

# Diseño de aplicación prototipo para la detección y clasificación de arritmia usando métodos de machine learning a partir de ECGs

1<sup>st</sup> Luis M. Cárcamo Morales

Departamento de Ingeniería de Sistemas  
Universidad del Norte  
Barranquilla, Colombia  
mlcarcamo@uninorte.edu.co

2<sup>nd</sup> Johnny Villegas García

Departamento de Ingeniería de Sistemas  
Universidad del Norte  
Barranquilla, Colombia  
jvillegasd@uninorte.edu.co

3<sup>rd</sup> Fabián Osorio Iriarte

Departamento de Ingeniería de Sistemas  
Universidad del Norte  
Barranquilla, Colombia  
fabiani@uninorte.edu.co

Tutor: Alfonso M. Mancilla Herrera

Tutor del curso: Wilson Nieto Bernal

**Abstract**—In this paper, we develop a stacked models machine learning approach for arrhythmia classification using the MIT-BIH arrhythmia database and applying different filters and noise-reduction methods. We perform the feature extraction by using RR intervals-related features and the wavelet transform. Our model uses both linear discriminant analysis and multilayer perceptrons. This approach shows an average accuracy of 89%. Sensitivity values of 93%, 94%, 97% and 98% were found for each of the classes and specificity values of 81%, 89%, 85% and 85% for each of the classes. The logical architecture of the project is based on the different services of Amazon Web Services (AWS), which offers us scalability according to the demand that the model requires. Using AWS EC2 to deploy the APIs in charge of supplying the information obtained from the model. In addition, we use a single page application made with Vue Js connected to these APIs and deployed in an AWS S3 bucket to display the information.

**Index Terms**—ecgs, data mining, wavelet transform, arrhythmia, linear discriminant analysis, signals processing

## I. INTRODUCCIÓN

Según la Organización Mundial de La Salud, Las enfermedades cardiovasculares representan alrededor del 31 % de las muertes en el mundo cada año. Estas enfermedades afectan el corazón y muchas veces la falta de un diagnóstico, y por tanto, de un tratamiento adecuado, conlleva a complicaciones. Una señal de alerta sobre posibles trastornos cardiacos es la presencia de arritmias. Las arritmias se dan por un latido irregular, o bien, muy rápido o muy lento. Para la detección de estas, existe un método no invasivo para observar la actividad eléctrica del corazón; este método se conoce como electrocardiograma (ECG).

En la actualidad, existen múltiples métodos basados en machine learning para la automatización del diagnóstico de arritmias, como Artificial Neural Networks (ANN) o K-nearest neighbors (KNN), Extreme Machine Learning [1], Support Vector Machine (SVM) [2]–[5].

En [3] se propone un método basado en Support Vector Machine, utilizando también un enfoque basado en Generalized Discriminant Analysis (GDA). En [4] se propone un

método que utiliza características dinámicas y proyectadas de los ECGs, estas son usadas para ajustar los parámetros de un clasificador SVM, el enfoque de este trabajo está orientado tanto a la clasificación a partir de *clases* como a la clasificación a partir de *sujetos*. [5] propone un método basado en el uso de un kernel radial basis function (RBF) con SVM para la clasificación y un método para la extracción de características de los ECGs basados en los modos variacionales de estos. [6] propone un enfoque bayesiano para la clasificación, usando un enfoque de maximización de la probabilidad posterior, este trabajo también hace uso de las características morfológicas de los ECGs. [1] usa un enfoque basado en Extreme Learning Machine (ELM), que es comparado con enfoques basados en k-Nearest Neighbors y SVM; en este trabajo, las características obtenidas de las señales de los ECG son obtenidas usando Wavelet Transform (WT). [2] propone un clasificador basado en SVM que recibe como datos de entrada los parámetros obtenidos de una proyección de los ECGs, mediante una variante de la transformada de Stockwell. En [7] se emplea un método basado en Linear Discriminant Classification; los datos utilizados corresponden a características morfológicas de los ECGs. La totalidad de los trabajos citados utilizan el dataset MIT-BIH arrhythmia databaset. Este dataset es bastante usado en este tipo de trabajos relacionados con la clasificación de arritmias [8].

Este documento presenta primero una revisión del marco teórico y de la literatura existente sobre el tema. Posteriormente, se presenta la metodología empleada para el desarrollo del modelo. Esto incluye el procesamiento de los datos, la extracción y selección de características. Finalmente, se muestran los resultados para distintas configuraciones del modelo y la mejor configuración, de la que se hace una comparación con trabajos existentes.

## II. DESCRIPCIÓN DEL PROBLEMA

En la actualidad muchos consultorios médicos y hospitales cuentan con sistemas especializados en el registros de

electrocardiogramas (ECG) para el análisis y diagnóstico de arritmias en pacientes, sin embargo no se ha evidenciado del uso de sistemas que buscan agilizar y apoyar a los doctores en el diagnóstico certero de la presencia de esta enfermedad en los ECG suministrados. En la Fig. 1 se presenta una representación en árbol de la descripción del problema.

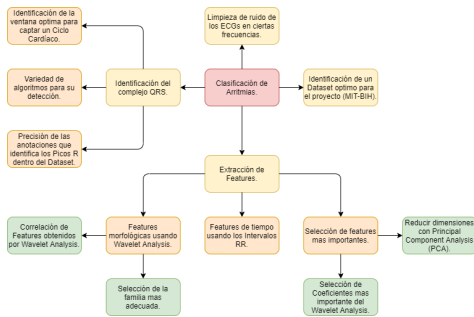


Fig. 1: Mapa del problema

### III. OBJETIVOS

#### A. Objetivo General

Desarrollar una aplicación prototipo que permita analizar y clasificar las arritmias a partir de la lectura de señales, en electrocardiogramas, disponibles en el dataset MIT-BIH arrhythmia database.

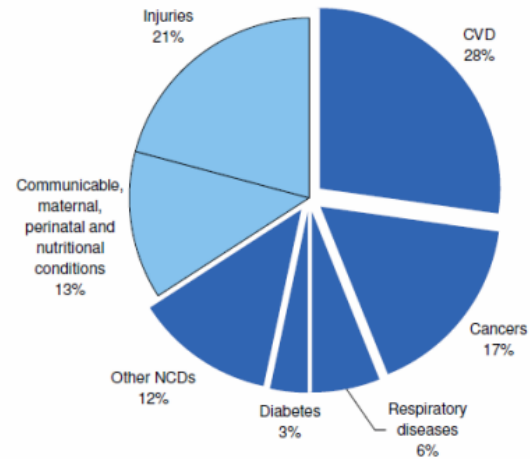
#### B. Objetivos Específicos

- Presentar los resultados de este proyecto para publicación en una revista ISI o SCOPUS.
- Crear un servicio basado en la nube, para ofrecer el modelo de clasificación como servicio por medio de AWS.
- Clasificación de arritmias mediante un modelos apilados usando multilayer perceptron y análisis de discriminantes lineales.
- Estudiar el funcionamiento de la librería WFDB y el dataset MIT-BIH arrhythmia database, diseñados para el estudio de arritmias.

### IV. JUSTIFICACIÓN

Las Enfermedades Cardiovasculares (ECV) son la primera causa de mortalidad en el mundo. Según las Estadísticas de Salud Mundial publicadas por la Organización Mundial de la Salud (OMS) en 2012, las enfermedades no transmisibles (ENT) fueron causantes del 63% (36 millones) de 57 millones de muertes ocurridas en el mundo durante 2008. Además, en el mismo período de tiempo, las ECV fueron la primera causa de muerte por ENT con 17 millones de muertes (48% del total de muertes causadas por ENT). En 2010, esta cifra llegó a 18,1 millones y se estima que aumentará a 25 millones en 2030 [9]. Como podemos ver en la figura 2 las enfermedades cardiovasculares en Colombia en el año 2011 fueron responsables de más de un cuarto de las muertes en el país en dicho año por lo tanto este tipo de enfermedades. Por lo tanto, una detección temprana de posibles riesgos de

Proportional mortality (% of total deaths, all ages)



NCDs are estimated to account for 66% of all deaths.

Fig. 2: [9] Enfermedades no transmisibles: perfil epidemiológico de Colombia, según Organización Mundial de la Salud, 2011.

sufrir enfermedades cardiovasculares es indispensable para tratar este tipo de enfermedades que tienen un papel muy importante en la mortalidad de la población mundial y una herramienta indispensable para la detección de este tipo de enfermedades ha sido el electrocardiograma.

El electrocardiograma (ECG) se ha utilizado como un indicador esencial e indispensable para la evaluación de la función cardiovascular más de un siglo. Estudios recientes sobre la variabilidad de la frecuencia cardíaca (VFC) revelan que también se puede obtener mucha más información para reflejar no solo las funciones cardiovasculares sino también muchas funciones no cardiovasculares a partir de firmas indiscernibles incluidas en el ECG/VFC [10]. Nuestro proyecto se basa en crear un modelo que permita automatizar el análisis y/o la lectura de los resultados de estos datos para así lograr una detección rápida y/o temprana de este tipo de enfermedades usando los electrocardiogramas como principal fuente de datos teniendo cuenta que estos nos brindan mucha información sobre el paciente sobre las posibilidades de llegar a tener este tipo de enfermedades logrando salvar vidas con la agilización de este proceso.

### V. MARCO TEÓRICO

#### A. Multilayer Perceptrons

El perceptrón multicapa es un tipo de Red Neuronal Artificial (RNA) que nace de la unión de varios perceptrones para poder resolver problemas que no son linealmente separables. La unidad de esta Red Neural es el perceptrón, que es una pequeña RNA que posee una capa de datos de entrada y otra de datos de salidas. El perceptrón tiene la limitación de poder

resolver problemas que son linealmente separables, debido a que el perceptrón busca hallar un Hiperplano tal que:

$$\overline{\mathcal{W}} * \overline{\mathcal{X}} = 0$$

Es decir, trabaja de manera similar como otra técnica de Machine Learning llamada Support Vector Machine (SVM), que es muy usada en la literatura. Estas técnicas tienen una relación más profunda evidenciada en su función de costo, las dos usan la *función Hinge*:

$$Li = \max\{-y_i(\overline{\mathcal{W}} * \overline{\mathcal{X}}), 0\}$$

La *función Hinge del Perceptrón* está desplazada en factor de 1.

Al trabajar con ARN, debemos definir que función de activación  $\Phi(\overline{\mathcal{W}} * \overline{\mathcal{X}})$  usar. Para este proyecto se usó la función de activación de Tangente Hiperbólica:

$$\Phi(v) = \frac{e^{2v} - 1}{e^{2v} + 1}$$

### B. Intervalos RR

El intervalo RR es la distancia que existe entre un pico R perteneciente al complejo QRS hasta el siguiente pico R. Los intervalos RR suelen ser de las características de tiempo más importantes al momento de crear modelos clasificadores de arritmias [3-9], debido a que están relacionados con el ritmo cardiaco.

### C. Linear Discriminant Analysis

El análisis de discriminantes lineales es tanto un método de clasificación como un método de reducción de dimensiones. Este método opera al buscar aumentar la separación entre clases.

Una generalización de este método para clasificación en más de dos clases fue propuesta por [11].

Este método deja una cantidad resultante de *features* igual al valor menor entre  $C - 1$  y  $N$  donde  $C$  es el número de clases y  $N$  es el número de *features*. Por lo que implica una reducción bastante importante en la dimensión del problema en el espacio de características. Estas características resultantes corresponden a una combinación lineal de las características originales.

El planteamiento de LDA como método de clasificación se puede formular de la siguiente manera [12]:

Sea  $K$  el número de clases, y sea  $\mathbf{x}$  un dato cualquiera. Entonces, dado este  $\mathbf{x}$  la clasificación estimada para este es

$$G_k(\mathbf{x}) = \arg \max_k \delta_k(\mathbf{x}),$$

$$\delta_k(\mathbf{x}) = \mathbf{x}^T \cdot \widehat{\Sigma}^{-1} \cdot \overline{\mathbf{x}}_k - \frac{1}{2} \cdot \overline{\mathbf{x}}_k^T \cdot \widehat{\Sigma}^{-1} \cdot \overline{\mathbf{x}}_k + \log \pi_k$$

Donde  $\widehat{\Sigma}$  es la matriz de covarianza estimada. En LDA se maneja el supuesto de que la matriz de covarianza es igual para todas las clases. Es decir,  $\widehat{\Sigma}_k = \widehat{\Sigma}$ ,  $k = 1 \dots K$ . Esta matriz de covarianza se calcula de la siguiente manera:

$$\widehat{\Sigma} = \sum_{k=1}^K \sum_{g_i=k} \frac{1}{(N - K)} \cdot (\mathbf{x}_i - \overline{\mathbf{x}}_k) \cdot (\mathbf{x}_i - \overline{\mathbf{x}}_k)^T,$$

$\overline{\mathbf{x}}_k$  corresponde a la media estimada de cada clase, calculada usando los datos de entrenamiento; y  $\pi_k$  corresponde a la prior de la clase  $k$ . Esta se calcula por defecto a partir de la proporción de datos de cada clase. En este trabajo, el número de clases en donde se aplica LDA es de 3, por lo que el número de características resultantes es de dos. Estas dos vienen a ser una combinación lineal de las características que resultan luego de la selección de las más relevantes para la clasificación.

## VI. REVISIÓN DE LA LITERATURA

En la siguiente tabla se muestra un listado exhaustivo de trabajos relacionados con el tema aquí presentado.

| Artículo  | Autores   | Año  | Palabras Clave  | Fuente     |
|---|---|------|---|------------|
| A Multi-label Learning Method to Detect Arrhythmia Based on 12-Lead ECGs [13]                                       | Zhu Jinjing, Xin Kaifa, Zhao Qingqing & Zhang Yue   | 2019 | Arrhythmia; DCGANs; ECG; Multi-label learning; SE-ResNet; WGAN-GP   | Springer   |
| Multi-stage SVM approach for cardiac arrhythmias detection in short single-lead ECG recorded by a wearable device   | Smisek Radovan, Hejc Jakub, Ronzhina Marina, Nemcova Andrea, Marsanova Lucie, Kolarova Jana, Smital Lukas & Vitek, Martin | 2018 | ECG; SNR estimation; SVM; atrial fibrillation; beat classification; feature selection; genetic algorithm  | IOPscience |
| Classification of hyperspectral remote sensing images with support vector machine [14]                              | Melgani Farid & Bruzzone Lorenzo  | 2004 | Classification; Feature reduction; Hughes phenomenon; Hyperspectral images; Multiclass problems; Remote sensing; Support vector machines (SVMs) | IEEE       |
| ECG beat classification method for ECG printout with Principle Components Analysis and Support Vector Machines [15] | Thanapatay Dusit, Suwansaroj Chaiwat & Thanawattano Chusak  | 2010 | Component; ECG classification; ECG printout; Principle components analysis; Support vector machines   | ICEIE 2010 |
| On the learnability and design of output codes for multiclass problems [16]   | Crammer Koby & Singer Yoram   | 2002 | Multiclass categorization; Output coding; SVM   | Springer   |
| Heartbeat classification using projected and dynamic features of ECG signal [4]                                     | Chen Shanshan, Hua Wei, Li Zhi, Li Jian & Gao Xingjiao  | 2017 | Compressed sensing; Heartbeat classification; Projection matrix; Support vector machine   | Elsevier   |
| Support vector machine based arrhythmia classification using reduced features [17]                                  | Song Mi Hye, Lee Jeon, Cho Sung Pil, Lee Kyoung Joung & Yoo Sun Kook  | 2005 | Arrhythmia classification; Linear discriminant analysis; Reduction of feature dimension; Support vector machine; Wavelet transform              | Springer   |
| A Diagnostic System for Detection of Atrial and Ventricular Arrhythmia Episodes from Electrocardiogram [5]          | Chetan Apoorv, Tripathy Rajesh K. & Dandapat Samarendra   | Año  | Cardiac arrhythmia; Classifiers; Distribution entropy; Electrocardiogram; Sample entropy; Variational mode decomposition                        | Springer   |
| Machine learning approach to recognize ventricular arrhythmias using VMD based features [18]                        | Mohanty Monalisa, Biswal Pradyut & Sabut Sukanta  | 2020 | Classification; ECG; Features; Machine learning; VMD; Ventricular arrhythmias   | Springer   |
| Automatic classification of heartbeats using ECG morphology and heartbeat interval features [6]                     | De Chazal Philip, O'Dwyer Maria & Reilly Richard B.   | 2004 | Electrocardiogram (ECG); Heartbeat classifier; Linear discriminant analysis; Statistical classifier model                                       | IEEE       |
| Support Vector Machines for Multi-Class Pattern Recognition [19]  | Weston J & Watkins C  | 1999 |   | DBLP       |
| The impact of the MIT-BIH arrhythmia database [8]   | Moody G. B. & Mark R. G.  | 2001 | Databases; Electrocardiography; History; Heart; Rhythm; Laboratories; Aging; Materials testing; Detectors; Manufacturing                        | IEEE       |

|  |   |      |   |  |
|--|---|------|---|--|
| An Algorithm for Detection of Arrhythmia [20]  | Rahman Mujeeb & Nasor Mohamed   | 2012 | Electrocardiogram; QRS Complex; Arrhythmia; Fiducial Point; Sampling Rate; Peak Valley Checker  | SAP (Scientific & Academic Publishing) |
| Artificial Neural Network based cardiac arrhythmia classification using ECG signal data [21]   | Jadhav S. M., Nalbalwar S. L. & Ghatol, Ashok   | 2010 | Accuracy; ECG arrhythmia; Momentum learning rule; Multilayer perceptron classification; Sensitivity; Specificity  | ICEIE 2010                             |
| ECG arrhythmia classification using spearman rank correlation and support vector machine [22]  | Khare Shreya, Bhandari Akshay, Singh Saurabh & Arora Anuja  | 2012 | Arrhythmia; PCA; SVM; Spearman Rank Correlation   | Springer                               |
| An Arrhythmia Classification Method Based on Selected Features of Heart Rate Variability Signal and Support Vector Machine-Based Classifier [23]                                 | Yaghouby F. & Ayatollahi A.   | 2009 | Heart rate variability signal; Arrhythmia classification; Support vector machine; Forward floating feature selection  | Springer                               |
| ECG feature extraction based on the bandwidth properties of variational mode decomposition [24]  | Mert Ahmet  | 2016 | Arrhythmia; classification; electrocardiogram; variational mode decomposition   | IOPscience                             |
| Performance comparison of Variational Mode Decomposition over Empirical Wavelet Transform for the classification of power quality disturbances using Support Vector Machine [25] | Aneesh C., Kumar Sachin, Hisham P. M. & Soman, K. P.  | 2015 | AM-FM signal; Alternate direction method of multipliers; Empirical mode decomposition; Empirical wavelet transform; Intrinsic mode function; Variational mode decomposition | Elsevier                               |
| Comparative study between artificial metaplasticity learning algorithm and other machine learning techniques: Application for cardiac arrhythmias classification [26]            | Benchaib Yasmine  | 2020 | AMMLP; Artificial Metaplasticity; Artificial neural network; Cardiac arrhythmias; Classification; MIT-BIH; Multi layer perceptron; Plasticity                               | Springer                               |
| Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification [27]  | Christov Ivaylo, Gómez-Herrero Gèrman, Krasteva Vessela, Jekova Irena, Gotchev Atanas & Egiazarian Karen                | 2006 | Automatic heartbeat classification; Holter ECG analysis; Matching Pursuits; Morphological ECG descriptors   | Elsevier                               |
| Variational mode decomposition [28]  | Dragomiretskiy Konstantin & Zosso Dominique   | 2014 | AM-FM; Fourier transform; Hilbert transform; Wiener filter; augmented Lagrangian; mode decomposition; spectral decomposition; variational problem                           | IEEE                                   |
| Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network [29]   | Hannun Awni Y., Rajpurkar Pranav, Haghpanahi Masoumeh, Tison Geoffrey H., Bourn Codie, Turakhia Mintu P. & Ng Andrew Y. | 2019 | Arrhythmias; Machine learning   | Nature                                 |
| Novel Methodology for Cardiac Arrhythmias Classification Based on Long-Duration ECG Signal Fragments Analysis  | Plawiak Pawel & Abdar Moloud  | 2020 | -   | Springer                               |

|  |  |      |   |               |
|--|--|------|---|---------------|
| Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal [3]       | Asl Babak Mohammadzadeh, Setarehdan Seyed Kamaledin & Mohebbi Maryam | 2008 | Arrhythmia classification; Generalized discriminant analysis; Heart rate variability; Nonlinear analysis; Support vector machine  | ScienceDirect |
| ECG arrhythmias recognition system based on independent component analysis feature extraction [30]                       | Jiang Xing, Zhang Liqing, Zhao Qibin & Albayrak Sahin                | 2006 | Electrocardiography; Independent component analysis; Feature extraction; Heart beat; Wavelet analysis; Wavelet transforms; Support vector machines; Support vector machine classification; Computer simulation; Spatial databases | IEEE          |
| Morphological heart arrhythmia detection using hermitian basis functions and kNN classifier [31]                         | Karimifard S., Ahmadian A., Khoshnevisan M. & Nambakhsh M. S.        | 2006 | ECG beat; Hermitian basis function; Morphological arrhythmia; kNN classifier  | IEEE          |
| Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine [32] | Yang Weiyi, Si Yujuan, Wang Di & Guo Buhao                           | 2018 | Arrhythmia recognition; Cardiovascular diseases; Deep learning; Noise robustness; Principal component analysis network  | Elsevier      |
| Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition [33]                                       | Osowski Stanislaw, Hoai Linh Tran & Markiewicz Tomasz                | 2004 | Combination of classifiers; Expert system; Heartbeat recognition; Support vector machine  | IEEE          |
| Computer-aided arrhythmia diagnosis with bio-signal processing: A survey of trends and techniques [34]                   | Dinakarrao Sai Manoj Pudukotai, Jantsch Axel & Shafique Muhammad     | 2019 | Arrhythmia detection; Computer-aided diagnosis; Electrocardiogram (ECG); Health-care; Machine learning; Neural networks; Support-vector machine   | ACM           |
| Robust algorithm for arrhythmia classification in ECG using extreme learning machine. [35]                               | Kim Jinkwon, Shin Hang Sik, Shin Kwangsoo & Lee Myoung-ho            | 2009 | Biomaterials; Biomedical Engineering and Bioengineering; Biomedical Engineering/Biotechnology; Biotechnology  | BMC           |
| Support vector machine for arrhythmia discrimination with wavelet-transform-based feature selection [36]                 | Millet-Roig J., Ventura-Galiano R., Chorro-Gascó F. J. & Cebrián A.  | 2000 | Support vector machines; Support vector machine classification; Logistics; Cancer; Pattern classification; Cardiology; Wavelet transforms; Rhythm; Electrocardiography; Neural networks   | IEEE          |
| Heartbeat classification using normalized RR intervals and morphological features [7]                                    | Lin Chun Cheng & Yang, Chun Min                                      | 2014 |   | Hindawi       |
| Electrocardiogram Beat Classification Using Support Vector Machine and Extreme Learning Machine [37]                     | Banupriya C. V. & Karpagavalli S.                                    | 2014 | Electrocardiogram; Extreme Learning Machine; Support Vector Machine; Wavelet  | Springer      |

|  |  |      |  |                            |
|--|--|------|--|----------------------------|
| A comparison of methods for multiclass support vector machines [38]  | Hsu Chih Wei & Lin Chih Jen  | 2002 | Decomposition methods; Multiclass classification; Support vector machines (SVMs)   | IEEE                       |
| Application of support vector machine for the detection of P- and T-waves in 12-lead electrocardiogram [39]                        | Mehta S. S. & Lingayat, N. S.  | 2009 | ECG; P-wave; QRS complex; SVM; T-wave  | Elsevier                   |
| A Novel Approach to Classify Cardiac Arrhythmia Using Different Machine Learning Techniques [40]                                   | Jain Parag, Arjun Babu C. S., Mohandoss Sahana, Anisham Nidhin, Gadade Shivakumar, Srinivas A. & Mohan Rajasekar | 2020 | Machine learning; ECG recordings; Cardiac Arrhythmia; Ensemble methods; Hard voting; Healthcare; Feature selection   | Springer                   |
| Optimized time–frequency features and semi-supervised SVM to heartbeat classification [2]  | Lekhal Redouane, Zidelmal Zahia & Ould-Abdesslam Djaffar   | 2020 | ECG beats classification; S-transform with compact support kernel; Semi-supervised SVM; Time–frequency features  | Springer                   |
| Fast multiclass SVM classification using decision tree based one-against-all method [41]   | Arun Kumar M. & Gopal M.   | 2010 | Decision tree; multiclass classification; one-against-all; support vector machines (svms)  | Springer                   |
| An introduction to support vector machines: and other kernel-based learning methods [42]   | Cristianini N & Shawe-Taylor J   | 2000 | Computer Science; Pattern Recognition and Machine Learning; Computational Biology and Bioinformatics   | Cambridge University Press |
| Wrapper method for feature selection to classify cardiac arrhythmia [43]   | Mustaqeem Anam, Anwar Syed Muhammad, Majid Muhammad & Khan Abdul Rashid  | 2017 | Arrhythmias; Classification; Feature selection; Random Forest; Wrapper method  | IEEE                       |
| Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features [44] | Khazae A. & Ebrahimzadeh, A.   | 2010 | ECG beat classification; Genetic algorithm; Multitaper method; Non-parametric PSD estimation methods; Parameter optimization; SVM  | Elsevier                   |
| Classification of electrocardiogram signals with support vector machines and extreme learning machine [1]                          | Karpagachelvi S., Arthanari M. & Sivakumar, M.   | 2012 | Electrocardiogram (ECG) signals classification; Extreme learning machine; Feature detection; Feature reduction; Generalization capability; Model selection issue; Support vector machine | Springer                   |
| ECG arrhythmia analysis by multicategory support vector machine [45]   | Khadtare Mahesh S. & Sahambi J. S.   | 2004 |  | Springer                   |
| QRS template matching for recognition of ventricular ectopic beats [46]  | Krasteva Vessela & Jekova Irena  | 2007 | Beat classification; Cross-correlation; ECG Holters; ECG diagnostic systems; Premature ventricular contractions PVC; QRS morphology; Time-frequency ECG analysis                         | Springer                   |
| Arrhythmia Recognition and Classification Using ECG Morphology and Segment Feature Analysis [13]                                   | Zhu Wenliang, Chen Xiaohe, Wang Yan & Wang Lirong  | 2019 | ECG morphology; SVM; classification; feature extraction  | IEEE                       |

|  |  |      |  |          |
|--|--|------|--|----------|
| Heartbeat classification using morphological and dynamic features of ECG signals [47]      | Ye Can, Vijaya Kumar B. V.K. & Coimbra Miguel Tavares  | 2012 | Heartbeat classification; independent component analysis; support vector machine; wavelet transform                            | IEEE     |
| Recognition of heartbeats using Support Vector Machine networks - A comparative study [48] | Osowski Stanislaw, Linh Tran Haoi & Markiewicz Tomasz  | 2005 |  | Springer |
| Classification of Arrhythmia [49]  | Samad Saleha, Khan Shoab A., Haq Anam & Riaz Amna  | 2014 | Accuracy, arrhythmia; decision tree classifier; k-NN classifier; naive bayes classifier  | IJOEE    |
| Arrhythmia classification using local hölder exponents and support vector machine [50]     | Joshi Aniruddha Rajshekhar, Chandran Sharat, Phadke Sanjay, Jayaraman V. K. & Kulka-rni, B. D. | 2005 | Support Vector Machine; Wavelet Transformation; Normal Sinus Rhythm; Multifractal Formalism; Multiclass Support Vector Machine | Springer |

TABLE I: Revisión de la literatura



## VII. METODOLOGÍA

En este trabajo realizamos una segmentación de cada uno de los registros del dataset, y se aplica a cada ECG filtros para la disminución del ruido. Para cada uno de los registros procesados, realizamos la extracción de características al aplicar la Wavelet Transform (WT), y además, al obtener los intervalos RR y algunas otras propiedades relacionadas con la forma de cada segmento. Para estas características obtenidas, se eliminan todas aquellas que son cuasi-constantes, es decir, aquellas con una varianza menor a cierta cota mínima establecida. También se hace una disminución del espacio de características al eliminar algunas de estas al aplicar la eliminación de las que están altamente correlacionadas. Esto también se hace con una cota, en este caso, máxima de correlación. Posterior a la reducción del espacio de características, entrenamos nuestro modelo y ajustamos los parámetros de estos. En este paso, usamos Linear Discriminant Analysis (LDA) y Multilayer perceptrons (MLP). En el modelo definitivo, primero se hace una clasificación entre Normal (N) y no normal (F, S, V), y posteriormente, para las clasificadas como no normales, aplicamos una nueva clasificación para definir a cuál de las 3 corresponde cada segmento.

En la Fig. 3 se muestra un vistazo general del proceso llevado a cabo en este trabajo.

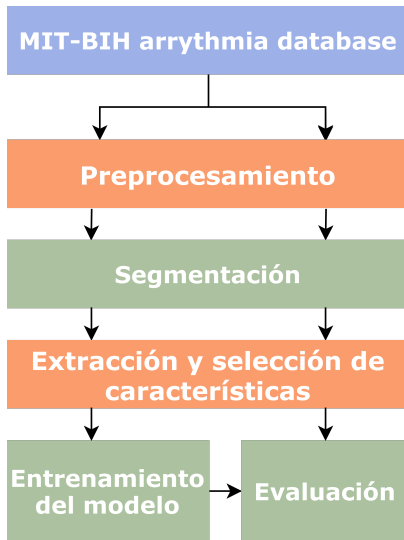


Fig. 3: Metodología empleada para el desarrollo del modelo clasificador de arritmias

### A. ECG Dataset

En este trabajo usamos el MIT-BIH arrhythmia dataset [8]. Este dataset consiste de 48 grabaciones de ECGs obtenidas de 47 pacientes. Cada uno de estos corresponde a extractos de media hora de grabaciones a dos canales de 24 horas. Uno de los dos canales es el modified limb II (MLII) y el otro canal varía de sujeto a sujeto, y en este caso, puede ser V1, V2, V4 o V5. Estos datos fueron grabados y digitalizados con una frecuencia de 360 Hz. Cada uno de estos registros tiene

sus respectivas anotaciones, de tal modo que cada latido tiene asignada una categoría (tipo de arritmia) de cada una de las clases presentes en el dataset.

En este trabajo sólo empleamos el canal MLII y cada registro es segmentado. Cada segmento corresponde a 300 puntos, lo que hace que cada uno de estos sea de aproximadamente 0.83 segundos. Cada segmento tiene su punto máximo en el punto central, de modo que a lado y lado está el resto de elementos que conforman cada segmento (onda P, T y S).

Similar a otros trabajos [4], creamos dos grupos a partir de los registros. Un subconjunto de los electrocardiogramas pasa a hacer parte del grupo de registros de entrenamiento, y los demás pasan a ser parte de los registros de prueba. Esta división se muestra en la Tabla II

|                  |  |
|------------------|--|
| Training Records | 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230 |
| Testing Records  | 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234 |

TABLE II: Training and testing records

Cada uno de los registros está debidamente anotado. Cada anotación corresponde a una de las 16 clases de arritmia presentes en este dataset. En este trabajo, usaremos las clases AAMI, que también han sido usadas en otros trabajos [4], [7], [13], [32], [46], [47].

La correspondencia entre las clases AAMI y las clases originales se presenta en la Tabla Table III.

| AAMI Class | MIT-BIH Class                             |
|------------|---|
| N          | N - Normal beat                           |
|            | L - Left bundle branch block beat (LBBB)  |
|            | R - Right bundle branch block beat (RBBB) |
|            | e - Atrial escape beat                    |
| S          | j - Nodal (junctional) escape beat        |
|            | S - Supraventricular premature beat       |
|            | A - Atrial premature beat                 |
| V          | a - Aberrated atrial premature beat       |
|            | J - Nodal (junctional) premature beat     |
|            | V - Premature ventricular contraction     |
| F          | E - Ventricular Escape                    |
|            | F - Fusion of ventricular and normal beat |

TABLE III: Heartbeat classifications

En la tabla IV se muestra la cantidad de datos de cada clase en los conjuntos de entrenamiento y prueba.

### B. Preprocesamiento

1) *Obtención de los picos*: Para la obtención de los picos se empleó el algoritmo *XQRS* disponible en la librería WFDB, que se usó para la lectura de los archivos de electrocardiogramas del dataset. Esta obtención de los picos se aplicó solamente a los datos de prueba, pues para los datos de entrenamiento se utilizaron los picos ya presentes en el dataset.

| Clase | Traning data | Testing data |
|-------|--------------|--------------|
| N     | 43945        | 42268        |
| S     | 3020         | 2146         |
| F     | 604          | 1700         |
| V     | 293          | 379          |

TABLE IV: Conteo de datos de cada clase para los datos de prueba y entrenamiento

2) *Reducción de ruido*: Antes de comenzar a segmentar los datos, se debe realizar una limpieza de las señales que no pertenecen al Electrocardiograma. Estas señales residuales están relacionadas con la respiración del paciente, el ruido que genera el dispositivo Holter usado para tomar los ECGs que conforman el dataset a usar y movimientos involuntarios de los pacientes. Además de estas señales residuales, los ECGs presentan un corrimiento del Baseline Wander, por lo cual también debemos corregir. En este proyecto usamos una combinación de filtros de señales que eliminan el ruido ubicados en alta y baja frecuencia y corrigen el Baseline Wander. Para el ruido de las bajas frecuencias se usó un Butterworth High Pass Filter de Orden 4 y Cutoff de 0.5 Hz, para el ruido ubicados en las frecuencias altas se usó un FIR Filter de Orden 12 y Cutoff de 35 Hz. Por último, se usó un Notch Filter a 360 Hz (la frecuencia de las señales de los ECGs del dataset) para corregir el Baseline Wander.

En las figuras 4c, 4a, 4b, 4d se muestran algunos ejemplos de segmentos una de los registros luego del debido preprocesamiento.

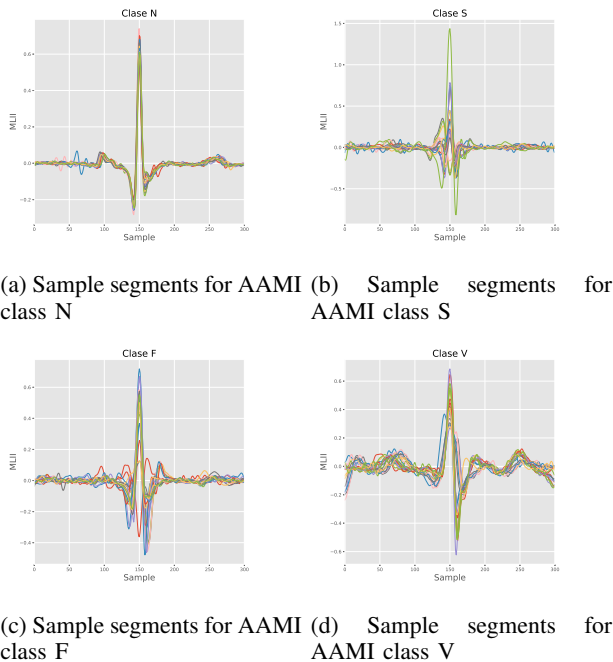


Fig. 4: Sample segments for each of the four AAMI classes

3) *Segmentación de los datos*: Cada uno de los registros del dataset contiene varios minutos de señales. Cada uno de

estos registros ha de contener, por lo tanto, muchos latidos. De modo que para proceder a entrenar un modelo, es necesario realizar en primer lugar una segmentación de los registros, de modo que los latidos contenidos en estos estén separados. En el dataset cada latido ya se encuentra etiquetado, por lo que este proceso recogerá el conjunto total de latidos con sus respectivas etiquetas. Es importante resaltar que al hablar de latidos nos referimos a las señales que lo componen. En la literatura el proceso de segmentación se hace a partir de una frecuencia de 360 Hz. De modo que la señal del electrocardiograma será segmentada por completo utilizando esta frecuencia. Esto corresponde a aproximadamente 0.83 segundos por segmento. A cada segmento se le hará un procedimiento de parametrización, mediante el cual se extraerán las características de estos para la posterior creación del clasificador.

### C. Extracción de características

1) *Wavelet Transform*: Wavelet Transform es una función que transforma un conjunto de señales a un dominio distinto. Es similar a Fourier Transform pero con la diferencia de que maneja tanto el Dominio de Fourier como el Dominio Real. Esta transformada viene dada por la siguiente función:

$$W_a(x, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi\left(\frac{t-b}{a}\right) dt, a \geq 1$$

Wavelet Transform se ha usado con diferentes fines, algunos trabajos [7], [15], [25], [30], [36], [37], [45], [47], [50]–[52] la han usado como herramienta para la detección del complejo QRS, otros trabajos lo han enfocado como extractor de características morfológicas para los ciclos cardiacos [17], [47]. En este proyecto usaremos Wavelet Transform como la fuente de características morfológicas de los ECGs.

En este trabajos empleamos la implemetación presentada en [53] para la obtención de estos coeficientes. Se aplica el método daubechies de orden 4 (db4) a nivel 7.

2) *Feature selection*: Posterior a la extracción de características usando Wavelet Transform y los intervalos RR, además de la curtosis y la oblicuidad, se realiza un proceso de selección de las características, con el objetivo de reducir el espacio de características inicial, que está conformado por 142 variables.

Primero se aplica una aliminación de las variables cuasi-constantes. Todas aquellas variables con una varianza menor que un límite inferior definido son eliminadas. El límite inferior usado en este trabajo es  $\sigma_{min} = 0.001$ . Luego de este proceso, se eliminan columnas a partir una cota de correlación máxiima. En este trabajo, la cota de correlación máxima permitid es  $max_{corr} = 0.8$ .

Luego de este proceso deja 39 características. En la Fig. 5 se muestra la matriz de correlación para cada una de las columnas restantes.

## VIII. MODELO O ARQUITECTURA LÓGICA

### A. Despliegue o Arquitectura física

Para el prototipo se recurrió a la Free Tier de Amazon Web Services para replicar la Arquitectura Lógica. Para el

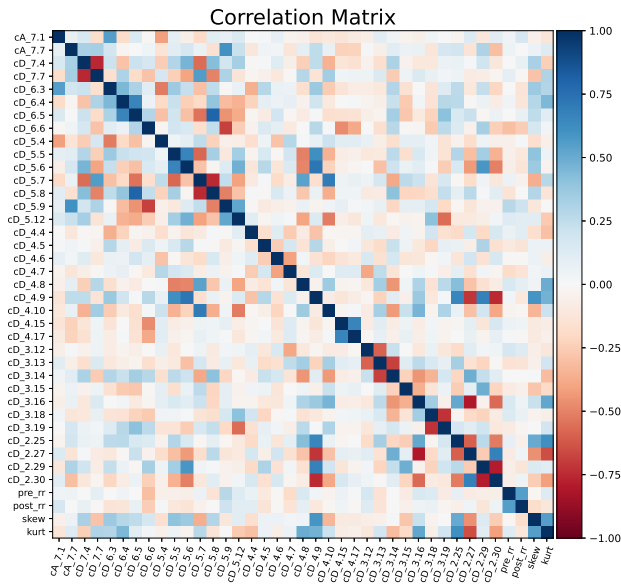


Fig. 5: Correlation matrix of the selected features

alojamiento tanto del Backend del proyecto como del modelo se utilizaron instancias EC2 t2.micro en Ohio (us-east-2). Para la base de datos, usamos una DB no relacional MongoDB, alojada en Mongo Atlas y para este componente también usamos la capa gratuita. En la base de datos se alojan las señales que son usadas para pruebas del prototipo y por ende, los usuarios finales puedan interactuar para ver el Clasificador en acción. Se requirió del servicio de AWS Route 53 para la obtención de un dominio que sería útil para la instancia EC2 que aloja el Backend del prototipo y a su vez, este servicio permite redirigir cualquier petición que apunte al dominio hacia la instancia EC2. Para el frontend, se usó un Bucket S3 de AWS para alojar los archivos estáticos del framework usado y así poder desplegarlo a Internet.

### B. Prototipo

El primero, es un Landing page en donde se expone de manera resumida información importante acerca del proyecto, así como los diferentes procesos llevados a cabo durante la creación del clasificador. También, se muestra información relacionada al puntaje del modelo y de los features a usar. Por otro lado, el segundo y último módulo es el de pruebas, donde el usuario final tiene acceso a un pequeño panel donde puede elegir que ECGs de pruebas va hacer que el modelo clasifique y el segmento. Cada segmento de un ECG consta de 16250 samples, que se traducen en 45 segundos, por lo cual, cada ECG se divide en 40 segmentos.

## IX. RESULTADOS

En esta sección se hace una muestra de los resultados obtenidos. Se muestra el proceso de selección del modelo y una comparación con los resultados de trabajos previos.

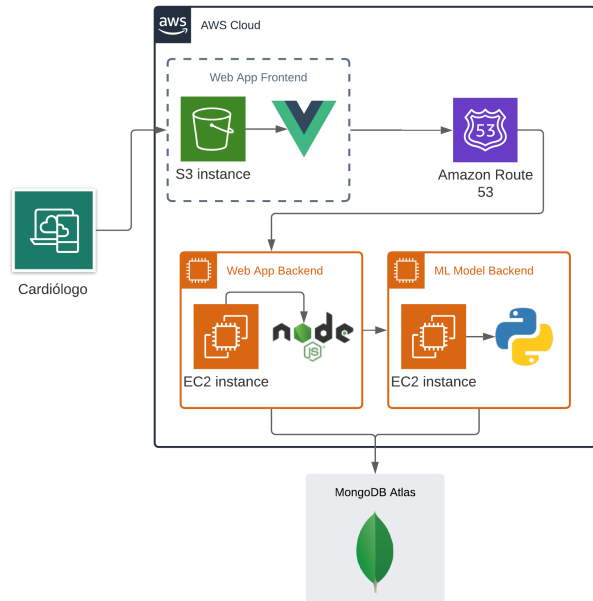


Fig. 6: Arquitectura lógica de la solución

### A. Métricas de evaluación

La sensibilidad se define como el cociente del total de verdaderos positivos (TP) y la suma de verdaderos positivos y falsos negativos (FN)

$$Se = \frac{TP}{TP + FN}, \quad (1a)$$

la especificidad se define como el cociente del lo verdaderos negativos y la suma de verdaderos negativos y falsos positivos

$$Sp = \frac{TN}{TN + FP}. \quad (1b)$$

### B. Selección del modelo

Aquí se presentan 3 alternativas para la selección del modelo. La primera consiste en clasificar los datos usando exclusivamente LDA, usando validación cruzada para elegir los mejores parámetros. Debido a que las clases presentan un desbalanceo importante, las *priors* de cada clase se eligen también en el proceso de validación cruzada. La segunda opción consiste en dividir la clasificación en dos pasos: Primero se clasifica si el segmento dado es normal (N) o si pertenece a alguna de las otras tres clases AAMI (S, F o V); esto se hace usando multilayer perceptron. En caso de pertenecer a una de las otras tres clases -que en este primer paso se toma como si fuera una sola-, se hace una segunda clasificación mediante LDA. En esto, presentamos dos variantes: La primera consiste en que las *priors* se dejan intactas, de modo que son calculadas a partir del número de elementos de cada una de estas tres clases. La segunda variante, que viene a ser la tercera opción, consiste en elegir mejores *priors* usando validación cruzada. El proceso de validación cruzada se realizó utilizando 10-folds estratificados.

El modelo escogido para la primera parte de la clasificación es multilayer perceptron. Este tiene 4 capas ocultas de 6, 5, 5 y 4 unidades respectivamente. La función de activación es tanh y el método de optimización que utiliza este es el descenso del gradiente estocástico (SGD). En el caso de LDA, las *priors* se dejan con sus valores por defecto, es decir, cada prior de cada clase corresponde al porcentaje de elementos de cada una. En la Fig. 7 se muestra este proceso gráficamente.

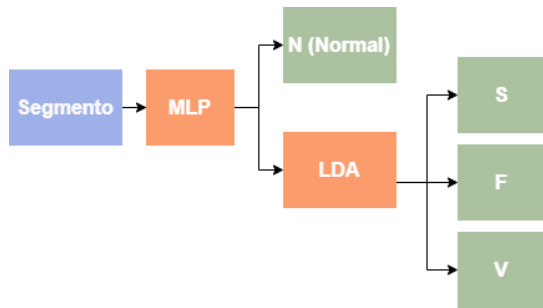


Fig. 7: Gráfica del proceso empleado por el modelo para la clasificación

### C. Evaluación del modelo

En la Fig. 8 se muestra la matriz de confusión para las predicciones obtenidas. Aquí se observa que para todas las clases hay una clasificación buena. La precisión calculada es del 89 %.

En la tabla V se muestran los resultados de especificidad (Sp), sensibilidad (Se) y la precisión

## X. EVALUACIÓN

En el marco de la asignatura de proyecto final, dos grupos realizaron una evaluación de nuestro prototipo utilizando el estándar ISO 9126 ISO 15504 + la *escala de likert*. En las figuras 9 y 10 se puede observar los resultados de los aspectos evaluados como los son comprensibilidad, documentación,

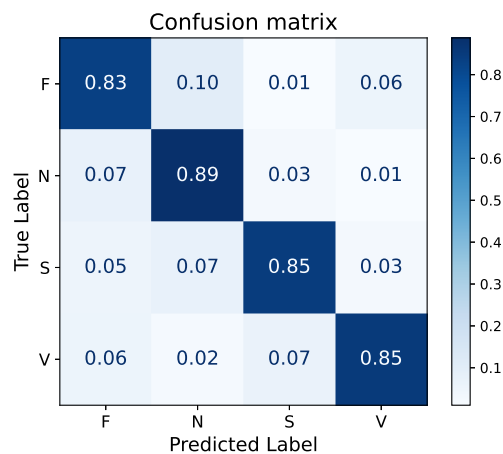


Fig. 8: Confusion matrix

| Clase | LDA + MLP |      |           |
|-------|-----------|------|-----------|
|       | Se        | Sp   | Bal. Acc. |
| F     | 0.81      | 0.93 | 0.89      |
| N     | 0.89      | 0.94 | 0.92      |
| S     | 0.85      | 0.97 | 0.93      |
| V     | 0.85      | 0.98 | 0.85      |

TABLE V: Resultados de la evaluación del modelo apilado para la clasificación de arritmias

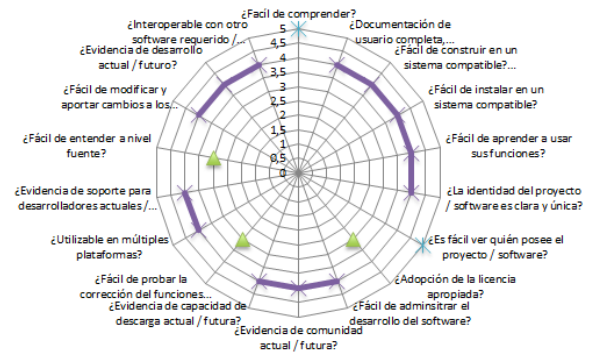


Fig. 9: Gráfico de evaluación 1

edificabilidad, instalabilidad, capacidad de aprendizaje, identidad, derechos de autor, licenciamiento, gobernancia entre otros aspectos tenidos en cuenta.

## XI. CONCLUSIONES

La clasificación y detección de arritmia es importante para un diagnóstico y un tratamiento adecuados. En este trabajo se presentó un modelo de machine learning basado en *stacked models*, empleando LDA y MLP para la clasificación de arritmias usando ECGs, usando las clases del estándar AAMI. Se evidencia que el proceso de extracción de características de las señales es un proceso clave para lograr clasificadores con buenos resultados. Los resultados mostrados evidencian una

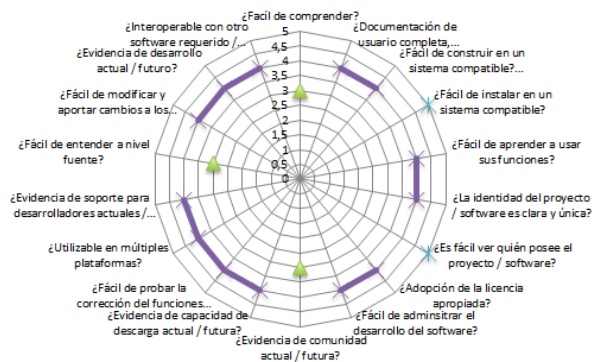


Fig. 10: Gráfico de evaluación 2

clasificación en general buena para las cuatro clases tratadas en este trabajo.

## REFERENCES

- [1] S. Karpagachelvi, M. Arthanari, and M. Sivakumar, "Classification of electrocardiogram signals with support vector machines and extreme learning machine," *Neural Computing and Applications*, vol. 21, pp. 1331–1339, sep 2012.
- [2] R. Lekhal, Z. Zidelmal, and D. Ould-Abdesslam, "Optimized time–frequency features and semi-supervised SVM to heartbeat classification," *Signal, Image and Video Processing*, pp. 1–8, may 2020.
- [3] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artificial Intelligence in Medicine*, vol. 44, no. 1, pp. 51–64, 2008.
- [4] S. Chen, W. Hua, Z. Li, J. Li, and X. Gao, "Heartbeat classification using projected and dynamic features of ECG signal," *Biomedical Signal Processing and Control*, vol. 31, pp. 165–173, jan 2017.
- [5] A. Chetan, R. K. Tripathy, and S. Dandapat, "A Diagnostic System for Detection of Atrial and Ventricular Arrhythmia Episodes from Electrocardiogram," *Journal of Medical and Biological Engineering*, vol. 38, pp. 304–315, apr 2018.
- [6] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1196–1206, jul 2004.
- [7] C. C. Lin and C. M. Yang, "Heartbeat classification using normalized RR intervals and morphological features," *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [8] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," 2001.
- [9] A. Bolívar-Mejía and B. E. Vesga-Angarita, *Burden of Cardiovascular Disease in Colombia*. 2013.
- [10] W. Chen, *Electrocardiogram*. 2018.
- [11] C. R. Rao, "The Utilization of Multiple Measurements in Problems of Biological Classification," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 10, pp. 159–203, nov 1948.
- [12] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning – data mining, inference, and prediction."
- [13] W. Zhu, X. Chen, Y. Wang, and L. Wang, "Arrhythmia Recognition and Classification Using ECG Morphology and Segment Feature Analysis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, pp. 131–138, jan 2019.
- [14] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 1778–1790, aug 2004.
- [15] D. Thanapatay, C. Suwansaroj, and C. Thanawattano, "ECG beat classification method for ECG printout with Principle Components Analysis and Support Vector Machines," in *ICEIE 2010 - 2010 International Conference on Electronics and Information Engineering, Proceedings*, vol. 1, 2010.
- [16] K. Cramer and Y. Singer, "On the learnability and design of output codes for multiclass problems," *Machine Learning*, vol. 47, pp. 201–233, may 2002.
- [17] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, and S. K. Yoo, "Support vector machine based arrhythmia classification using reduced features," *International Journal of Control, Automation and Systems*, vol. 3, no. 4, pp. 571–579, 2005.
- [18] M. Mohanty, P. Biswal, and S. Sabut, "Machine learning approach to recognize ventricular arrhythmias using VMD based features," *Multidimensional Systems and Signal Processing*, vol. 31, pp. 49–71, jan 2020.
- [19] J. Weston and C. Watkins, "Support Vector Machines for Multi-Class Pattern Recognition," tech. rep.
- [20] M. Rahman and M. Nasor, "An Algorithm for Detection of Arrhythmia," *International Journal of Biological Engineering*, vol. 2, pp. 44–47, dec 2012.
- [21] S. M. Jadhav, S. L. Nalbalwar, and A. Ghatol, "Artificial Neural Network based cardiac arrhythmia classification using ECG signal data," in *ICEIE 2010 - 2010 International Conference on Electronics and Information Engineering, Proceedings*, vol. 1, 2010.
- [22] S. Khare, A. Bhandari, S. Singh, and A. Arora, "ECG arrhythmia classification using spearman rank correlation and support vector machine," in *Advances in Intelligent and Soft Computing*, vol. 131 AISC, pp. 591–598, Springer, New Delhi, 2012.
- [23] F. Yaghouby and A. Ayatollahi, "An Arrhythmia Classification Method Based on Selected Features of Heart Rate Variability Signal and Support Vector Machine-Based Classifier," pp. 1928–1931, Springer, Berlin, Heidelberg, 2009.
- [24] A. Mert, "ECG feature extraction based on the bandwidth properties of variational mode decomposition," *Physiological Measurement*, vol. 37, pp. 530–543, mar 2016.
- [25] C. Aneesh, S. Kumar, P. M. Hisham, and K. P. Soman, "Performance comparison of Variational Mode Decomposition over Empirical Wavelet Transform for the classification of power quality disturbances using Support Vector Machine," in *Procedia Computer Science*, vol. 46, pp. 372–380, Elsevier B.V., jan 2015.
- [26] Y. Benchaib, "Comparative study between artificial metaplasticity learning algorithm and other machine learning techniques: Application for cardiac arrhythmias classification," in *Advances in Intelligent Systems and Computing*, vol. 1103 AISC, pp. 72–82, Springer, jul 2020.
- [27] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical Engineering and Physics*, vol. 28, pp. 876–887, nov 2006.
- [28] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, pp. 531–544, feb 2014.
- [29] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, pp. 65–69, jan 2019.
- [30] X. Jiang, L. Zhang, Q. Zhao, and S. Albayrak, "ECG arrhythmias recognition system based on independent component analysis feature extraction," in *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, Institute of Electrical and Electronics Engineers Inc., 2006.
- [31] S. Karimifard, A. Ahmadian, M. Khoshnevisan, and M. S. Nambakhsh, "Morphological heart arrhythmia detection using hermitian basis functions and kNN classifier," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 1367–1370, 2006.
- [32] W. Yang, Y. Si, D. Wang, and B. Guo, "Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine," *Computers in Biology and Medicine*, vol. 101, pp. 22–32, oct 2018.
- [33] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 582–589, apr 2004.
- [34] S. M. P. Dinakarrao, A. Jantsch, and M. Shafique, "Computer-aided arrhythmia diagnosis with bio-signal processing: A survey of trends and techniques," *ACM Computing Surveys*, vol. 52, pp. 1–37, may 2019.
- [35] J. Kim, H. S. Shin, K. Shin, and M. Lee, "Robust algorithm for arrhythmia classification in ECG using extreme learning machine," *Biomedical engineering online*, vol. 8, p. 31, oct 2009.
- [36] J. Millet-Roig, R. Ventura-Galiano, F. J. Chorro-Gascó, and A. Cebrián, "Support vector machine for arrhythmia discrimination with wavelet-transform-based feature selection," *Computers in Cardiology*, pp. 407–410, 2000.
- [37] C. V. Banupriya and S. Karpagavalli, "Electrocardiogram Beat Classification Using Support Vector Machine and Extreme Learning Machine," in *Advances in Intelligent Systems and Computing*, vol. 248 VOLUME I, pp. 187–193, Springer Verlag, 2014.
- [38] C. W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, pp. 415–425, mar 2002.
- [39] S. S. Mehta and N. S. Lingayat, "Application of support vector machine for the detection of P- and T-waves in 12-lead electrocardiogram," *Computer Methods and Programs in Biomedicine*, vol. 93, pp. 46–60, jan 2009.
- [40] P. Jain, C. S. Arjun Babu, S. Mohandoss, N. Anisham, S. Gadade, A. Srinivas, and R. Mohan, "A Novel Approach to Classify Cardiac Arrhythmia Using Different Machine Learning Techniques," pp. 517–526, Springer, Singapore, 2021.

- [41] M. Arun Kumar and M. Gopal, "Fast multiclass SVM classification using decision tree based one-against-all method," *Neural Processing Letters*, vol. 32, pp. 311–323, dec 2010.
- [42] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines: and other kernel-based learning methods," 2000.
- [43] A. Mustaqeem, S. M. Anwar, M. Majid, and A. R. Khan, "Wrapper method for feature selection to classify cardiac arrhythmia," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 3656–3659, Institute of Electrical and Electronics Engineers Inc., sep 2017.
- [44] A. Khazaee and A. Ebrahimzadeh, "Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features," *Biomedical Signal Processing and Control*, vol. 5, pp. 252–263, oct 2010.
- [45] M. S. Khadtare and J. S. Sahambi, "ECG arrhythmia analysis by multiclass support vector machine," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3285, pp. 100–107, oct 2004.
- [46] V. Krasteva and I. Jekova, "QRS template matching for recognition of ventricular ectopic beats," *Annals of Biomedical Engineering*, vol. 35, pp. 2065–2076, dec 2007.
- [47] C. Ye, B. V. Vijaya Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.
- [48] S. Osowski, T. H. Linh, and T. Markiewicz, "Recognition of heartbeats using Support Vector Machine networks - A comparative study," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3697 LNCS, pp. 637–642, Springer, Berlin, Heidelberg, 2005.
- [49] S. Samad, S. A. Khan, A. Haq, and A. Riaz, "Classification of Arrhythmia," *International Journal of Electrical Energy*, pp. 57–61, 2014.
- [50] A. Joshi, Rajshekhar, S. Chandran, S. Phadke, V. K. Jayaraman, and B. D. Kulkarni, "Arrhythmia classification using local hölder exponents and support vector machine," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3776 LNCS, pp. 242–247, Springer, Berlin, Heidelberg, dec 2005.
- [51] P. S. Addison, "Wavelet transforms and the ECG: a review.," *Physiological measurement*, vol. 26, pp. R155–99, oct 2005.
- [52] S. Poornachandra, "Wavelet-based denoising using subband dependent threshold for ECG signals," *Digital Signal Processing: A Review Journal*, vol. 18, no. 1, pp. 49–55, 2008.
- [53] G. R. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O'Leary, "Pywavelets: A python package for wavelet analysis," *Journal of Open Source Software*, vol. 4, no. 36, p. 1237, 2019.