rendimiento.

Los resultados obtenidos demuestran que el marco propuesto MIFH (Machine Intelligence Framework for Heart Disease Diagnosis) supera a varios métodos de referencia en términos de precisión, sensibilidad y especificidad. MIFH se destaca por ofrecer la mejor solución entre todos los modelos predictivos ingresados, considerando los criterios de rendimiento y mejorando la eficacia del sistema.

En cuanto a los resultados, el marco logra distintos niveles de precisión, sensibilidad, especificidad, puntuación F1, coeficiente de correlación de Matthew (MCC) y área bajo la curva (AUC) dependiendo del algoritmo de aprendizaje automático aplicado. Por ejemplo, el modelo predictivo Random Forest (RF) alcanzó una precisión del 93.44%, una sensibilidad del 89.28% y una especificidad del 96.96%, subrayando su potencial para asistir a doctores y radiólogos en

el diagnóstico preciso de pacientes con enfermedades cardíacas.

### III. EXPERIMENTOS

En este estudio se hace uso de la base de datos "Heart Attack Analysis & Prediction Dataset", disponible en Kaggle [7]. y originalmente procedente del UCI Machine Learning Repository. Específicamente, se emplea la base de datos de Cleveland, que fue recopilada del V.A. Medical Center en Long Beach y la Cleveland Clinic Foundation en California, Estados Unidos. Esta base de datos ha sido ampliamente reconocida y utilizada en la comunidad de investigación médica para el análisis y predicción de enfermedades cardíacas [8].

La base de datos utilizada contiene 303 instancias y 14 características seleccionadas entre un total de 76 atributos disponibles. El valor de la clase de salida original es una variable de múltiples clases con un rango de valores de 0 a 4. Sin embargo, la versión de la base de datos en Kaggle ha sido transformada para que la variable de salida solo tome los valores de 0 y 1. En esta transformación, el valor 0 representa la ausencia de enfermedad cardíaca, mientras que los valores de 1 a 4 se agrupan bajo el valor 1, indicando la presencia de enfermedad cardíaca según el nivel de estenosis basado en el estado angiográfico del paciente.

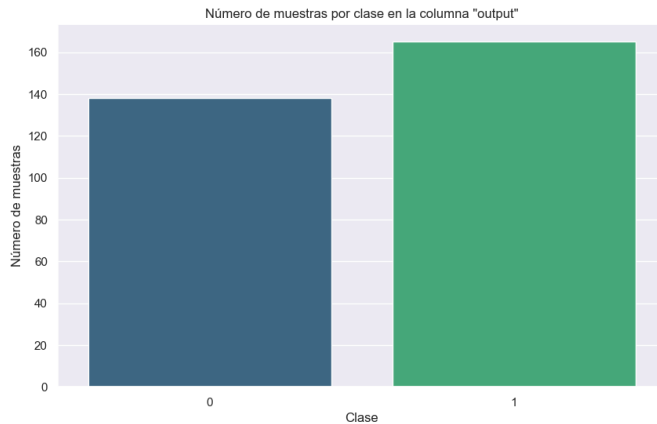


Fig. 2. Diagrama de barras clasificado por ausencia (0) o presencia (1) de enfermedad cardíaca

La base de datos presenta una distribución relativamente equilibrada entre las clases de la variable objetivo 'output', con 165 muestras correspondientes a la clase 1 y 138 muestras a la clase 0. Esta distribución cercana al equilibrio es crucial para el entrenamiento de modelos de aprendizaje automático, ya que ayuda a prevenir el sesgo hacia la clase mayoritaria, asegurando que el modelo sea igualmente efectivo para detectar ambas condiciones. Este balance facilita una evaluación más fiable de la precisión del modelo y minimiza la necesidad de aplicar técnicas de reequilibrio de clases, tales como el sobremuestreo o el submuestreo, que a menudo se requieren en conjuntos de datos con una distribución desigual.

Atributo	Descripción	Tipo	Rango	Valores Únicos	Media	Mediana	Moda
Age	Edad del paciente en años completos	Continuo	29 a 77	41	54,4	55	58
Sex	Sexo del paciente (1 = masculino; 0 = femenino)	Categorico	0 a 1	2	0,68	1	1
CP	Tipo de dolor torácico (0 a 3) Valor 0: Angina Típica Valor 1: Angina Atípica Valor 2: Dolor No Anginoso Valor 3: Asintomático	Categorico	0 a 3	4	0,97	1	0
Trtbps	Presión arterial en reposo (mm Hg al ingreso al hospital)	Continuo	94 a 200	49	131,62	130	120
Chol	Colesterol sérico en mg/dl	Continuo	126 a 564	152	246,26	240	197
FBS	Azúcar en sangre en ayunas Valor 0: <= 120 mg/dL Valor 1: > 120 mg/dL	Categorico	0 a 1	2	0,15	0	0
Restecg	Resultados electrocardiográficos en reposo Valor 0: Normal Valor 1: Anormalidad de la onda ST-T Valor 2: Hipertrofia ventricular izquierda probable o definida	Categorico	0 a 2	3	0,53	1	1
Thalachh	Máxima frecuencia cardíaca alcanzada	Continuo	71 a 202	91	149,64	153	162
Exng	Angina inducida por el ejercicio Valor 0: No Valor 1: Si	Categorico	0 a 1	2	0,33	0	0
Oldpeak	Depresión del ST inducida por el ejercicio respecto al reposo	Continuo	0 a 6,20	40	1,04	0,8	0
Slp	Pendiente del segmento ST del ejercicio	Categorico	0 a 2	3	1,4	1	2
Ca	Número de vasos principales (0-3) vistos en fluoroscopia	Categorico	0 a 4	5	0,73	0	0
Thall	Estado del corazón Valor 0: Ninguno (Normal) Valor 1: Defecto fijo Valor 2: Defecto reversible Valor 3: Talasemia	Categorico	0 a 3	4	2,31	2	2
Num (Objetivo)	Diagnóstico de enfermedad cardíaca (estado angiográfico) Valor 0: No Valor 1: Si	Categorico	0 a 1	2	0,54	1	1

Tabla I. Descripción de las variables de la base de datos HD

### IV. MODELOS DE PREDICCIÓN

Para abordar la clasificación de muestras en el contexto de enfermedades cardíacas, se ha implementado cinco modelos de aprendizaje automático: Funciones Discriminantes Gaussianas, K vecinos más cercanos, Redes Neuronales Feed-Forward, Random Forest y Máquinas de Vectores de Soporte. Esta selección de modelos tiene como objetivo optimizar la precisión en la predicción de condiciones cardíacas, adaptándose meticulosamente a las particularidades de los datos clínicos como se mostrará a continuación.

#### A. Análisis con Funciones Discriminantes Gaussianas

El modelo Quadratic Discriminant Analysis (QDA) es un clasificador que establece un límite de decisión cuadrático. Este enfoque se genera ajustando densidades condicionales a cada clase según los datos y aplicando la regla de Bayes. El modelo ajusta una densidad gaussiana para cada clase, lo que permite que cada grupo tenga su propia matriz de covarianza y medios, facilitando la adaptación a estructuras más complejas en los datos.

Modelo	Precisión	Recall	F1-score	Tiempo de ejecución
QDA reg_param=0.3	0.843524	0.835161	0.833376	0.000641
QDA reg_param=0.4	0.836857	0.828495	0.826717	0.000725
QDA reg_param=0.6	0.839287	0.828387	0.825439	0.000853
QDA reg_param=0.7	0.839287	0.828387	0.825439	0.000758
QDA reg_param=0.5	0.835332	0.825161	0.822501	0.000836
QDA reg_param=0.1	0.828302	0.822258	0.821242	0.000815
QDA reg_param=0.2	0.829658	0.822258	0.820748	0.000756
QDA reg_param=0.8	0.831520	0.818387	0.814649	0.000772
QDA reg_param=0.9	0.828440	0.815269	0.811712	0.000761
QDA reg_param=1	0.823150	0.811828	0.808685	0.000773
QDA reg_param=0.05	0.814720	0.808925	0.808235	0.003100
QDA reg_param=0.01	0.807936	0.802366	0.801787	0.001178
QDA reg_param=0.02	0.807936	0.802366	0.801787	0.001354
QDA reg_param=0.0	0.800524	0.795806	0.795237	0.001453

Tabla II. Rendimiento de QDA según el parámetro de regularización

Los resultados obtenidos al aplicar QDA (Tabla II) con diversos valores del parámetro de regularización muestran una mejora progresiva en las métricas de Precisión, Recall y F1-score conforme aumenta este parámetro, como se detalla en la Tabla II. Este aumento en la regularización parece contribuir a la estabilidad de las estimaciones del modelo en el conjunto de datos, posiblemente reduciendo el sobreajuste y mejorando su capacidad de generalización. El desempeño óptimo se observa cuando el parámetro de regularización se establece en 0.3, donde se alcanza un F1-score de 0.83. Sin embargo, al seguir incrementando se observa, que el rendimiento tiende a estabilizarse y luego a disminuir, indicando que una excesiva regularización podría perjudicar la habilidad del modelo para capturar las relaciones fundamentales en los datos.

### B. Análisis con el Modelo K-Nearest Neighbors (KNN)

Para refinar la evaluación del modelo KNeighborsClassifier, se experimentó con una gama de configuraciones, variando el número de vecinos desde 3 hasta 43.

El análisis revela que el rendimiento óptimo se da en torno a 11 vecinos, donde el F1-score alcanza un valor de aproximadamente 0.832 (Tabla III). Más allá de este punto, aumentar el número de vecinos no resulta en mejoras significativas de F1-score y puede incluso comenzar a deteriorar el rendimiento del modelo, lo que sugiere un posible sobreajuste con configuraciones de menor número de vecinos o subajuste con demasiados vecinos. Este patrón subraya la importancia de seleccionar cuidadosamente el número de vecinos en KNN para asegurar una generalización efectiva, especialmente en el contexto de diagnósticos médicos donde la precisión de la predicción es vital.

Modelo	Precisión	Recall	F1-score	Tiempo de ejecución
KNN n_neighbors=11	0.848241	0.834946	0.832498	0.000574
KNN n_neighbors=15	0.850125	0.835161	0.831915	0.000602
KNN n_neighbors=13	0.851659	0.835161	0.831634	0.000563
KNN n_neighbors=7	0.838965	0.828495	0.826528	0.001266
KNN n_neighbors=21	0.849924	0.828602	0.823132	0.000526
KNN n_neighbors=31	0.848080	0.828602	0.823055	0.000517
KNN n_neighbors=17	0.840954	0.825269	0.821390	0.000582
KNN n_neighbors=19	0.842948	0.825376	0.820818	0.000536
KNN n_neighbors=29	0.844398	0.825269	0.819891	0.000544
KNN n_neighbors=9	0.832849	0.821720	0.819357	0.001093
KNN n_neighbors=3	0.821741	0.818602	0.817755	0.001861
KNN n_neighbors=5	0.823688	0.815376	0.813836	0.000806
KNN n_neighbors=37	0.836852	0.818710	0.813331	0.000531
KNN n_neighbors=27	0.836441	0.818710	0.813262	0.000525
KNN n_neighbors=41	0.838157	0.818710	0.813046	0.000528
KNN n_neighbors=23	0.839872	0.818817	0.812882	0.000543
KNN n_neighbors=25	0.838289	0.818710	0.812827	0.000506
KNN n_neighbors=33	0.839993	0.818602	0.812714	0.000535
KNN n_neighbors=35	0.834049	0.815376	0.809979	0.000544
KNN n_neighbors=43	0.831843	0.812151	0.806515	0.000534

Tabla III. Rendimiento del modelo KNN en función del número de vecinos

### C. Análisis con el Modelo Random Forest

En el estudio de la detección de enfermedades cardíacas, se evaluó exhaustivamente el modelo Random Forest utilizando una amplia gama de configuraciones de hiperparámetros para optimizar su rendimiento. Se probaron múltiples combinaciones que incluyen:

- Número de árboles ('n\_estimators'): 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50.
- Profundidad máxima de los árboles ('max\_depth'): 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 y 'None' (sin límite).
- Criterio de división ('criterion'): 'gini' y 'entropy'.

Con estas configuraciones, se realizaron un total de 420 combinaciones posibles. Este enfoque metódico asegura un análisis exhaustivo sobre cómo cada configuración afecta la precisión, el recall y el F1-score del modelo.

Entre todas las configuraciones probadas, dos en particular mostraron resultados prometedores. La primera configuración con una profundidad de nueve ('depth=9') y el criterio de entropía alcanzó un F1-score de 0.846163, demostrando ser altamente efectiva. Sin embargo, una configuración más simple con una profundidad de dos ('depth=2') y el mismo criterio de entropía produjo resultados casi idénticos en términos de F1-score (0.845859), pero con una ejecución ligeramente más rápida. Este hallazgo sugiere que ajustes en la profundidad pueden tener un impacto significativo en la eficiencia del

modelo sin comprometer la calidad del mismo.

Modelo	Precisión	Recall	F1-score	Tiempo de ejecución
RF_n=30_depth=9_criterion=entropy	0.858345	0.847957	0.846163	0.033536
RF_n=30_depth=2_criterion=entropy	0.858703	0.848280	0.845859	0.029150
RF_n=50_depth=50_criterion=entropy	0.855491	0.844731	0.843051	0.077160
RF_n=50_depth=None_criterion=entropy	0.855491	0.844731	0.843051	0.054466
RF_n=50_depth=40_criterion=entropy	0.855491	0.844731	0.843051	0.080417
RF_n=50_depth=30_criterion=entropy	0.855491	0.844731	0.843051	0.069704
RF_n=50_depth=20_criterion=entropy	0.855491	0.844731	0.843051	0.059277
RF_n=50_depth=2_criterion=entropy	0.854198	0.844839	0.842871	0.048506
RF_n=40_depth=10_criterion=entropy	0.849055	0.841290	0.840060	0.047563
RF_n=10_depth=6_criterion=entropy	0.852759	0.841505	0.840010	0.011133
RF_n=30_depth=3_criterion=gini	0.848921	0.841398	0.839623	0.029963
RF_n=50_depth=2_criterion=gini	0.851087	0.841505	0.839563	0.047617
RF_n=40_depth=2_criterion=entropy	0.852976	0.841613	0.839432	0.043481
RF_n=30_depth=10_criterion=entropy	0.849437	0.841290	0.839423	0.033127
RF_n=30_depth=2_criterion=gini	0.854095	0.841720	0.839154	0.032691
RF_n=6_depth=7_criterion=entropy	0.845240	0.838495	0.837397	0.007092
RF_n=50_depth=9_criterion=entropy	0.850832	0.838602	0.836816	0.055422
RF_n=20_depth=8_criterion=entropy	0.847282	0.837957	0.836774	0.024646
RF_n=20_depth=3_criterion=gini	0.845465	0.838065	0.836657	0.027759
RF_n=30_depth=8_criterion=entropy	0.847871	0.838065	0.836383	0.032772
RF_n=20_depth=50_criterion=entropy	0.849688	0.837957	0.836234	0.027089
RF_n=20_depth=None_criterion=entropy	0.849688	0.837957	0.836234	0.034226
RF_n=20_depth=20_criterion=entropy	0.849688	0.837957	0.836234	0.027784
RF_n=20_depth=40_criterion=entropy	0.849688	0.837957	0.836234	0.023744

Tabla IV. Parte inicial de la tabla de resultados del modelo Random Forest según diferentes configuraciones de profundidad, número de árboles y criterios

Dada la similitud en el rendimiento entre los dos modelos destacados, se seleccionó el modelo con menor profundidad. Este modelo no solo es más rápido y eficiente en términos computacionales, sino que también ofrece ventajas en términos de simplicidad y generalización. Los árboles más simples son más fáciles de interpretar y tienden a generalizar mejor a nuevos datos, lo que los hace menos susceptibles al sobreajuste. Este enfoque equilibra efectivamente entre obtener un modelo potente y mantener la capacidad de generalizar más allá del conjunto de datos de entrenamiento, lo que es crucial en aplicaciones médicas donde la precisión y la robustez son esenciales.

#### D. Análisis con Feed-Forward Neural Networks

En el estudio de la detección de enfermedades cardíacas utilizando redes neuronales de tipo feed-forward, se implementaron dos configuraciones distintas para abordar el problema de clasificación biclase. La primera configuración utilizó una capa de salida sigmoideal con una única neurona, donde la salida representa la probabilidad de pertenecer a la clase positiva. Se empleó la función de costo logística, que es eficaz para entrenamientos que requieren una clasificación precisa al comparar la salida predicha con el resultado binario esperado. Se exploraron diferentes tamaños de la capa oculta, probando 10, 20 y 50 neuronas, y se variaron las épocas de entrenamiento en 10, 30 y 50 para evaluar cómo la complejidad y el número de iteraciones afectan el desempeño del modelo.

Neurons	Epochs	Precision	Recall	F1-Score	Average Time
10	30	0.847504	0.841398	0.839946	1.775288
50	10	0.849752	0.841398	0.839575	1.171429
20	30	0.846087	0.837849	0.835947	1.774375
50	50	0.841052	0.834731	0.833666	3.967645
20	50	0.844103	0.834731	0.833110	2.876412
50	30	0.838289	0.831505	0.830145	2.084879
10	50	0.827514	0.821398	0.819914	2.652928
20	10	0.826290	0.821290	0.819783	1.409872
10	10	0.816458	0.811720	0.810972	1.160498

Tabla V. Resultados de la evaluación del modelo de Red Neuronal Feed-Forward utilizando una única neurona con función de activación sigmoideal.

En la evaluación de la red neuronal con función de activación sigmoideal, la configuración que utilizó 10 neuronas en una capa oculta y se entrenó durante 30 épocas sobresalió al proporcionar los mejores resultados (Tabla V).

La segunda configuración optó por una capa de salida softmax con dos neuronas, proporcionando una probabilidad para cada una de las clases del problema biclase, con la suma de probabilidades siendo 1. Aquí se utilizó la función de costo de entropía cruzada categórica, adecuada para modelos cuyas salidas son probabilidades. Esta función es particularmente útil para configuraciones con múltiples clases de salida. Al igual que en la configuración sigmoideal, se experimentó con la misma gama de neuronas en la capa oculta y se variaron las épocas de entrenamiento para observar las diferencias en rendimiento.

Neurons	Epochs	Precision	Recall	F1-Score	Average Time
50	10	0.847912	0.838065	0.836216	1.712431
50	30	0.839365	0.834731	0.833383	2.101617
20	50	0.835641	0.831183	0.830412	3.360874
20	30	0.835421	0.828172	0.826530	2.070739
10	30	0.832365	0.824624	0.822930	1.706691
50	50	0.828682	0.821828	0.820519	2.925104
10	50	0.824398	0.818172	0.817101	2.416097
20	10	0.788216	0.781290	0.778531	1.082019
10	10	0.789534	0.778495	0.776016	1.068244

Tabla VI. Resultados de la evaluación del modelo de RNFF utilizando dos neuronas con función de activación softmax en la capa de salida.

En la evaluación de la segunda configuración del modelo de redes neuronales utilizando la función de activación softmax y la pérdida de entropía categórica, los mejores resultados se obtuvieron utilizando una configuración de 50 neuronas en la capa oculta con un entrenamiento de 10 épocas. Esta configuración logró una precisión del 84.79%, un recall del 83.81% y un F1-score del 83.62%, con un tiempo promedio de ejecución de aproximadamente 1.17 segundos (Tabla V).

Al comparar las dos configuraciones de la red neuronal para

la clasificación de enfermedades cardíacas, la primera configuración, empleando una función de activación sigmoideal y pérdida logística, mostró en general mejores resultados que la segunda configuración que utilizó la función de activación softmax con pérdida de entropía categórica.

#### E. Análisis con Modelo de Máquinas de Vectores de Soporte

Continuando con la exploración de modelos de aprendizaje automático para la clasificación de enfermedades cardíacas, se realizaron experimentos adicionales utilizando el modelo de Máquina de Vectores de Soporte (SVM). En estos experimentos, se ajustaron cuidadosamente los hiperparámetros C y gamma, evaluando combinaciones de C = 0.1, 1, 10, 100, 200, 300, 900, 1000, y con valores de gamma desde 0.0001 hasta 0.01.

Modelo	Precisión	Recall	F1-score	Tiempo de ejecución
SVM_C=300_gamma=0.001	0.855171	0.841613	0.838212	0.003508
SVM_C=1_gamma=0.01	0.854675	0.838495	0.834368	0.003104
SVM_C=900_gamma=0.0001	0.850092	0.835054	0.831500	0.003014
SVM_C=1000_gamma=0.0001	0.850092	0.835054	0.831500	0.012895
SVM_C=100_gamma=0.001	0.850092	0.835054	0.831500	0.003027
SVM_C=200_gamma=0.001	0.850546	0.835054	0.831184	0.003378
SVM_C=10_gamma=0.001	0.850751	0.835161	0.831031	0.003058
SVM_C=300_gamma=0.0001	0.852687	0.835161	0.830851	0.003185
SVM_C=100_gamma=0.0001	0.848350	0.831935	0.827393	0.003192
SVM_C=200_gamma=0.0001	0.845608	0.828602	0.824093	0.003143
SVM_C=0.1_gamma=0.1	0.848545	0.828602	0.823641	0.004284
SVM_C=10_gamma=0.01	0.833152	0.818602	0.814617	0.002974
SVM_C=1_gamma=0.1	0.825557	0.815054	0.812939	0.003480
SVM_C=900_gamma=0.001	0.830629	0.815269	0.811228	0.005094
SVM_C=1000_gamma=0.001	0.830629	0.815269	0.811228	0.005425
SVM_C=10_gamma=0.0001	0.838874	0.815376	0.808660	0.003916
SVM_C=1_gamma=0.001	0.835540	0.812043	0.805422	0.003953
SVM_C=0.1_gamma=0.01	0.837565	0.812043	0.805155	0.006682

Tabla VII. Mejores resultados de la evaluación del modelo SVM según diferentes configuraciones de C y Gamma.

Estos experimentos meticulosos permitieron medir el impacto directo de cada configuración en las métricas de precisión, recall y el puntaje F1 del modelo. Los mejores resultados se obtuvieron con C = 300 y gamma = 0.001, donde se alcanzó un F1-score de 0.838212.

Este conjunto de pruebas subraya la relevancia de la afinación precisa de hiperparámetros en SVM, destacando cómo ajustes aparentemente menores pueden influir significativamente en la efectividad del modelo. Estos resultados son prometedores para futuras aplicaciones y optimizaciones en el campo de la detección y diagnóstico de condiciones cardíacas, asegurando que los modelos no solo sean precisos, sino también prácticos para uso clínico.

En el análisis de los modelos aplicados a la detección de enfermedades cardíacas, se observa un destacado desempeño de

tres técnicas según el F1-score. El modelo Random Forest, configurado con un criterio de entropía y una profundidad máxima de dos, lideró en eficacia con un F1-score de 0.845859 y demostró ser el más rápido, con un tiempo de ejecución de solo 0.029150 segundos. Las Redes Neuronales Feed-Forward, con una única neurona en la capa de salida y 10 neuronas en la capa oculta durante 30 épocas, también mostraron resultados sólidos con un F1-score de 0.839946, aunque su tiempo de ejecución fue significativamente mayor, 1.775288 segundos. Por otro lado, el modelo de Máquina de Vectores de Soporte (SVM) con parámetros C=300 y gamma=0.001, aunque ligeramente por debajo en F1-score con 0.838212, se destacó por su rapidez, con un tiempo de apenas 0.003508 segundos.

#### V. SELECCIÓN DE CARACTERÍSTICAS

En el proceso de selección de características para la optimización de modelos de clasificación de enfermedades cardíacas, se eligió emplear un enfoque de selección secuencial ascendente utilizando como criterio la métrica de puntuación *f1\_weighted*. Este método es particularmente efectivo para equilibrar precisión y recall, elementos cruciales en aplicaciones médicas donde es fundamental tanto la correcta identificación de casos positivos como la minimización de falsos negativos.

Se seleccionó como función criterio el clasificador K Neighbors (KNN) con 11 vecinos para guiar la selección de características. La elección de esta función wrapper se justifica por su robustez y simplicidad, ya que KNN es conocido por su capacidad de generalizar bien a través de diferentes conjuntos de datos sin necesidad de suposiciones complicadas sobre la distribución de estos. Adicionalmente, en el apartado anterior, KNN no solo demostró ser efectivo en términos de precisión, sino también eficiente en cuanto a tiempo de ejecución, facilitando iteraciones rápidas durante la selección de características. También es sabido que las funciones wrapper tienen como ventaja sobre las funciones filtro, la selección de subconjuntos más acotados.

Los resultados del proceso (Tabla VIII) de selección revelaron que ciertas combinaciones de características ofrecen un rendimiento notablemente superior en comparación con otras. Por ejemplo, el subconjunto de las características CP, Exang, Oldpeak, Slope y Caa, mostró un desempeño destacado, sugiriendo que estos atributos poseen una alta relevancia para predecir la presencia de enfermedades cardíacas con efectividad. Este hallazgo es esencial, ya que identifica los indicadores más influyentes que pueden ser cruciales para diagnósticos más precisos y tratamientos efectivos.



feature_idx	cv_scores	avg_score
(2, 8, 9, 10, 11)	[0.8260869565217391, 0.8260869565217391, 0.8, ...	0.828213
(0, 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12)	[0.8260869565217391, 0.8478260869565217, 0.8, ...	0.828116
(1, 2, 3, 7, 8, 9, 10, 11)	[0.8478260869565217, 0.8478260869565217, 0.777...	0.823575
(1, 2, 3, 6, 7, 8, 9, 10, 11)	[0.8695652173913043, 0.8260869565217391, 0.777...	0.823575
(1, 2, 3, 5, 6, 7, 8, 9, 10, 11)	[0.8043478260869565, 0.8260869565217391, 0.755...	0.81942
(1, 2, 8, 9, 10, 11)	[0.8478260869565217, 0.8260869565217391, 0.8, ...	0.819227
(1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12)	[0.8260869565217391, 0.8695652173913043, 0.777...	0.81913
(1, 2, 7, 8, 9, 10, 11)	[0.8260869565217391, 0.8043478260869565, 0.844...	0.814976
(2, 8, 9, 11)	[0.7608695652173914, 0.782608695652174, 0.8, 0...	0.806473
(8, 9, 11)	[0.8043478260869565, 0.782608695652174, 0.8, 0...	0.801836
(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)	[0.782608695652174, 0.8478260869565217, 0.7555...	0.801643
(9, 11)	[0.8043478260869565, 0.7608695652173914, 0.711...	0.766377
(11,)	[0.8043478260869565, 0.6739130434782609, 0.755...	0.75343

Tabla VIII. Índices de las características seleccionadas junto con el fi promedio generado por esta combinación.

La aparición repetida de "Ca" (número de vasos principales coloreados por fluoroscopia), en los conjuntos óptimos puede indicar su relevancia en el contexto de la detección de enfermedades cardíacas, donde la presencia de vasos sanguíneos significativos puede ser un indicador crítico de riesgo. Por otro lado, algunas características como "Chol" (colesterol sérico) y la edad rara vez aparecen en los subconjuntos seleccionados, lo que puede sugerir que estas variables tienen menos impacto o redundancia informativa en presencia de otras características más discriminativas. Esta observación podría guiar decisiones futuras sobre simplificación de modelos o enfocarse en datos que proporcionan la mayor utilidad diagnóstica.

La implementación de este método no solo proporciona un modelo optimizado con un conjunto reducido de características, sino que también mejora la eficiencia computacional al eliminar atributos redundantes o menos informativos. Este enfoque garantiza que el modelo final sea más sencillo y rápido, sin comprometer la capacidad de generalización, lo cual es vital para aplicaciones prácticas en entornos clínicos donde el tiempo y la precisión son fundamentales.

La comparación entre los resultados del modelo Random Forest utilizando un conjunto de características seleccionadas frente a la utilización de todas las características muestra la efectividad de la selección de características. A pesar de reducir la cantidad de características a solo cinco, el modelo mantiene un rendimiento robusto, logrando un F1-score de 0.845859 con un tiempo de ejecución notablemente reducido de 0.038297 segundos. Este resultado subraya la eficiencia en el rendimiento sin un compromiso en la precisión, demostrando la utilidad de omitir características redundantes o menos informativas, lo cual es vital para la optimización de recursos en entornos de producción. Así mismo, el modelo SVM exhibió un F1-score de 0.838212 utilizando estas características seleccionadas, idéntico al obtenido con un conjunto más amplio, evidenciando una consistencia en la capacidad predictiva a pesar de la simplificación del modelo.

En contraste a los resultados anteriores, la evaluación del modelo de red neuronal feed-forward con una neurona en la capa de salida, tras la selección de características, mostró un F1-

score de 0.807886 con un tiempo de ejecución de 2.003435 segundos. Este resultado es notablemente inferior al obtenido utilizando todas las características disponibles, donde se alcanzó un F1-score de 0.839946 en un tiempo ligeramente menor de 1.775288 segundos. Este comportamiento sugiere que, aunque la reducción de características puede simplificar el modelo y reducir la complejidad computacional, en este caso particular ha disminuido la capacidad del modelo para capturar la variabilidad en los datos y realizar predicciones precisas. Es evidente que la omisión de ciertas características ha impactado negativamente en la efectividad del modelo, subrayando la importancia de una cuidadosa selección de características que no comprometa la capacidad predictiva del modelo en aras de la eficiencia.

## VI. EXTRACCIÓN DE CARACTERÍSTICAS

### A. Análisis con Modelo de Máquinas de Vectores de Soporte

En el análisis de clasificación para optimizar los modelos de Support Vector Machine (SVM), se implementó una metodología de extracción de características utilizando el Análisis de Componentes Principales (PCA). Este enfoque consistió en explorar diversas configuraciones del número de componentes principales, variando desde 1 hasta 13. Cada configuración se evaluó mediante un pipeline que integró PCA y SVM, donde PCA redujo la dimensionalidad de los datos y SVM actuó sobre estos datos transformados para la clasificación.

PCA Components	Precision	Recall	F1-Score	Average Execution Time
12	0.848628	0.840968	0.838692	0.004197
6	0.845256	0.837742	0.835881	0.003653
11	0.846019	0.837634	0.835359	0.003522
5	0.841642	0.834409	0.832689	0.003299
8	0.838512	0.831290	0.829765	0.003700
7	0.839550	0.831290	0.829635	0.003637
13	0.839530	0.830968	0.828433	0.004289
9	0.837613	0.828065	0.825574	0.003535
4	0.831035	0.824409	0.822304	0.003517
10	0.834723	0.824731	0.822022	0.003730
2	0.821038	0.814409	0.812253	0.003527
3	0.815715	0.807957	0.805683	0.003637
1	0.804914	0.794516	0.789980	0.004762

Tabla IX. Resultados del SVM con C=300 y gamma=0,001 con extracción de características.

Los resultados obtenidos del análisis de componentes principales (PCA) aplicado antes de entrenar el modelo SVM (Tabla IX) revelan cómo la variación en el número de componentes impacta en las métricas de rendimiento del modelo. La configuración con 12 componentes PCA resultó ser la más efectiva, alcanzando un F1-Score de aproximadamente

0.8387, con un tiempo de ejecución relativamente bajo de 0.004197 segundos.

Sin embargo, no siempre más componentes significa mejor desempeño. Por ejemplo, con todos los 13 componentes, el rendimiento decrece y se acompaña de un tiempo de ejecución más alto. Esto puede indicar un punto de rendimientos decrecientes, donde aumentos adicionales en la cantidad de componentes no traducen en mejoras de rendimiento y sólo contribuyen a una mayor carga computacional.

En resumen, la configuración óptima en este escenario involucra utilizar 12 de los 13 componentes disponibles, ofreciendo un buen equilibrio entre rendimiento del modelo y eficiencia computacional.

La comparación de los resultados del modelo SVM con y sin la aplicación de PCA (reducción de dimensionalidad a 12 componentes) ofrece una perspectiva interesante sobre la efectividad de la extracción de características. Al observar los resultados, el modelo con 12 componentes PCA logra un F1-Score de 0.8387, mientras que el modelo sin extracción de características presenta un F1-Score ligeramente inferior de 0.8382. Aunque la diferencia en el rendimiento en términos de F1-Score es mínima, destaca la eficacia de la PCA en la conservación de información relevante con menos variables.

Un aspecto notable es el tiempo de ejecución, donde el modelo SVM sin PCA tiene un tiempo de 0.003508 segundos, que es comparable al tiempo de 0.004197 segundos para el modelo con PCA y 12 componentes. Este aumento en el tiempo de ejecución con PCA es el costo de transformar los datos, lo cual es un factor a considerar cuando la eficiencia de tiempo es crucial.

### B. Análisis con Feed-Forward Neural Networks

Al evaluar el rendimiento de una red neuronal con una capa de salida sigmoideal y variando el número de componentes principales (PCA) en la entrada, se observan tendencias importantes en el comportamiento del modelo. La configuración que utiliza 11 componentes PCA destaca por alcanzar el mejor equilibrio entre las métricas de precisión, recall y F1-score, con un F1-score de aproximadamente 0.8366. Este resultado sugiere que esta cantidad de componentes es óptima para capturar la información esencial para la clasificación, sin incluir ruido innecesario o información redundante.

Precision	Recall	F1-score	Average Time	PCA Components
0.842326	0.837957	0.836641	7.400122	11
0.834121	0.828065	0.826783	3.150169	5
0.832979	0.827849	0.826332	6.082858	7
0.833238	0.827742	0.825686	2.968058	12
0.824657	0.821183	0.819319	2.669995	9
0.831639	0.821075	0.818120	3.679107	8
0.825071	0.818280	0.816478	2.582938	10
0.823565	0.817742	0.815861	4.868959	6
0.824471	0.817849	0.815716	7.546278	3
0.817071	0.811398	0.809406	2.932113	4
0.816612	0.808065	0.806378	2.340835	2
0.808805	0.801935	0.800671	2.335832	13
0.794113	0.788065	0.786398	1.973515	1

Tabla X. Resultados del mejor modelo RNA con extracción de características.

El impacto del número de componentes en el rendimiento es considerable. Con 11 componentes, se logran los valores más altos en todas las métricas evaluadas, lo que indica una captura efectiva de características significativas para el proceso de decisión del modelo. En contraste, el uso de un solo componente resulta en el rendimiento más bajo, ya que es insuficiente para encapsular la complejidad de los datos.

En cuanto al tiempo de ejecución, este se incrementa con el número de componentes, alcanzando su punto más alto con 11 componentes a 7.40012 segundos. Esto refleja la carga computacional adicional derivada de hacer la transformación de las características.

Los extremos en la cantidad de componentes muestran resultados menos eficientes. Utilizar todas las 13 componentes disponibles no mejora significativamente las métricas en comparación con usar un número menor pero más estratégico de componentes, como los 11 mencionados. Esto subraya la importancia de una selección cuidadosa del número de componentes PCA para mejorar la eficiencia y efectividad del modelo, facilitando una generalización robusta mientras se optimiza el uso de recursos computacionales.

Al analizar los resultados de la Red Neuronal con una capa de salida sigmoideal, utilizando 11 componentes de PCA, se observa un F1-score de 0.836641. Este rendimiento es inferior al F1-score de 0.839946 alcanzado por el modelo sin utilizar extracción de características. La ligera disminución en el rendimiento puede indicar que, aunque la PCA ayuda a reducir la dimensionalidad y enfocar el modelo en características más relevantes, la eliminación de algunas dimensiones podría estar excluyendo información útil para la clasificación en este caso específico.

Es notable también que el tiempo de ejecución para el modelo con 11 componentes PCA es significativamente más alto (7.40012 segundos) comparado con el modelo sin PCA (1.775288 segundos). Esto podría reflejar el costo adicional en términos de cálculo y tiempo que implica la transformación de PCA antes del entrenamiento del modelo.



En este contexto, aunque la PCA es una herramienta poderosa para manejar la alta dimensionalidad y mejorar la eficiencia de los modelos en algunos casos, su aplicación debe ser evaluada cuidadosamente dependiendo de la estructura y naturaleza de los datos específicos, asegurando que el beneficio en la reducción de dimensionalidad no se haga a expensas de perder información crítica para la tarea de clasificación.

### C. Análisis con Random Forest

En el análisis de Random Forest aplicado junto con la extracción de características por PCA, se observa cómo varía el rendimiento del modelo dependiendo del número de componentes principales utilizados. La tabla muestra que, en general, la precisión y el recall fluctúan ligeramente en función del número de componentes PCA, lo que implica una variabilidad moderada en cómo la reducción de dimensionalidad afecta la capacidad del modelo para clasificar correctamente las observaciones.

Modelo	Precisión	Recall	F1-score
Random Forest with PCA 8	0.803969	0.791398	0.784725
Random Forest with PCA 1	0.809778	0.791505	0.784131
Random Forest with PCA 4	0.808628	0.791290	0.784011
Random Forest with PCA 2	0.810917	0.791505	0.783721
Random Forest with PCA 10	0.810926	0.791398	0.783681
Random Forest with PCA 11	0.804708	0.788280	0.781513
Random Forest with PCA 6	0.812859	0.788065	0.777764
Random Forest with PCA 5	0.802074	0.784731	0.776731
Random Forest with PCA 9	0.800060	0.781505	0.771732
Random Forest with PCA 7	0.793962	0.778065	0.769864
Random Forest with PCA 12	0.794078	0.778280	0.768794
Random Forest with PCA 3	0.784964	0.768172	0.761095
Random Forest with PCA 13	0.778150	0.765161	0.758525

Tabla XI. Resultados del Random Forest con 30 árboles y profundidad= 2 con extracción de características.

Los resultados indican que con 8 componentes PCA, el modelo logra mayor F1-score de 0.7847, que es relativamente alto comparado con otras configuraciones. Sin embargo, cuando la cantidad de componentes es mínima (por ejemplo, 1 y 2 componentes PCA), el F1-score es casi constante, lo que indica que un número excesivo de componentes puede no proporcionar beneficios adicionales significativos y puede reflejar sobreajuste o ruido en los datos.

Al comparar el desempeño del modelo Random Forest sin PCA frente al mejor modelo con PCA (utilizando 8 componentes), se observan diferencias significativas en los resultados. Sin realizar extracción de características, el Random Forest alcanza un F1-score de 0.8459, con un tiempo de ejecución de 0.02915 segundos.

Por otro lado, con la extracción de características mediante PCA utilizando 8 componentes, el modelo logra una precisión

de 0.8039, un recall de 0.7914 y un F1-score de 0.7847, con un tiempo de ejecución de 0.003637 segundos. Este último escenario muestra una reducción en todas las métricas de desempeño, pero con una mejora significativa en el tiempo de ejecución.

## VII. ANÁLISIS DE RESULTADOS

En la evaluación del desempeño de diversos modelos y técnicas aplicadas a la predicción de enfermedades cardíacas, se observan variaciones significativas en las métricas de precisión, recall y F1-Score, que son indicativos esenciales del rendimiento de los modelos de clasificación.

Los estudios previos como HDPM (2020) [3] y Hybrid CNN-LSTM (2023) [4] se destacan por lograr puntuaciones F1 superiores al 98%, lo que refleja un equilibrio excelente entre precisión y sensibilidad y una capacidad destacada para generalizar efectivamente sobre datos no vistos. Estos modelos integran técnicas de aprendizaje profundo y combinan métodos que explotan tanto características temporales como espaciales de los datos, demostrando la eficacia de las metodologías avanzadas en contextos complejos.

	Study Reference	Accuracy	Precision	Recall	F1-Score	Specificity
0	[3] HDPM (2020)	98.40%	98.57%	98.33%	98.32%	N/A
1	[6] MIFH (2020)	93.44%	N/A	89.28%	N/A	96.96%
2	[4] Hybrid CNN-LSTM (2023)	97.75%	99.13%	98.74%	99.83%	99.42%
3	[5] HRFLM (2019)	88.47%	87.5%	92.8%	90.0%	82.6%
4	Current work QDA reg_param=0.3	N/A	84.35%	83.52%	83.33%	N/A
5	Current work KNN n_neighbors=11	N/A	84.82%	83.49%	83.25%	N/A
6	RF n=30, depth=2, criterion=entropy	N/A	85.87%	84.83%	84.59%	N/A
7	Current work Neural Network Neurons=10, Epochs...	N/A	84.75%	84.14%	83.99%	N/A
8	Current work Neural Network Neurons=50, Epochs...	N/A	84.79%	83.81%	83.62%	N/A
9	Current work SVM C=300, gamma=0.001	N/A	85.55%	84.16%	83.82%	N/A
10	Current work Random Forest with Selected Features	N/A	85.87%	84.82%	84.59%	N/A
11	Current work Neural Network with Selected Feat...	N/A	81.67%	81.07%	80.79%	N/A
12	Current work SVM with Selected Features	N/A	85.52%	84.16%	83.82%	N/A
13	Current work SVM with PCA, 12 Components	N/A	84.86%	84.10%	83.87%	N/A
14	Current work Neural Network with 11 Components	N/A	84.23%	83.80%	83.66%	N/A
15	Current work Random Forest with 8 Components	N/A	80.40%	79.14%	78.47%	N/A

Tabla XII. Comparación del rendimiento de los modelos de aprendizaje propuestos frente al trabajo existente.

Comparando esto con el rendimiento de técnicas más convencionales utilizadas en el trabajo actual, como SVM, RNA, Random Forest sin reducción de características, los F1-Scores obtenidos rondan el 83.82%, 83.99% y el 84.59% respectivamente. Estos resultados, si bien no alcanzan los máximos obtenidos con métodos más sofisticados, son robustos y validan la eficacia de los enfoques tradicionales para tareas de clasificación.

La implementación de PCA para la reducción de la dimensionalidad antes de aplicar modelos como SVM y Random Forest mostró una leve disminución en el desempeño en comparación con los modelos que utilizan todas las características. Por ejemplo, el Random Forest con 8 componentes PCA mostró un descenso en el F1-Score,

sugiriendo que la reducción de dimensiones, aunque beneficia en la simplificación del modelo y la reducción del tiempo computacional, puede conllevar la pérdida de información crítica para la clasificación.

En el contexto de aplicaciones médicas, podría ser preferible evitar la extracción de características, ya que la interpretabilidad es fundamental en este campo. Los modelos de aprendizaje automático en el ámbito médico no solo deben ser precisos en sus predicciones, sino también transparentes y explicables. Los profesionales de la salud necesitan comprender los fundamentos de las decisiones tomadas por los modelos para poder confiar en ellos y utilizarlos de manera efectiva en el diagnóstico y el tratamiento de pacientes. Aunque las técnicas de extracción de características pueden mejorar el rendimiento predictivo de los modelos al eliminar características irrelevantes o redundantes, también pueden dificultar la interpretabilidad de los resultados. Esto se debe a que las características resultantes pueden ser combinaciones lineales o no lineales de las características originales, lo que dificulta la comprensión de su significado clínico.

Los modelos de redes neuronales, evaluados tanto con todas las características como con las seleccionadas a través de PCA, exhiben variabilidad en su eficacia, con F1-Scores que varían significativamente. Esto ilustra cómo la elección del número de componentes PCA y la selección de características afectan directamente la efectividad del modelo.

A diferencia de la reducción de dimensionalidad con PCA, los resultados obtenidos con la selección de características fueron sobresalientes al aplicarse en las mejores configuraciones de los modelos Random Forest y SVM. En ambos casos, el rendimiento se mantuvo constante con las 5 características seleccionadas: Chest Pain Type (Tipo de Dolor de Pecho), Exercise Induced Angina (Angina Inducida por el Ejercicio), ST Depression Induced by Exercise Relative to Rest (Depresión del Segmento ST Inducida por el Ejercicio en Relación al Reposo), Slope of the Peak Exercise ST Segment (Pendiente del Segmento ST en el Pico del Ejercicio) y Number of Major Vessels Colored by Fluoroscopy (Número de Vasos Mayores Coloreados por Fluoroscopia). La selección de características no afectó el desempeño en términos de F1-score. Por lo tanto, se puede afirmar que los dos mejores modelos son Random Forest y SVM con selección de características, obteniendo un F1-score de 84.59% y 83.82% respectivamente. Este enfoque resulta beneficioso, ya que, a pesar de mantener un rendimiento constante, se logra ahorrar en costos de dinero y tiempo asociados a la recolección de las 8 características restantes, lo cual es fundamental desde una perspectiva de negocio.

Por otro lado, para las Redes Neuronales, la selección de características resultó en una disminución del rendimiento. Esto resalta la importancia de evaluar individualmente cada modelo al aplicar técnicas de selección de características

Estos resultados subrayan la importancia de seleccionar adecuadamente las características y las técnicas de reducción de dimensionalidad para optimizar el rendimiento de los modelos de aprendizaje automático. La decisión de utilizar todas las características disponibles frente a un subconjunto reducido debe ponderar cuidadosamente el equilibrio entre precisión, complejidad del modelo y costos computacionales, ajustándose a las necesidades específicas del análisis y las limitaciones inherentes a los datos.

## REFERENCES

- [1] "The top 10 causes of death," Organización Mundial de la Salud, 9 dic. 2020. [En línea]. Disponible en: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>. [Accedido: 1 de mayo de 2024].
- [2] [1] "Heart Disease Risk Factors In Young Adults," Sprint Medical, 19 enero 2023. [Online]. Disponible: <https://www.sprintmedical.in/blogs/heart-and-vascular/heart-disease-risk-factors-in-young-adults>. [Accedido: 1 de mayo de 2024].
- [3] N. L. Fitriyani, M. Syafrudin, G. Alfian and J. Rhee, "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System," in *IEEE Access*, vol. 8, pp. 133034-133050, 2020, doi: 10.1109/ACCESS.2020.3010511.
- [4] M. S. A. Reshan, S. Amin, M. A. Zeb, A. Sulaiman, H. Alshahrani and A. Shaikh, "A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks," in *IEEE Access*, vol. 11, pp. 121574-121591, 2023, doi: 10.1109/ACCESS.2023.3328909.
- [5] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in *IEEE Access*, vol. 7, pp. 81542-81554, 2019, doi: 10.1109/ACCESS.2019.2923707.
- [6] A. Gupta, R. Kumar, H. Singh Arora and B. Raman, "MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis," in *IEEE Access*, vol. 8, pp. 14659-14674, 2020, doi: 10.1109/ACCESS.2019.2962755.
- [7] R. Rahman, "Heart Attack Analysis & Prediction Dataset," 2021. [Online]. Disponible: <https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset/data>. [Accedido: 1 de mayo de 2024].
- [8] Janosi, Andras, Steinbrunn, William, Pfisterer, Matthias, and Detrano, Robert. (1988). Heart Disease Data Set. [Online]. Disponible: <https://archive.ics.uci.edu/dataset/45/heart+disease> [Accedido: 1 de mayo de 2024].