

PROYECTO INTEGRADOR EN INTELIGENCIA ARTIFICIAL

PRESENTACIÓN DEL PROYECTO

Grupo 6:

Luis Andress Bustamante Jiménez

Damaris Irene Alarcón Vallejo

Docente: Gladys Maria Villegas Rugel

19 de septiembre de 2025

Contenido:

1. MAPA DEL ESTADO DEL ARTE.....	3
• Revisión sistemática:.....	3
• Análisis comparativo de 5 enfoques principales existentes.....	5
• Timeline de avances en el área	6
• Matriz de fortalezas/debilidades de enfoques actuales.....	7
• Identificación clara del gap	8
2. ANÁLISIS DE DATASETS DISPONIBLES	9
• Comparación entre datasets: ventajas y desventajas	11
• Recomendación fundamentada del dataset principal: MIT-BIH	12
• Dataset secundario (plan b)	13
3. DEFINICIÓN DE MÉTRICAS DE ÉXITO.....	15
• Métricas técnicas	15
• Métricas de eficiencia	15
• Métricas de impacto	15
• Umbrales de éxito	16
4. ANÁLISIS DE STAKEHOLDERS.....	16
• Matriz de influencia / interés	17
• Análisis de necesidades, expectativas y barreras	17
• Plan de comunicación diferenciado por stakeholder.....	19
5. DOCUMENTO DE ALCANCE DEL PROYECTO	21
• Objetivo general:.....	21
• Objetivos específicos:	22
• Alcance incluido:	22
• Alcance excluido:.....	22
• Criterios de aceptación:	22
6. CRONOGRAMA ÁGIL	23
Roles Scrum Definidos:	23
7. PLAN DE RECURSOS.....	26
• Recursos Humanos.....	26
Estimación de Horas por Sprint:.....	26
• Recursos Técnicos.....	26
• Recursos Financieros.....	26
Presupuesto Estimado:	26
• Plan de Adquisición y Gestión de Recursos	27
8. HITOS Y ENTREGABLES.....	27
9. Referencias Bibliográficas:	28

Clasificación explicable de arritmias cardíacas a partir de electrocardiogramas transformados en espectrogramas mediante redes neuronales convolucionales

1. MAPA DEL ESTADO DEL ARTE.

- Revisión sistemática:

Nº	Año	Título / Referencia	Enfoque principal / qué aporta	Link
1	2025	<i>Arrhythmia classification based on multi-input convolutional neural network with attention mechanism — Bin Zheng et al.</i>	Usan representaciones STFT duales de distintas resoluciones, CNN con bloques SE de atención; evaluado en MIT-BIH y SPH.	https://doi.org/10.1371/journal.pone.0326079
2	2023	<i>Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals — YD Daydulo et al.</i>	Representaciones tiempo-frecuencia, modelos preentrenados (ResNet/AlexNet), clasifican ARR / CHF / NSR entre otras clases con alta precisión.	https://doi.org/10.1186/s12911-023-02326-w
3	2024	<i>Intra- and Interpatient ECG Heartbeat Classification Based on Multimodal Convolutional Neural Networks with an Adaptive Attention Mechanism — IF Di Paolo & A. R. G. Castro</i>	Transformaciones tipo HSFC y RP, utilizan 2 derivaciones (MLII y V1), método multimodal + atención adaptativa, evaluaciones intra- e inter-paciente.	https://doi.org/10.3390/app14209307
4	2024	<i>Multimodal ECG heartbeat classification method based on a convolutional neural network embedded with FCA — Feiyan Zhou & Duanshu Fang et al.</i>	Transforman señales ECG en imágenes modales usando RP, GAF, MTF, modelo CNN con módulo FCA, alta precisión (~99.6%) en MIT-BIH con cinco clases de arritmia.	https://doi.org/10.1038/s41598-024-59311-0
5	2023	<i>Multi-classification method of arrhythmia based on multi-scale residual neural network and multi-channel data fusion — F. Zhang et al.</i>	Residual network multiescala + fusión de datos multi-canal, imágenes de ECG, precisión y generalización alta.	https://doi.org/10.3389/fphys.2023.1253907
6	2024	<i>Res-BiANet: A Hybrid Deep Learning Model for Arrhythmia Detection Based on PPG Signal — Wu Y. et al.</i>	Modelo híbrido (ResNet + BiLSTM) + mecanismo self-attention, señales PPG (no ECG), clasifica varias arritmias + ritmo sinusal, buen desempeño.	https://doi.org/10.3390/electronics13030665
7	2023	<i>HARDC: A novel ECG-based heartbeat classification method to detect arrhythmia using hierarchical attention</i>	Arquitectura híbrida: CNN dilatada + RNN dual estructurada + atención jerárquica, mejoras en extracción de características	https://arxiv.org/abs/2303.06020

		<i>based dual structured RNN with dilated CNN</i>	locales y globales, buen rendimiento en MIT-BIH / desafío CinC etc.	
8	2022	<i>Arrhythmia Classifier Using Convolutional Neural Network with Adaptive Loss-aware Multi-bit Networks Quantization — Hanshi Sun et al.</i>	Clasificador 1-D CNN quantizado/adaptativo para reducir consumo de memoria/poder; muchas clases de ritmo; trade-off entre eficiencia y precisión.	https://arxiv.org/abs/2202.12943
9	2020	<i>Multi-Lead ECG Classification via an Information-Based Attention Convolutional Neural Network — Hao Tung, Chao Zheng, Xinsheng Mao, Dahong Qian</i>	Atención por canal para múltiples derivaciones, mejoran sensibilidad y precisión vs modelos sin atención, usan MIT-BIH.	https://arxiv.org/abs/2003.12009
10	2025	<i>A novel hybrid CNN-transformer model for arrhythmia detection without R-peak identification using stockwell transform</i>	Aplica la transformada de Stockwell para extraer características de ECG, combina CNN con Transformer, y lo hace <i>sin necesidad de detección previa de picos R</i> . Logra alta exactitud tanto en MIT-BIH como en conjuntos más recientes. Útil para sistemas en tiempo real y uso práctico	https://www.nature.com/articles/s41598-025-92582-9?utm_source=chatgpt.com
11	2025	<i>EXGnet: a single-lead explainable-AI guided multiresolution network with train-only quantitative features — T. Talukder Showrav et al.</i>	Modelo fiable para ECG single-lead; multiresolución; incorpora Grad-CAM y guía XAI durante entrenamiento; datasets Chapman, Ningbo.	https://arxiv.org/abs/2506.12404
12	2025	<i>Classification of multi-lead ECG based on multiple scales and hierarchical feature CNNs — F Zhou et al.</i>	Multi-lead, múltiples escalas, jerarquía de características, evalúa en MIT-BIH.	https://www.nature.com/articles/s41598-025-94127-6
13	2024	<i>Deep learning-assisted arrhythmia classification using 2-D ECG spectrograms — Pinjala N. Malleswari et al.</i>	Transforman señales ECG a imágenes con CWT, usan transfer learning con CNNs preentrenadas; métricas altas.	https://www.researchgate.net/publication/387580665_Deep_learning-assisted_arrhythmia_classification_using_2-D_ECG_spectrograms
14	2024	<i>Comparative study of time-frequency transformation methods for ECG signal classification — MS Song et al.</i>	Comparan STFT, CWT, etc.; analizan el efecto de la transformación en atención / Grad-CAM / qué regiones activan; usan MIT-BIH.	https://www.frontiersin.org/journals/signal-processing/articles/10.3389/frsip.2024.1322334/full
15	2025	<i>Enhanced ECG Arrhythmia Diagnosis with Deep Learning and Selective Attention</i>	DCNN + mecanismo de atención selectiva; logran altas precisiones multclasificación.	https://www.researchgate.net/publication/391592637_Enhanced_Electrocardiogram_Arrhythmia_Diagnosis_with_De

		<i>Mechanism — Hasanain S. Mansour et al.</i>		ep Learning and Selective Attention Mechanism
16	2025	<i>Analysis of Cardiac Arrhythmias Based on ResNet-ICBAM — C. Wang et al.</i>	Propuesta dual canal: uno con ECG en crudo, otro transformado en imágenes (GADF, MTF, RP etc.), fusiones y atención.	https://www.mdpi.com/1424-8220/25/3/661
17	2025	<i>Automated ECG Arrhythmia Classification Using Feature Images with Common Matrix Approach-Based Classifier — A. Kirkbas et al.</i>	Generan representaciones espectrales / tiempo-frecuencia; usan varios datasets; precisión alta.	https://www.mdpi.com/1424-8220/25/4/1220
18	2024	<i>A hybrid cardiovascular arrhythmia disease detection using ConvNeXt-X models — MA Talukder et al.</i>	Usan variantes de ConvNeXt, técnicas de balanceo de clases; evalúan en MIT-BIH; muy alta precisión.	https://www.nature.com/articles/s41598-024-81992-w
19	2025	<i>Explainable AI (XAI) for Arrhythmia detection from electrocardiograms — Joschka Beck & Arlene John</i>	Comparan técnicas de XAI (SHAP, DeepLIFT, gradientes) sobre modelo CNN; evalúan preferencia de médicos, qué tipo de explicación funciona mejor.	https://arxiv.org/abs/2508.17294
20	2025	<i>Deep learning-based prediction of atrial fibrillation from polar transformed time-frequency electrocardiogram— D Kwon et al.</i>	Usan STFT espectrogramas transformados a dominio polar; CNNs para detectar FA; dataset CinC-2017.	https://pubmed.ncbi.nlm.nih.gov/40063554/

• **Análisis comparativo de 5 enfoques principales existentes.**

Enfoque	Representación de entrada	Arquitecturas típicas	Datasets usados	Métricas reportadas	Fortalezas	Limitaciones	Referencias
1. Modelos 1D end-to-end	Señal ECG cruda 1D (latido/ventanas)	CNN 1D, RNN (LSTM/GRU), híbridos CNN+RNN con atención jerárquica	MIT-BIH, CinC, SPH	Acc. 94–98%, F1 macro >0.90	Simplicidad, baja latencia, despliegue en edge devices (con quantization)	Sensibles a ruido, poca interpretabilidad, dependen de segmentación precisa	[7], [8], [9], [15]
2. Modelos 2D tiempo-frecuencia	Espectrogramas STFT, CWT, Stockwell, transformadas polares	CNN 2D, ResNet, AlexNet, transfer learning	MIT-BIH, CinC-2017, Chapman	Acc. 96–99%, AUC >0.95	Captura patrones robustos, aprovecha visión por computador, evita extracción manual	Coste computacional alto, elección de parámetros TF crítica, pérdida de info temporal fina	[1], [2], [10], [13], [14], [20]
3. Enfoques multimodales / multi-lead	Varias derivaciones (MLII, V1...), combinaciones de imágenes (RP, GAF, MTF)	CNN multimodal, ResNet multi-escala, fusión de canales	MIT-BIH, SPH, Ningbo	Acc. ≈99%, alta generalización intra/interpaciente	Integra info complementaria, mayor robustez interpaciente	Complejos, riesgo de overfitting, alto costo de entrenamiento	[3], [4], [5], [12], [16], [17]
4. Modelos con atención y explicabilidad	1D o 2D + mapas de atención/Grad-CAM	CNN con SE-blocks, FCA, ICBAM, XAI	MIT-BIH, Chapman, Ningbo	Acc. 97–99%, mejoras en F1; médicos prefieren	Explicaciones visuales útiles, foco en morfología ECG, sube	Atención no siempre ≠ explicación fiel,	[1], [3], [4], [11], [14], [15], [16], [19]

		(Grad-CAM, SHAP, DeepLIFT)		heatmaps interpretables	métricas al filtrar ruido	mayor costo computacional	
5. Híbridos y eficiencia (CNN+Transformer/ConvNeXt)	1D/2D dual-stream, PPG como alternativa	CNN+Transformer, ConvNeXt, quantized CNN, BiLSTM+attention	MIT-BIH, CinC, Chapman, PPG datasets	Acc. >99% en SOTA, latencia reducida con quantization	Arquitecturas modernas, robustas, despliegue eficiente en tiempo real	Transformers requieren más datos, PPG menos estandarizado que ECG	[6], [8], [10], [16], [18]

- **Timeline de avances en el área**

Año	Avance principal	Detalles y contribuciones clave
2020	Atención en multi-lead ECG	Se emplean módulos de atención por canal (channel-wise attention) para derivaciones múltiples, mejorando sensibilidad y precisión frente a modelos que usan single-lead o sin atención. Ej: <i>Multi-Lead ECG Classification via an Information-Based Attention Convolutional Neural Network</i> .
2022	Eficiencia y cuantización	Trabajos orientados a reducir uso de memoria/poder y hacer modelos aptos para hardware limitado. Por ejemplo, CNN 1D quantizado/adaptativo, "loss-aware multi-bit networks quantization". [8]
2023	Transfer learning y arquitecturas híbridas con atención jerárquica	Uso de modelos preentrenados (ResNet, AlexNet) sobre representaciones tiempo-frecuencia; además, arquitecturas híbridas tipo CNN dilatada + RNN dual estructurada + atención jerárquica (por ejemplo HARDC). Contribuciones en precisión, generalización.
2024	Multimodalidad, comparativas de transformadas TF, consolidación de nuevos enfoques arquitectónicos	Se ven métodos multimodales que combinan RP, GAF, MTF, múltiples derivaciones ("multi-lead"), también comparaciones entre STFT, CWT, etc. Aparición de variantes modernas de arquitecturas como ConvNeXt adaptadas al problema de ECG.
2025	Explicabilidad, multiresolución, modelos SOTA, nuevos tipos de representación como polar TF	Avances importantes hacia modelos más explicables clínicamente (uso de Grad-CAM, SHAP, bloques SE, ICBAM, FCA), modelos multiresolución como EXGnet; modelos que fusionan la señal cruda y la imagen; representación polar para detectar FA, etc. Ejemplos confirmados: <i>A novel hybrid CNN-transformer model</i>

		using Stockwell transform (PubMed), EXGnet (arXiv), trabajo XAI reciente (arXiv).
--	--	---

• **Matriz de fortalezas/debilidades de enfoques actuales**

Enfoque	Fortalezas	Debilidades
1. Modelos 1D end-to-end (ECG crudo)	<ul style="list-style-type: none"> - Arquitecturas simples y rápidas. - Baja latencia, aptos para despliegue en dispositivos médicos portátiles. - No requieren transformaciones previas. 	<ul style="list-style-type: none"> - Sensibles a ruido y artefactos. - Menor interpretabilidad. - Alta dependencia de segmentación precisa y detección de R-peaks.
2. Modelos 2D tiempo-frecuencia (espectrogramas CNN)	<ul style="list-style-type: none"> - Capturan patrones robustos tiempo-frecuencia. - Permiten uso de arquitecturas preentrenadas (transfer learning). - Buen desempeño en métricas (Acc > 96%). 	<ul style="list-style-type: none"> - Coste computacional elevado. - Posible pérdida de información temporal fina. - Resultados muy sensibles a la elección de parámetros TF.
3. Enfoques multimodales / multi-lead	<ul style="list-style-type: none"> - Combinan información de varias derivaciones y vistas (RP, GAF, MTF). - Mayor generalización interpaciente. - Reportan precisiones cercanas al 99%. 	<ul style="list-style-type: none"> - Complejidad alta de entrenamiento y modelos. - Requieren grandes volúmenes de datos. - Riesgo de overfitting si no hay regularización adecuada.
4. Modelos con atención y explicabilidad (XAI)	<ul style="list-style-type: none"> - Mejor interpretabilidad clínica con mapas de calor y mecanismos de atención. - Capacidad de resaltar regiones críticas del ECG (ondas P, QRS, T). 	<ul style="list-style-type: none"> - La atención no siempre equivale a una explicación fiel. - Alto costo de cómputo. - Escasas validaciones con médicos reales.

	- Suelen mejorar métricas al filtrar ruido irrelevante.	
5. Híbridos y eficiencia (CNN+Transformer/ConvNeXt, cuantización)	<ul style="list-style-type: none"> - SOTA en precisión y robustez (>99%). - Transformers capturan dependencias globales. - Cuantización y distilación permiten despliegue eficiente en tiempo real. 	<ul style="list-style-type: none"> - Mayor complejidad de implementación y tuning. - Transformers requieren más datos para evitar sobreajuste. - PPG como alternativa al ECG no siempre estandarizada.

- **Identificación clara del gap**

Limitación detectada en el estado actual	Evidencia en la literatura	Implicación práctica	Estrategia propuesta en el proyecto
Explicabilidad insuficiente y poco validada	Uso de Grad-CAM/SHAP sin pruebas de fidelidad ni evaluación con médicos ([1], [11], [14], [19])	Baja confianza clínica; explicaciones pueden ser engañosas	Incorporar módulos de atención + XAI con validaciones de fidelidad (occlusion, ROAR, sanity checks) y evaluación piloto con cardiólogos
Dependencia de un solo dataset (MIT-BIH)	Validación centrada en splits por latido/paciente único ([2], [4], [9])	Modelos no generalizan bien a datos clínicos reales	Uso de múltiples bases (MIT-BIH, CinC, Chapman) y validación inter-paciente
Pérdida de información temporal fina al pasar a 2D	Espectrogramas suavizan detalles morfológicos relevantes (QRS, onda P) ([13], [14], [20])	Riesgo de perder señales críticas para diagnóstico	Aplicar representaciones multi-resolución (ventanas STFT/CWT combinadas) y complementar con features temporales

Robustez insuficiente frente a ruido/artefactos	Pocos estudios evalúan en condiciones clínicas reales ([6], [8], [18])	Degradación fuerte en entornos reales	Diseñar un banco de pruebas de ruido/artefactos, augmentación fisiológica y métricas de robustez
Calibración e incertidumbre no consideradas	La mayoría reporta solo Accuracy/F1, sin métricas de confianza ([7], [15], [16])	Difícil uso clínico seguro; riesgo de decisiones erróneas	Implementar calibración probabilística (ECE, Brier) y mecanismos de detección OOD

2. ANÁLISIS DE DATASETS DISPONIBLES

- **Dataset A: MIT-BIH Arrhythmia Database**

Característica	Detalle
Tamaño / estructura técnica	48 grabaciones de ECG de media hora cada una, con dos canales (dual-lead), muestreo = 360 samples/sec por canal, resolución de 11 bits sobre rango ± 10 mV. (~104.3 MB descomprimido) (PhysioNet)
Procedencia / licencia / uso	PhysioNet; "Open Data Commons Attribution License v1.0". Cualquiera puede acceder bajo términos de licencia.
Calidad (completitud, consistencia, representatividad)	Anotaciones hechas por al menos dos cardiólogos; variedad de tipos de arritmias incluidas (algunas comunes, otras menos comunes). Sin embargo: pequeño número de sujetos (47), solo 2 canales, datos no recientes, quizás poca diversidad demográfica. También algunas anotaciones "non-beat" y tipos de anotaciones múltiples que generan confusión (se ha reportado error de conteo de latidos en algunos trabajos)
Idoneidad para tu problema	Muy buena para prototipo, benchmarking, para espectrogramas, para comparar modelos CNN. Limitaciones: 2-derivaciones limita información espacial multi-lead; señal de muestreo 360 Hz es aceptable pero quizás menor que otros; pocos pacientes, lo que podría perjudicar generalización.
Accesibilidad	Completamente accesible públicamente via PhysioNet; sin costos.

- **Dataset B: Chapman-Shaoxing-Ningbo “12-lead ECG database for arrhythmia study”**

Característica	Detalle
Tamaño / estructura técnica	45,152 pacientes; ECG de 12 derivaciones; muestreo de 500 Hz; múltiples clases de arritmia y otras condiciones cardiovasculares incluidas. (PhysioNet)
Procedencia / licencia / uso	PhysioNet; uso académico, abierta; hay publicación asociada (Scientific Data) que lo describe. Credenciales revisadas por profesionales.
Calidad	Alta: muchas derivaciones, tamaño grande, muestreo moderno, diversidad de pacientes. Posible variabilidad en etiquetas (errores humanos / lectura automática), necesidad de limpieza de datos y estandarización de etiquetas.
Idoneidad	Muy adecuada para tu problema: riqueza de información espacial (12 derivaciones), buena resolución, tamaño grande facilita entrenamiento de CNN, posibilidad de espectrogramas multi-lead; buen candidato para generalización inter-paciente.
Accesibilidad	Dataset público en PhysioNet; parece no tener costos. Debes revisar condiciones específicas de uso/licencia, pero generalmente uso académico permitido.

- **Dataset C: PhysioNet / CinC Challenge 2021 + fuentes múltiples (PTB-XL, CPSC, etc.)**

Característica	Detalle
Tamaño / estructura técnica	>100.000 ECGs de 12 derivaciones combinadas; entrenamientos, validaciones, tests; duraciones variadas (6–144 segundos algunas, otros 10 segundos); muestreo generalmente 500 Hz; archivos en formato WFDB y otros formatos de ECG estándar. (PhysioNet)
Procedencia / licencia / uso	PhysioNet Challenge; uso público para entrenamiento (parte), para validación/test; condición de uso académica; algunas partes privadas para test; posibles restricciones para test.

Calidad	Alta diversidad de datos demográficos y clínicos; varios orígenes geográficos; buena cantidad; posibles inconsistencias entre bases (etiquetas, duración, muestreo); algunas señales cortas lo que puede limitar transformaciones de espectrograma si ventana grande.
Idoneidad	Excelente para generalización; buen tamaño para CNN + XAI; diversidad de derivaciones; útil para evaluar robustez inter-paciente; posibilidad de usar subset para entrenamiento, validation, test; ideal para evaluar efectos de ruido / artefactos.
Accesibilidad	La parte de entrenamiento es pública; test puede estar restringido si es challenge; quizás algunos metadatos demográficos limitados; posiblemente hay límites de uso para publicación si se usan los sets privados.

- **Comparación entre datasets: ventajas y desventajas**

Dataset	Ventajas clave	Desventajas / limitaciones
MIT-BIH	Fácil de manejar; ya ampliamente usado → buenos benchmarks; buen nivel de anotaciones; bajo costo/recursos; señal de buena calidad; accesibilidad garantizada.	Sólo 2 derivaciones; pocos pacientes; datos antiguos; poca diversidad demográfica; menor muestreo comparado a otros; posible sesgo de selección; limitado para espectrogramas que requieren ventanas largas o múltiples leads.
Chapman-Shaoxing-Ningbo (12-lead)	Derivaciones múltiples → espacio para modelos espectrograma multi-lead; gran tamaño; resolución alta; diversidad clínica; moderno.	Mayor tamaño podría requerir más recursos de cómputo; limpieza de etiquetas puede ser necesaria; duraciones fijas o variables pueden requerir estandarización; algunas clases de arritmia pueden estar poco representadas; posibles condiciones de licencia/licencias específicas de uso.

PhysioNet / CinC Challenge (multi-origen)	Gran diversidad; variedad de duraciones; disponibilidad de entrenamiento/validación/test; buenos para robustez; múltiples derivaciones; datos recientes; posibilidad de generalización realista.	Parte del test puede no estar disponible libremente para todos; variabilidad en muestreo o estructura que requiere normalización; documentos pueden tener metadatos incompletos; señales cortas limitan transformaciones con ventanas largas; posiblemente mayor costo en procesamiento; heterogeneidad que hace difícil modelado uniforme.
---	--	---

- **Recomendación fundamentada del dataset principal: MIT-BIH**

Criterio	Ventajas del MIT-BIH	Limitaciones que reconoce la literatura para ese criterio
Benchmark histórico / comparación	Es muy usado en la literatura de clasificación de arritmias con CNN y espectrogramas, hay muchos trabajos con los cuales comparar resultados. Permite posicionarte respecto al estado del arte.	Fortalece la reproducibilidad, pero también genera un sesgo de “todos usan MIT-BIH”, lo que puede limitar la novedad si no combinas con otros datasets.
Accesibilidad	Muy accesible: está disponible públicamente vía PhysioNet, licencia abierta (Open Data Commons Attribution License v1.0), sin costos.	Ninguna barrera económica grande, aunque hay que tener cuidado con los términos de citación / uso; no siempre hay metadatos demográficos muy completos.
Formato y estructura adecuados para espectrogramas y CNN	Buena calidad de señal: muestreo 360 Hz, dos canales, resoluciones estables; muchas anotaciones por latido realizadas por cardiólogos; consistente	Con sólo dos canales limita información espacial multi-derivación; el tamaño (48 registros) es modesto para arquitecturas muy profundas si no

	estructura de archivos; cantidad moderada de anotaciones ($\approx 110,000$ latidos anotados) que permite entrenar modelos CNN de espectrogramas.	se usa augmentación u otros datos adicionales. También, los datos son relativamente antiguos (años 70s), lo que puede implicar características de adquisición diferentes a equipos modernos.
Relevancia clínica	Incluye varios tipos de arritmias (comunes y algunas menos comunes) para evaluación; ampliamente aceptado como estándar de evaluación en literatura; cumple con estándares de anotaciones (beat-by-beat) lo que favorece precisión y evaluación rigurosa.	Puede tener sesgo demográfico (localización geográfica, edad, género) poco variado; derivaciones limitadas podrían no capturar todos los matices para ciertos tipos de arritmias menos comunes; artefactos modernos/noise actuales pueden diferir.
Costo computacional / facilidad para prototipo	Al tener tamaño moderado y estructura relativamente simple, entrenamiento inicial es manejable, ideal para prototipos, experimentación, pruebas rápidas de arquitectura y explicabilidad.	Para modelos muy grandes, redes multimodales o espectrogramas múltiples puede requerir augmentación y CPU/GPU significativos; limitaciones en generalización si no se incorpora variabilidad externa.

- **Dataset secundario (plan b)**

Escenario de fallo	Dataset alternativo sugerido	Ventajas de ese dataset como sustituto	Qué ajustes o cuidados extra requerirías

1. Problemas legales/licencia (por ejemplo, restricciones de uso, cambios en disponibilidad)	Chapman-Shaoxing-Ningbo “12-lead ECG arrhythmia”	Tiene muchas derivaciones (12 canales), muestreo moderno (500 Hz), gran cantidad de sujetos; permite espectrogramas multi-lead con más riqueza espacial y representatividad clínica alta.	Se deberá ajustar el modelo para manejar múltiples derivaciones; más recursos de cómputo; más tiempo de preprocesamiento; verificación cuidadosa de etiquetas y formato; asegurarse de licencia académica y citación adecuada.
2. Problemas técnicos / falta de diversidad de clases en MIT-BIH	PhysioNet / CinC Challenge 2021 (o bases como PTB-XL + otros integrados)	Proporcionan mucha más diversidad de datos, varias duraciones, múltiples fuentes; permiten validación inter-paciente y robustez; mayor cantidad de clases diagnósticas.	Requiere unificación de formatos; normalización de datos (muestreo, derivaciones); limpieza de etiquetas; dividir bien por paciente para evitar fuga; posiblemente filtros de calidad pues algunos segmentos son cortos o con ruido.
3. Requerimientos de espectrogramas largos o multi-ventana si MIT-BIH no lo alcanza	Combinación de MIT-BIH + dataset secundario (por ejemplo usar MIT-BIH para prototipo, luego entrenar/fine-tune con Chapman o CinC)	Permite mantener continuidad con tu dataset principal para comparaciones, pero aumentar diversidad y generalización; puede mejorar robustez y resultados finales.	Alineación de sampling rate; armonización de etiquetas; balance de clases; potencial mayor costo de cómputo; manejo de datos de distintas calidades; posibles diferencias en señal/noise que haya que ajustar con augmentación o preprocesamiento específico.

3. DEFINICIÓN DE MÉTRICAS DE ÉXITO

- **Métricas técnicas**

- Métrica primaria: F1-score macro

Se utilizará como métrica principal, ya que permite equilibrar las clases mayoritarias y minoritarias del MIT-BIH, importante para arritmias poco frecuentes.

- Métricas secundarias

- Precision y Recall por clase: para ver cuántos falsos positivos (precisión) y falsos negativos (recall) hay, especialmente en clases críticas.
- Accuracy global: para comparaciones históricas / con otros estudios.
- AUC-ROC / AUC-PR si el modelo produce probabilidades, para evaluar desempeño con umbrales distintos.

- **Métricas de eficiencia**

- Tiempo de inferencia por ventana / latido completo (incluyendo preprocesamiento + generación de espectrograma + clasificación).
- Uso de recursos (memoria RAM, CPU/GPU) durante la inferencia, para estimar viabilidad en uso clínico o dispositivo.

- **Métricas de impacto**

- Reducción estimada de errores diagnósticos comparado con métodos actuales
Estimar porcentaje de falsos negativos reducidos en clases críticas (ej. ventricular, fibrilación).
- Tiempo de diagnóstico reducido
Estimar cuánto se acelera el proceso si el modelo sirve como herramienta de apoyo al personal médico.
- Confianza / aceptación clínica
Nivel de confianza que los profesionales (cardiólogos o técnicos) tienen en las explicaciones producidas; porcentaje de casos donde la explicación se considera útil/relevante.
- Métricas de usabilidad
- SUS (System Usability Scale) u otra escala breve de usabilidad, aplicada a un grupo pequeño de profesionales para evaluar interfaz de visualización de explicaciones.

- Tiempo de interpretación / revisión
Cuánto se tarda el usuario médico en revisar la salida explicable frente a un método que no proporcione explicación.
- Tasa de aceptación de explicación
En la prueba con usuarios, cuántos aceptan confiar en la explicación generada claramente, cuántas veces la usan como base para decisión diagnóstica.

- **Umbrales de éxito**

Métrica	Umbral mínimo aceptable	Objetivo de excelencia
F1-score macro (MIT-BIH)	≥ 0.80	≥ 0.97
Accuracy global	≥ 0.90	≥ 0.98
Recall en clases críticas (VEB / SVEB / fibrilación, etc.)	≥ 0.80	≥ 0.90

4. ANÁLISIS DE STAKEHOLDERS

Los stakeholders que participan o se ven afectados por el proyecto “Clasificación explicable de arritmias cardíacas mediante espectrogramas + CNN + XAI” se dividen en tres niveles:

- **Stakeholders primarios**
 - Usuarios finales directos: cardiólogos, técnicos en ECG, personal clínico que va a usar la herramienta.
 - Beneficiarios: pacientes con sospechas de arritmias, cuyo diagnóstico puede mejorar con ayuda del sistema.
 - Decisores de implementación: administradores hospitalarios, directores de departamentos de cardiología, gerentes de tecnología médica.
- **Stakeholders secundarios**
 - Reguladores / entes de control sanitario: organismos que regulan dispositivos médicos, estándares de salud, protección de datos.
 - Proveedores de datos y de tecnología: fabricantes de equipos ECG, software de anotación, proveedores de hardware y otros.
 - Competidores indirectos: otras herramientas o soluciones de IA para ECG / diagnóstico automático, comerciales o investigativas.

- **Stakeholders clave**

- Sponsors o financiadores: universidad, agencias de investigación, fondos de salud o tecnológicos.
- Expertos del dominio clínico/técnico: cardiólogos, electrofisiólogos, ingenieros biomédicos, investigadores en ML aplicado a salud.
- Implementadores técnicos: equipo de ML / IA, ingenieros de software, ingenieros de datos, especialistas en infraestructura.

- **Matriz de influencia / interés**

Stakeholder	Nivel de interés	Nivel de influencia
Usuarios finales (cardiólogos, técnicos ECG)	Alto	Medio - Alto
Pacientes	Alto	Medio
Decisores de implementación	Medio-Alto	Alto
Reguladores sanitarios	Medio	Medio
Proveedores de datos / tecnología	Medio	Medio
Competidores indirectos	Bajo-Medio	Bajo-Medio
Sponsors / financiadores	Alto	Alto
Expertos del dominio	Alto	Medio-Alto
Implementadores técnicos	Alto	Alto

- **Análisis de necesidades, expectativas y barreras**

Stakeholder	Necesidades y expectativas	Potenciales resistencias o barreras
Usuarios finales (cardiólogos, técnicos ECG)	Esperan que la herramienta sea fiable, interpretaciones claras, detección temprana de arritmias, minimizar errores críticos, integración al flujo clínico, interfaz usable.	Desconfianza hacia modelos de IA, resistencia al cambio, preocupación porque el modelo "se equivoque" en casos particulares, dificultad si las explicaciones son poco claras o demasiado técnicas.

Pacientes	Diagnósticos más rápidos, mejores resultados clínicos, menos exámenes invasivos, transparencia del proceso, seguridad y privacidad de los datos.	Preocupaciones sobre el uso de datos personales, falta de comprensión del sistema, temor a errores, aceptación cultural de tecnología médica.
Decisores de implementación (hospitales, clínicas)	Evidencia de eficacia clínica, impacto en reducción de costes y errores, mejoras en eficiencia, cumplimiento regulatorio, retorno de la inversión, interoperabilidad con otros sistemas hospitalarios.	Costos iniciales, infraestructura necesaria, formación del personal, riesgos legales, resistencia institucional si no se ve claro el beneficio inmediato.
Reguladores / entes de control sanitario	Conformidad con normas de seguridad y salud, transparencia, auditoría, pruebas clínicas, protección de datos, validez, robustez frente a sesgos.	Regulaciones estrictas, necesidad de evidencia clínica, responsabilidad legal, posibles demoras burocráticas, estándares locales que pueden diferir.
Proveedores de datos / tecnología	Colaboración, acuerdos claros de licencia/datos, reconocimiento, compatibilidad técnica, acceso a datos limpios, estándares de interoperabilidad.	Preocupaciones sobre propiedad intelectual, licencias restrictivas, calidad de datos, compatibilidad de hardware/software, costes de mantenimiento.
Competidores indirectos	Seguimiento técnico de avances, adaptaciones, mejoras de su propia oferta, reconocimiento.	Competencia en mercado/investigación, posible incertidumbre sobre sus propias ventajas frente al nuevo sistema.
Sponsors / financiadores	Retorno esperado (científico, clínico, reputacional), cumplimiento de plazos, buenas publicaciones, impacto medible, visibilidad.	Riesgo de no alcanzar objetivos, sobrecostos, demoras, resultados no tan robustos, problemas de adopción.

Expertos del dominio (cardiólogos, ingenieros biomédicos)	Veracidad clínica de los resultados, que las explicaciones correspondan a patrones fisiológicos reales, que puedan contribuir al modelo, colaboración, validaciones externas, publicaciones de calidad.	Escepticismo si explicaciones no satisfacen estándares clínicos, dificultad técnica para interpretar salidas del modelo, resistencia si sienten que su juicio clínico queda reemplazado, barreras de comunicación técnica-médica.
Implementadores técnicos	Datos accesibles y documentados, herramientas/librerías adecuadas, infraestructura suficiente, pruebas, soporte, claridad metodológica, reproducibilidad.	Problemas con hardware, datos ruidosos, etiquetas poco claras, compatibilidad técnica, plazos ajustados, sobrecarga de trabajo si es mucho la parte técnica para 5 semanas.

- **Plan de comunicación diferenciado por stakeholder**

Stakeholder	Contenido clave del mensaje	Canales / método de comunicación	Frecuencia / momento de contacto
Usuarios finales (cardiólogos, técnicos ECG)	Demostración de funcionamiento; ejemplos de explicación visual; comparativas con diagnósticos existentes; seguridad y limitaciones; interfaz de usuario.	Talleres / sesiones prácticas, presentaciones clínicas, demostraciones interactivas, feedback directo, vídeos o capturas de pantalla.	Al inicio del desarrollo (recolección de requerimientos), durante prototipos intermedios, antes de la evaluación piloto, en la entrega final.
Pacientes	Qué hace la herramienta; cómo ayuda; qué beneficios reales esperan; privacidad y seguridad; qué no hace.	Consentimientos informados, charlas informativas, folletos, material visual sencillo,	Antes del uso clínico o piloto; durante la prueba piloto; al finalizar

		entrevistas si se hace estudio piloto.	con resumen de resultados.
Decisores de implementación	Datos de rendimiento técnico; comparativas de costo-beneficio; requerimientos de hardware/infraestructura; impacto clínico y operacional; riesgos y mitigaciones; plan de adopción.	Informes ejecutivos, reuniones con dirección hospitalaria, presentaciones de ROI, dashboard de métricas, documentación clara.	Al inicio (planificación), cuando haya resultados intermedios que apoyen el valor, antes de escalado / adopción institucional.
Reguladores sanitarios	Evidencia de validación clínica; cumplimiento normativo; auditoría; privacidad; transparencia; seguridad; estándares de salud.	Informes técnicos formales; documentación metodológica; cumplimiento de regulaciones locales/nacionales; reuniones si necesario; certificaciones si aplican.	En la fase de diseño del sistema, antes de pruebas clínicas, en momentos de entrega para cumplimiento legal, en instancias regulatorias.
Proveedores de datos / tecnología	Especificaciones de datos necesarias; calidad esperada; acuerdos de licencia; estándares técnicos; compatibilidad; posibles colaboraciones.	Reuniones técnicas; contratos / acuerdos; documentación técnica; workshops si aplican; feedback mutuo.	Al definir datasets o fuentes de datos, durante la integración de hardware/software, cuando haya necesidades de cambio.
Expertos del dominio	Validación clínica; participación en revisión de casos y explicaciones;	Grupos focales, revisiones clínicas de los modelos,	Durante prototipos, durante evaluaciones

	aportes metodológicos; feedback científico; posibles co-autores o colaboración académica.	talleres interdisciplinarios, seminarios, revisiones de literatura conjunta.	intermedias y validaciones, al presentar resultados finales, en publicación si aplica.
Sponsors / financiadores	Estado del proyecto; cumplimiento de hitos; evidencia de impacto técnico y clínico; uso de recursos; visibilidad / publicaciones; riesgos/retos.	Reportes de progreso periódicos, reuniones de seguimiento, presentaciones ejecutivas, resúmenes visuales, presupuestos.	Con regularidad definida (por ejemplo, al inicio, a mitad, al finalizar; también cuando haya hitos técnicos clave).
Implementadores técnicos	Requisitos técnicos; acceso a datos; documentación y estándares; herramientas de desarrollo; soporte técnico/herramientas; pruebas de rendimiento; integración.	Repositorios de código, documentación técnica, reuniones de equipo, code reviews, sprints de desarrollo, testings.	Continuamente durante el desarrollo; antes de integraciones principales; cuando haya nuevas versiones o cambios grandes; antes de versiones de evaluación.

5. DOCUMENTO DE ALCANCE DEL PROYECTO

- **Objetivo general:**

Desarrollar y evaluar un prototipo de inteligencia artificial capaz de clasificar arritmias cardíacas en cinco clases AAMI a partir de espectrogramas de señales de ECG, utilizando CNNs preentrenadas y técnicas de explicabilidad, logrando un rendimiento mínimo de F1 macro ≥ 0.80 en conjunto de prueba independiente por paciente.

- **Objetivos específicos:**

1. Preparar y segmentar el dataset MIT-BIH Arrhythmia Database según protocolo AAMI, garantizando separación por paciente para evitar fuga de información.
2. Transformar señales crudas de ECG en espectrogramas reproducibles, asegurando uniformidad y normalización para entrenamiento.
3. Entrenar un modelo baseline con ResNet50 y un modelo ligero con MobileNetV2, empleando transferencia de aprendizaje y regularización, con reportes de métricas de validación.
4. Incorporar Grad-CAM para explicar las decisiones de los modelos.
5. Evaluar el prototipo en conjunto de prueba independiente por paciente, con métricas F1 macro, verificando que se cumplen los criterios de éxito establecidos.
6. Desarrollar una aplicación web para la implementación y uso de esta herramienta que permita el diagnóstico médico.

- **Alcance incluido:**

- Procesamiento de señales de ECG del MIT-BIH Arrhythmia Database.
- Transformación de señales 1D en espectrogramas 2D mediante STFT.
- Implementación y entrenamiento de dos CNNs preentrenadas: ResNet50 (baseline) y MobileNetV2 (modelo ligero).
- Clasificación en cinco clases AAMI: N, SVEB, VEB, F, Q.
- Incorporación de explicabilidad con Grad-CAM, generando mapas de activación sobre espectrogramas.
- Evaluación cuantitativa con métricas estándar: F1 macro y sensibilidad por clase.
- Documentación final e informe técnico con resultados, interpretaciones y limitaciones.

- **Alcance excluido:**

- Uso de otros datasets distintos al MIT-BIH.
- Inclusión de técnicas de ensemble, optimización exhaustiva de hiperparámetros o arquitecturas fuera de ResNet50/MobileNetV2.
- Validación clínica en entornos hospitalarios o pruebas con pacientes reales.

- **Criterios de aceptación:**

- Correctitud funcional: el pipeline debe transformar ECGs en espectrogramas y clasificarlos automáticamente en las cinco clases AAMI.

- Calidad mínima de desempeño: alcanzar F1 macro ≥ 0.80 y recall ≥ 0.70 en clases minoritarias (SVEB y VEB) en el conjunto de prueba independiente por paciente.
- Explicabilidad: visualización de casos de Grad-CAM deben mostrar coherencia clínica, resaltando regiones vinculadas a otras anomalías relevantes.
- Reproducibilidad: el código debe estar documentado, versionado y ejecutable con las dependencias especificadas.
- Entrega documental: informe técnico con metodología, resultados, limitaciones, junto con slides de presentación.

6. CRONOGRAMA ÁGIL



Roles Scrum Definidos:

Product Owner (PO):	Luis Andress Bustamante Jimenez
Scrum Master:	Damaris Irene Alarcón Vallejo
Development Team:	Luis Andress Bustamante Jimenez / Damaris Irene Alarcón Vallejo

Ceremonias:

- **Sprint Planning:** Inicio de cada sprint, definición de alcance.
- **Daily Scrum:** 15 min diarios.
- **Sprint Review:** Demo al finalizar sprint.
- **Sprint Retrospective:** Reflexión y mejoras.

	Sprint Goal:	Historia de Usuario:	Definition of Done:	Riesgos & Mitigación:
SPRINT 1 – Investigación y Diseño	Tener el entorno de desarrollo listo, dataset procesado y la arquitectura técnica definida	Como Data Scientist, quiero explorar y procesar el dataset MIT-BIH para	Dataset procesado: Anotaciones correctamente mapeadas a clases AAMI,	Riesgo: Dataset con ruido o clases desbalanceadas. Mitigación: Usar técnicas de

	para comenzar el modelado.	<p>extraer clases AAMI.</p> <p>Como ML Engineer, quiero definir y documentar la arquitectura de red neuronal (CNN) que usaré.</p> <p>Como Investigador, quiero revisar papers y benchmarks para fundamentar la metodología.</p> <p>Como Médico, quiero poder dar un diagnóstico robusto y detectado oportunamente.</p>	<p>estructurado por paciente.</p> <p>Revisión bibliográfica: referencias clave resumidas.</p>	rebalanceo y explorar otros registros manualmente.
SPRINT 2 – Desarrollo Core	Preprocesar señales, generar espectrogramas STFT .	<p>Como Data Scientist, quiero segmentar las señales en ventanas de ECG (5–10 s).</p> <p>Como Ingeniero de datos, quiero transformar las ventanas en espectrogramas STFT en 2D.</p> <p>Como ML Engineer, quiero entrenar un modelo ResNet50 sobre espectrogramas como baseline.</p>	<p>Segmentación: Base de datos de pacientes dividida en train y test.</p> <p>Espectrogramas: 100% de ventanas convertidas a imágenes.</p> <p>Modelo ResNet50: Entrenamiento completado, F1 inicial ≥ 0.80 en validación.</p> <p>Pruebas: Al menos 3 pruebas unitarias para verificar</p>	<p>Riesgo: Fallas al generar espectrogramas válidos.</p> <p>Mitigación: Ajustar STFT y revisar espectros visualmente.</p> <p>Riesgo: Rendimiento inicial pobre (<60% F1).</p> <p>Mitigación: Aplicar técnicas de optimización.</p>

			consistencia de procesamiento.	
SPRINT 3 – Optimización y Validación	Optimizar el modelo, entrenar versión eficiente (MobileNetV2), integrar Grad-CAM y diseñar interfaz básica.	<p>Como ML Engineer, quiero entrenar un modelo ligero MobileNetV2 para mejorar eficiencia.</p> <p>Como Científico de datos, quiero integrar Grad-CAM para visualizaciones explicables.</p> <p>Como UX Designer, quiero diseñar una interfaz simple para mostrar espectros y resultados.</p>	<p>MobileNetV2: Entrenado con resultados comparables al baseline ($F1 \geq 0.75$).</p> <p>Tuning: Grid o búsqueda bayesiana completada, mejoras documentadas.</p> <p>Grad-CAM: Mapa de calor sobre al menos 20 casos explicando decisiones del modelo.</p> <p>UI básica: Notebook o web simple que muestre imagen, predicción y visualización Grad-CAM.</p>	<p>Riesgo: Sobreajuste en clases frecuentes.</p> <p>Mitigación: Regularización, early stopping y validación inter-paciente.</p> <p>Riesgo: Grad-CAM produce visualizaciones poco interpretables.</p> <p>Mitigación: Probar distintas capas y compararlas con morfología de ECG.</p>
SPRINT 4 – Finalización y Documentación	Evaluar el sistema final, documentar resultados científicamente y preparar presentación final y el prototipo de implementación.	<p>Como Data Scientist, quiero evaluar el modelo en test set con métricas clínicas.</p> <p>Como Tester, quiero realizar pruebas integrales de performance y estabilidad.</p> <p>Como Product Owner, quiero</p>	<p>Evaluación: $F1$ macro ≥ 0.80; recall ≥ 0.70 en clases SVEB y VEB.</p> <p>Informe final: Documento técnico.</p> <p>Prototipo y slides: Notebook ejecutable o video explicativo y presentación.</p>	<p>Riesgo: Informe y demo se retrasan por fallos en última etapa.</p> <p>Mitigación: Congelar versión estable al final de semana 5 para asegurar tiempo de redacción y edición.</p>

		generar un prototipo reproducible y una presentación clara.		
--	--	---	--	--

7. PLAN DE RECURSOS

• Recursos Humanos

Equipo Core:

Project Manager (PM): 1 persona (20 h/semana).

Responsabilidades: planificación general del proyecto, seguimiento mediante metodologías ágiles (Scrum) y gestión de la relación con los stakeholders.

Data Scientist: 1 persona (20 h/semana cada una).

Responsabilidades: desarrollo de modelos, extracción de características, ajuste de hiperparámetros y análisis de métricas.

Estimación de Horas por Sprint:

Sprint	Horas totales
Sprint 1	240 h totales
Sprint 2	300 h totales
Sprint 3	280 h totales
Sprint 4	220 h totales

• Recursos Técnicos

Categoría	Detalle
Hardware	Equipos locales: laptops/PCs para análisis exploratorio.
Software	Licencias y Herramientas: Python + librerías open source, Jupyter Notebook, Google Colab, PyCharm
Colaboración	GitHub
Datos	Datasets Públicos: MIT-BIH, 12-lead ECG arrhythmia.

• Recursos Financieros

Presupuesto Estimado:

Este proyecto es de carácter exclusivamente investigativo, por lo tanto, no implica costos significativos más allá de los recursos básicos ya disponibles.

Categoría	Costo Estimado (USD)
<i>Equipos computacionales (existentes)</i>	\$0 (ya disponibles)
<i>Infraestructura cloud (uso puntual)</i>	\$65 (máximo estimado)
<i>Conectividad / Internet</i>	\$100 (uso doméstico)
Total estimado del proyecto	\$165 USD

- **Plan de Adquisición y Gestión de Recursos**

Recursos Humanos: Participación de investigadores asignados al proyecto con bajo esquema académico y voluntario.

Infraestructura Técnica: Se utilizará infraestructura local y servicios cloud gratuitos o escalables bajo demanda. No se contempla compra de hardware adicional.

Software: Herramientas basadas en tecnologías open source y plataformas gratuitas. No se adquieren licencias comerciales.

Datos: Se trabajará exclusivamente con datasets abiertos y validados en la literatura científica. No se requiere compra de bases de datos ni contratación para anotación manual.

8. HITOS Y ENTREGABLES

Cada hito será revisado por el equipo mediante validación cruzada y control de calidad de entregables.

Hito	Entregables	Formato	Criterios de Aceptación	Riesgo	Mitigación
Hito 1	Dataset MIT-BIH descargado y organizado. Mapeo de anotaciones a clases AAMI. Revisión bibliográfica.	Archivos .csv. Informe técnico .pdf	Dataset estructurado correctamente con clases AAMI balanceadas.	Inconsistencia o ruido en los datos.	Revisión manual de los datos, de ser necesario.
Hito 2	Ventanas de ECG segmentadas y normalizadas. Partición de datos en train, test y val.	Archivos .csv, scripts .ipynb	Pipeline reproducible.	Segmentación incorrecta.	Comparación contra anotaciones originales.

Hito 3	Espectrogramas STFT generados	Archivos .png. Scripts Python	Espectrogramas generados para el 100% de la muestra	Parámetros STFT inadecuados	Ajuste de ventana y solapamiento
Hito 4	Modelo basado en ResNet50 entrenado	Cuaderno de entrenamiento. ipynb	F1-score ≥ 0.70 en validación	Bajo rendimiento en clases minoritarias	Técnicas de balanceo y data augmentation
Hito 5	Entrenamiento de MobileNetV2	Modelos entrenados, tabla comparativa .csv	F1-score ≥ 0.70 ; inferencia <1 segundo por muestra	Diferencias de rendimiento no justificadas	Análisis detallado de precisión y eficiencia
Hito 6	Implementación de Grad-CAM	Imágenes .png, notebooks .ipynb	Explicabilidad coherente con morfología ECG en al menos 20 casos	Mapas poco interpretables	Selección de capas adecuadas y ajuste de resolución
Hito 7	Evaluación final sobre test set	Reporte .pdf, métricas en .csv, matriz de confusión	F1 macro ≥ 0.80 ; recall ≥ 0.75 en SVEB y VEB	Sobreajuste tras optimización	Regularización, early stopping y validación cruzada
Hito 8	Informe técnico final, presentación, manual de uso, prototipo ejecutable	Informe .pdf, slides .pptx, código + demo .ipynb/.mp4	Documentación clara y completa, demo funcional	Falta de tiempo para pulido final	Congelamiento del modelo en semana 5, redacción paralela

9. Referencias Bibliográficas:

- [1] B. Zheng, W. Luo, M. Zhang, H. Jin et al., "Arrhythmia classification based on multi-input convolutional neural network with attention mechanism," 2025. [Online]. Available: <https://doi.org/10.1371/journal.pone.0326079>
- [2] Y. D. Daydulo et al., "Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals," 2023. [Online]. Available: <https://doi.org/10.1186/s12911-023-02326-w>
- [3] Í. F. Di Paolo & A. R. G. Castro, "Intra- and Interpatient ECG Heartbeat Classification Based on Multimodal Convolutional Neural Networks with an Adaptive Attention Mechanism," 2024. [Online]. Available: <https://doi.org/10.3390/app14209307>

- [4] F. Zhou & D. Fang et al., "Multimodal ECG heartbeat classification method based on a convolutional neural network embedded with FCA," 2024. [Online]. Available: <https://doi.org/10.1038/s41598-024-59311-0>
- [5] F. Zhang, M. Li, L. Song et al., "Multi-classification method of arrhythmia based on multi-scale residual neural network and multi-channel data fusion," *Frontiers in Physiology*, vol. 14, 2023. [Online]. Available: <https://doi.org/10.3389/fphys.2023.1253907>
- [6] Y. Wu et al., "Res-BiANet: A Hybrid Deep Learning Model for Arrhythmia Detection Based on PPG Signal," 2024. [Online]. Available: <https://doi.org/10.3390/electronics13030665>
- [7] "HARDC: A novel ECG-based heartbeat classification method to detect arrhythmia using hierarchical attention based dual structured RNN with dilated CNN," 2023. [Online]. Available: <https://arxiv.org/abs/2303.06020>
- [8] H. Sun et al., "Arrhythmia Classifier Using Convolutional Neural Network with Adaptive Loss-aware Multi-bit Networks Quantization," 2022. [Online]. Available: <https://arxiv.org/abs/2202.12943>
- [9] H. Tung, C. Zheng, X. Mao, D. Qian, "Multi-Lead ECG Classification via an Information-Based Attention Convolutional Neural Network," 2020. [Online]. Available: <https://arxiv.org/abs/2003.12009>
- [10] "A novel hybrid CNN-transformer model for arrhythmia detection without R-peak identification using Stockwell transform," 2025. [Online]. Available: <https://www.nature.com/articles/s41598-025-92582-9>
- [11] T. Talukder Showrav et al., "EXGnet: a single-lead explainable-AI guided multiresolution network with train-only quantitative features," 2025. [Online]. Available: <https://arxiv.org/abs/2506.12404>
- [12] F. Zhou et al., "Classification of multi-lead ECG based on multiple scales and hierarchical feature CNNs," 2025. [Online]. Available: <https://www.nature.com/articles/s41598-025-94127-6>
- [13] N. M. Malleswari, et al., "Deep learning-assisted arrhythmia classification using 2-D ECG spectrograms," 2024. [Online]. Available: https://www.researchgate.net/publication/387580665_Deep_learning-assisted_arrhythmia_classification_using_2-D_ECG_spectrograms
- [14] M. S. Song et al., "Comparative study of time-frequency transformation methods for ECG signal classification," 2024. [Online]. Available:

<https://www.frontiersin.org/journals/signal-processing/articles/10.3389/frsip.2024.1322334/full>

[15] H. S. Mansour et al., “Enhanced ECG Arrhythmia Diagnosis with Deep Learning and Selective Attention Mechanism,” 2025. [Online]. Available:

https://www.researchgate.net/publication/391592637_Enhanced_Electrocardiogram_Arrhythmia_Diagnosis_with_Deep_Learning_and_Selective_Attention_Mechanism

[16] C. Wang et al., “Analysis of Cardiac Arrhythmias Based on ResNet-ICBAM,” 2025.

[Online]. Available: <https://www.mdpi.com/1424-8220/25/3/661>

[17] A. Kirkbas et al., “Automated ECG Arrhythmia Classification Using Feature Images with Common Matrix Approach-Based Classifier,” 2025. [Online]. Available:

<https://www.mdpi.com/1424-8220/25/4/1220>

[18] M. A. Talukder et al., “A hybrid cardiovascular arrhythmia disease detection using

ConvNeXt-X models,” 2024. [Online]. Available: <https://www.nature.com/articles/s41598-024-81992-w>

[19] J. Beck & A. John, “Explainable AI (XAI) for Arrhythmia detection from

electrocardiograms,” 2025. [Online]. Available: <https://arxiv.org/abs/2508.17294>

[20] D. Kwon et al., “Deep learning-based prediction of atrial fibrillation from polar transformed time-frequency electrocardiogram,” 2025. [Online]. Available:

<https://pubmed.ncbi.nlm.nih.gov/40063554/>