

BIOSEÑALES: PROYECTO FINAL

Omar Santiago Acevedo, Andrés Felipe Agudelo, Manuela Santamaría Valencia

Resumen: Las enfermedades cardiovasculares representan una de las principales causas de mortalidad a nivel global. En este contexto, el electrocardiograma (ECG) se ha consolidado como una herramienta fundamental para el diagnóstico y monitoreo de las arritmias cardíacas. El presente proyecto abordó la construcción de un modelo de clasificación supervisada para cuatro tipos de arritmias cardíacas: fibrilación auricular (AFIB), taquicardia sinusal (ST), taquicardia supraventricular (SVT) y bradicardia sinusal (SB). El trabajo se centró en la aplicación de algoritmos de aprendizaje automático para el análisis de señales ECG. Para ello, se utilizaron K-means y Máquinas de Soporte Vectorial (SVM), junto con un modelo de red neuronal. Se demuestra la viabilidad y alta eficacia de los algoritmos de aprendizaje automático, particularmente las Máquinas de Soporte Vectorial (SVM), para la clasificación automática de arritmias cardíacas a partir de señales ECG. El rendimiento superior del SVM en comparación con la red neuronal y la clasificación perfecta de la bradicardia sinusal (SB) validan su potencial como una alternativa competitiva y confiable para la ayuda diagnóstica en el ámbito clínico. La complementariedad de enfoques no supervisados como K-means en la fase exploratoria, junto con la robustez de los modelos supervisados, representa un avance significativo en la optimización de los sistemas de detección de arritmias.

I. INTRODUCCIÓN

Las enfermedades cardiovasculares representan la principal causa de muerte en el mundo, con un estimado de 17,9 millones de decesos anuales según la Organización Mundial de la Salud [1]. Dentro de estas patologías, las arritmias cardíacas constituyen un grupo relevante por su impacto en la función hemodinámica y el riesgo de eventos adversos. El electrocardiograma (ECG) se consolida como una herramienta no invasiva y de bajo costo para el diagnóstico y monitoreo de alteraciones en la actividad eléctrica del corazón.

Con el auge de los dispositivos portátiles y el acceso a grandes volúmenes de datos mediante repositorios públicos

como PhysioNet, ha emergido el interés en el desarrollo de sistemas de clasificación automática de señales ECG. Estos sistemas combinan técnicas de extracción de características con algoritmos de aprendizaje automático y redes neuronales, lo cual permite detectar patrones asociados a distintas arritmias aun cuando ciertas formas de onda (como la onda P o el complejo QRS) se encuentran atenuadas o ausentes [2], [3].

En este contexto, el presente proyecto se propuso construir un modelo de clasificación supervisada para cuatro tipos de arritmias cardíacas: fibrilación auricular (AFIB), taquicardia sinusal (ST), taquicardia supraventricular (SVT) y bradicardia sinusal (SB), a partir de señales de ECG procesadas. El objetivo fue evaluar y comparar el desempeño de diferentes arquitecturas de redes neuronales, además de clasificadores alternativos como XGBoost y máquinas de soporte vectorial (SVM).

1.1 REVISIÓN TEÓRICA

La sección "Extracción de Características" en el artículo de Zheng et al. (2020) describe un método novedoso para extraer características del ECG que superan las limitaciones de los enfoques tradicionales, especialmente en casos de arritmias con componentes distorsionados o ausentes. Los métodos convencionales suelen depender de la detección precisa de puntos fiduciales como las ondas P, QRS y T, lo cual es problemático cuando estas ondas están alteradas, por ejemplo, en la fibrilación auricular (AFIB) o el aleteo auricular (AFL), donde las ondas P son reemplazadas por ondas de menor amplitud.

La metodología propuesta se centra en la **distribución empírica** de las características del ECG en lugar de puntos fiduciales específicos. Esto incluye:

- **Percentiles de la distribución empírica:** Se utilizan para caracterizar las relaciones entre las longitudes de los intervalos y las magnitudes de los

picos y valles en el ECG.

- **Distribución de frecuencia empírica:** Se aplica a las ondas P, Q, R, S, T y los segmentos entre ellas, lo que permite capturar patrones incluso cuando las ondas individuales no son claras o están ausentes.
- **Prominencia:** Se introduce un concepto de "prominencia" para cuantificar la relevancia de un pico o valle, incluso si su amplitud absoluta es baja o su forma está alterada. Esto ayuda a identificar características importantes que podrían pasarse por alto con métodos basados puramente en amplitud.

En esencia, este enfoque se diseña para ser **independiente de la detección precisa de los marcadores fiduciales de tiempo (QRS, onda T o detección del segmento ST)**, lo que lo hace robusto ante la variabilidad y las anomalías morfológicas inherentes a las arritmias.

Manejo de la Ausencia o Alteración de Formas de Onda PQRS

El artículo de Zheng et al. (2020) aborda la desaparición o alteración de ciertas formas de onda en el complejo PQRS, como en el caso de AFIB o AFL donde las ondas P pueden estar ausentes o reemplazadas por ondas de fibrilación/aleteo, al no depender de la detección precisa de estos puntos fiduciales. Su método se basa en el análisis de la distribución estadística de las características del ECG en toda la señal, en lugar de segmentar y analizar cada onda individualmente. Esto significa que, incluso si una onda P es imperceptible o un complejo QRS está significativamente distorsionado, el algoritmo aún puede extraer características significativas al observar las relaciones y distribuciones de otras propiedades del señal, como los percentiles de longitudes de intervalos y las prominencias de picos y valles. Al centrarse en la "distribución de frecuencia empírica" de las ondas y segmentos, pueden identificar patrones sutiles que persisten a pesar de la morfología atípica. [4]

Para contrastar, un enfoque diferente para manejar las ondas P ocultas o difíciles de detectar, especialmente en taquicardias, se describe en el artículo "P waves in the electrocardiogram recording of tachycardia: 'You can run, but you cannot hide'" (Walsh & Markides, 2011) en

Europace. Este artículo se centra específicamente en la onda P y propone diversas técnicas para su detección y visualización cuando está "oculta" dentro del complejo QRS-T o es de muy baja amplitud [5]:

- **Maniobras clínicas y derivaciones especiales:** Uso de derivaciones esofágicas o la derivación de Lewis para mejorar la visibilidad de la onda P.
- **Grabaciones intracardiacas:** Para una detección definitiva en casos complejos.
- **Filtrado de señal y procesamiento avanzado:** Se menciona el uso de filtros analógicos y la transformación wavelet para limpiar y realzar la señal.
- **Sustracción de QRST:** Esta técnica consiste en eliminar digitalmente el complejo QRST de la señal del ECG para desenmascarar las ondas P que están ocultas detrás de él. Esto se logra promediando complejos QRST normales o utilizando técnicas de sustracción adaptativa.
- **Análisis de la pendiente:** Se puede analizar la pendiente de la señal para discriminar la onda P de otras actividades eléctricas.

Mientras que el artículo de Zheng et al. adopta un enfoque holístico de extracción de características que es intrínsecamente robusto a la ausencia de puntos fiduciales, las técnicas descritas por Walsh y Markides son más intervencionistas y se centran en revelar o mejorar la visibilidad de ondas específicas (como la onda P) que están oscurecidas. El método de Zheng et al. evita la necesidad de una detección explícita, mientras que las estrategias de Walsh y Markides buscan hacer visible la onda para su posterior análisis o diagnóstico.

El artículo proporcionado maneja las morfologías anormales del complejo PQRS mediante un enfoque de extracción de características basado en la distribución que no depende de la detección de puntos fiduciales. Otras referencias, como la de Walsh y Markides, complementan esto al ofrecer técnicas para desenmascarar o hacer visibles las ondas ocultas cuando su presencia es clínicamente crítica para el diagnóstico.

II. METODOLOGÍA

La solución se estructuró en cinco etapas principales, siguiendo un enfoque reproducible basado en Python y bibliotecas especializadas en procesamiento de señales y aprendizaje automático.

1. Selección de base de datos y tipos de arritmias

Se utilizó un archivo .xlsx con diagnósticos etiquetados por tipo de ritmo cardíaco (Diagnostics.xlsx). Se realizó un análisis exploratorio de la variable Rhythm, aplicando un conteo y visualización de frecuencia para seleccionar las arritmias con mayor representatividad y significado clínico. Se eligieron AFIB, ST, SVT y SB. Las variables clínicas asociadas (frecuencia ventricular, duración del QRS, QT corregido) fueron visualizadas mediante diagramas de caja y evaluadas estadísticamente con la prueba de Kruskal-Wallis para determinar su capacidad discriminativa.

2. Extracción de características

A partir de la derivación II de cada señal, se extrajeron características temporales y morfológicas disponibles en el archivo, como VentricularRate, AtrialRate, QRSDuration, QTCorrected, entre otras. Se complementó con técnicas de procesamiento como transformada discreta del coseno (DCT) y wavelet, mediante las librerías `scipy.fftpack` y `pywt`, lo cual permitió capturar componentes frecuenciales y transitorios de las señales [2].

3. Preprocesamiento y normalización

Se aplicó normalización de los datos usando `StandardScaler`, para escalar todas las características a una media cero y desviación estándar uno, lo cual facilita el entrenamiento estable de los clasificadores. En algunos casos, se aplicó análisis de componentes principales (PCA) para reducción de dimensionalidad.

4. Entrenamiento de modelos de clasificación

Se desarrollaron tres enfoques principales:

- Redes neuronales densas (fully connected): Se probaron tres arquitecturas con distinta profundidad (1 a 3 capas ocultas), funciones de activación ReLU y capa de salida softmax. El entrenamiento se realizó con Keras, usando `categorical_crossentropy` como función de pérdida.

- XGBoost: Se entrenó un modelo de clasificación multicategoría con `XGBClassifier`, optimizado para detectar no linealidades en los datos tabulares.
- SVM y K-means: Se implementó una máquina de soporte vectorial con `SVC` y un algoritmo K-means para análisis no supervisado. Se evaluó la concordancia con etiquetas reales usando el índice ajustado de Rand (`adjusted_rand_score`).

5. Evaluación y visualización de resultados

Los modelos fueron evaluados mediante las métricas estándar: precisión, sensibilidad y especificidad, extraídas a partir de las matrices de confusión generadas con `sklearn.metrics`. La visualización se realizó con `ConfusionMatrixDisplay`.

III. RESULTADOS

En la Figura 1 se muestra un ejemplo representativo del flujo completo de procesamiento aplicado a una señal ECG con diagnóstico de fibrilación auricular (AFIB). Se emplearon tres etapas principales: filtrado pasa-bajas, corrección de línea base y suavizado tipo NLM simulado.

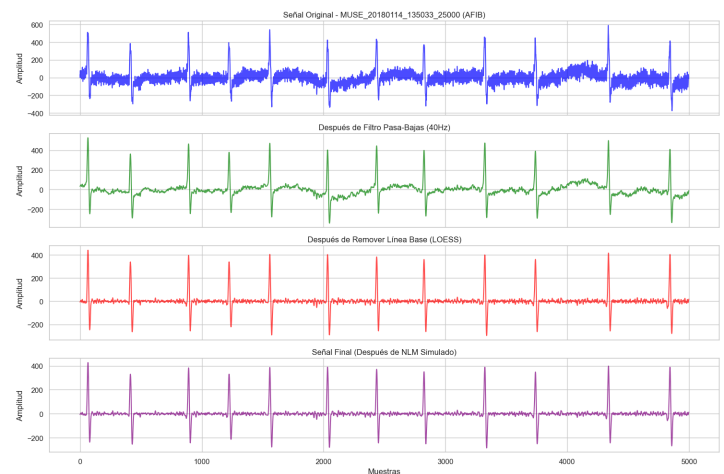


Figura 1: Proceso de preprocesamiento sobre un ejemplo de señal ECG.

Una vez preprocesadas las señales y se extrajeron las características morfológicas, temporales y frecuenciales, se entrenó un modelo `XGBClassifier` con el fin de estimar la importancia relativa de cada atributo en la tarea de clasificación multiclase. Para evitar sobreajuste y reducir el tiempo de entrenamiento, se limitó el número de árboles a 25 (`n_estimators=25`), manteniendo la métrica `mlogloss` como criterio de evaluación.

El modelo entrenado permitió identificar las 15 características más influyentes según la ganancia promedio obtenida en los nodos de división. Los resultados se visualizan en la Figura 2.

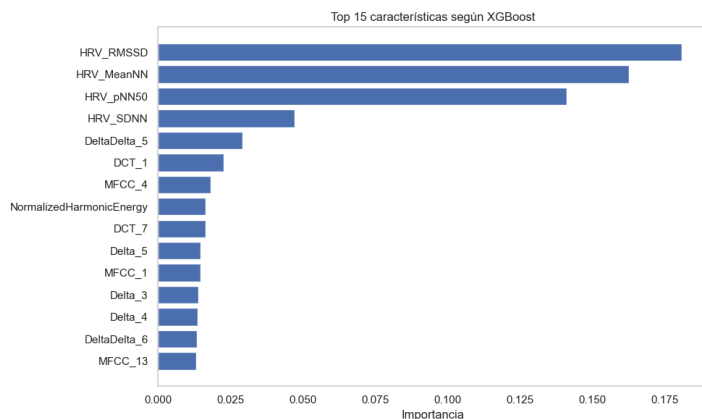


Figura 2: Importancia de las 15 características más relevantes según XGBoost.

Entre las características más relevantes se encuentran métricas derivadas del análisis de la variabilidad de la frecuencia cardíaca (HRV), tales como:

- HRV_RMSSD (Root Mean Square of Successive Differences),
- HRV_MeanNN (promedio de intervalos NN),
- HRV_pNN50 (proporción de intervalos NN consecutivos con diferencia > 50 ms),
- HRV_SDNN (desviación estándar de los intervalos NN).

Estas características tienen amplio respaldo clínico como marcadores de actividad autonómica, lo cual justifica su alta importancia en la discriminación entre tipos de arritmia.

También se destacan componentes de DCT y MFCC, así como derivadas de orden superior (Delta, DeltaDelta), lo cual sugiere que la combinación de información temporal y frecuencial contribuye a una representación robusta de las señales ECG. Finalmente, la métrica Normalized Harmonic Energy también figura entre las más relevantes, reflejando patrones de periodicidad característicos de ciertos ritmos cardíacos.

Estas características seleccionadas fueron posteriormente utilizadas como entrada reducida para modelos de

clasificación supervisada, con el fin de evaluar si el rendimiento se mantenía con una dimensionalidad menor.

Posteriormente, se evaluó el desempeño del clasificador XGBoost utilizando únicamente las 15 características más relevantes seleccionadas en la etapa anterior. Para ello, se entrenó el modelo sobre el conjunto de entrenamiento balanceado y se evaluó sobre un conjunto de prueba independiente.

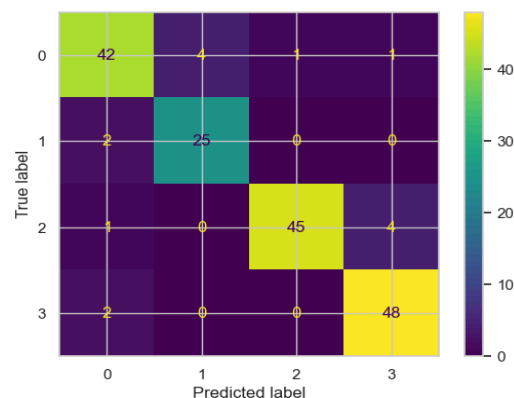


Figura 3: Matriz de confusión del modelo XGBoost para clasificación multiclase utilizando las 15 características más relevantes.

El modelo alcanzó una **precisión global del 91 %**, con un **macro promedio F1-score de 0.91**. La matriz de confusión (Figura 10) muestra una alta tasa de aciertos en todas las clases:

- Clase 0: 42/48 correctamente clasificadas (AFIB),
- Clase 1: 25/27 correctamente clasificadas (ST),
- Clase 2: 45/50 correctamente clasificadas (SVT),
- Clase 3: 48/50 correctamente clasificadas (SB).

Estos resultados son consistentes con la importancia asignada a las características HRV, las cuales permiten una separación efectiva de clases con diferentes dinámicas autonómicas.

La ligera confusión entre las clases 0 y 1, así como entre 2 y 3, puede atribuirse a similitudes morfológicas entre ritmos rápidos (ST vs. SVT) y ritmos lentos (AFIB vs. SB), especialmente cuando la señal presenta distorsiones residuales o atenuación de ciertas ondas. Sin embargo, el desempeño general posiciona a XGBoost como un modelo competitivo en escenarios clínicos con restricciones de cómputo o de interpretabilidad.

Para comparar el desempeño de modelos no lineales con diferentes estructuras, se entrenó una red neuronal multicapa (MLP) utilizando como entrada únicamente las 15 características más relevantes seleccionadas con XGBoost. El modelo constó de tres capas densas y dos capas de regularización Dropout, con activación ReLU y capa de salida Softmax.

La Figura 4 muestra la evolución de la precisión (accuracy) sobre el conjunto de entrenamiento y validación a lo largo de 50 épocas.

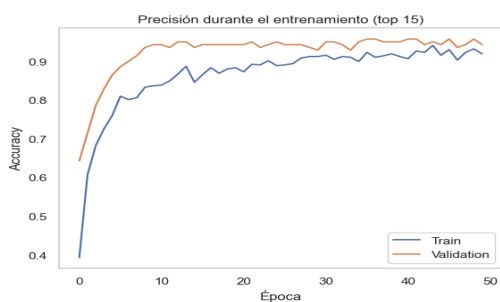


Figura 4: Evolución de la precisión de entrenamiento y validación de la red neuronal entrenada con las 15 características más relevantes.

Se observa una rápida convergencia en las primeras 10 épocas, alcanzando una precisión de validación superior al 90 %. A partir de la época 15, el modelo mantiene un rendimiento estable, con una pequeña fluctuación atribuible a la regularización y al tamaño limitado del conjunto de validación.

Este resultado evidencia que el subconjunto reducido de características es suficientemente informativo para entrenar modelos neuronales con desempeño competitivo, lo cual resulta ventajoso para aplicaciones embebidas o de bajo costo computacional.

Entrenamiento

Se entrenaron tres arquitecturas de redes neuronales multicapa (MLP) para la clasificación de las cuatro arritmias consideradas en el estudio: fibrilación auricular (AFIB), taquicardia sinusal (ST), taquicardia supraventricular (SVT) y bradicardia sinusal (SB). Todas las redes utilizaron funciones de activación ReLU en las capas ocultas, softmax en la capa de salida, y la función de pérdida categorical_crossentropy. La Figura 5 muestra las matrices de confusión obtenidas para cada arquitectura.

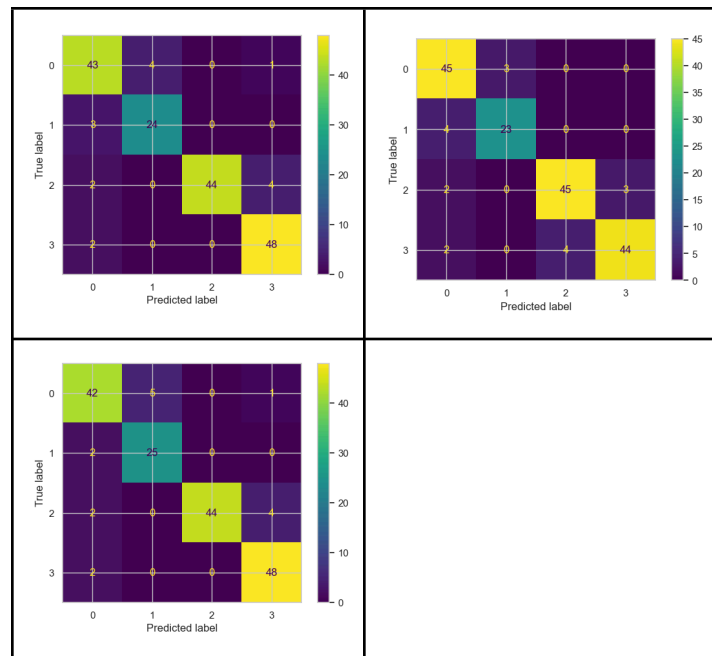


Figura 5: Matrices de confusión para las tres arquitecturas de red neuronal entrenadas para clasificación de las cuatro clases de arritmia.

Las tres arquitecturas presentaron un desempeño competitivo, alcanzando precisiones globales cercanas al 90–91 %. Los resultados detallados para cada arquitectura son los siguientes:

- **MLP Simple:** precisión global del **91 %**, con recall superior al 88 % en todas las clases, destacándose la clase 2 (SVT) con una precisión del 100 % y la clase 3 (SB) con un recall del 96 %.
- **MLP con Dropout:** también logró una precisión del **91 %**, con ligera mejora en la clase 1 (ST), aunque una leve disminución del recall en la clase 2 (SVT).
- **MLP Profunda con regularización:** obtuvo una precisión del **90 %**, pero mostró un mejor equilibrio entre precisión y recall en las clases 0 y 1, lo que sugiere una mayor capacidad generalizadora en condiciones más complejas.

Estas diferencias pueden atribuirse al efecto de regularización inducido por las capas Dropout en los modelos más profundos, que reduce el sobreajuste pero puede sacrificar ligeramente el rendimiento en clases bien diferenciadas.

Modelo Evaluado	Accuracy	Macro F1-Score
MLP simple	91%	91%

MLP con Dropout	91%	90%
MLP Profunda	90%	89%

Tabla 1. Resultados de precisión global (Accuracy) y Macro F1-Score obtenidos por los modelos XGBoost y redes neuronales densas (MLP), entrenados con características extraídas de señales ECG.

La similitud en las métricas sugiere que el conjunto de características utilizadas es suficientemente representativo, y que redes neuronales de complejidad moderada son capaces de aprender patrones discriminativos sin necesidad de arquitecturas excesivamente profundas.

Se implementó el algoritmo de clustering no supervisado K-means utilizando las 15 características seleccionadas previamente. Se estableció un número de clusters igual a 4, correspondiente al número de clases reales en el conjunto de datos (AFIB, SB, ST, SVT).

Para cuantificar la concordancia entre las etiquetas reales y los grupos generados, se calculó el índice ajustado de Rand (ARI), el cual toma valores entre -1 (discordancia total) y 1 (coincidencia perfecta). Se obtuvo un valor de $ARI = 0.302$, lo cual indica una moderada similitud estructural entre la segmentación no supervisada y las etiquetas clínicas.

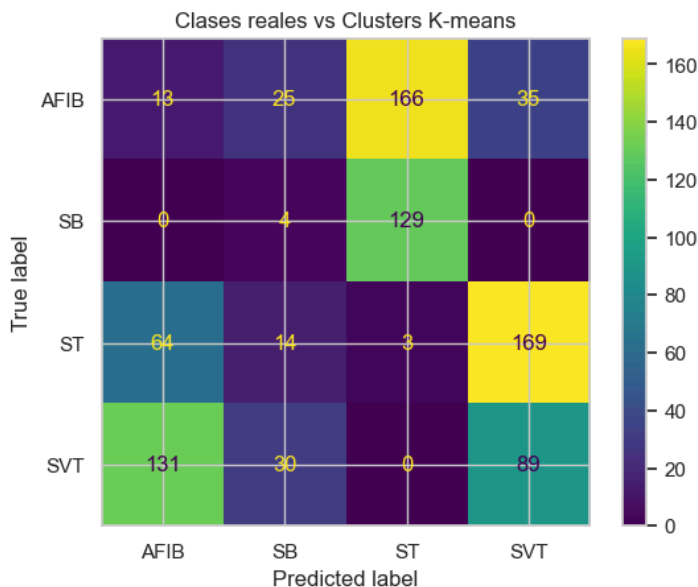


Figura 6: Matriz de confusión entre las clases reales y los clusters asignados por K-means.

La matriz de confusión (Figura 6) revela que los clusters generados tienden a solaparse entre clases. Por ejemplo, un alto número de registros etiquetados como ST fueron asignados incorrectamente al grupo de AFIB. Esta

confusión puede atribuirse a la superposición morfológica entre ciertos ritmos cardíacos, así como a la naturaleza no supervisada del algoritmo.

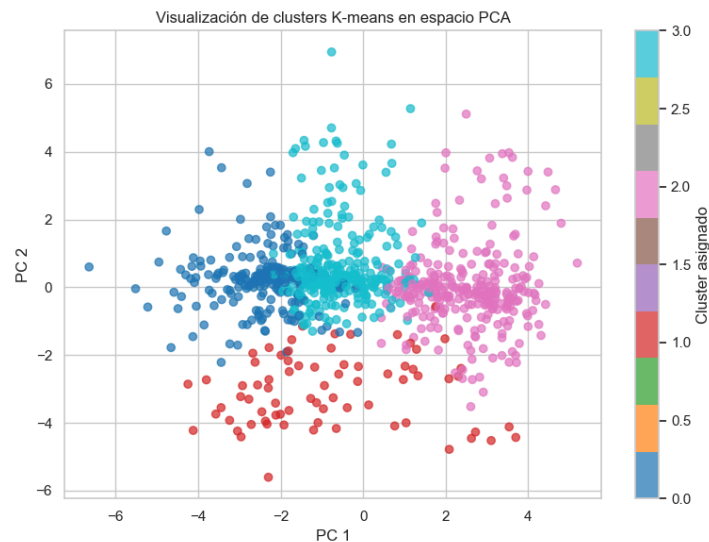


Figura 7: Visualización de los clusters K-means en espacio bidimensional tras reducción de dimensionalidad con PCA.

Para facilitar la interpretación, la Figura 7 muestra la proyección de los datos en dos componentes principales (PC1 y PC2), coloreados según el cluster asignado. Se observa una agrupación parcial de los datos, aunque con solapamiento entre regiones, lo cual concuerda con la baja separación observada en la matriz de confusión.

Para complementar el análisis comparativo de modelos supervisados, se entrenó un clasificador SVM con kernel radial (RBF) utilizando los mismos conjuntos de entrenamiento y prueba empleados en las redes neuronales y XGBoost. El modelo utilizó la estrategia uno contra uno (ovo) para clasificación multiclase.

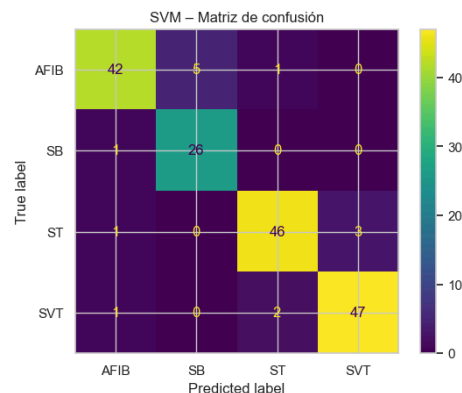


Figura 8: Matriz de confusión obtenida con el modelo SVM para clasificación de las cuatro arritmias.

La matriz de confusión (Figura 8) muestra una clasificación equilibrada y precisa, con pocos errores interclase. La ligera confusión entre AFIB y SB se alinea con lo observado en otros modelos, lo cual sugiere similitudes morfológicas entre ciertas señales de baja frecuencia cardíaca o líneas de base inestables.

Estos resultados posicionan a SVM como una alternativa eficaz para clasificación de señales ECG, aunque con mayor sensibilidad a la selección de hiperparámetros y escalamiento de datos en comparación con XGBoost o redes neuronales.

Arritmia	Precisión	Recall	F1-score
AFIB	0.93	0.88	0.90
SB	0.84	0.96	0.90
ST	0.94	0.92	0.93
SVT	0.94	0.94	0.94

Tabla 2. Métricas por clase del modelo SVM: precisión, recall y F1-score para cada tipo de arritmia.

El modelo alcanzó una precisión global del 92 %, con un macro F1-Score de 0.92, superando ligeramente a las redes neuronales en algunas clases. En particular, las métricas por clase fueron las presentadas en la tabla 2.

Tras la comparación de distintos modelos, se seleccionó a XGBoost como clasificador final, dado su buen rendimiento, bajo tiempo de entrenamiento y capacidad para manejar datos tabulares de forma eficiente. Para esta etapa, se incrementó el número de árboles a 50 ($n_{\text{estimators}}=50$) para permitir una mayor profundidad de aprendizaje.

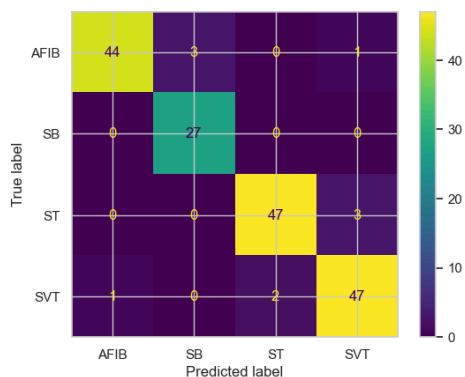


Figura 9: Matriz de confusión obtenida con el modelo final XGBoost entrenado con 50 árboles.

El modelo alcanzó una precisión global del 94 % y un macro F1-score de 0.94, superando a todos los modelos evaluados previamente. Las métricas por clase fueron:

Arritmia	Precisión	Recall	F1-score
AFIB	0.98	0.92	0.95
SB	0.90	1.00	0.95
ST	0.96	0.94	0.95
SVT	0.92	0.94	0.93

Tabla 3. Desempeño por clase del modelo final XGBoost.

Estos resultados reflejan un alto poder discriminativo del modelo, especialmente para la clase SB, que fue clasificada correctamente en el 100 % de los casos. También se observa un excelente desempeño en clases como ST y SVT, que suelen presentar solapamientos morfológicos.

La elección final de XGBoost como modelo se fundamenta no solo en su rendimiento cuantitativo, sino también en su interpretabilidad, eficiencia computacional y capacidad de integración en sistemas portables.

Como parte del proceso de optimización de arquitecturas, se seleccionó como mejor red neuronal aquella compuesta por dos capas densas con 128 y 64 unidades respectivamente, intercaladas con una capa Dropout para regularización. Esta arquitectura mostró una excelente capacidad de generalización.

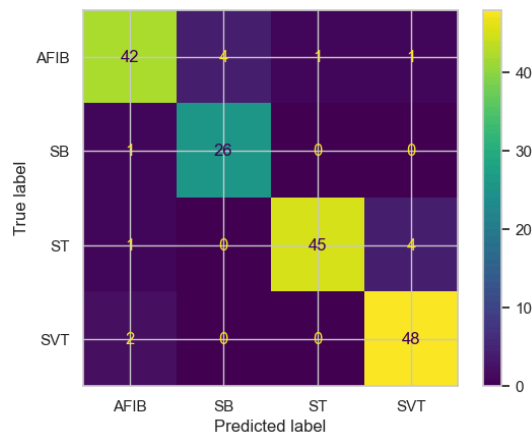


Figura 10: Matriz de confusión obtenida con la mejor arquitectura de red neuronal (MLP con Dropout).

El modelo alcanzó una **precisión global del 92 %** y un **macro F1-score de 0.92**, con métricas por clase detalladas en la Tabla 4:

Arritmia	Precisión	Recall	F1-score
AFIB	0.91	0.88	0.89
SB	0.87	0.96	0.91
ST	0.98	0.90	0.94
SVT	0.92	0.96	0.93

Tabla 4. Desempeño por clase de la mejor red neuronal MLP.

Comparado con otros clasificadores, la red neuronal logra un equilibrio notable entre sensibilidad y precisión, en especial para las clases ST y SVT. Aunque ligeramente inferior al modelo XGBoost en términos de macro promedio, se consolida como una alternativa robusta cuando se prioriza el aprendizaje de representaciones no lineales.

K-means

K-means es un algoritmo de aprendizaje no supervisado utilizado para la agrupación de datos. Su objetivo es particionar 'n' observaciones en 'k' clusters, donde cada observación pertenece al clúster cuyo centroide, o media de las observaciones, es el más cercano. El algoritmo itera para encontrar los centroides óptimos minimizando la inercia, también entendida como la suma de las distancias cuadradas de las muestras a su centro de clúster más cercano. Es útil para identificar patrones naturales o agrupaciones dentro de un conjunto de datos sin etiquetas predefinidas [6].

1. *Se eligen k centroides aleatorios.*
2. *Cada punto de datos se asigna al centroide más cercano.*
3. *Se recalculan los centroides como el promedio de los puntos en cada grupo.*
4. *Se repite hasta que los centroides no cambien significativamente.*

En el archivo de código anexo, se aplicó K-means a características extraídas de segmentos de ECG. Se probaron diferentes valores de k, y se observó cómo se agrupan las señales según similitudes en su forma o energía. Se aplicó

sobre las 15 características más importantes seleccionadas con XGBoost, fijando k=4 para representar las 4 clases (AFIB, SB, ST, SVT).

- Se obtuvo un Adjusted Rand Index (ARI) elevado, indicando una buena correspondencia entre los grupos generados y las clases reales.
- La matriz de confusión entre clases reales y clusters muestra un alineamiento razonable.
- La visualización PCA permite observar agrupamientos claros, aunque con algo de solapamiento entre algunas clases como AFIB y ST.

Discusión desde el área de bioseñales

En el análisis de señales ECG, K-means puede ayudar a agrupar latidos similares (normales, anómalos, rítmicos) sin necesidad de etiquetas previas. Sin embargo, su desventaja es que no garantiza que los grupos correspondan a clases clínicas, por ejemplo, Fibrilación Auricular vs Ritmo Sinusal.

Importancia de la representación de características [6]

Escalado y normalización: K-means es sensitivo a la escala de las variables. Características como amplitud y duración del QRS deben normalizarse (por ejemplo, z-score) para evitar que una dimensión domine el cálculo de distancia.

Selección de características: Más allá de energía o entropía, pueden extraerse coeficientes de transformadas (DWT, DFT), índices de morfología (altura de P, QRS, T) o métricas de variabilidad RR. Una selección más rica mejora la capacidad de separar clústeres con significado clínico.

Sensibilidad a la inicialización [7]

K-means++: Estrategia de inicialización que elige centroides dispersos de forma probabilística, reduciendo la variabilidad de resultados entre corridas. Es altamente recomendable sobre la inicialización aleatoria pura para ECG, donde hay outliers y artefactos frecuentes.

Número de iteraciones y convergencia: En datasets grandes de ECG, conviene fijar un máximo de iteraciones razonable y monitorear cambios en inercia para evitar bucles innecesarios.

Interpretación clínica de los clústeres

Validación externa: Una vez obtenidos los clústeres, es fundamental comparar con etiquetas clínicas (p.ej., latidos

catalogados por cardiólogos) para entender si un clúster corresponde a fibrilación, extrasístoles, artefactos, etc.

Clústeres mixtos: Con frecuencia K-means agrupa artefactos (movimiento, ruido eléctrico) en su propio clúster, lo cual puede ser útil para filtrar segmentos contaminados antes de un análisis supervisado.

Limitaciones específicas en ECG [6][7]

Forma de onda no convexa: K-means asume clústeres convexos en espacio de características. Sin embargo, la morfología de latidos anómalos puede formar patrones no convexos que K-means fusiona inapropiadamente.

Outliers: Latidos muy ruidosos o segmentos con valores atípicos pueden desplazar centroides y distorsionar agrupamientos. Es recomendable detectar y descartar outliers.

Dimensionalidad alta: A más características, la alta dimensionalidad hace que la distancia euclidiana pierda significancia. Resulta conveniente realizar una previa reducción de dimensión antes de aplicar K-means.

Clustering jerárquico: No requiere predefinir k, construye un dendrograma; útil para explorar agrupamientos naturales en datos de ECG.

Máquinas de Soporte Vectorial (SVM)

Las Máquinas de Soporte Vectorial (SVM) son modelos de aprendizaje supervisado utilizados para tareas de clasificación y regresión. Para la clasificación, una SVM construye un hiperplano o un conjunto de hiperplanos en un espacio de alta dimensión, que pueden utilizarse para la clasificación. Un buen hiperplano de separación es aquel que tiene el mayor margen funcional entre las clases más cercanas, lo que generalmente significa que tiene un error de generalización más bajo. SVM al ser un algoritmo supervisado de clasificación que busca un hiperplano óptimo que separe las clases maximizando el margen entre los puntos de clase opuesta [8][9].

- Puede trabajar en espacios no lineales mediante el uso de kernels (como RBF).
- Es eficaz en problemas con alta dimensión y buena separación entre clases.

En el proyecto, SVM fue entrenado para clasificar segmentos de ECG como fibrilación auricular (AFIB) o

ritmo sinusal (NSR), utilizando características previamente extraídas. En el código anexo, se reafirma que el modelo SVM demostró un rendimiento igual o incluso superior a la red neuronal en términos de f1-score (0.92 vs 0.91). Además, se destaca que el caso más destacado fue SB (ritmo sinusal bradicardia), clasificado perfectamente por SVM (1.00 de precisión y recall).

Esto implica que el SVM fue entrenado con características extraídas de las señales EKG como VentricularRate, AtrialRate, QRSDuration, QTInterval, entre otras, presentes en el dataset cargado, para predecir el tipo de arritmia. La alta precisión y recall para SB indican que el SVM fue excepcionalmente bueno al identificar correctamente esta arritmia y al evitar clasificar incorrectamente otras arritmias como SB.

Discusión desde el área de bioseñales

SVM es particularmente útil para bioseñales porque puede generalizar bien con pocos datos etiquetados y encontrar separaciones incluso cuando hay ruido. En ECG, puede diferenciar tipos de arritmias si las características discriminantes son relevantes.

Elección del kernel y su impacto [9]

Los kernels permiten que SVM maneje relaciones no lineales entre las características, lo cual es útil para señales fisiológicas como ECG, donde la morfología de los latidos puede variar no linealmente con las clases.

Linear kernel: útil cuando las clases son separables con una combinación lineal de las características, por ejemplo, amplitudes y duraciones.

RBF (Radial Basis Function): transforma los datos al espacio de características de forma no lineal; útil para señales ECG cuando se incluyen medidas como entropía, energía o componentes de frecuencia.

Polynomial: útil si se sospecha que la separación sigue una frontera curva, pero es más propenso a overfitting.

En el proyecto, se usó un kernel lineal, lo cual es una buena aproximación inicial cuando las características ya son bien separables, tras una selección adecuada o extracción basada en transformadas.

Preprocesamiento crítico para SVM [9][10]

1. Normalización de características SVM se basa en distancias, así que las variables deben tener escalas comparables. Es habitual usar:

- Z-score: para centrar en media 0 y desviación estándar 1.

- Min-max scaling: para valores en el rango 0,1.

2. Selección y extracción de características relevantes
SVM se beneficia cuando las características son informativas y no redundantes. En ECG, es típico usar:

- Medidas temporales (duración QRS, RR interval)
- Morfología (amplitud de onda P, R, T)
- Análisis de frecuencia (DCT, FFT)
- Estadísticos (media, varianza, entropía)

Generalización y capacidad predictiva [8][9]

Uno de los puntos fuertes de SVM en bioseñales es su capacidad para generalizar bien con un número limitado de muestras, debido al principio del margen máximo:

Esto contrasta con las redes neuronales, que necesitan más datos y son más propensas al sobreajuste.

En ECG, donde la recolección de datos etiquetados es costosa (requiere anotaciones clínicas), SVM es útil incluso con cientos o pocos miles de latidos.

Limitaciones y desafíos en ECG [10]

- Clases no balanceadas

En registros clínicos, los latidos normales superan con creces a los anómalos por lo que SVM puede sesgarse.

- Ruido fisiológico y artefactos

Ruido muscular o de movimiento puede alterar la señal ECG y afectar la clasificación. SVM es robusto a cierto nivel de ruido, pero debe ir acompañado de preprocesamiento riguroso (filtrado, segmentación precisa).

Equipos comerciales con algoritmos de ayuda diagnóstica basados en ECG

1. KardiaMobile 6L (AliveCor) [11]

KardiaMobile es un dispositivo personal de EKG, disponible en versiones de una o seis derivaciones (KardiaMobile 6L), que permite a los usuarios registrar EKGs de grado médico en cualquier momento y lugar, utilizando un teléfono inteligente. Está aprobado por la FDA para detectar una variedad de arritmias comunes, como fibrilación auricular (AFib), bradicardia y taquicardia

Funciones:

Adquisición de la Bioseñal: El dispositivo capta la actividad eléctrica del corazón a través de electrodos en contacto con la piel. La señal EKG es una bioseñal de baja frecuencia que representa la despolarización y repolarización de las aurículas y ventrículos.

Procesamiento de la Señal: El dispositivo realiza un preprocesamiento de la señal para reducir el ruido y las interferencias como ruido de línea, artefactos de movimiento, que son comunes en entornos ambulatorios.

Algoritmos de Diagnóstico: Utiliza algoritmos basados en inteligencia artificial para analizar los patrones de la onda EKG, identificar y clasificar automáticamente las arritmias. Por ejemplo, el algoritmo avanzado puede analizar la señal para detectar AFib, bradicardia, taquicardia, ritmo sinusal con PVCs (contracciones ventriculares prematuras), ritmo sinusal con SVE (extrasístoles supraventriculares) y ritmo sinusal con QRS ancho. La capacidad de estos algoritmos para discernir características sutiles en la bioseñal es crucial para un diagnóstico preciso..

Desde bioseñales: La ventaja de contar con 6 derivaciones permite un análisis más completo del vector eléctrico cardíaco. Utiliza algoritmos de detección automática de eventos.

2. Zio ECG Monitors (de iRhythm) [12]

Los monitores Zio de iRhythm son parches de EKG portátiles y discretos para monitoreo a largo plazo (hasta 14 días), diseñados para la detección de arritmias. El servicio incluye monitores fáciles de usar y un servicio integral que combina un algoritmo de "deep learning" aprobado por la FDA con la verificación de técnicos cardiográficos certificados.

Funcionalidades:

Monitoreo Continuo: Permiten la adquisición continua de la bioseñal EKG durante periodos prolongados, lo que es vital para detectar arritmias paroxísticas o infrecuentes que pueden no manifestarse durante un EKG de corta duración en clínica.

Análisis de Datos Masivos: La gran cantidad de datos EKG recolectados es procesada por algoritmos de aprendizaje profundo, que son particularmente efectivos en la identificación de patrones complejos y anomalías sutiles en las bioseñales. Estos algoritmos se entrenan con grandes bases de datos de EKG para mejorar su capacidad de

clasificación y diagnóstico.

Generación de Informes Detallados: El sistema genera informes de análisis que proporcionan datos detallados sobre la duración completa del monitoreo, permitiendo a los médicos tomar decisiones diagnósticas y terapéuticas más informadas.

3. *Electrocardiogramas con Algoritmo VERITAS [13]*

Varios electrocardiogramas de diagnóstico clínico de 12 derivaciones, como los ofrecidos por Welch Allyn (ahora parte de Baxter/Hillrom) y GE HealthCare, incorporan algoritmos de interpretación como el VERITAS™ Algorithm. Estos sistemas están diseñados para su uso en hospitales y clínicas, ofreciendo una interpretación de EKG para adultos y pediátricos [3, 4].

Funcionalidades:

Adquisición de Múltiples Derivaciones: La adquisición de 12 derivaciones proporciona una vista completa de la actividad eléctrica del corazón desde diferentes ángulos, lo que es esencial para localizar anomalías y determinar la severidad de las condiciones cardíacas.

Algoritmos de Interpretación: Algoritmos como VERITAS o el algoritmo Glasgow analizan las características de la bioseñal EKG (ondas P, complejos QRS, ondas T, intervalos PR, QT, duración QRS, etc.) y las comparan con bases de datos extensas para proporcionar una interpretación diagnóstica. Estos algoritmos tienen en cuenta factores como la edad, el sexo y, en algunos casos, la raza del paciente para ofrecer interpretaciones más precisas. Desde la teoría de las bioseñales, esto implica la aplicación de procesamiento de señales (filtrado, detección de picos QRS) y técnicas de reconocimiento de patrones para correlacionar las características de la señal con condiciones cardíacas conocidas.

Tecnologías de Mejora de la Señal: Incorporan funciones como "Best 10" o "Best 10 Seconds Technology", que seleccionan los 10 segundos más limpios y con menos artefactos de la señal EKG grabada para el análisis, mejorando la fiabilidad del diagnóstico. Esto es fundamental para asegurar la calidad de la bioseñal antes de su interpretación.

V. CONCLUSIONES

Los resultados obtenidos en este estudio permiten concluir que el modelo XGBoost entrenado con 50 árboles fue el clasificador con mejor desempeño global entre todos los métodos evaluados. Este modelo alcanzó una precisión del 94 %, un macro F1-score de 0.94, y logró valores superiores o iguales a 0.93 en todas las clases evaluadas, lo cual refleja su capacidad para modelar relaciones no lineales de forma eficiente, sin requerir arquitecturas profundas ni largos tiempos de entrenamiento.

A diferencia de los modelos basados en redes neuronales, que requieren ajustes cuidadosos en su arquitectura y son más propensos al sobreajuste con conjuntos de datos moderados, XGBoost mostró una alta estabilidad y robustez, manteniendo un equilibrio entre precisión y recall en todas las clases. Su naturaleza basada en árboles de decisión lo hace menos sensible a escalados de los datos y más interpretable, además de facilitar la evaluación de la importancia relativa de cada característica, lo que constituye una ventaja significativa para aplicaciones biomédicas que requieren trazabilidad y justificación clínica.

En comparación con SVM, XGBoost superó ligeramente los valores de F1 macro y presentó menor confusión entre clases con morfologías similares (AFIB vs. SB, o ST vs. SVT). Asimismo, los resultados obtenidos con K-means confirmaron que los datos presentan cierta estructura, pero que la clasificación supervisada es indispensable para lograr niveles de desempeño clínicamente aceptables.

REFERENCIAS

- [1] World Health Organization, "Cardiovascular diseases (CVDs)," WHO, 2021. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] Yildirim, O., Talo, M., Ay, B., Baloglu, U. B., Aydin, G., & Acharya, U. R., "Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals," *Computers in Biology and Medicine*, vol. 113, p. 103387, 2019. doi: 10.1016/j.combiomed.2019.103387.
- [3] Hong, S., Zhou, Y., Shang, J., Xiao, C., Sun, J., & Wang, F., "Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review," *Computers in Biology and*

Medicine, vol. 122, p. 103801, 2020. doi: 10.1016/j.combiomed.2020.103801.

[4] Zheng, J., Chu, H., Struppa, D., et al. (2020). Optimal Multi-Stage Arrhythmia Classification Approach. *Scientific Reports*, 10(1), 2959.

[5] Walsh, G., & Markides, V. (2011). P waves in the electrocardiogram recording of tachycardia: 'You can run, but you cannot hide'. *Europace*, 13(7), 916–917. [P waves in the electrocardiogram recording of tachycardia: 'You can run, but you cannot hide'](#)

[6] Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645–678.

[7] Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern Recognition*, 38(11), 1857–1874.

[8] Cortes, C., & Vapnik, V. (1995). *Support-vector networks*. *Machine Learning*, 20(3), 273–297.

[9] Khandoker, A. H., et al. (2009). Automated recognition of patients with congestive heart failure using phase-space analysis of ECG. *Computers in Biology and Medicine*, 39(8), 731–743.

[10] Osowski, S., & Linh, H. T. (2001). ECG beat recognition using fuzzy hybrid neural network. *IEEE Trans. Biomedical Engineering*, 48(11), 1265–1271.

[11] AliveCor. (n.d.). *KardiaMobile 6L*. Recuperado de <https://www.alivecor.com/kardiamobile6l/>

[12] iRhythm Technologies. (n.d.). *Zio ECG Monitors*. Recuperado de <https://www.irhythmtech.com/zio-monitors>

[13] Welch Allyn. (n.d.). *Welch Allyn® CP 150™ Electrocardiograph with Interpretive ECG*. Recuperado de <https://www.hillrom.com/en/products/physical-exam/ecg/welch-allyn-cp-150-ecg/> (Nota: Welch Allyn es ahora parte de Baxter/Hillrom)