



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



**DESARROLLO DE UN SISTEMA PARA CLASIFICACIÓN DE PATOLOGÍAS
CARDIOVASCULARES EN SEÑALES ELECTROCARDIOGRÁFICAS (ECG)
APLICANDO INTELIGENCIA ARTIFICIAL Y CLOUD COMPUTING**

AUTOR: DARÍO FERNANDO BRAVO TUAY

**PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA, ELÉCTRICA, SISTEMAS Y
TELECOMUNICACIONES
FACULTAD DE INGENIERÍAS Y ARQUITECTURA**



UNIVERSIDAD DE PAMPLONA

PAMPLONA, 16 DE MAYO DEL 2021



SC-CER96940

"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



**DESARROLLO DE UN SISTEMA PARA CLASIFICACIÓN DE PATOLOGÍAS
CARDIOVASCULARES EN SEÑALES ELECTROCARDIOGRÁFICAS (ECG)
APLICANDO INTELIGENCIA ARTIFICIAL Y CLOUD COMPUTING**

DARÍO FERNANDO BRAVO TUAY

**Trabajo de grado presentado como requisito para optar al título de
INGENIERO EN TELECOMUNICACIONES**

**Director: GERMÁN ARLEY PORTILLA GONZÁLEZ
ING. ELECTRÓNICO**

**PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA, ELÉCTRICA, SISTEMAS Y
TELECOMUNICACIONES
FACULTAD DE INGENIERÍAS Y ARQUITECTURA
UNIVERSIDAD DE PAMPLONA
PAMPLONA, 16 DE MAYO DEL 2021**

**UNIVERSIDAD DE PAMPLONA
FACULTAD DE INGERÍAS Y ARQUITECTURA
DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA, SISTEMAS Y
TELECOMUNICACIONES**



SC-CER96940

"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



**PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES
TRABAJO PRESENTADO PARA OPTAR POR EL TÍTULO DE INGENIERO EN
TELECOMUNICACIONES**

TEMA:

**DESARROLLO DE UN SISTEMA PARA CLASIFICACIÓN DE PATOLOGÍAS
CARDIOVASCULARES EN SEÑALES ELECTROCARDIOGRÁFICAS (ECG)
APLICANDO INTELIGENCIA ARTIFICIAL Y CLOUD COMPUTING**

**FECHA DE INICIO DEL TRABAJO:
FECHA DE TERMINACIÓN DEL TRABAJO:**

NOMBRES Y FIRMAS DE AUTORIZACIÓN PARA LA SUSTENTACIÓN:

**DARÍO FERNANDO BRAVO TUAY
AUTOR**

**GERMÁN ARLEY PORTILLA
DIRECTOR**

**HERNANDO JOSÉ VELANDIA
DIRECTOR DE PROGRAMA**

JURADO CALIFICADOR:

Mag. LUIS ENRIQUE MENDOZA

**SERGIO ESTIVENSON PINTO
ING. ELECTRÓNICA**



SC-CER96940

"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



DEDICATORIA

Doy gracias primeramente a Dios por permitirme culminar un logro más de mi vida, por darme la salud, sabiduría y fortaleza para seguir en pie a pesar de las adversidades que se han presentado en mi vida.

Trabajo de grado dedicado:

A mi madre, Ana Lía Tuay Méndez por fomentar la educación desde casa, por darme el amor más grande, puro y sincero del mundo, por ser una madre trabajadora y ejemplar que a pesar de las circunstancias que se presentan en el transcurso de la vida sigue en pie de lucha, por quitarse el pan de la boca para que sus hijos estén bien y no le falte nada, por ello y mucho más estoy dispuesto a dar todo por ella.

En memoria de mi padre Darío Bravo, han pasado 23 años de tu partida y por más de que nunca te conocí me duele tanto saber cuándo las personas que te conocían me comentan lo especial y maravilloso que eras, sé que me cuidas desde el reino celestial.

En conmemoración de mi tía Betulia Tuay Méndez; las adversidades de la vida se han llevado a una de las personas más queridas de mi vida, sé que estés donde estés, estás llena de felicidad, gozas de las grandes bendiciones que Dios ha preparado para ti, pues él te ha llamado y no debemos ser egoístas, me permitió pasar momentos llenos de felicidad y risa con alguien tan especial como lo fuiste tu. Te recordaré con amor.

En conmemoración de mi tío Gerardo Tuay Méndez, fuiste alguien especial conmigo y no olvido los pequeños consejos que me dabas cuando pasaba por tu casa, las risas y chistes que me contabas, los pequeños gestos que hacías conmigo y lo poco que tenías lo compartías conmigo.

En conmemoración de mi tío Juan Francisco Tuay Méndez, fuiste como un padre para mi madre, por eso el cariño que siento hacía a ti nunca se apaciguará, recuerdo cada momento contigo.

A mis compañeros, Diego Javier Leal Díaz y Darwin Manuel Mercado Cerpa, gracias a ello por infundir sus conocimientos y darme la motivación para no decaer en la decepción, es bueno saber que existen personas de buen corazón y que se puede contar con ellas.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



A mi novia, Valentina González por ser una persona especial y de gran apoyo motivacional en mi vida, gracias por los consejos que me das las veces en las que no me siento con ganas de realizar las cosas.

A los Ingenieros, Gustavo Adolfo Quijada, Germán Arley Portilla, Luis Enrique Mendoza, Hernando José Velandia, Edwin Mauricio Sequeda por compartir sus grandes conocimientos y por la guía que han realizado para mi construcción profesional.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



RESUMEN

Este trabajo se presenta como una alternativa tecnológica para aquellas personas que padecen de enfermedades cardiovasculares tales como disritmia cardíaca, hipertrofia ventricular, bloqueo de rama y afección valvular ya sea de forma congénita o mal cuidado de la salud por el mismo.

En primer lugar, se adquieren las señales electrocardiográficas (ECG) de pacientes sanos y pacientes con dichas patologías mediante la base de datos “Physiobank”, en lo cual, se aplican técnicas de tratamiento y procesamiento de señales tales como ajuste de nivel, Transformada Rápida de Fourier (FFT), Wavelet (DCT), entropía y energía, para la posterior búsqueda de características o patrones que demuestren una diferencia de las señales y otorgar una correcta clasificación mediante la implementación Machine Learning bajo el lenguaje de programación Python empleado en Linux en relación a las patologías presentes y dar una mayor facilidad en cuanto al diagnóstico por parte del especialista, con base en lo anterior se selecciona la mejor técnica en relación a los resultados de optimización y clasificación.

En segundo lugar, una vez finalizada la caracterización y clasificación de las ondas ECG se procede a aplicar estrategias de Cloud Computing para gestionar los datos almacenados en servidores remotos, procesarlos en la nube y poder obtener una respuesta como tal de la señal ingresada al sistema, por tanto, cualquier entidad podrá tener acceso al sistema de clasificación de forma gratuita para prestar el servicio. Para este caso, se trabaja en la plataforma “Pythonanywhere” el cual permite alojar código Python y ser ejecutado como un aplicativo web.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



CONTENIDO

CAPÍTULO I: INTRODUCCIÓN	14
RESUMEN	14
JUSTIFICACIÓN	14
DELIMITACIÓN	15
OBJETIVO GENERAL	15
OBJETIVOS ESPECÍFICOS	15
ACOTACIONES	15
CAPÍTULO II: MARCO TEÓRICO	17
2.1. ELECTROCARDIOGRAMA	18
2.1.1. ECG normal.....	19
2.1.2. Derivaciones en el ECG.....	21
2.1.3. Errores en la obtención del electrocardiograma.....	21
2.2. ANATOMÍA DEL CORAZÓN	22
2.2.1. Válvulas cardíacas.....	22
2.2.2. Potencial de acción de las células miocárdicas.....	23
2.2.3. Problemas que surgen por el mal funcionamiento del aparato circulatorio y el corazón.....	23
2.2.4. Sistema de conducción.....	24
2.3. ELECTROCARDIOGRAMA EN LAS CARDIOPATIAS CONGÉNITAS	26
2.3.1. Hipertrofia auricular derecha.....	26
2.3.2. Hipertrofia auricular izquierda.....	26
2.3.3. Hipertrofia atrial combinada.....	27
2.3.4. Hipertrofia ventricular derecha.....	28
2.3.4.1. Criterios de hipertrofia ventricular derecha.....	28
2.3.4.2. Patrones de sobrecarga de presión correspondientes al ventrículo derecho.....	29



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



2.3.5. Hipertrofia ventricular izquierda.....	29
2.3.5.1 Criterios de hipertrofia ventricular izquierda.....	30
2.3.5.2 Patrón de sobrecarga correspondientes al ventrículo izquierdo... 30	
2.4. ELECTROCARDIOGRAMA EN DEFECTOS DE LA CONDUCCIÓN INTRAVENTRICULAR.	30
2.4.1. Bloqueo de rama derecha del haz de His.	31
2.4.2. Bloqueo de rama izquierda del haz de His.	31
2.5. ANORMALIDADES VENTRICULARES (VÁLVULO PATÍAS).....	32
2.5.1. Estenosis aórtica.....	32
2.5.2. Estenosis mitral.....	33
2.5.3. Insuficiencia aórtica.	34
2.5.4. Insuficiencia mitral.	34
2.6. DISRITMIAS (ARRITMIAS).	34
2.6.1 Bradiarritmias.	34
2.6.2 Taquiarritmias.....	35
2.6.2.1. Supraventriculares.....	35
2.6.2.2. Ventriculares.	35
2.7. PROCESAMIENTO DE SEÑALES.....	36
2.8. EXTRACCIÓN DE CARACTERÍSTICAS.....	36
2.8.1. Transformada discreta de Fourier (DFT).	36
2.8.2. Transformada wavelet discreta (DWT).....	37
2.8.2.1. Descomposición Wavelet.....	38
2.8.3 Entropía.....	39
2.8.4. Energía.	39
2.8.5. Normalización.....	39
2.9. TÉCNICAS DE CLASIFICACIÓN DE MACHINE LEARNING.....	40
2.9.1. Aprendizaje Supervisado.....	40
2.9.1.1. Máquina de soporte vectorial (SVM).....	41





ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



2.9.1.2. Clasificador bayesiano (Naïve Bayes).....	44
2.9.1.3 Árboles de decisión.....	45
2.9.1.3.1. Índice de Gini.....	46
2.9.1.3.2. Ganancia de información.....	46
2.10. CLOUD COMPUTING.....	47
2.10.1. Tipos de Cloud Computing.....	48
2.11. ESTADO DEL ARTE.....	49
2.11.1. Trabajos realizados a nivel internacional.....	49
2.11.2. Trabajos realizados a nivel nacional.....	52
2.11.3. Trabajos realizados a nivel regional.....	54
CAPÍTULO III: METODOLOGÍA.....	57
3.1. CONCEPTOS TEÓRICOS.....	58
3.2. ADQUISICIÓN DE SEÑALES ELECTROCARDIOGRÁFICAS.....	59
3.3. PREPROCESAMIENTO.....	63
3.3.1. Ajuste de amplitud (estandarización de la señal).....	64
3.3.2. Centrando señal en cero.....	64
3.3.3. Agrupando como datagrama.....	65
3.4. PROCESAMIENTO.....	67
3.4.1. Primera línea de procesamiento.....	67
3.4.2. Segunda línea de procesamiento.....	72
3.5. CLASIFICACIÓN APLICANDO MACHINE LEARNING.....	77
3.5.1. Support Vector Machine (VSM).....	78
3.5.2. Naïve Bayes (NB).....	79
3.5.3. Árboles de decisión.....	81
3.6. PYTHONANYWHERE (CLOUD).....	81
CAPÍTULO IV: RESULTADOS.....	86
4.1. BASE DE DATOS PHYSIOBANK.....	86
4.2. TÉCNICAS DE TRATAMIENTO Y PROCESAMIENTO DE SEÑALES.....	86



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



4.3. CLASIFICACIÓN.....	88
4.4. INTERFAZ GRÁFICA CON TKINTER.....	97
4.5. IMPLEMENTACIÓN DE CLOUD COMPUTING	105
CAPÍTULO V: CONCLUSIONES	117
REFERENCIAS.....	118



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ÍNDICE DE FIGURAS

Figura 1 Registro de señal ECG típico [2].....	18
Figura 2 ECG normal [3].....	19
Figura 3 Anatomía del corazón [5]	22
Figura 4 Esquema del sistema de conducción cardíaca [7]	25
Figura 5 Hipertrofia Auricular Derecha [3]	26
Figura 6 Hipertrofia Auricular Derecha [3]	27
Figura 7 Hipertrofia atrial combinada [7].....	27
Figura 8 Hipertrofia ventricular derecha [3].....	28
Figura 9 Patrones de sobrecarga de presión del ventrículo derecho [7]	29
Figura 10 Hipertrofia ventricular izquierda [3].....	30
Figura 11 Bloqueo de rama derecha [3]	31
Figura 12 Bloqueo de rama izquierda [3]	32
Figura 13 Afección Valvular. A, estenosis aórtica. B, estenosis mitral. C, insuficiencia aórtica. D, insuficiencia mitral [8].....	33
Figura 14 Bradicardia sinusal [9]	35
Figura 15 Proceso árbol jerárquico de filtrado [15]	38
Figura 16 Diagrama de flujo del aprendizaje supervisado [20]	41
Figura 17 Hiperplano de separación de una SVM [21]	42
Figura 18 Diagrama de flujo para árboles de decisión [24]	45
Figura 19 Tipos de Cloud Computing [27]	48
Figura 20 Estructura metodológica	57
Figura 21 Sitio web Physiobank.....	59
Figura 22 Banco de descarga de señales	60
Figura 23 Descarga de archivos	61
Figura 24 Código de módulos para carga de paquetes .mat	62
Figura 25 Tipos y estructura del archivo .mat	62
Figura 26 Derivación i señal con afección valvular	63
Figura 27 Estandarización de la señal.....	64
Figura 28 Centrando señales en cero	65
Figura 29 Agrupamiento mediante DataFrame.....	66
Figura 30 Señal con ajuste de amplitud y posicionamiento en cero	66
Figura 31 Aplicación de FFT a señal ECG	67
La Figura 32 Eliminación de ruido	68
Figura 33 Cambio de dominio de la señal ECG	69
Figura 34 Entropía de la señal ECG con ventana de 10 puntos	70
Figura 35 Entropía de la señal ECG con ventana de 20 puntos	71



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



Figura 36 Entropía de la señal ECG con ventana de 30 puntos	72
Figura 37 Wavelet con db4 a la señal ECG	74
Figura 38 Energía de la señal ECG con ventana de 10 puntos.....	75
Figura 39 Energía de la señal ECG con ventana de 20 puntos.....	76
Figura 40 Energía de la señal ECG con ventana de 30 puntos.....	77
Figura 41 Página principal de Pythonanywhere	82
Figura 42 Creando nueva App web	83
Figura 43 Agregando nueva web app.....	83
Figura 44 App web flask creada.....	84
Figura 45 Aplicativo web	85
Figura 46 Porcentaje de estimación empleando VSM.....	93
Figura 47 Porcentaje de estimación empleando Naïve Bayes.....	95
Figura 48 Porcentaje de estimación empleando Árboles de decisión.....	96
Figura 49 Configurando archivo .desktop.....	98
Figura 50 Aplicación Sistema clasificador de patologías cardíacas	99
Figura 51 Ventana principal GUI.....	99
Figura 52 Ventana de procesos	100
Figura 53 Señales ECG para clasificación	100
Figura 54 Información de las señales ECG	101
Figura 55 Carga de señales a la GUI.....	101
Figura 56 Preprocesamiento de las derivadas del corazón	102
Figura 57 Patrones característicos	103
Figura 58 Proceso de Clasificación	104
Figura 59 Resultado de la predicción	104
Figura 60 Directorios y archivos en Pythonanywhere	105
Figura 61 Archivos que conforma Flask.....	106
Figura 62 Configuración de variable de entorno	107
Figura 63 Archivo WSGI	107
Figura 64 Configuración del archivo WSGI	108
Figura 65 Aplicativo web	109
Figura 66 Campos de texto vacías	110
Figura 67 Archivo no cargado.....	111
Figura 68 Cargando archivo que no es .mat	112
Figura 69 Validación de archivos .mat.....	113
Figura 70 Comprobación del sistema	114
Figura 71 Resultado de predicción	115



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



ÍNDICE DE TABLAS

Tabla 1 Señales bioeléctricas típicas [2]	19
Tabla 2 Asignación de recursos en modelos de Cloud Computing [25]	48
Tabla 3 Etiquetado de datos.....	79
Tabla 4 Etiqueta de datos entre paciente con afección valvular y paciente sano.....	80
Tabla 5 Etiqueta de datos entre paciente con bloqueo de rama y paciente sano	80
Tabla 6 Etiqueta de datos entre paciente con disritmia y paciente sano.....	80
Tabla 7 Etiqueta de datos entre paciente con hipertrofia y paciente sano	80
Tabla 8 Longitud de datos de las líneas de procesamiento	89
Tabla 9 Matriz completa.....	91



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



CAPÍTULO I: INTRODUCCIÓN

RESUMEN.

Según las estadísticas del Instituto Nacional de Salud, en Colombia las enfermedades cardiovasculares son las primeras causas de muerte. A pesar de esto las muertes por estas enfermedades cada año presentan una mayor proporción del total de las muertes en Colombia, por ello, en base a las nuevas herramientas computacionales que se han desarrollado a lo largo del tiempo, estas se deben integrar en la búsqueda de una alternativa para un diagnóstico médico más rápido y eficiente por parte del especialista o entidad de salud y así mejorar la condición de vida del paciente.

JUSTIFICACIÓN.

La idea central del trabajo de investigación se basa en tener un sistema o entorno capaz de clasificar patologías cardíacas aplicando herramientas de procesamiento de señales e inteligencia artificial como Machine Learning, que permita dar una mayor agilidad en cuanto a diagnóstico de ciertas patologías como afección valvular, bloqueo de rama, derrame de hipertrofia y disritmia cardíaca. El proyecto cuenta con conocimientos bases en cuanto a la aplicación de herramientas de procesamiento de señales aplicados en la carrera, conceptualización biomédica por parte del semillero de investigación, recursos educativos en línea, documentos de investigación y la aplicación de inteligencia artificial en base a cursos realizados de forma individual.

La aplicación de este proyecto brinda un mayor rendimiento para la entidad de salud con el hecho de ahorrar tiempo en cuanto al diagnóstico por parte del especialista y a su vez la probabilidad de mejoría en un tiempo más preciso para una mayor calidad de la salud a los pacientes debido a un servicio más eficaz.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



DELIMITACIÓN.

OBJETIVO GENERAL.

Desarrollar un sistema para clasificación de patologías cardiovasculares en señales electrocardiográficas (ECG) aplicando Inteligencia Artificial y Cloud Computing.

OBJETIVOS ESPECÍFICOS.

- Caracterizar morfológicamente las enfermedades cardiovasculares a tratar.
- Consolidar la base de datos de las señales electrocardiográficas tanto de pacientes sanos como de pacientes con patologías cardiovasculares por medio de la base de datos Physiobank.
- Aplicar técnicas de tratamiento y procesamiento sobre las señales ECG para la extracción de patrones característicos.
- Clasificar las señales ECG aplicando Machine Learning.
- Aplicar estrategias de Cloud Computing para la carga de datos y procesamiento en la nube.

ACOTACIONES.

- El desarrollo del trabajo de investigación se realizará bajo el lenguaje de programación Python el cual tiene como ventaja ser un lenguaje de licencia gratuito, flexible, ordenado y limpio, a su vez es adecuado para grandes cantidades de información en cuanto a procesamiento y análisis estadístico, para ello es necesario el uso de librerías y módulos.
- Se cuenta con una base de datos de señales ECG de pacientes sanos y pacientes con antecedentes cardiovasculares
- En base a la caracterización morfológica, las patologías cardíacas a analizar en este proyecto de investigación se desglosan en otras, por tanto, el tiempo que conlleva analizar cada una de estas es superior al estimado, así que se trabajan con las afecciones cardíacas de forma general.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



- Este proyecto busca realizar la clasificación de señales ECG aplicando una de las muchas técnicas de inteligencia artificial como lo es Machine Learning mediante la aplicación cursos realizados bajo la plataforma de educación virtual Udemy, Platzi, videos de aprendizaje sobre Machine Learning en línea; para lograr una detección y diagnóstico más eficaz de las cardiopatías ya mencionadas.
- Se implementa el uso de estrategias Cloud Computing bajo la plataforma Pythonanywhere que permite la carga de archivos, almacenamiento y ejecución de código Python en la web mediante la implementación del *micro-framework* denominado **Flask**.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



CAPÍTULO II: MARCO TEÓRICO

Este capítulo aborda temáticas relacionadas con las afecciones cardíacas las cuales son las responsables de la mayoría de muertes de seres humanos en Colombia, por ello, se debe mostrar las bases teóricas con respecto a las señales electrocardiográficas que ayuden a entender cuando y donde se presenta la falla cardíaca, parte de la anatomía del corazón, seguidamente la adquisición de las señales ECG desde la base de datos Physiobank y finalmente se trabajan las señales aplicando las técnicas de tratamiento y procesamiento convencionales para la extracción de patrones característicos de las mismas, posteriormente se procede a la aplicación de técnicas de Machine Learning para la respectiva clasificación y aplicación de Cloud Computing para la carga de datos a la nube y su correspondiente procesamiento.

Gracias a los nuevos avances científicos y tecnológicos hoy en día se han desarrollado diversos sistemas para el mejoramiento de la salud de las personas, por tanto, este trabajo presenta una alternativa de diagnóstico rápido y preciso para aquellos pacientes que padecen dichas cardiopatías con la finalidad de mejorar su calidad de salud y vida por medio de la realización correcta de procedimientos asignados por el especialista con respecto a la patología presente.

El método de adquisición de las señales ECG se denomina electrocardiografía y se encarga de registrar la actividad eléctrica del corazón mediante el posicionamiento de electrodos los cuales se colocan en la superficie corporal; se presentan los fenómenos de despolarización y repolarización mediante la diferencia de potencial que se genera en el corazón. [1]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

2.1. ELECTROCARDIOGRAMA.

Los electrocardiogramas comprenden una serie de pasos secuenciales; en primer lugar se adquieren las señales (se incluye el filtrado), algunos de estos se les denomina electrocardiógrafos de 12 derivadas; en segundo lugar se realiza la transformación de los datos, es decir, se obtiene el complejo promedio por cada derivada, se clasifican los latidos como *dominantes* y *no dominantes*; el tercer paso es donde se caracteriza cada una de las ondas y complejos; se identifica el inicio y fin de cada deflexión, luego se procede a medir la amplitud de las deflexiones e intervalos y por último se clasifica el diagnóstico. [1]

El registro de la actividad eléctrica del corazón es de gran importancia en el área cardiovascular por el hecho de entregar información importante en relación al funcionamiento del corazón; su representación es de forma gráfica y consiste en una línea base y varias deflexiones en formas de ondas variables a lo largo del eje del tiempo. [2]

La figura 1 muestra la representación del análisis de la señal con respecto al tiempo.

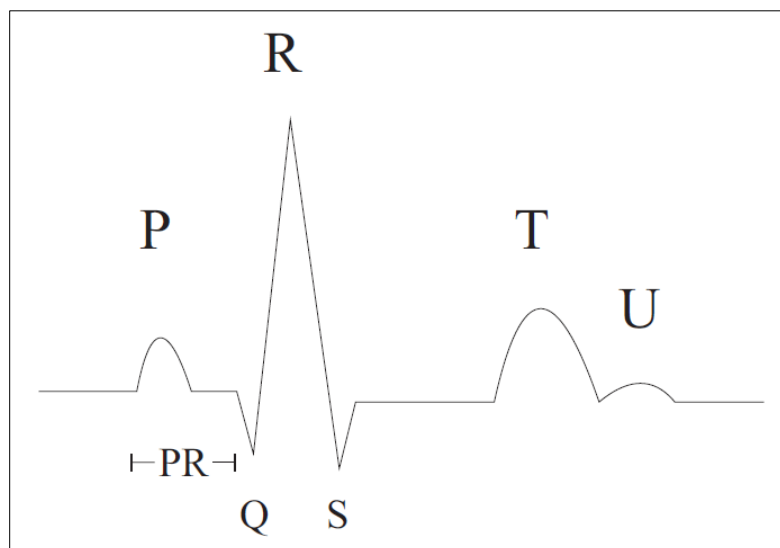


Figura 1 Registro de señal ECG típico [2]

Estas ondas están presentes en todo el registro ECG cada cierto período de tiempo, todas las señales ECG tiene una forma de onda característica denominada complejo QRS. La frecuencia de un registro de señal ECG se determina mediante la medición

temporal de dos complejos QRS sucesivos, dicha medición es de gran utilidad en cuanto al diagnóstico médico y registro electrocardiográfico. [2]

El rango espectral de la señal ECG va desde los 0 Hz a los 100 Hz, tal y como se aprecia en la tabla 1.

Tabla 1 Señales bioeléctricas típicas [2]

Señal	Amplitud (mV)	Rango Frecuencia (Hz)
ECG	0,02 – 5,0	0,05 – 100
EEG	0,0002 – 0,3	DC – 150
EMG	0,1 – 5,0	DC - 10000

2.1.1. ECG normal.

Un ECG normal consta de varias deflexiones de ondas que se alternan con respecto a la línea de base, en la figura 2 se muestra con detalle un electrocardiograma normal. [3]

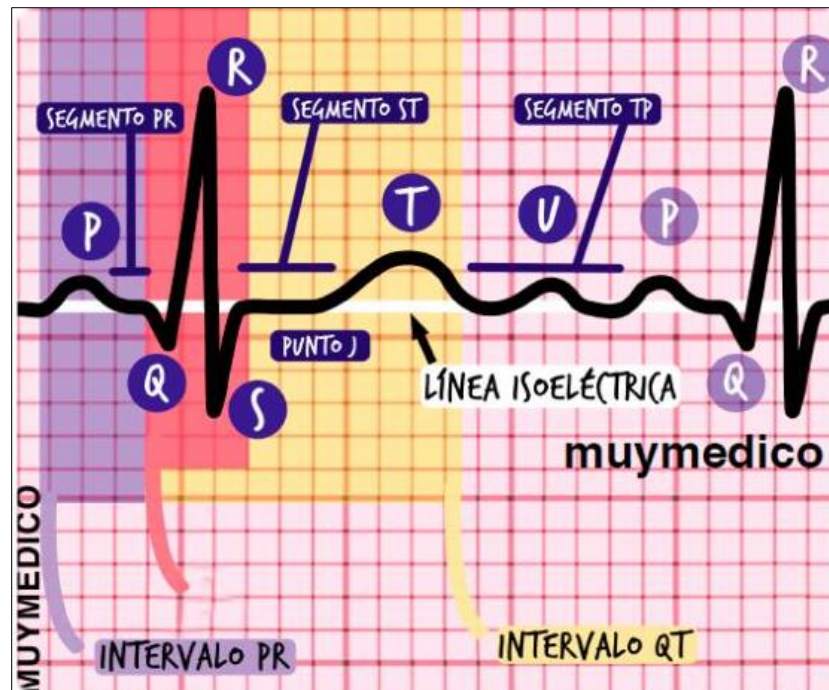


Figura 2 ECG normal [3]



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



1. **Onda P.** Es la primera deflexión con respecto a la línea de base que aparece en el ECG. Representa la contracción de las aurículas (despolarización auricular) y a su vez el envío de la sangre hacia los ventrículos.
2. **Intervalo PR.** El impulso viaja por el nodo auriculoventricular (AV), Haz de His.
3. **Segmento PR.** Se le conoce como la línea isoelectrica, es cuando las aurículas terminan de vaciarse y allí se produce una desaceleración en la transmisión de la corriente eléctrica que se presenta en el corazón un poco antes de dar inicio a la contracción de las cavidades inferiores (ventrículos).
4. **Complejo QRS.** Los ventrículos se contraen y eyectan toda la sangre. Representa la despolarización ventricular. Este complejo se compone de las ondas Q, las cuales indican si el complejo inicia con una onda negativa, onda R, son aquellas ondas positivas, si hay más ondas positivas se les denomina R', para aquellas que son ondas negativas tras la onda R se les llamará onda S, en caso de que se presente un voltaje mucho menor respecto al normal, se pone en minúscula.
5. **Punto J.** Indica el fin del complejo QRS e inicio del segmento ST.
6. **Segmento ST.** En caso de que se presente una elevación o descenso en relación a la línea de base puede indicar que se padece de insuficiencia en el riego del corazón.
7. **Intervalo QT.** Abarca la despolarización y repolarización ventricular. En parte, incluye la repolarización auricular, pero no se suele ver porque coincide con el complejo QRS
8. **Onda T.** Consiste en una deflexión que generalmente es positiva, esta onda representa el momento en el que corazón entra en un período de relajación una vez que ya ha expulsado la sangre que se estaba en los ventrículos, se conoce como la repolarización ventricular.
9. **Onda U.** No es común verla y se desconoce su origen.
10. **Segmento TP.** Línea isoelectrica del corazón (sin actividad eléctrica presente allí), este segmento sirve como referencia para valorar desniveles del segmento ST (punto J).



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



2.1.2. Derivaciones en el ECG.

Las derivaciones se denominan del posicionamiento de los electrodos en el paciente. Generalmente se realiza un registro ECG de 12 derivaciones, por tanto, esto capta el registro de potencial del corazón desde 12 puntos de vista distintos lo cual permite situar el corazón en el espacio mediante la captación de la señal bioeléctrica. [3]

Se definen 2 tipos de derivaciones:

- *Derivaciones en las extremidades*, otorgan información sobre el plano frontal del corazón [3]. Estas se clasifican en:
 - *Estándar (bipolares)*
 1. I
 2. II
 3. III
 - *Aumentadas (Monopolares)*
 1. aVF
 2. aVR
 3. aVL
- *Derivaciones en el tórax*, representan un corte transversal del corazón, van desde la derivada V1 hasta la derivada V6.

2.1.3. Errores en la obtención del electrocardiograma.

La realización correcta de un electrocardiograma permite una correcta interpretación del mismo; es necesaria la preparación del paciente antes de realizarse dicho examen, pues este debe estar calmado, sin agitación para evitar perturbaciones en la lectura, a su vez una correcta colocación de los electrodos para la adquisición de las distintas derivadas del corazón y una excelente aplicación de los parámetros técnicos. [4]

Uno de los errores de diagnósticos más comunes es debido a la presencia de artefactos en el cuerpo del paciente, como lo son celular, anillos, relojes, tarjetas de crédito, entre otros, a causa de lo anterior se conlleva a la toma de decisión de consumo de medicamentos antiarrítmicos o el uso de marcapasos mediante el implante. [4]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

2.2. ANATOMÍA DEL CORAZÓN.

El corazón con un tamaño un poco más grande que una mano cerrada y con un peso entre 200 y 425 gramos, se encuentra en el centro del pecho, entre los pulmones. Está envuelto en una membrana de dos capas como especie de bolsa denominada *pericardio*; la capa externa del pericardio rodea los vasos sanguíneos del corazón y esta se une con la espina dorsal y el diafragma por medio de ligamentos, por otro lado, la capa interna del pericardio está unida al músculo cardíaco. El corazón puede latir entre las dos capas que le envuelven gracias a una capa de líquido que las separa. [5]

El corazón tiene cuatro cavidades, se dividen en cavidades superiores denominadas *aurícula* tanto derecha como izquierda, y las cavidades inferiores denominadas *ventrículos*, tanto derecho como izquierdo. Estas cuatro cavidades se encuentran separadas por una pared muscular que lleva por nombre *tabique*. [5]

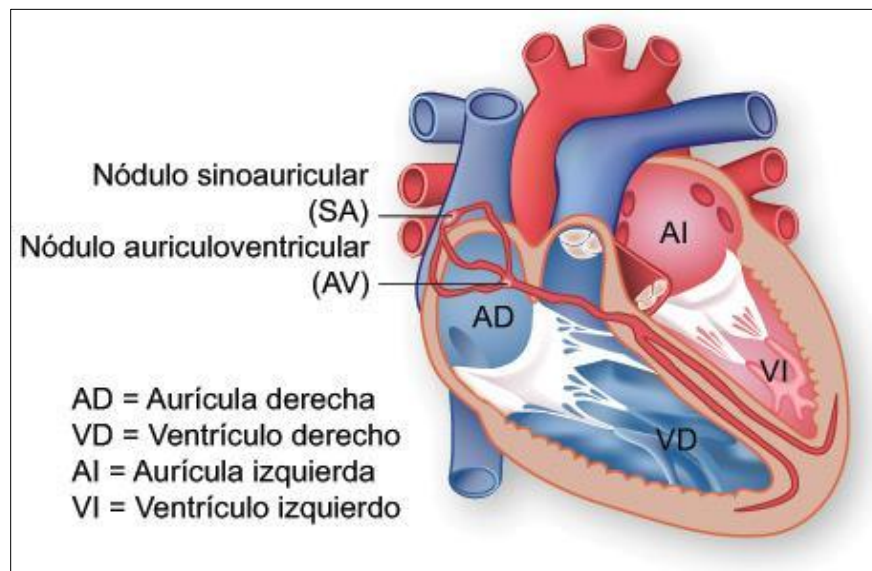


Figura 3 Anatomía del corazón [5]

2.2.1. Válvulas cardíacas.

Las válvulas cardíacas se encargan de controlar el flujo sanguíneo mediante la apertura y cierre de las mismas; hace que el flujo de sangre sea unidireccional y suficiente para abastecer cada uno de los órganos del cuerpo humano. [5]



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Estás se clasifican en cuatro:

- *Válvula tricúspide*, se encargar de controlar el flujo sanguíneo entre la aurícula y ventrículo derecho respectivamente.
- *Válvula pulmonar*, su principal función es la de transportar la sangre hacía los pulmones para ser oxigenada. Esta controla el flujo sanguíneo entre el ventrículo derecho y las arterias pulmonares.
- *Válvula mitral*, permite el paso de la sangre oxigenada de la aurícula izquierda hacía el ventrículo izquierdo.
- *Válvula aórtica*, permite el paso de la sangre oxigenada del ventrículo izquierdo a la aorta, esta es la arteria la cual se encarga de transportar la sangre al resto del organismo.

2.2.2. Potencial de acción de las células miocárdicas.

Se clasifica en 5 fases diferentes y se caracterizan por los diferentes iones que participan en cada una de ellas. [5], [6]

- *Fase 0 (despolarización)*, apertura de canales de sodio (Na), estos dependen del voltaje.
- *Fase 1 (potencial de acción cardíaco)*, la salida de potasio (K) provoca una repolarización rápida y corta.
- *Fase 2 (meseta)*, es la responsable de la contracción del corazón, es debido a la apertura de los canales lentos de sodio y calcio.
- *Fase 3 (repolarización rápida)*, donde se abren los canales para el potasio, la permeabilidad de la membrana aumenta.
- *Fase 4 (exacerbación de la bomba de sodio potasio)*, es la encargada de bombear iones sodio hacia el exterior de la célula por medio de la membrana, al tiempo que introduce potasio del exterior hacía el interior.

2.2.3. Problemas que surgen por el mal funcionamiento del aparato circulatorio y el corazón.

Las afecciones cardíacas afectan a niños, jóvenes y ancianos tanto en hombres como en mujeres. Los problemas circulatorios del corazón se pueden clasificar en dos



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



categorías, *congénitos* y *adquiridos*; para el primero, el problema persiste desde el nacimiento de la persona y el segundo, las afecciones se desarrollaron en algún momento de su ciclo de vida. [5]

- *Defectos congénitos*, se define como anomalías que se presentan en la estructura del corazón desde el nacimiento. Comúnmente estos defectos se propician durante el desarrollo del feto por causa de alteraciones genéticas.
- *Defectos cardíacos adquiridos*, se tienen algunos de los defectos cardíacos adquiridos:
 - *Arritmia (disritmia)*, se conocen como anomalías en el ritmo del corazón.
 - *Cardiomiopatía*, enfermedad crónica conocida por debilitar el músculo cardíaco (miocardio).
 - *Enfermedad arteriocoronaria*, provocada por la *arterosclerosis*. Taponamiento de las arterias coronarias debido a depósitos de grasa, calcio y células muertas.
 - *Hiperlipidemia (hipercolesterolemia)*, debido al exceso de colesterol en la sangre, este es una sustancia cerosa que se concentra en las células del cuerpo, sangre y alimentos.
 - *Hipertensión (presión sanguínea elevada)*, con el tiempo puede provocar daños al corazón y arterias.
 - *Enfermedad de Kawasaki*, afecta las membranas de la mucosa, la piel y los ganglios, esta puede provocar *vasculitis* lo cual causa inflamación en los vasos sanguíneos.
 - *Enfermedad cardíaca reumática*, puede provocar daño permanente al corazón e incluso la muerte debido a la fiebre reumática que no se trata a tiempo.

2.2.4. Sistema de conducción.

El sistema de conducción cardíaco, conocido como *cardioconector* está conformado por un conjunto de fibras miocárdicas que se encargan de la conducción del corazón, su principal función es la de producir y transmitir los impulsos eléctricos los cuales permiten que se generen las contracciones del corazón. En la figura 4 se aprecia el esquema del sistema de conducción cardíaca. [7]



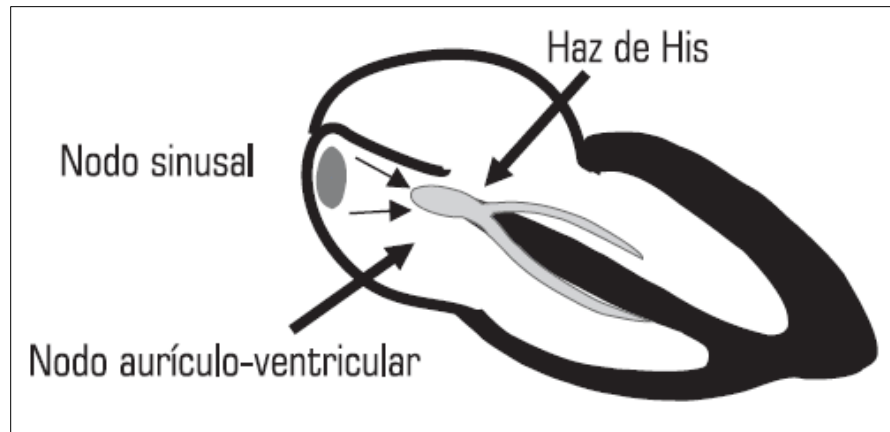


Figura 4 Esquema del sistema de conducción cardíaca [7]

Se agrupa en nódulos y haces, se pueden dividir en dos segmentos.

- *Nodo sinusal*, el nodo sinusal, también conocido como *Nodo de Keith*, es la estructura donde se origina el impulso eléctrico el cual activa todo el corazón. Se posiciona entre la aurícula derecha cerca de la vena cava superior. [7]
Sus características de funcionamiento se originan por:
 - *Automatismo*
 - *Autoexcitación*
- *Nodo aurículo-ventricular*, el nodo aurículo-ventricular se encarga de retrasar la conducción y protege el ventrículo de frecuencias cardíacas rápidas. Se encarga de formar el impulso (automatismo). [7]
- *Haz de His*, se encarga de transmitir los impulsos a los ventrículos del corazón para contraerlos. La velocidad de conducción que se genera por el Haz de His (estructura) es demasiado rápida debido al tipo de organización que tienen las fibras. [7]

Se divide en dos ramas:

- *Rama izquierda del haz de His*
- *Rama derecha del haz de His*



2.3. ELECTROCARDIOGRAMA EN LAS CARDIOPATIAS CONGÉNITAS.

2.3.1. Hipertrofia auricular derecha.

Una hipertrofia auricular derecha se caracteriza por alteraciones en la onda P con medidas por encima de 2,5 mm en el ECG denominadas P pulmonar en las derivadas II, III y aVR. Otra de las alteraciones que presenta esta cardiopatía son los QRS con un voltaje menor en V1. El comportamiento de la señal ECG se aprecia en la figura 5. [3], [7]

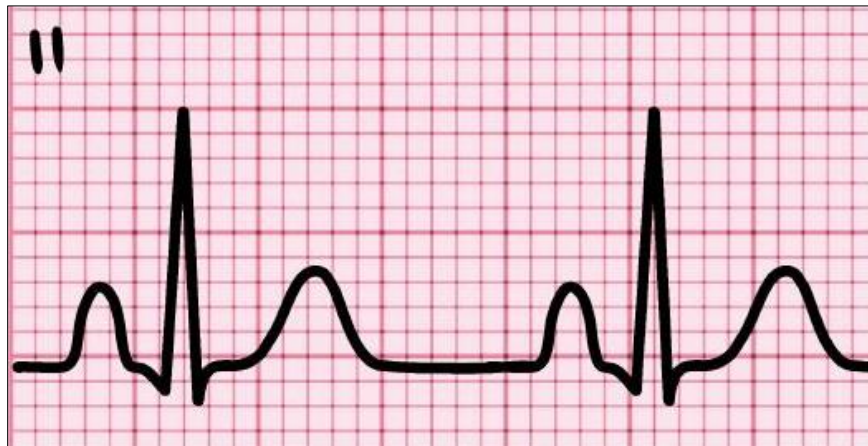


Figura 5 Hipertrofia Auricular Derecha [3]

2.3.2. Hipertrofia auricular izquierda.

En la hipertrofia auricular izquierda existen falsos negativos sin ritmo sinusal, la onda P es bimodal en sus extremidades. Presenta alteraciones en las derivadas I, II y V1. La figura 6 presenta las alteraciones de la onda P para la hipertrofia auricular izquierda.[3], [7]

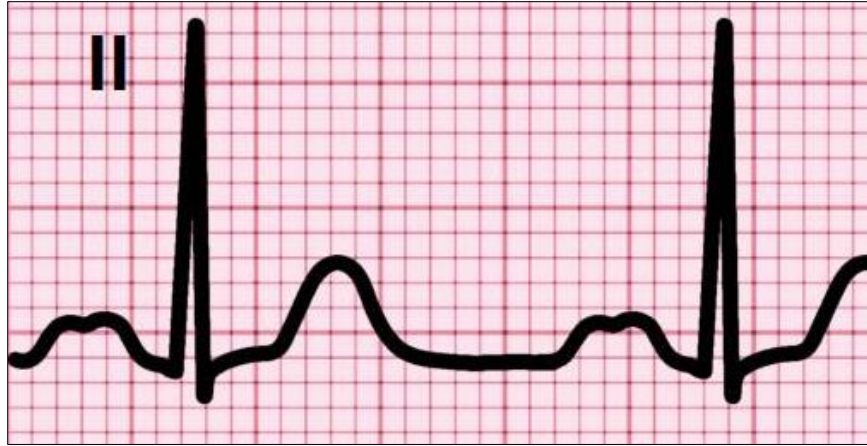


Figura 6 Hipertrfia Auricular Derecha [3]

2.3.3. Hipertrfia atrial combinada.

Presenta duraciones considerables y aumento de voltaje en la onda P con amplitudes mayores a 2,5 mm en el ECG. [7]

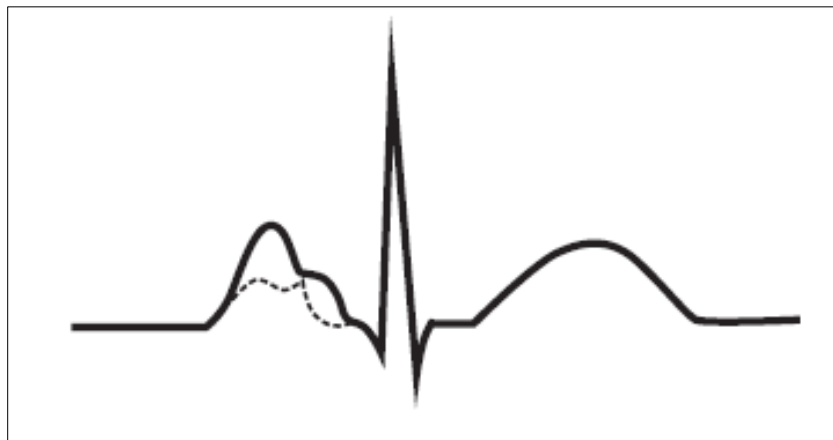


Figura 7 Hipertrfia atrial combinada [7]

2.3.4. Hipertrofia ventricular derecha.

La hipertrofia en sus inicios solo produce aumento de voltaje, con el paso del tiempo afecta la repolarización cuando se presentan cambios en la onda T. Se presencian alteraciones en la onda R con un aumento de voltaje en las derivadas V1 y V2, el eje se desplaza a la derecha y la onda S es más negativa para las derivadas V5 y V6. [3], [7]

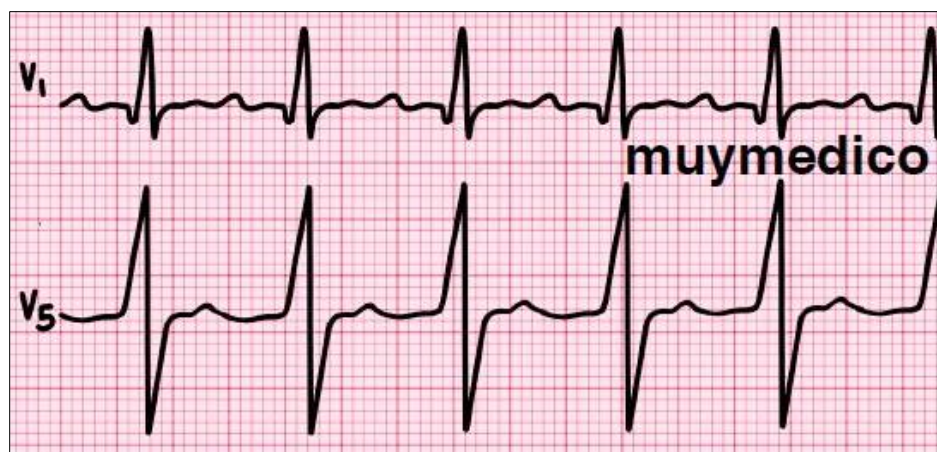


Figura 8 Hipertrofia ventricular derecha [3]

2.3.4.1. Criterios de hipertrofia ventricular derecha.

1. Eje desviado a la derecha de la pared.
2. Se presencia un aumento de voltaje en el complejo QRS en las derivaciones derechas y anteriores.
3. Patrón QR, donde la onda Q está entre 0.5 a 1 mm y la onda R es mayor a 10mm en V1 y V4. Esto implica una presión sistólica del ventrículo derecho en personas de cualquier edad.
4. RSR' en V1.
5. La relación R/S anormal a favor del ventrículo derecho en ausencia de bloqueo de rama.
6. Se presentan cambios en la onda T, es positiva en la derivada V1 en pacientes mayores de tres días. Si la onda T está invertida de forma asimétrica en V1, significa que hay un patrón de sobrecarga de presión del ventrículo derecho.

7. Onda q (Voltaje menor respecto al normal, presenta un patrón qR o qRs).

2.3.4.2. Patrones de sobrecarga de presión correspondientes al ventrículo derecho.

- Sobrecarga de volumen, causa bloqueo incompleto de la rama derecha del haz de His.
- Sobrecarga de presión, se presentan ondas R altas en precordiales derechas con alteración en la morfología del complejo QRS (Rs, rS, qRs o qR).

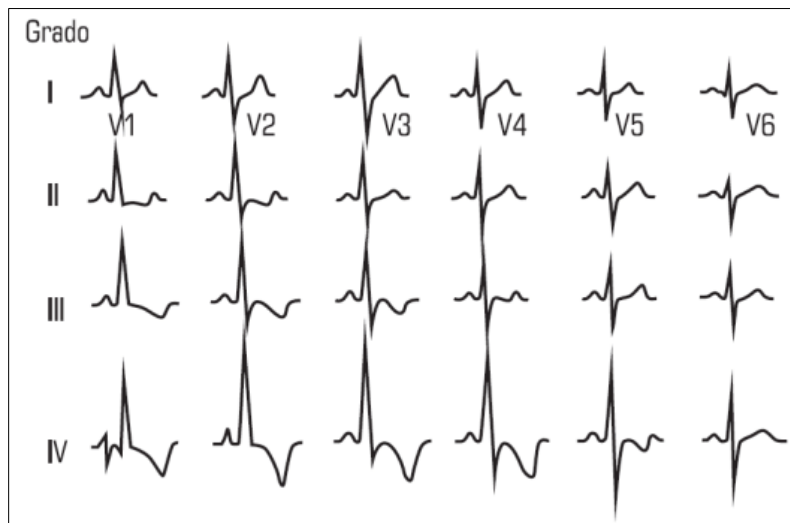


Figura 9 Patrones de sobrecarga de presión del ventrículo derecho [7]

2.3.5. Hipertrofia ventricular izquierda.

La hipertrofia ventricular izquierda (HVI) es la más común debido a la sobrecarga de presión, en la HVI leve se presenta un aumento en la onda R con respecto a la derivada V6 con ondas T positivas, por el contrario, cuando son severas, hay un retraso sobre la conducción y alteración de la repolarización, debido a esto se observa una alteración en base a la onda T y depresión del segmento ST. Las alteraciones que se presentan allí son, desviación del eje del complejo QRS hacia la izquierda, posiblemente QRS puede presentar un voltaje alto, la onda S se puede presentar de forma negativa tanto en la derivada V1 y V2. [3], [7]

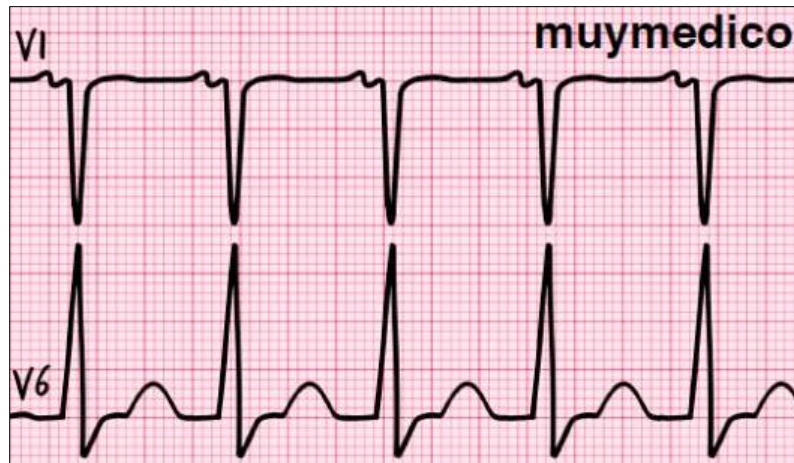


Figura 10 Hipertrofia ventricular izquierda [3]

2.3.5.1 Criterios de hipertrofia ventricular izquierda.

1. Desviación del eje hacia la izquierda.
2. Predominio del ventrículo izquierdo en los voltajes del QRS.
3. Relación R/S anormal a favor del ventrículo izquierdo.
4. La onda Q presenta anomalías para las derivadas II, III, aVF, V5 y V6.
5. Se presentan cambios en la onda T.

2.3.5.2 Patrón de sobrecarga correspondientes al ventrículo izquierdo.

- Sobrecarga de volumen, la onda R presenta aumentos en V5 y V6 con retardos.
- Sobrecarga de presión, se presenta bloqueo de rama derecha del haz de his incompleto.

2.4. ELECTROCARDIOGRAMA EN DEFECTOS DE LA CONDUCCIÓN INTRAVENTRICULAR.

El nodo auriculoventricular continúa de manera distal con el haz de His, este se divide en dos ramas, *rama derecha* y *rama izquierda del haz de His*. Los bloqueos de rama del Haz de His son causados por alteraciones en la conducción que se presentan en

una de sus ramas, esto conlleva a que el impulso se propague por la rama no bloqueada. [7]

2.4.1. Bloqueo de rama derecha del haz de His.

El bloqueo de rama derecha se debe a un menor grosor y mayor recorrido de la misma, esto causa que la rama izquierda sea más vulnerable. Se pueden presentar alteraciones como, complejo QRS con una onda r de baja amplitud inicial en V1 seguido de una onda S seguida de otra onda R, denominada R', pues se presentan dos ondas R en la señal ECG, por tanto, se denomina como rSR', a su vez el segmento ST y la onda T pueden presentar polaridad inversa al complejo QRS. Estas variaciones se pueden detallar en la figura 11. [3], [7]



Figura 11 Bloqueo de rama derecha [3]

2.4.2. Bloqueo de rama izquierda del haz de His.

El impulso eléctrico desciende por la rama derecha, lo cual causa la estimulación del ventrículo derecho y el septum interventricular de forma sincrónica en dirección contraria y el ventrículo izquierdo a través del septum interventricular con retardo. Las alteraciones cuando se presenta un bloqueo de rama izquierda causan una onda R positiva elevada en las derivadas V5 y V6 y ondas negativas en el segmento QS en V1 y V2, a su vez, en V1 no se presenta onda r inicial de forma positiva y en V6 no se halla

la primera onda q (negativa), la onda T y el segmento ST presenta polaridad inversa al complejo QRS. [3], [7]

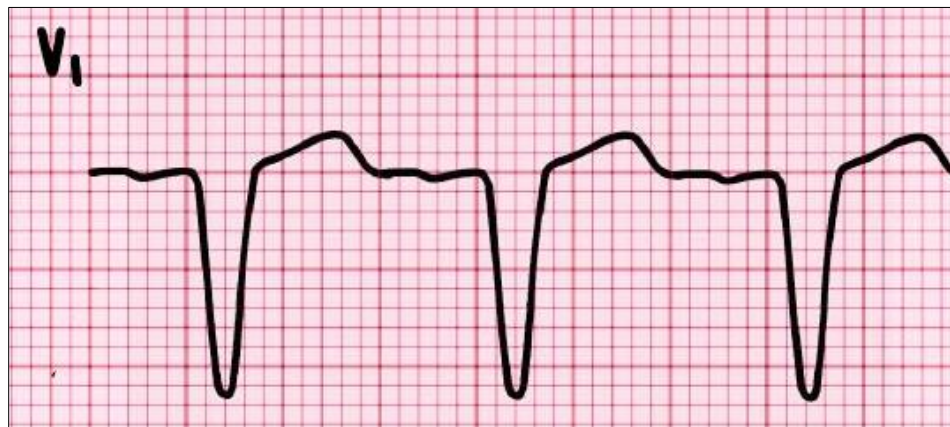


Figura 12 Bloqueo de rama izquierda [3]

2.5. ANORMALIDADES VENTRICULARES (VÁLVULO PATÍAS).

Cuando las válvulas no funcionan de forma adecuada, se causa daño a la acción de bombeo del corazón. En la figura 13 se muestra la vista general de 4 defectos más comunes que influyen función ventricular izquierda.

2.5.1. Estenosis aórtica.

La figura 13A visualiza el comportamiento de la señal cuando se presenta la afección valvular de estenosis aórtica. Se presentan alteraciones con signos de hipertrofia ventricular izquierda y a su vez las ondas R tienen un alto voltaje que se pueden notar en las derivadas V5, V6, I y aVL. [3], [8]

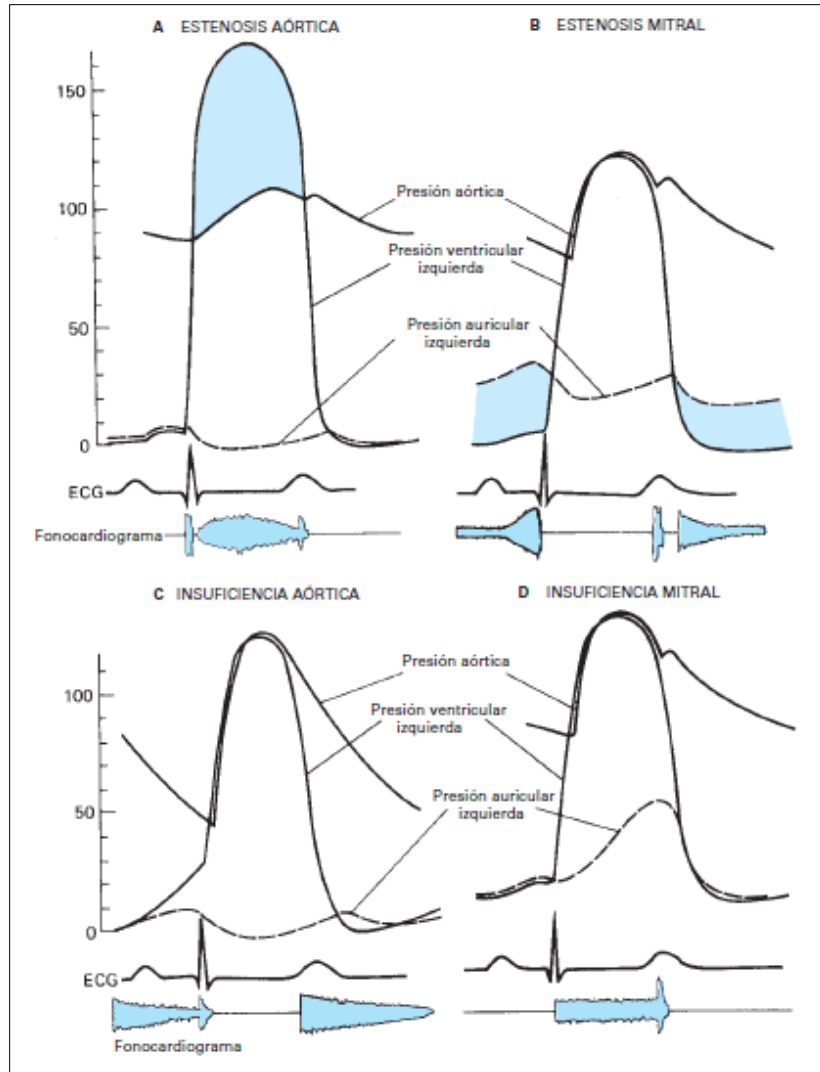


Figura 13 Afección Valvular. A, estenosis aórtica. B, estenosis mitral. C, insuficiencia aórtica. D, insuficiencia mitral [8]

2.5.2. Estenosis mitral.

La figura 13B visualiza el comportamiento de la señal cuando se presenta la afección valvular de estenosis mitral. La presión auricular izquierda incrementa con la estenosis mitral. La hipertrofia auricular derecha y la hipertrofia ventricular derecha es avanzada



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



a causa de una hipertensión pulmonar, generada por la elevación de presión auricular izquierda. [3], [8]

2.5.3. Insuficiencia aórtica.

La figura 13C visualiza el comportamiento de la señal cuando se presenta la afección valvular de insuficiencia aórtica. Las hojuelas de la válvula aórtica no se cierran de manera correcta causando que la sangre regurgite desde la aorta hacia el ventrículo izquierdo durante el período diastólico, esto refleja signos de hipertrofia ventricular izquierda grave. [3], [8]

2.5.4. Insuficiencia mitral.

La figura 13C visualiza el comportamiento de la señal cuando se presenta la afección valvular de insuficiencia mitral. Presenta un ECG normal leve y crecimiento ventricular izquierdo con una onda R de alto voltaje en las derivadas V5-V6, I y aVL. Cuando la válvula mitral es insuficiente implica que la sangre regurgita desde dentro del ventrículo izquierdo hacia la aurícula izquierda durante el período sistólico. [3], [8]

2.6. DISRITMIAS (ARRITMIAS).

Las disritmias también conocidas como arritmias, se clasifican en dos grupos: *Bradiarritmias*, aquellas que presentan una frecuencia cardíaca por debajo de 60 pulsaciones por minutos y las *Taquiarritmias*, las cuales presentan frecuencia cardíaca por encima de las 100 pulsaciones por minuto. [9]

2.6.1 Bradiarritmias.

Esta cardiopatía se presenta más que todo en atletas, son consideradas como patología cuando la frecuencia cardíaca son menores a 60 o 40 pulsaciones por minuto. Su origen de este problema radica en afecciones en el nodo sinusal, nodo auriculoventricular y el sistema de His-Purkinje. En la figura 14 se tiene observa una señal ECG que presenta



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



bradicardia sinusal con pulsaciones alrededor de 44 latidos por minuto, cabe resaltar que dicha patología pertenece a las enfermedades del nodo sinusal. [9]



Figura 14 Bradicardia sinusal [9]

2.6.2 Taquiarritmias.

Las taquiarritmias comúnmente son causadas por anemia, ejercicio, fiebre, entre otros. Cualquier ritmo cardíaco por encima de las 100 pulsaciones por minuto se declara como taquiarritmia [9]. Se pueden clasificar mediante los siguientes criterios:

1. Sitio de origen.
2. Mecanismo.
3. Electrocardiograma superficial con presencia de QRS ancho y estrecho.
4. Presentación clínica donde presente o no colapso hemodinámico.

2.6.2.1. Supraventriculares.

- Fibrilación auricular.
- Flutter auricular.
- Taquicardia atrial ectópica.

2.6.2.2. Ventriculares.

- Taquicardia en corazón sano.
- Taquicardias en corazón cardio miopático.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



2.7. PROCESAMIENTO DE SEÑALES.

El procesamiento de señales se aplica con el objetivo de leer la información que contienen las mismas, principalmente se refiere a la manipulación de estas mediante la implementación de diversas técnicas especializadas, con el fin de extraer ciertas características o patrones importantes de señales unidimensionales o imágenes 2D por medio de herramientas o algoritmos computacionales.

2.8. EXTRACCIÓN DE CARACTERÍSTICAS.

La aplicación de técnicas en los conjuntos de señales que se han consolidado para la extracción de patrones o características es útil para el proceso de clasificación generando como resultados propiedades de separación óptimas. En el siguiente apartado se muestra con más detalles las técnicas empleadas.

2.8.1. Transformada discreta de Fourier (DFT).

En la adquisición de señales de tipo bioeléctricas surgen muchos factores de interferencias, tales como, el ruido de la red eléctrica, ruidos de fuentes naturales y/o artificiales, instrumentos en uso, contracciones corporales y demás señales bioeléctricas que se pueden captar durante el proceso de adquisición se consideran señales con frecuencia distinta, por tanto, es útil filtrar estas frecuencias por medio de la aplicación de la transformada discreta de Fourier.

La DTF es una de las principales técnicas de procesamiento aplicada a señales digitales. Permite transformar señales 1D y 2D del dominio espacial o también denominado temporal al dominio de las frecuencias o dominio espectral permitiendo un análisis más detallado con respecto a sus componentes de frecuencia. [10], [11]

La DTF se satisface por la ecuación (1) que se describe a continuación.

$$G(k) = \sum_{n=0}^{N-1} g(n)e^{-j\frac{2\pi knT}{N}}, \quad k = 0, \dots, N-1 \quad (1)$$



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



La información de la señal es la misma tanto en el dominio temporal como en el dominio de frecuencia, ningún dominio proporciona mayor información que otro, es decir, contiene el mismo número de muestras. La inversa de la transformada discreta de Fourier (IDFT), se le conoce como *ecuación de síntesis* y está expresada por la ecuación (2). [12]

$$g(n) = \frac{1}{N} \sum_{k=1}^{N-1} G(k) e^{j \frac{2\pi knT}{N}}, \quad n = 0, \dots, N - 1 \quad (2)$$

Con respecto a la ecuación 1 y 2, existe una similitud entre estas, debido a esto se causan características duales en las señales de ambos dominios discretos. [12]

El algoritmo para calcular la DFT se denomina Transformada Rápida de Fourier (FFT). [11]

2.8.2. Transformada wavelet discreta (DWT).

La transformada Wavelet es una técnica de procesamiento esencial para señales unidimensionales e imágenes, la gran ventaja que proporciona esta herramienta es que tiene diversas capacidades que otras transformaciones no tienen. La transformada wavelet discreta acepta señales continuas, por tanto, solamente aplica escalas que sean discretas para la formación de la transformada. La DWT se usa para describir el comportamiento o propiedades de las señales con respecto al tiempo, contrario a la DTF la cual divide en secciones de tiempo la señal, más bien la DWT divide las formas de ondas en segmentos de escala. [12], [13]

La transformada wavelet discreta presenta una gran ventaja con respecto a la transformada de wavelet continua (CWT), pues la segunda proporciona sobre muestreo en la forma de onda de la señal original, es decir, genera demasiados coeficientes con respecto a los que se necesitan para especificar de forma única la señal. La DWT ahorra dicha redundancia al restringir la variación de las escalas. [13]

La siguiente ecuación de síntesis (3) describe la transformada wavelet discreta.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

$$x(t) = c \sum_{j=0}^{N-1} \sum_{k=0}^{M-1} W_{jk} \Psi_{jk}(t) \quad (3)$$

2.8.2.1. Descomposición Wavelet.

Para filtrar el ruido en señales aplicando DWT, se aplica en varios niveles; se realizan dos procesos conocidos como los coeficientes de wavelet, se tienen filtros pasa bajos y pasa altos de descomposición que corresponden al coeficiente de aproximación y coeficiente de detalle respectivamente; lo anterior implica una separación de la señal en dos grupos, el primero donde contiene las frecuencias altas y el segundo donde contiene las frecuencias bajas. Los niveles de descomposición se implementan mediante un árbol jerárquico de filtros, la estructura del árbol se compone pasando a la siguiente etapa los coeficientes de aproximación o coeficientes de detalles y así sucesivamente según convenga tal y como se observa en la figura 15. [14], [15]

- Las características de los coeficientes son los siguientes:
 - *Filtro mitad banda*, este permite el paso solo para la mitad de la banda superior o inferior de X señal de entrada.
 - *Ganancia en DC*, con un factor de valor $\sqrt{2}$.
 - *Los filtros de descomposición y reconstrucción don de cuadratura*, el cual permite la recuperación de la señal, lo anterior es posible cuando se aplica un nivel de descomposición que no degrade la señal.

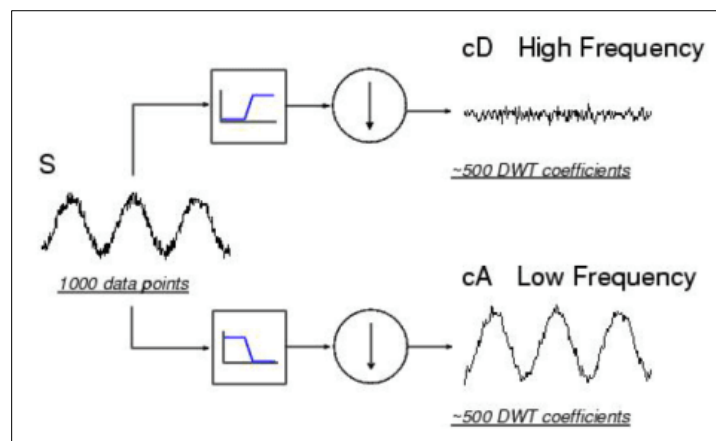


Figura 15 Proceso árbol jerárquico de filtrado [15]



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



2.8.3 Entropía.

La entropía es otra medida física que se usa comúnmente para el procesamiento de señales biomédicas e inclusive imágenes. Se basa en la teoría de información de Shannon; útil para el análisis y comparación de distribuciones de probabilidad [16]. La entropía mide la cantidad de información ya sea de un enunciado o variable, en otras palabras, permite calcular parte de la energía que no se emplea para producir trabajo. Se representa mediante la ecuación (4). [12]

$$H = - \sum_{i=1}^n p_i \ln(p_i) \quad (4)$$

Donde, p_i es la probabilidad de los i -estados.

2.8.4. Energía.

Las señales ECG se descomponen en subbandas necesarias para la eliminación de ruido usando DWT, luego, las subbandas seleccionadas se usan en la construcción del espacio de características para la posterior clasificación a partir de estos coeficientes o subbandas se puede calcular la energía relativa para cada nivel de descomposición; la energía es de gran importancia para la extracción de información de señales y se define con la ecuación (5). [17], [18]

$$ER = \frac{E_j}{E_{total}}, \quad j = 1 \text{ hasta } n \quad (5)$$

Donde E_j se define con la ecuación (6).

$$E_j = \sum_k |C_{j,k}^i|^2 \quad (6)$$

2.8.5. Normalización.

La normalización transforma el conjunto de datos de modo en que el mínimo de la señal se transforma en cero y el máximo se transforma en 1 y el resto es escalado de forma



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



lineal, en otras palabras, normalizar una señal es ajustar dicha señal a una curva normal, lo más importantes es que mantiene la misma forma de la señal sin que se pierda información. [19]

La normalización se satisface por la ecuación (7).

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (7)$$

2.9. TÉCNICAS DE CLASIFICACIÓN DE MACHINE LEARNING.

Con la clasificación el usuario puede realizar una toma de decisión o una acción con la información teniendo en cuenta el uso que se le va a dar; se compone de todo tipo de técnicas o herramientas computacionales avanzadas para clasificar datos de interés por medio de sus características o patrones.

El Machine Learning o aprendizaje automático permite a las máquinas mejorar en una tarea determinada con experiencia, en otras palabras, es un método científico que se usa para la extracción de patrones y relaciones que hay en los datos por sí solos, esos patrones son útiles para realizar predicciones de comportamientos y para tomas de decisiones futuras. Todas las técnicas del aprendizaje automático se clasifican como de inteligencia artificial (IA), estas técnicas permiten a las máquinas analizar datos pasados denominados *datos de entrenamiento*, crear su propio modelo para después realizar predicciones con datos de prueba. [20]

2.9.1. Aprendizaje Supervisado.

Este tipo de aprendizaje realiza el entrenamiento de algoritmos usando etiquetas en los datos de entrada y salida, es decir, el aprendizaje de las máquinas se realiza por medio de datos etiquetados previamente por el programador, el objetivo de este tipo de aprendizaje es aprender la forma de mapear las entradas en salidas dado los valores correctos por medio de un supervisor. [20]

En la figura 16 se aprecia el diagrama de flujo del aprendizaje supervisado.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

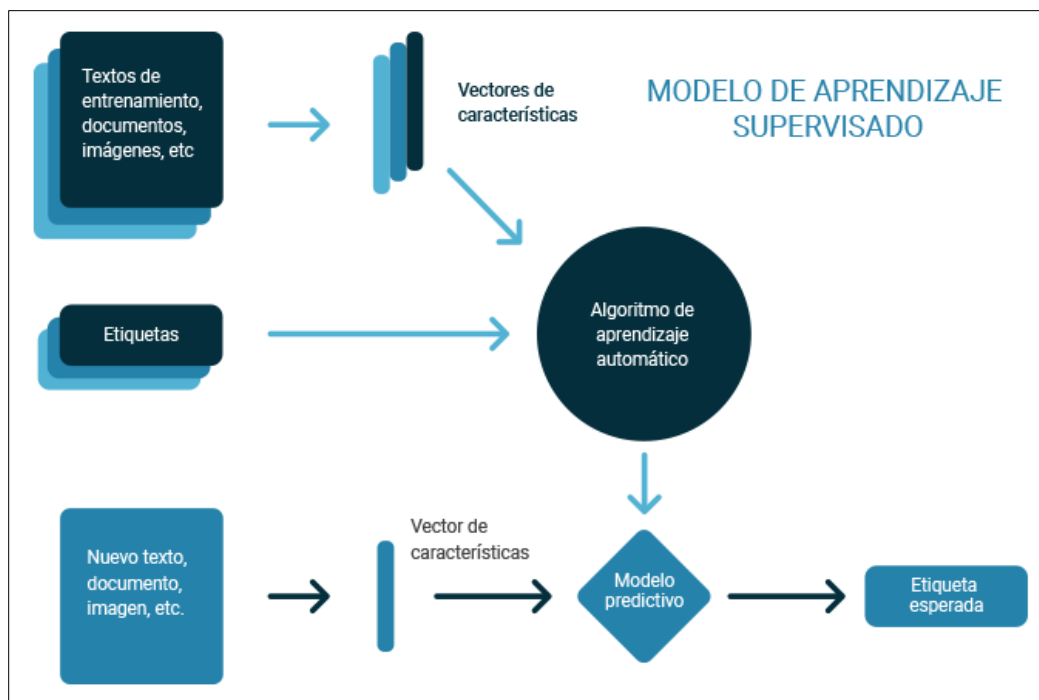


Figura 16 Diagrama de flujo del aprendizaje supervisado [20]

2.9.1.1. Máquina de soporte vectorial (SVM).

Las máquinas de soporte vectorial ofrecen una precisión muy amplia a comparación de otras técnicas de clasificación, es conocido por su truco de kernel para manejar los espacios de entrada no lineales. SVM (Support Vector Machine por sus siglas en inglés) es una nueva técnica de clasificación, basado en la idea de minimización de riesgo estructural, dados los dato de entrenamiento etiquetados (aprendizaje supervisado) el algoritmo genera un hiperplano óptimo (clasificador de maximización de los márgenes) entre las clases, en otras palabras, este hiperplano es una línea que divide un plano en dos partes donde cada clase o etiqueta se encuentra en cada lado, primero mapea los datos de entrada a un espacio de dimensión mayor, es decir, si los datos de entrada se encuentra en una dimensión \mathbb{R}^2 estos son mapeados por la máquina de soporte vectorial a \mathbb{R}^3 . [21]

La separación por el hiperplano se observa en la figura 17.

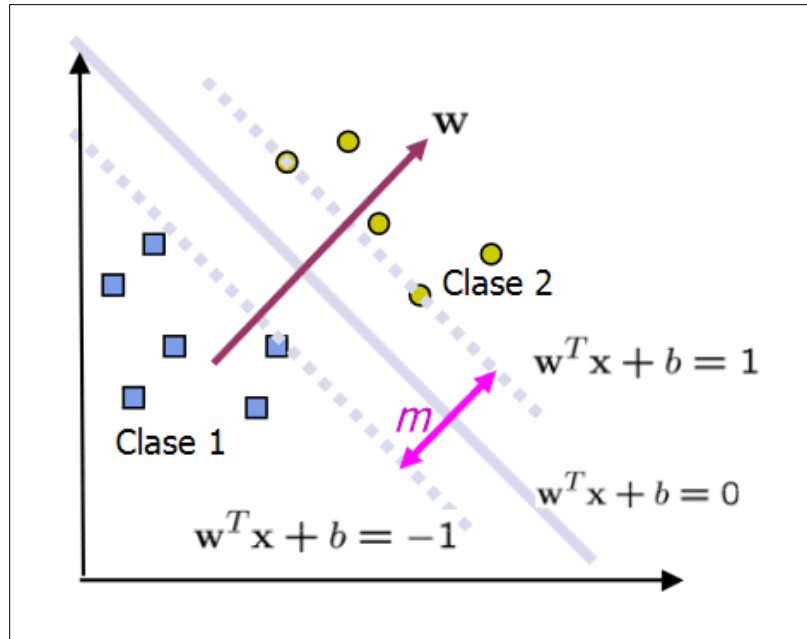


Figura 17 Hiperplano de separación de una SVM [21]

Los elementos que se encuentran en la figura 17 se definen de la siguiente manera:

- *Vector soporte*, puntos de datos más cercanos al hiperplano, son relevantes para la construcción del clasificador.
- *Hiperplano de separación*, plano de decisión que separa entre un conjunto de objetos que tienen membresía de clases diferentes
- *Margen (m)*, es el espacio que hay entre las dos líneas en los puntos más cercanos de la clase. Se calcula como la distancia perpendicular desde los puntos más cercanos al hiperplano
- *Peso (W)*, peso del vector.

El límite de decisión está dado cuando $y = 0$ [13], por tanto, usando la ecuación (8),

$$y = \sum_{i=1}^N w_i x_i + b = x_i w + b = 0 \quad (8)$$



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Donde:

x_i es el patrón de entrada

w es el peso del vector

b es una constante la cual mide que tanto el plano está separado del hiperplano

Se supone que las clases identificar como ± 1 en el análisis de una SVM, lo cual simplifica la parte matemática, debido a lo anterior, el valor de y debe ser ± 1 en los puntos más cercanos al hiperplano, por tanto, las ecuaciones para las rectas que forman el margen deben ser [13],

$$x_i w + b \geq 1 \text{ cuando } y = +1 \quad (8.1)$$

y

$$x_i w + b \leq 1 \text{ cuando } y = -1 \quad (8.2)$$

Por tanto, las ecuaciones (8.1) y (8.2) pueden ser combinadas de la siguiente forma,

$$y_i(x_i w + b) \geq 1 \quad (8.3)$$

La ecuación anterior establece que w y b realiza que las clases caigan al lado apropiado de la línea del vector de soporte del hiperplano.

La ecuación (8.4) define la distancia al origen que separa la clase 1 dese la línea,

$$d_0 = \frac{(1 - b)}{\|w\|}, \text{ si } y_i \geq 1 \quad (8.4)$$

Y la ecuación (8.5) define la distancia de separación de la clase -1 desde la línea al origen,

$$d_0 = \frac{(-1 - b)}{\|w\|}, \text{ si } y_i \leq 1 \quad (8.5)$$

La resta ente las ecuaciones (8.4) y (8.5) permiten obtener el margen definido por la ecuación (8.6).



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



$$M = \frac{(1 - b)}{\|w\|} - \frac{(-1 - b)}{\|w\|} = \frac{2}{\|w\|} \quad (8.6)$$

El margen máximo se obtiene minimizando $\|w\|^2$.

Donde,

$\|w\|^2$ es igual a $w^T w$.

2.9.1.2. Clasificador bayesiano (Naïve Bayes).

Naïve Bayes es uno de los algoritmos más poderosos y simples para la clasificación basado en el Teorema de Bayes, es muy útil para conjuntos de datos muy grandes. Uno de sus principales inconvenientes es el desconocimiento de la distribución de probabilidad de clases involucradas, sin embargo, si se tiene un gran número de muestras de datos, se podría estimar la probabilidad. [12], [22]

La ecuación (9), describe el teorema de Bayes.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (9)$$

Donde,

$P(h)$, es la *probabilidad a priori* o *probabilidad previa* de h , a groso modo es la probabilidad de que la hipótesis h sea cierta.

$P(D)$, conocida como la *probabilidad marginal*, describe que tan probable es que un individuo presente dichas características (probabilidad de los datos independientemente de la hipótesis).

$P(D|h)$, se conoce como *probabilidad condicionada*, es la probabilidad de los datos D dado que la hipótesis h era cierta.

$P(h|D)$, conocida como la *probabilidad a posteriori*, es la probabilidad de la hipótesis h dada los datos D .

2.9.1.3 Árboles de decisión.

Los árboles de decisión es un algoritmo perteneciente a Machine Learning, siendo el una de las técnicas más populares por la razón de gráficamente un ser humano puede interpretar lo que sucede. Esta técnica se divide en dos:

- *Árboles de clasificación*, ayuda a predecir una categoría en base a las observaciones, por ejemplo, Hombre – Mujer, Color, estado de la persona, comprar o no comprar, entre otras.
- *Árboles de regresión*, trata con datos cuantitativos.

Gracias a su estructura, ofrece diversas elecciones adecuadas entre muchas posibilidades, es decir, permite la selección una y otra vez de las distintas opciones para explorar las diferentes alternativas que se presentan para una toma de decisión. Esta técnica es jerárquica con múltiples vías donde el valor de la característica es el criterio del diagnóstico que permita evaluar la toma de decisión correcta. Su estructura se representa mediante la figura 18. [23], [24]

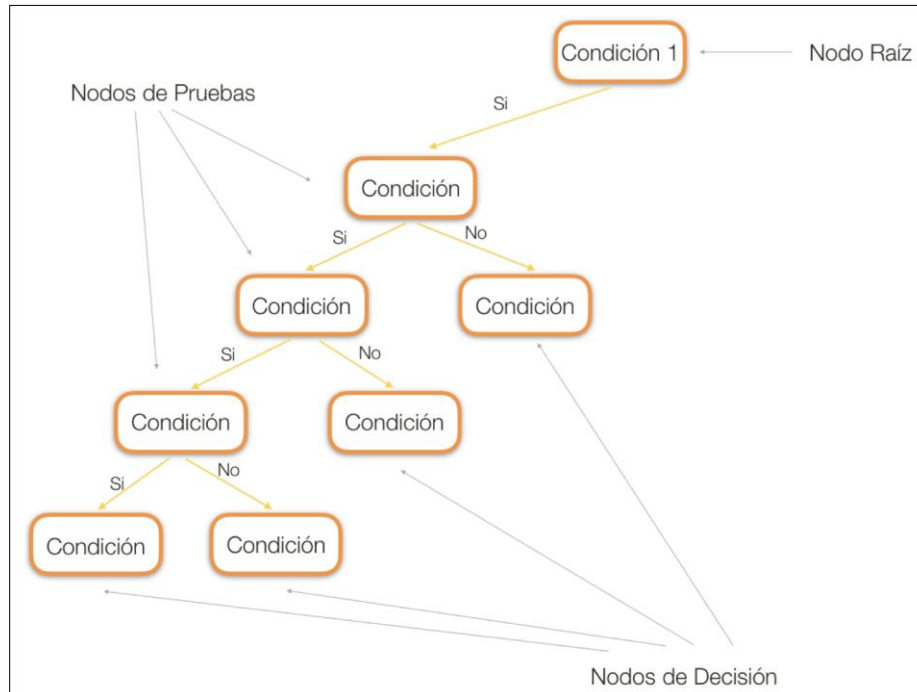


Figura 18 Diagrama de flujo para árboles de decisión [24]



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Donde,

Nodo raíz, es el nodo superior del árbol de decisión

Nodos de prueba, representan una característica

Las ramas, simbolizan una regla de decisión.

Nodos de decisión u hoja, es el resultado.

La heurística conocida como la regla de partición, es una medida de selección de los atributos los cuales permiten determinar puntos de ruptura para un conjunto en un nodo específico. [24]

2.9.1.3.1. Índice de Gini.

Métrica para medición de frecuencia con que un elemento es elegido al azar y es identificado como de forma incorrecta. [24]

2.9.1.3.2. Ganancia de información.

La ganancia de información mide el cambio en la entropía cuando se particionan las instancias en subconjunto más pequeños a partir de un nodo. [24]

Consideraciones para la construcción de un árbol de decisión empleando ganancia de información:

- Todas las instancias de formación se deben asociar al nodo raíz.
- Se emplea la ganancia de información para la elección de etiquetar atributos con nodos correspondientes.
- Cada subárbol se construye recursivamente en el conjunto de instancias que se clasificarán en ese camino.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



2.10. CLOUD COMPUTING.

Cloud Computing ofrece una gran flexibilidad y disponibilidad de recursos a menor costo, consiste en servicios que se ofrecen por medio de la red tales como aplicaciones, almacenamiento, acceso remoto, correo electrónico, entre otros, el único requerimiento para el uso de la computación en la nube es una computadora no necesariamente con buenas características, sino más bien con buena conexión a internet y que esta cuente con un navegador web. [25]

Los servicios generales que se ofrecen en distintas áreas se ofrecen a los clientes y estos pueden acceder a ellos por medio de un navegador web, por otro lado, los datos se almacenan en los servidores de la nube, estos servicios se clasifican en tres modelos de servicio en la nube:

- *Software as a Service (SaaS)*, se puede acceder a las aplicaciones desde diversos dispositivos por medio de un navegador web. Los proveedores de servicio en la nube (CSP) se encargan de ejecutar y mantener a flote a aplicación, software, sistema operativo (SO), etc. Del lado del cliente, el modelo SaaS se ve como una aplicación de interfaz web accesible desde un navegador donde ofrece servicios y software completo a través de la red. Algunos ejemplos de estos son correo electrónico, Google Docs, Gmail, Zoho, Salesforce, Google apps, entre otros. [25], [26]
- *Platform as a Service (PaaS)*, el consumidor puede desplegar aplicaciones ya sea adquiridas o creadas por el mismo, este solamente tiene el control sobre las aplicaciones que despliega y el alojamiento de las aplicaciones bajo el SO y recursos virtuales que proporciona el CP. El modelo PaaS ofrece al cliente la capacidad de poder diseñar, modelar, desarrollar y probar aplicaciones de forma directa en la nube, este a su vez tiene el control del ciclo de vida del software. Algunos ejemplos son Windows azure, Google app engine, Aptana Cloud, Developer tools, entre otros. [25], [26]
- *Infrastructure as a Service (IaaS)*, el CSP proporciona un conjunto de recursos informáticos virtualizados como ancho de banda, memoria de procesamiento, almacenamiento y otros recursos en la nube de forma que el consumidor/cliente despliegue y ejecute software arbitrario. En este modelo el cliente no tiene la capacidad de gestionar ni controlar la infraestructura de nube subyacente, pero este tiene control sobre el SO, almacenamiento, aplicaciones desplegadas y a su vez el control limitado de componentes de



red seleccionados. La tecnología en IaaS es virtual, pues este convierte los recursos físicos necesarios en recursos lógicos que pueden ser lanzados por los clientes. Algunos ejemplos de IaaS son Drop Box, Amazon EC2, AWS, Akamai, Servers, Storage, Bandwidth, Networking.

En la tabla 2 se muestra la asignación de recursos en los modelos de servicio en la nube para los CPS y clientes.

Tabla 2 Asignación de recursos en modelos de Cloud Computing [25]

	Software de aplicación	Sistema operativo	Recursos virtuales / Hardware
SaaS	CSP	CSP	CSP
PaaS	Cliente	CSP	CSP
IaaS	Cliente	Cliente	CSP

2.10.1. Tipos de Cloud Computing.

Independientemente del modelo en el que se realice el despliegue, se tiene cuatro principales tipos de Cloud Computing con distintos requisitos. En la figura 19 se aprecian los distintos tipos de nubes.



Figura 19 Tipos de Cloud Computing [27]



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



- *Nube pública*, la infraestructura y demás servicios están disponibles para el público por medio de la red, esta es propiedad de una organización y se gestiona por un CSP que vende los servicios que ofrece a los consumidores. La seguridad y privacidad son grandes problemas en esta nube. Algunos de los servicios en la nube públicos son Google App Engine, Salesforce.com y Amazon EC2. [25], [26]
- *Nube privada*, su infraestructura y recursos informáticos es operada únicamente por una organización, esta puede ser gestionada por la misma organización o un CSP, tiene un nivel de seguridad más seguro a comparación de una nube pública. [25], [26]
- *Nube híbrida*, resulta de la combinación entre dos o más nubes públicas, privadas o comunitarias. Los componentes están unidos por medio de tecnología estandarizada y son administrados como una sola unidad permitiendo la portabilidad de los datos y aplicaciones. [25], [26]
- *Nube comunitaria*, son semejantes a las nubes públicas, a diferencia de estas, su infraestructura y recursos informáticos son compartidos por varias organizaciones que tiene preocupaciones similares (misión, política, requisitos de seguridad y ciertas consideraciones acerca del cumplimiento normativo). Ejemplos: Nube educativa usada por universidades de todo el mundo para ofrecer servicios de educación, investigación y demás, a esto también se suman las instituciones. [25], [26]

2.11. ESTADO DEL ARTE.

En los últimos 5 años se han realizado diversos proyectos de investigación en el ámbito de la medicina haciendo referencia a cardiología. Se tiene un listado de 1.111 publicaciones de las cuales 673 se destacan en función de arritmias, cardiopatías, insuficiencia cardíaca, imagen cardíaca, entre otros. Dichos artículos se publicaron entre el 2014 y 2020. [28]

2.11.1. Trabajos realizados a nivel internacional.

En el año 2017 se realizó la publicación de una investigación que lleva por nombre *Clasificación de los latidos del ECG mediante la descomposición en modo empírico y la*



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



combinación de características (Santanu Sahoo, Monalisa Mohanty, Suresh Behera y Sukanta Kumar Sabut – Journal Of Medical Engineering & Technology) se basa en la extracción de características informativas para la clasificación de seis tipos de latidos cardíacos de señales ECG obtenidas de la base de datos de arritmias MIT-BIH aplicando herramientas de procesamiento de señales como Transformada de ondículas discretas (DWT), umbralización adaptativa, y análisis asistido por computadora. [29]

Se aplican las herramientas de procesamiento de señales mencionadas anteriormente para la eliminación de ruido y la detección de los picos R precisos y complejos QRS, las características propias de cada señal ECG son características temporales, morfológicas y estadísticas que se extraen de dichas señales ya procesadas las cuales se combinan para formar un conjunto de características donde se clasifican con una red neuronal probabilística y una red neuronal en base radial para reconocer los latidos de la arritmia. [29]

En el año 2014 se realizó la publicación y aceptación del proyecto de investigación Un enfoque de aprendizaje automático para la clasificación de la calidad de la señal de ECG multinivel (Li Qiao, Rajagopalan Cadathur, Clifford Gari D. – Journal Of Medical Engineering & Technology) dicho trabajo describe un algoritmo de clasificación de la calidad de la señal ECG de cinco niveles. [30]

Se realiza un entrenamiento con máquinas de soporte vectorial (VSM) para realizar la clasificación de señales ECG adquiridas desde la base de datos de arritmias MIT-BIH. Se tiene dos conjuntos, prueba y entrenamiento, por medio de la simulación se crean segmentos limpios del ECG y se agregan tres tipos de ruidos de ECG a diferentes niveles de relación señal-ruido (SNR). En dicho trabajo se derivaron 13 métricas de calidad de señal a partir de segmentos de formas de ondas de ECG, se entrenaron y se probaron diferentes combinaciones de las 13 métricas ya mencionadas en los conjuntos de datos simulados y se seleccionó y validó de la base de datos la mejor combinación que produjo la mayor precisión de clasificación. [30]

En el año 2017 se realizó Un modelo de red neuronal convolucional profunda para clasificar los latidos del corazón (U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, Ru San Tan - Journal Of Medical



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Engineering & Technology), donde se desarrolla una red neuronal convolucional profunda (CNN) de 9 capas para la identificación de 5 categorías diferentes de latidos en las señales de ECG. [31]

Este trabajo se lleva a cabo en un conjunto de señales originales y ruido, este conjunto fue incrementando artificialmente para emparejar el número de instancias de las 5 clases de latidos y se filtra para eliminar el ruido de alta frecuencia presente allí. La metodología consta de un preprocesamiento; se realiza una correcta eliminación del ruido y eliminación de la línea de base de todos los ECG, se aplica filtros Daubechies Wavelet 6, en segundo lugar, segmentación de los latidos del ECG; todas las señales de ECG son segmentados en segmentos que tienen 260 muestras de largo, posteriormente cada segmento se normaliza mediante la normalización de puntuación Z para abordar el problema de escala de amplitud y eliminar el efecto de compensación antes de la alimentación en la red de CNN para entrenamiento y pruebas. [31]

En el año 2018 se implementó el artículo de investigación *Detección de arritmias mediante una red neuronal convolucional profunda con señales de ECG de larga duración* (Özal Yıldırima, Paweł Pławiakb, Ru-San Tanc, U. Rajendra Acharya - Journal Of Medical Engineering & Technology) dentro de este artículo se a conocer un nuevo enfoque de aprendizaje profundo para la detección de 17 clases de arritmias cardíacas de larga duración con base en el análisis de señales electrocardiográficas (ECG). [32]

La investigación se basa en 1000 fragmentos de señales ECG de la base de datos de arritmias MIT-BIH para una derivación (MLII) de 45 personas. Se diseñó una estructura completa de extremo a extremo en lugar de la extracción y selección de características hechas a mano utilizadas en los métodos tradicionales. Se diseña un nuevo modelo de red neuronal convolucional 1D (1D-CNN). El método propuesto es ser eficiente, rápido (tiempo real para la clasificación), no complejo y fácil de usar (extracción y selección de características combinadas). [32]

En el año 2019 se realiza un proyecto basado en *Características geométricas para el reconocimiento de contracciones ventriculares prematuras con selección de algoritmos de aprendizaje automático basados en procesos de jerarquía analítica* (Bruno Rodríguez de Oliveira a, Caio Cesar Ensidede Abreu b, Marco Aparecido Queiroz



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Duarte c, Jozue Vieira Filho - Journal Of Medical Engineering & Technology) el cual presenta un nuevo método de reconocimiento de contracciones ventriculares prematuras basado en un conjunto simplificado de características, las cuales son extraídas de figuras geométricas construidas sobre el complejo QRS. [33]

El método en primera etapa consiste en el preprocesamiento basado en el escalado de la señal ECG de para la eliminación de ruido de la misma, después la señal se segmenta teniendo en cuenta el tiempo de despolarización ventricular y se extrae un nuevo conjunto de características geométricas. Se presentan nuevas simulaciones que abarcan ocho clasificadores diferentes con el motivo de la validación de este enfoque. Para la selección de los clasificadores se basa en el proceso de jerarquía analítica. [33]

2.11.2. Trabajos realizados a nivel nacional.

En el año 2017 se publica un artículo de investigación en la ciudad de Bogotá – Colombia, que lleva por nombre Diagnóstico automático del síndrome coronario agudo utilizando un sistema multiagente basado en redes neuronales (John Jaime Sprockel Díaz, Juan José Diaztagle Fernández y Enrique González Guerrero – Revista Colombiana de Cardiología), su objetivo es desarrollar un sistema multiagente que ensamble las decisiones de varias redes neuronales para el diagnóstico del dolor torácico enfocado a los síndromes coronarios agudos. [34]

Su metodología consiste en el estudio de pruebas diagnósticas en el que se entrenan un conjunto de redes neuronales con una precisión aproximada al 70% para después ser ensambladas mediante tres sistemas de votación y luego adicionar el resultado de redes especiales en poblaciones particulares y seleccionar la mejor configuración que hará parte de un sistema multiagente para el diagnóstico del dolor torácico. El diagnóstico del síndrome coronario agudo representa una tarea muy compleja. Su principal forma de presentación, el dolor torácico, implica tener un gran número de consideraciones en los cuales la mayoría son benignas que ponen en riesgo la vida. Una de la herramienta de mayor uso en el campo de la IA son las redes neuronales artificiales las cuales se definen como un paradigma computacional basado en modelos matemáticos los cuales tienen la capacidad de reconocer patrones. [34]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Para el año 2015 se publicó el artículo de investigación *Modelo computacional para la identificación de endofenotipos y clasificación de pacientes con artritis reumatoide a partir de datos genéticos, serológicos y clínicos, utilizando técnicas de inteligencia computacional* (Luis Morales Muñoz, Gerardo Quintana y Luis Fernando Niño – Revista Colombiana de Reumatología) donde utiliza modelos de inteligencia computacional para la clasificación de identificación de endofenotipos en pacientes con artritis reumatoide y controles sanos, a partir de información genética y la teoría del epítoto compartido. [35]

Los autores desarrollaron modelos computacionales para clasificación aplicando las técnicas de inteligencia computacional (rama de la inteligencia artificial) tal como las redes neuronales, redes bayesianas y métodos como k-means. El conjunto de datos seleccionado empleado por los modelos tecnológicos obtuvo información de pacientes con artritis reumatoide (AR) y controles sanos. El estudio a sus inicios contaba con 135 pacientes y 140 los cuales pertenecen a un control sano, para un total de 275 registros. Al realizar la etapa de preprocesamiento de los datos, varios registros fueron eliminados por presentar información inconsistente. [35]

Para el año 2014 se desarrolló una investigación denominada *Redes neuronales en el diagnóstico del infarto agudo de miocardio* (John J. Sprockel, Juan J. Diaztagle, Wilson Alzate y Enrique González – Revista Colombiana de Cardiología) donde se toman datos de pacientes mayores de 18 años que consultaron el servicio de urgencias del Hospital San José por dolor torácico como motivo de consulta principal, se emplea NEUROSOLUTIONS, un ambiente de desarrollo de redes neuronales creado por NeuroDimension Inc. [36]

Una de las primeras causas de muerte en el mundo por enfermedades no transmisibles se debe al infarto agudo de miocardio. Las redes neuronales se emplean como herramienta de soporte para un diagnóstico preciso. En dicha investigación se realizó el entrenamiento y prueba de varias redes neuronales, empleando diferentes arquitecturas para el diagnóstico del infarto agudo de miocardio, a partir de los datos de escala de clasificación del grupo de pacientes que ingresaron por dolor torácico al servicio de urgencias del Hospital San José de Bogotá. [36]



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



2.11.3. Trabajos realizados a nivel regional.

En el año 2016 se desarrolló una tesis de investigación que lleva por nombre Desarrollo de un sistema para clasificación de caracteres obtenidos de señales electroencefalográficas (EEG) aplicado en pacientes con afasia motriz (Germán Arley Portilla González – Universidad de Pamplona) el cual presenta una alternativa de comunicación para aquellas personas que presenten afasia motriz causada por accidentes cardiovasculares que afecten las áreas cerebrales importantes en el proceso del habla. [37]

Se adquieren las señales mediante puntos específicos de la cabeza del paciente, se acondicionan dichos datos y se llevan a un sistema de filtrado digital donde se tomarán los parámetros con base a los impulsos eléctricos debido a la interacción neuronal presente, luego se realiza la búsqueda de características o patrones determinados para cada actividad cerebral generada por medio de un movimiento reflejado sobre las áreas de interés, cabe tener en cuenta que la afasia motriz causa diversos problemas en las personas tales como problemas psicológicos como consecuencia de la pérdida del lenguaje y el déficit de comunicación con otras personas. [37]

Para el año 2020 se realizó la publicación del artículo de investigación Procesamiento de señales cerebrales provenientes de estímulos visuales y auditivos utilizando análisis wavelet y redes neuronales artificiales (Alberto Lizcano-Portilla, Luis Enrique Mendoza, Zulmary Nieto-Sánchez – Revista UIS Ingenierías), presenta el diseño y desarrollo de un prototipo portátil para la adquisición, procesamiento y clasificación de señales EEG con el objetivo de caracterizar estímulos visuales y auditivos. [38]

El avance tecnológico, enfocándose al estudio del cerebro humano se ha desarrollado en gran medida durante los últimos años, las señales EEG, señales bioeléctricas y demás señales relacionadas con la parte de la actividad eléctrica del cerebro es importante para poder comprender el funcionamiento del cerebro humano. La importancia de dicha investigación radica en el número de pacientes con diferentes problemas tipo discapacidad motriz, problemas de concentración, entre otros., y la necesidad que provoca priorizando el costo y la disponibilidad de desplazamiento, por tanto, este trabajo ofrece una solución a esta problemática. Se aplican técnicas de



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



procesamiento de señales e inteligencia artificial para tratar las señales EEG, lo cual reduce los costos en cuanto al diseño y fabricación de filtros complejo. [38]

Por último, en el año 2003 se publicó la investigación denominada Redes neuronales artificiales en conducción de calor multidimensional transitorio (E. Liliana Arias, Lenny G. García, Rodrigo Correa – Revista UIS Ingenierías), este artículo ilustra el uso de las redes neuronales para la predicción de la temperatura de fenómenos de conducción de calor multidimensional transitorio. [39]

Se plantean dos casos en una dimensión con condición inicial constante y condiciones de frontera, para uno, de Dirichlet y para el otro, convectivas. Con estas mismas condiciones, se abordan geometrías en dos y tres dimensiones y se desarrollan sus soluciones analíticas para obtener los patrones de entrada y salida utilizados en el posterior entrenamiento, verificación y generalización de las redes neuronales artificiales. Para predecir la temperatura de los casos estudiados a partir de variables espaciales y temporales mediante la inteligencia artificial, Redes Neuronales Artificiales, se empleó el Perceptrón multicapa con conexiones hacia adelante. [39]

El desarrollo de los proyectos de investigación mencionados anteriormente influye de forma precisa y directa en el avance tecnológico y aplicación de las nuevas herramientas como lo es la inteligencia artificial. En el área de la medicina el uso de nuevas tecnologías es de mucha importancia puesto que permite una mejor calidad de servicio en cuanto a la salud para aquellas personas que padecen de ciertas patologías, no solamente se emplean para campos como cardiología, también se emplean para la parte cerebral, discapacidad, así como también para otras áreas.

La realización del presente proyecto con lleva como ventaja para las entidades de salud una disminución del tiempo de espera en cuanto al diagnóstico por parte del especialista de ciertas patologías cardíacas poco comunes. La implementación de la interfaz de clasificación es de mucha utilidad dado que el tiempo estimado para la detección y diagnóstico de patologías cardíacas es de una semana por mucho, tiempo el cual el paciente podría avanzar a una etapa más riesgosa debido a un tratamiento tardío puesto que no es conveniente la aplicación de procedimientos sin el conocimiento de la falla cardíaca en el paciente, por lo tanto, dicho proyecto permite un rendimiento más



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



rápido y preciso en cuanto a la detección de la patología presente una vez que se obtiene la señal ECG y permite una mejora de la estabilidad de la salud de los pacientes por el hecho de ser diagnosticada la patología en tiempo preciso.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

CAPÍTULO III: METODOLOGÍA

En el presente capítulo se da a conocer el paso a paso de la metodología implementada para el desarrollo del sistema de clasificación de señales ECG basándose en conceptos generales con respecto a las temáticas a tratar, la adquisición de señales mediante una base de datos, el preprocesamiento para el acondicionamiento de las señales, procesamiento de las mismas para la adquisición de patrones característicos, su respectiva clasificación mediante la implementación de técnicas de IA basados en Machine Learning, su respectiva interfaz gráfica y el alojamiento de código en la nube por medio de Pythonanywhere. En la figura 20 da a conocer la estructura metodológica de cada uno de los aspectos más relevantes del trabajo de investigación.

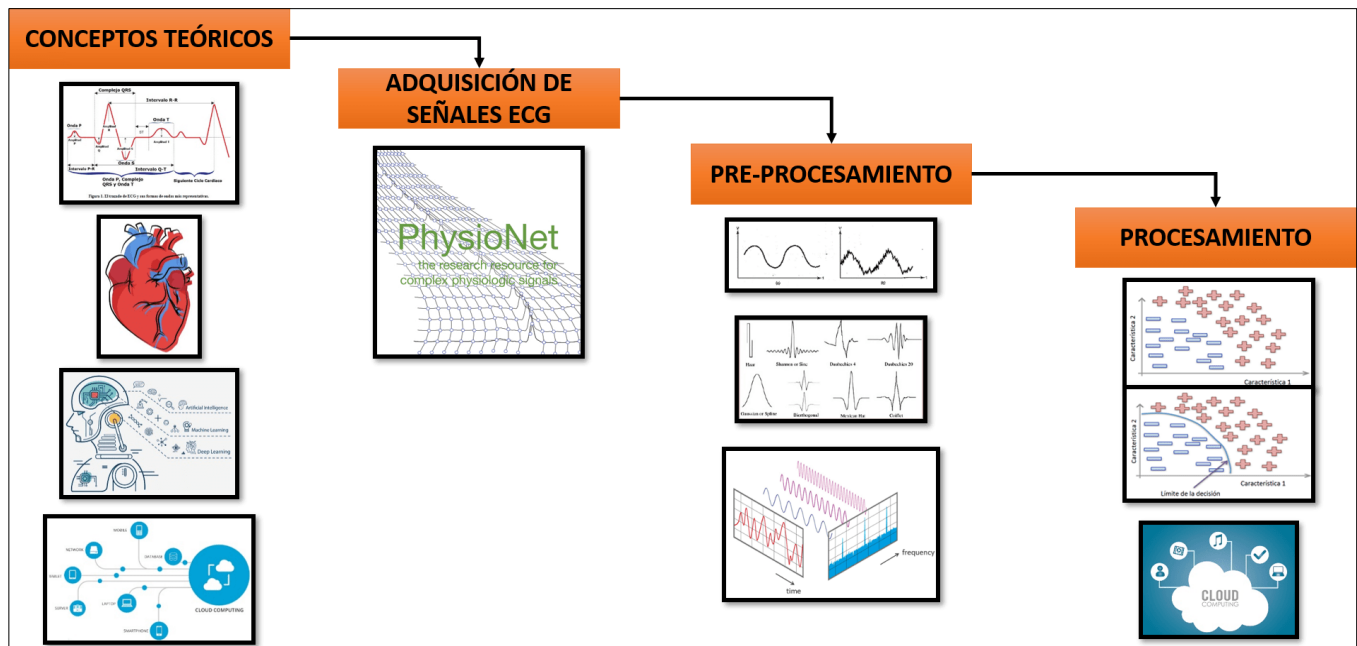


Figura 20 Estructura metodológica



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



3.1. CONCEPTOS TEÓRICOS.

La conceptualización teórica es de gran importancia puesto que se obtendrá conocimiento a lo referente, por tanto, se dará una mejor legibilidad al desarrollo del trabajo, los conceptos a tener en cuenta son todos aquellos que se han escrito en el capítulo 2, por tanto, es necesario cada uno de estos.

Los electrocardiogramas registran la actividad eléctrica del corazón y gracias a estos es posible determinar si una persona presenta patologías ya sea fisiológicas o que haya adquirido por alguna otra afección durante las etapas de su vida, por ello, el saber en qué momento y en qué zona del corazón se presenta alguna alteración en la funcionalidad del mismo permite que la persona se someta a exámenes médicos que mejoren su calidad de vida gracias a la detección de la afección con tiempo anticipado, con respecto a lo anterior, las afecciones cardiovasculares han sido estudiadas más a fondo y gracias a la tecnología que se fomenta hoy en día es posible tratar la mayoría de estas de una mejor manera y detectarlas lo antes posible, por ello, la importancia del conocer sobre los síntomas que causan estas enfermedades y las consecuencias que pueden generar es de gran ayuda para los expertos puesto que reducen la mortalidad de la población con respecto a dicho fenómeno.

Las herramientas computacionales tales como la IA y técnicas para el tratamiento de señales bioeléctricas es de gran ventaja para el área de medicina puesto que con ayuda de estas es posible analizar las señales y generar resultados óptimos para el procesamiento de las mismas y así poder solucionar problemas de diagnósticos más anticipado y lograr que la persona no padezca de afecciones más graves o dado el caso sufra de muerte severa a causa de una afección cardiovascular, de otro modo, un inconveniente que pueden presentar las entidades de salud es la poca gestión de sistemas que ayuden a agilizar los procesos de diagnósticos en el menor tiempo posible; por tanto, hoy en día es posible realizar monitoreos en tiempo real desde cualquier parte del mundo a través de la red gracias al alojamiento de softwares y/o aplicaciones en la web por medio de Cloud Computing, es allí donde entra esta temática la cual ofrece diversas ventajas para la carga de datos, almacenamiento, procesamiento y demás para la gestión de los mismos y generar resultados precisos y dar una solución a tiempo con respecto a la problemática que se esté tratando.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



3.2. ADQUISICIÓN DE SEÑALES ELECTROCARDIOGRÁFICAS.

La consolidación de las señales ECG se realiza por medio de la base de datos denominada Physiobank, la cual consta de pequeños números de registros que pueden ser demasiado grandes con un tamaño alrededor del gigabyte. La gran mayoría de las datas se desarrollan en el MIT y el hospital de Beth Israel de Boston las cuales son de acceso gratuito para la comunidad investigadora. [40]

En la figura 21 se puede observar el sitio web Physiobank.

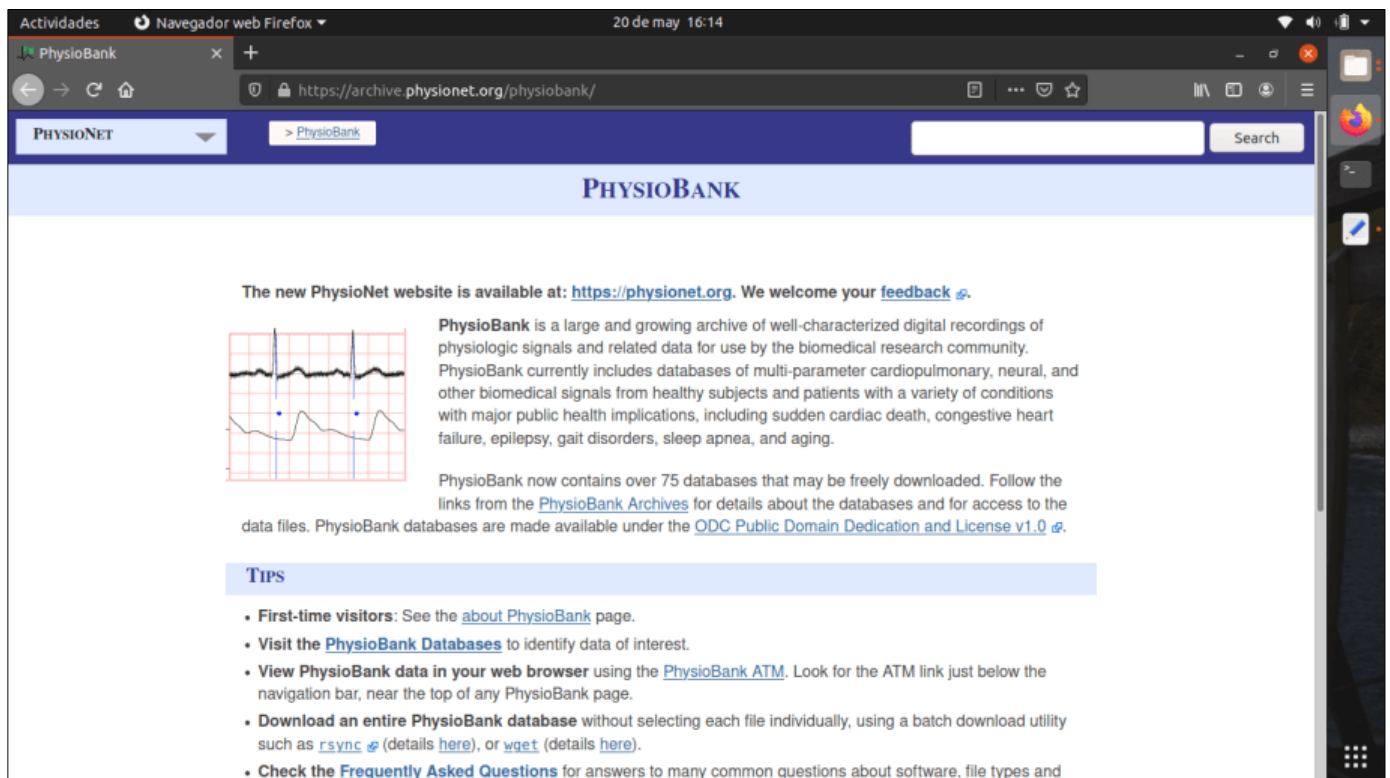


Figura 21 Sitio web Physiobank

Physiobank permite ver y exportar sus datos mediante *Physiobank ATM* denominado cajero automático que permite la exploración mediante el navegador web, tiene como herramienta la opción de explorar y exportar datos convirtiendo los archivos de señal



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



WFDB en archivos de texto, CSV, EDF o .mat, este último tipo de archivo es en el que nos centraremos a descargar. [40]

En la figura 22 se muestra con detalle la selección de una de las bases más comunes para patologías cardíacas (existen muchos más bandos de patologías) denominada *PTB Diagnostic ECG Database (ptbdb)*, donde se puede seleccionar la grabación de la señal por medio de la elección con la que se ha etiquetado los datos, a su vez se pueden seleccionar las señales, ya sean las 12 derivaciones del corazón o tan solo una derivada, la opción de anotaciones se omite, como salida del dato se seleccionan grabaciones de 1 minuto, y el formato del dato será estándar, y como exportación del dato se selecciona el archivo como .mat.

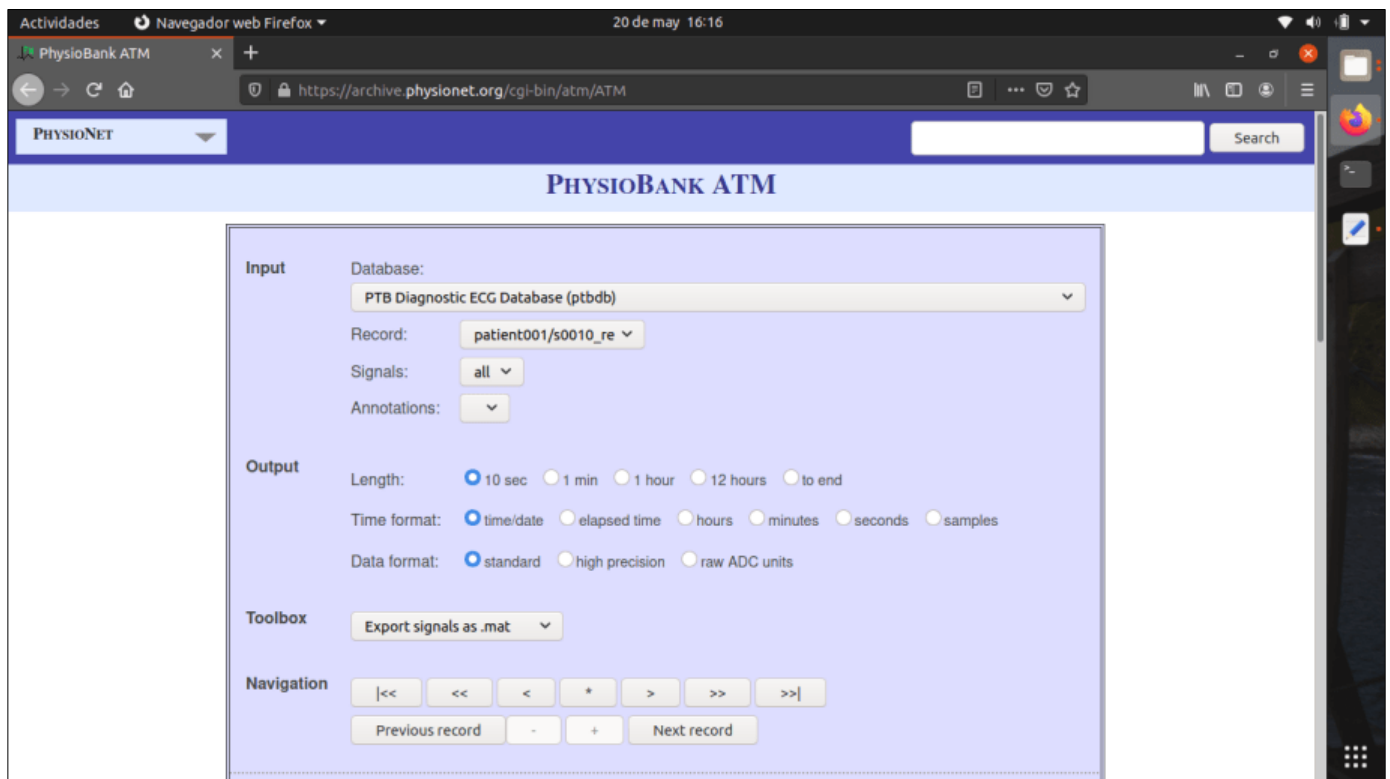


Figura 22 Banco de descarga de señales

Una vez realizado el proceso de exportación de los datos para poder descargar la señal o señales. En la figura 23 se aprecia el tipo de archivo a descargar, su cabecera y su



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



información. El archivo .mat es la señal de interés a descargar, el archivo .info es un archivo .txt donde contiene información de grabación de la señal (se necesita la información que contiene más adelante para la etapa de procesamiento) y el archivo .hea, describe con detalle las características de las señales, datos del paciente, como fecha, edad, sexo, diagnóstico y demás.

Actividades Navegador web Firefox 20 de may 16:16

PhysioBank ATM

https://archive.physionet.org/cgi-bin/atm/ATM

Selected input: record ptbdb/patient001/s0010_re , from 0:00.000 to 0:10.000 [PTB Diagnostic ECG Database \(ptbdb\)](#)

The output below was prepared using this command:

```
vfdb2mat -r ptbdb/patient001/s0010_re -f 0 -t 10 -l s1000000 ->s0010_rem.info
```

Download these files:

- [s0010_rem.mat](#) (binary, 300192 bytes; the matrix of raw signal values)
- [s0010_rem.info](#) (text, 618 bytes; signal names and other information about s0010_rem.mat)
- [s0010_rem.hea](#) (text, 2861 bytes; needed to read s0010_rem.mat using applications in the [WFDB Software Package](#) or functions in the [WFDB Toolbox for MATLAB](#))
- [plotATM.m](#) (m-code text; a function that reads s0010_rem.mat and s0010_rem.info and plots the converted data.)

How to read these files in MATLAB or Octave:
Download both s0010_rem.mat and s0010_rem.info. Also download plotATM.m if you have not done so previously.

Each row of s0010_rem.mat contains the samples of one signal. Each column contains a sample of each signal observed at the same time. The time intervals between consecutive columns are equal and specified in s0010_rem.info.

In MATLAB or Octave, run the command

```
plotATM('s0010_rem')
```

Figura 23 Descarga de archivos

Una vez que han descargado los archivos necesarios se procede a cargar en Python, se analiza la estructura del archivo .mat y se desglosan las señales (derivaciones), para ello es necesarios de ciertos módulos en Python donde se cargan los tipos de datos y que permitan realizar la descomposición de las derivadas, puesto que el archivo es tipo celda, y contiene un encabezado que no se requiere. En la figura 24 se aprecia el código de los módulos que se requieren en Python y de la carga de los datos, con respecto a lo anterior, la lectura de los archivos .mat en Python se realiza mediante la función *loadmat* la cual se encuentra en la librería *scipy.io*.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



```
4 import numpy as np
5 import matplotlib.pyplot as plt
6 import scipy.io as sio
7 from scipy.fftpack import fft, ifft, dct, idct
8 import pywt
9 import math
10
11
12 #####
13 # CARGANDO SEÑAL Y AJUSTANDO #
14 #####
15
16 #Cargando Señal:
17 Senales = sio.loadmat("s0030_ren.mat") # Cargando Señal .mat
18 print(type(Senales))
```

Figura 24 Código de módulos para carga de paquetes .mat

Al cargarse el archivo en Python, la función `type` permite saber qué tipo de archivo es, para este caso, se trata de un diccionario en el que se puede localizar cada una de las variables que se encuentra en dichos archivos, a su vez se muestra el contenido o la estructura del archivo el cual la variable de interés es `val` (tipo array) por el hecho de contener los puntos de muestras de la señal o derivada i para este caso, la figura 25 muestra con detalle lo explicado anteriormente.

```
handoathlete@FernandoB:~/Documentos/Sistema Clasificador de Señales ECG/1. Procesamiento/1. Valvular/PI$ ipython3 Procesamiento.py
Tipo de dato de la señal: <class 'dict'>

Estructura del paquete .mat: {'__header__': b'MATLAB 5.0', '__version__': '1.0', '__globals__': [], 'val': array([[ -54,  -39,   2, ...,
438,  454,  449],
[ -1557, -1573, -1537, ..., 1465, 1456, 1439],
[ -1503, -1534, -1538, ..., 1026, 1001,  989],
...,
[  409,  422,  414, ..., 194, 197, 195],
[ -642, -661, -648, ..., 458, 467, 472],
[  -77,  -69,  -75, ..., 151, 152, 151]], dtype=int16)}
```

Figura 25 Tipos y estructura del archivo .mat

En la figura 26 se puede observar la señal de derivación i una vez que se descompuesto la celda, dicha señal contiene alrededor de 30000 puntos de muestra.

NOTA: La figura inferior es un acercamiento a los tres primeros picos de la derivada i .

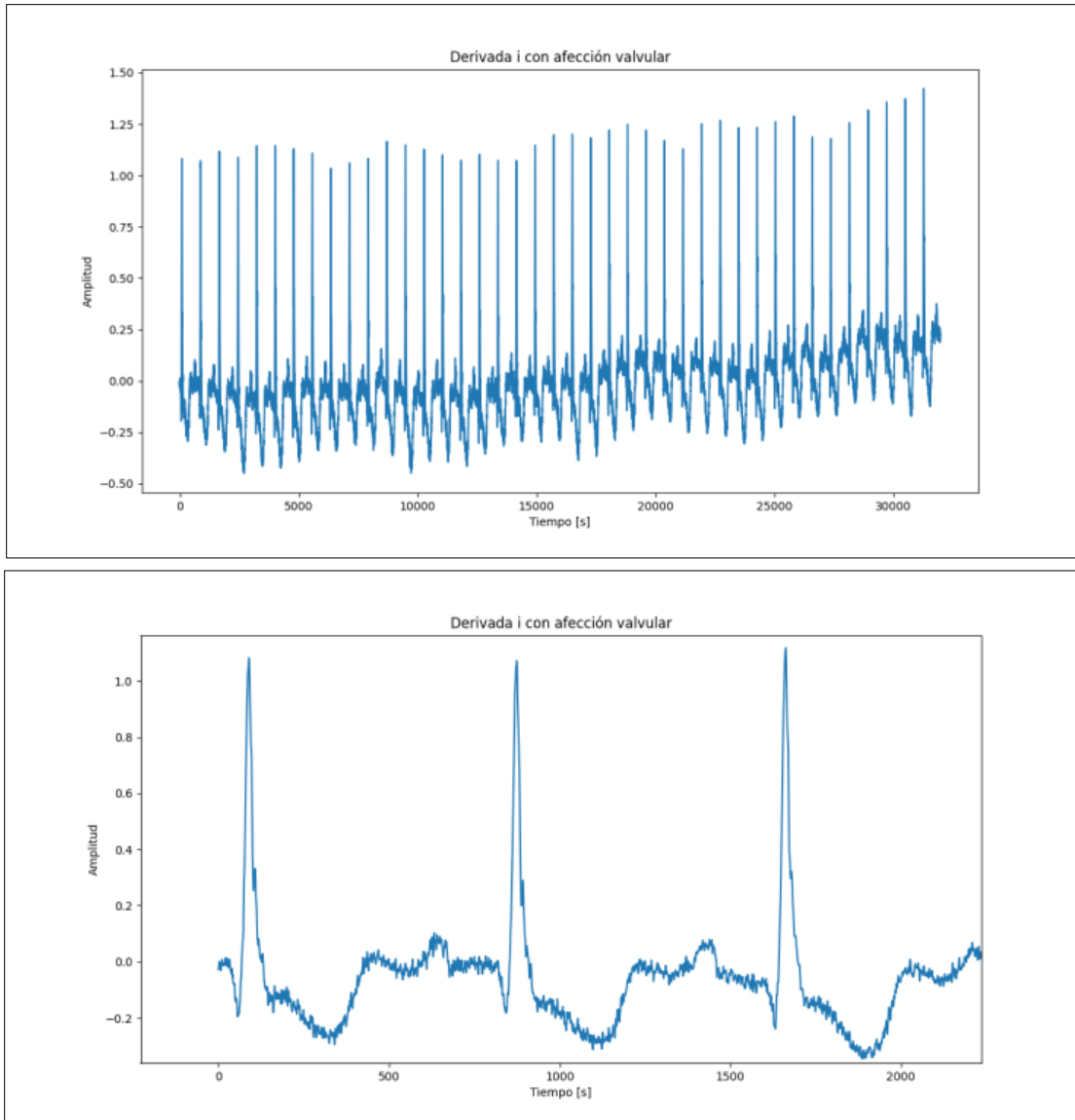


Figura 26 Derivación i señal con afección valvular

3.3. PREPROCESAMIENTO.

La etapa de preprocesamiento tiene como finalidad el acondicionamiento de las señales que se usarán para las siguientes etapas. Por tanto, para ello se debe realizar un ajuste



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



de amplitud (estandarización de la señal) con respecto a la información que se otorga en el archivo .info, centrar la señal en cero, ajustar la visualización de las señales.

3.3.1. Ajuste de amplitud (estandarización de la señal).

La estandarización de la señal es el proceso de llevar la señal al estado original en la que ha sido grabada, por tanto, para lograr esto es necesaria la información que trae consigo el archivo .info donde ofrece la frecuencia, ganancia, base y tiempo de grabación de la señal, con lo anterior se procede a ejecutar la diferencia entre la variable *val* con la base, y se divide sobre la ganancia de la misma, la figura 27 muestra parte del código para realizar dicho paso.

```
19 #Base:  
20 Base = 0  
21 #Frecuencia de muestreo:  
22 Fs = 1000  
23 #Ganancia:  
24 Ganancia = 2000  
25 #Ajustando amplitud (Estandarizando la señal):  
26 Senal_Final = (Senales["val"] - Base)/Ganancia
```

Figura 27 Estandarización de la señal

3.3.2. Centrando señal en cero.

La gran mayoría de señales que se han descargado desde la base de datos no están centradas en cero, por ello, es necesario aplicar este paso para obtener un mejor procesamiento de las mismas, la implementación de este paso se realiza a cada una de las derivadas que trae consigo la señal mediante un ciclo for, restando a sus mismos valores la media de cada señal y dividiendo entre la desviación estándar de las mismas datos, dicho paso se muestra en la figura 28. Es necesario dar a conocer que para la correcta visualización de los datos se debe crear un vector tiempo que contenga la misma longitud de la señal, generando así el tiempo total de grabación, la cual es de aproximadamente 31 segundos para dicha señal y se puede ver con más detalle en la figura 30.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



```
30 #Centrar señal en cero:
31 Senal_Cero = []
32 for Posicion, Valor in enumerate(Senal_Final):
33     Senal_Cero.append((Valor - np.mean(Valor))/np.std(Valor))
34 Senal_Cero = np.array(Senal_Cero)
35
36 #Creando Vector Tiempo
37 Ts = 1/Fs
38 N = Senal_Cero.shape[1]
39 Vector_Tiempo = [i*(1/Fs) for i in range(N)]
40 Vector_Tiempo = (np.array(Vector_Tiempo))
41
42 plt.plot(Vector_Tiempo, Senal_Cero[0])
43 plt.title("Derivada i - Ajuste de amplitud - Posicionamiento en cero")
44 plt.xlabel("Tiempo [s]")
45 plt.ylabel("Amplitud")
46 plt.xlim([0, Vector_Tiempo[-1]])
47 plt.show()
```

Figura 28 Centrando señales en cero

3.3.3. Agrupando como datagrama.

El proceso de agrupación tipo datagrama permite ordenar grandes cantidades de datos de una mejor forma tanto para su respectiva visualización como para el tratamiento de los mismos al aplicar las demás técnicas, la figura 29 refleja dicho agrupamiento. Esto se logra por medio del módulo *Pandas* que trae la función *DataFrame*, su estructura es de tipo diccionario. Cabe resaltar que a las señales se les realiza un recorte hasta los 15420 datos, debido a que la mayoría de las demás señales descargadas no tienen la misma cantidad longitudinal de datos, por tanto, se aplica esta simetría para mayor facilidad de trabajo, como consecuencia del recorte de datos de la señal, esta no tendrá el mismo tiempo de grabación, por tanto, su tiempo cambiará debido al proceso anterior.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

```

39 #Convirtiendo a DataFrame las derivadas desglosadas:
40 Senales_Derivadas = pd.DataFrame(
41     {
42         'Senal_i' : Senal_Cero[0, :15420],
43         'Senal_ii' : Senal_Cero[1, :15420],
44         'Senal_iii' : Senal_Cero[2, :15420],
45         'Senal_avr' : Senal_Cero[3, :15420],
46         'Senal_avl' : Senal_Cero[4, :15420],
47         'Senal_avf' : Senal_Cero[5, :15420],
48         'Senal_v1' : Senal_Cero[6, :15420],
49         'Senal_v2' : Senal_Cero[7, :15420],
50         'Senal_v3' : Senal_Cero[8, :15420],
51         'Senal_v4' : Senal_Cero[9, :15420],
52         'Senal_v5' : Senal_Cero[10, :15420],
53         'Senal_v6' : Senal_Cero[11, :15420],
54         'Senal_vx' : Senal_Cero[12, :15420],
55         'Senal_vy' : Senal_Cero[13, :15420],
56         'Senal_vz' : Senal_Cero[14, :15420],
57     }
58 )

```

Figura 29 Agrupamiento mediante DataFrame

Una vez que se finalizan los pasos anteriores, se procede a visualizar la misma señal que se ha mostrado en la figura 26 con los respectivos resultados, esto se aprecia en la figura 30.

NOTA: En este proceso no se ha aplicado el recorte a la señal, por tanto, se mantiene su tiempo de grabación con la que se ha descargado a través de la base de datos.

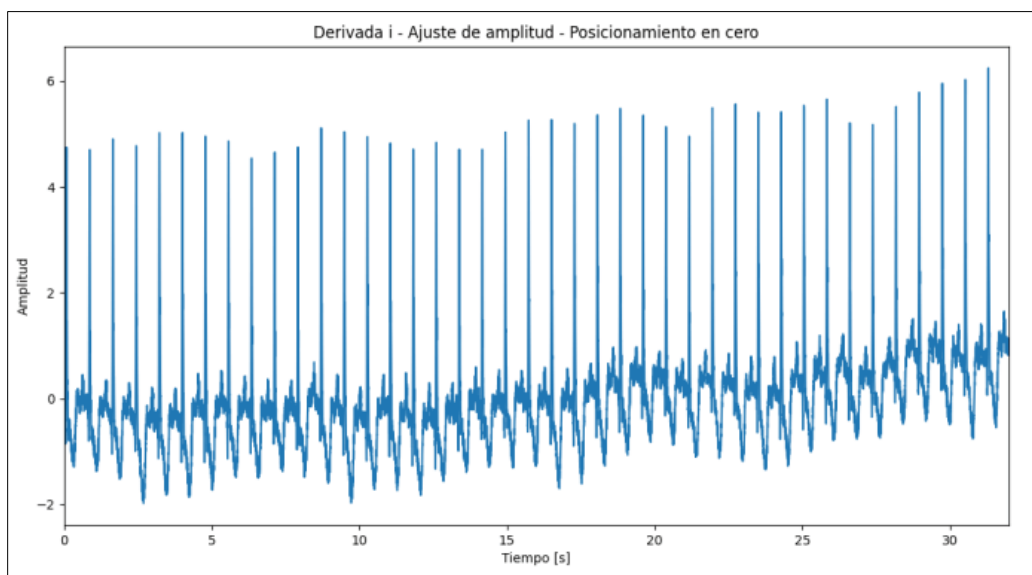


Figura 30 Señal con ajuste de amplitud y posicionamiento en cero



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



3.4. PROCESAMIENTO.

La etapa de procesamiento consiste en la aplicación de las técnicas para el tratamiento y procesamiento de señales junto con la extracción de características para el reconocimiento y clasificación de las mismas y así otorgar un diagnóstico correcto. En esta etapa se han aplicado diversas técnicas de procesamiento, las cuales se dividen en dos líneas de procesamiento, y estas a su vez se dividen en tres sublíneas respectivamente.

3.4.1. Primera línea de procesamiento.

Inicialmente se aplica la Transformada Discreta de Fourier mediante el algoritmo de Transformada Rápida de Fourier que proporciona Python por medio del módulo de *Pandas* con el llamado de la función *np.fft.fft*. (*np* es el nombre alias convencional que se proporciona al módulo para su llamado). La figura 31 muestra el resultado de aplicar FFT a la señal ECG.

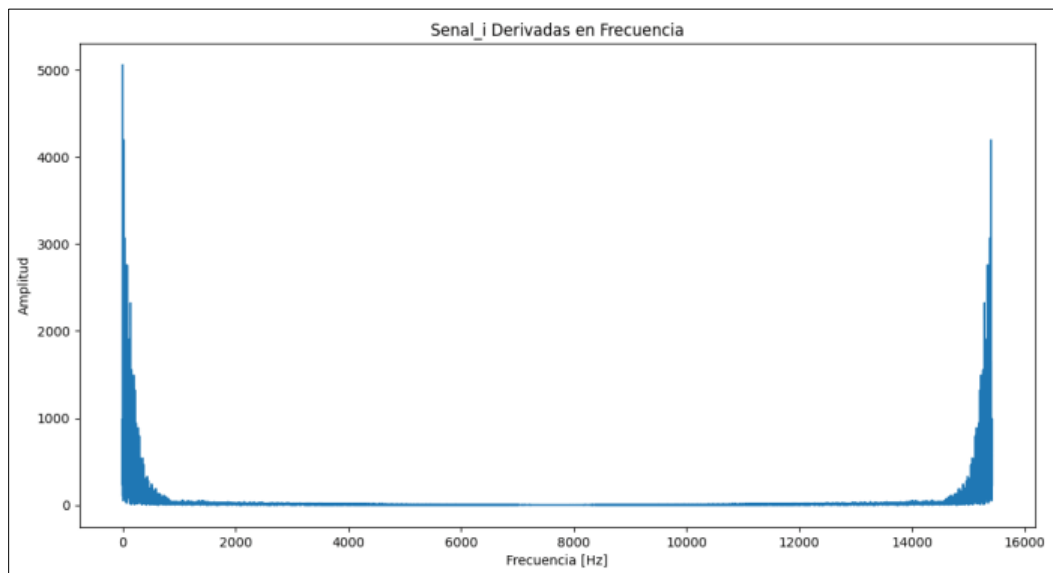


Figura 31 Aplicación de FFT a señal ECG



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

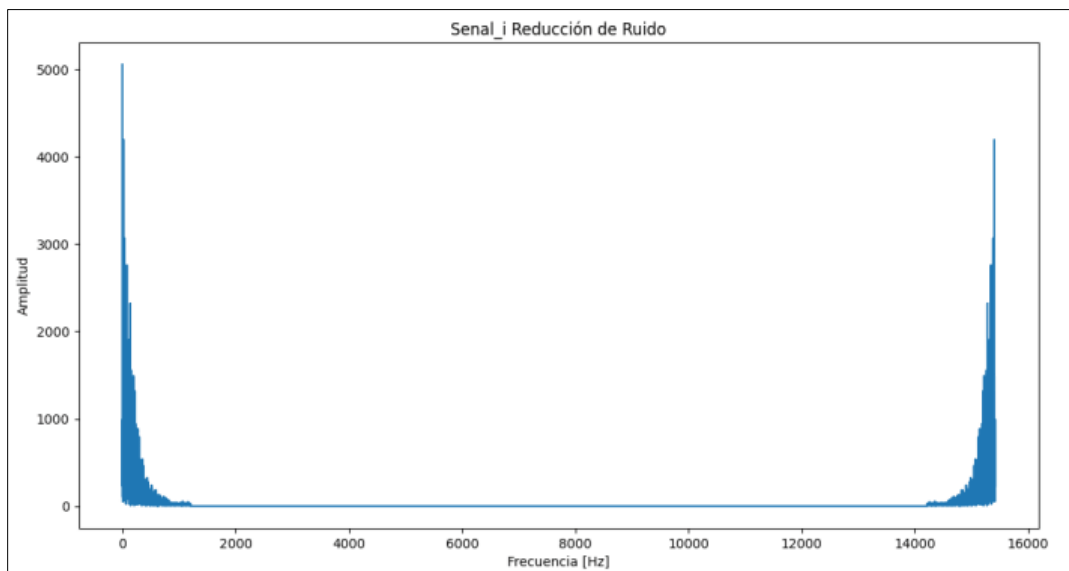
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Un detalle importante de resaltar es que las señales que proporciona *Physiobank* no da a conocer la frecuencia de corte de las señales, por tanto, para la reducción de ruido sin tener presente la frecuencia de corte se realiza un filtro poco convencional, en la figura 32 se muestra la eliminación o filtrado de ruido desde un punto de muestra inicial hasta un punto de muestra final. Todas las señales se reducen a una longitud de 15420 antes pasar por la etapa de procesamiento, por tanto, el filtrado se realiza desde los 1200 puntos hasta la longitud final de la señal restando los 1200 puntos como tal, es decir, $\text{fft_senal}[1200: (\text{Senales_Derivadas.shape}[0] - 1200)] = 0$.



La Figura 32 Eliminación de ruido

Aplicado lo que es FFT se procede a aplicar la transformada inversa (IFFT), esto es posible mediante el llamado de la función *ifft* que proporciona el módulo *pandas* de la forma *np.fft.ifft*, por tanto, una vez filtrado el ruido de la señal a partir de esta se podrá aplicar las siguientes técnicas para su respectiva extracción de los patrones característicos. Debido a que se ha realizado el recorte de longitud de la señal, su tiempo de grabación se reduce. La figura 33 muestra el resultado de llevar nuevamente al dominio temporal la señal con su respectiva eliminación de ruido, la figura inferior es un acercamiento a los tres primeros picos de la señal.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

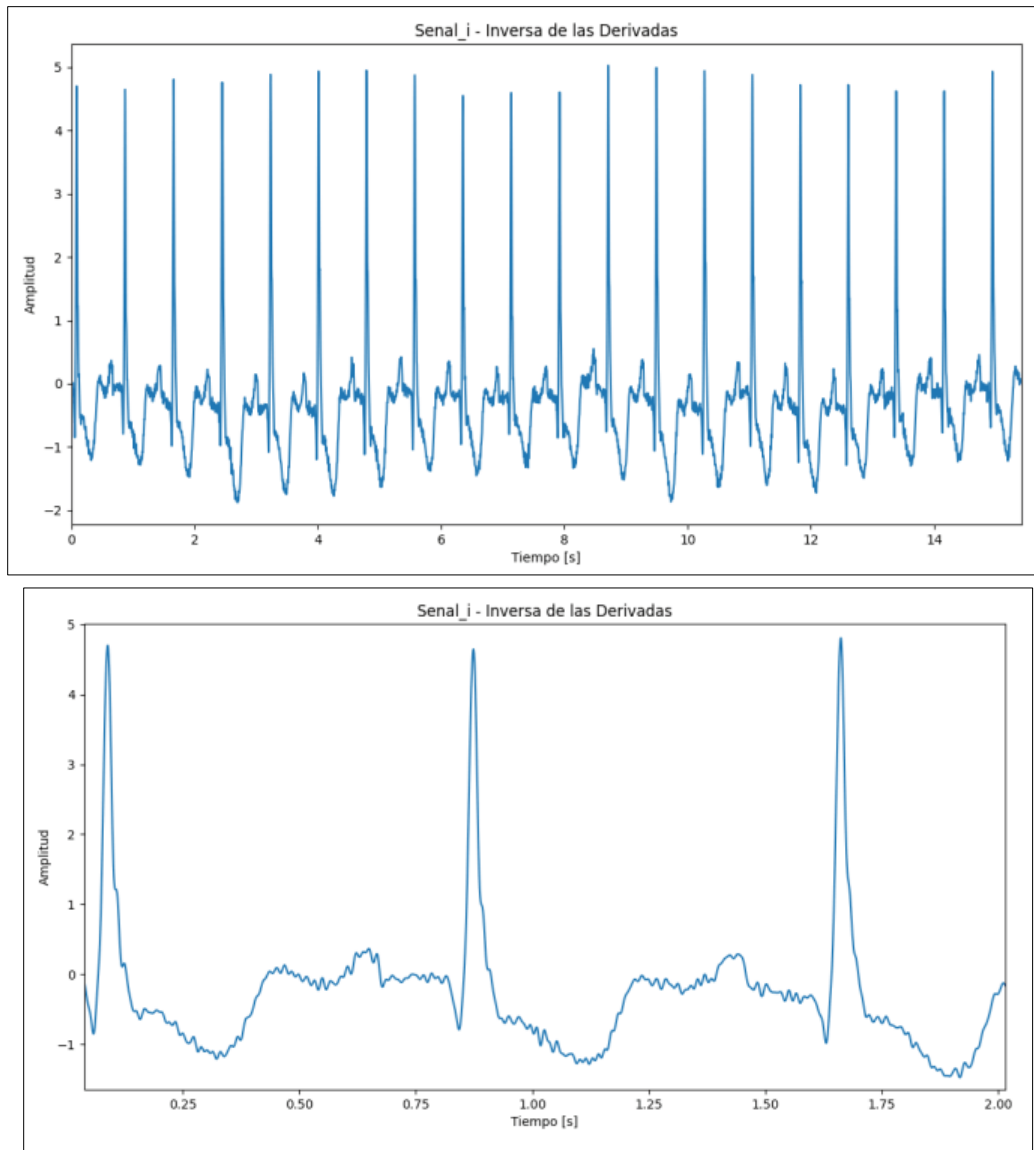


Figura 33 Cambio de dominio de la señal ECG

En el paso siguiente se aplica entropía en base a la ecuación (10), la cual consiste en elevar al cuadrado cada uno de los datos de la señal y estos se multiplican por el logaritmo en base dos del mismo dato elevado al cuadrado mediante sumatorias con ventanas de 10, 20 y 30 puntos; dichos puntos seleccionados permiten escoger un



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



margen de prueba y extraer las características de cada señal, para la elección de la ventana correcta a dicho procesamiento se da a partir de los resultados y la información que suministren.

Cada línea de procesamiento se compone de sublíneas referente a las ventanas que se emplean.

$$\text{Entropía} = \sum_{k=1}^N x(n)_k^2 * \log_2(x(n)_k^2) \quad (10)$$

Donde:

k es la cantidad de puntos de la ventana a recorrer hasta la longitud total del dato.

$x(n)$ son los puntos del vector señal elevados al cuadrado cada uno.

La figura 34 muestra el resultado de aplicar entropía con ventanas de 10 puntos para este caso.

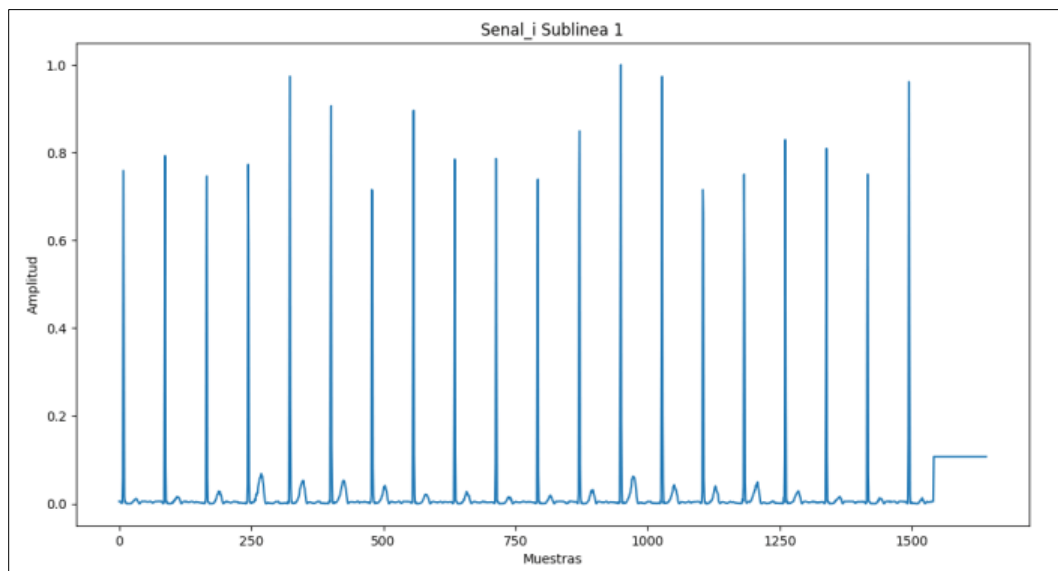


Figura 34 Entropía de la señal ECG con ventana de 10 puntos



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



La sublínea dos de la primera línea de procesamiento corresponde a la aplicación de entropía con ventanas de 20 puntos la cual se aprecia en la figura 35. Se obtiene como resultado una mayor compresión de las muestras por el hecho de sumar más datos, con respecto a la figura 34 se obtuvieron 1641 muestras con ventana de 10 puntos y para la ventana de 20 puntos se obtiene un total de 870 muestras.

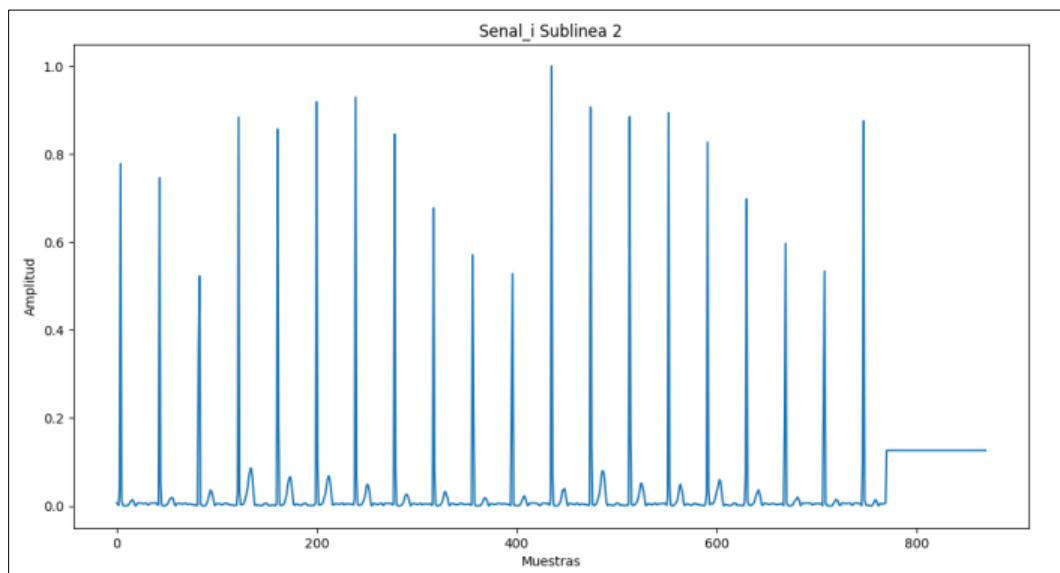


Figura 35 Entropía de la señal ECG con ventana de 20 puntos

Del mismo modo se aplica una tercera sublínea con ventana de 30 puntos, donde se genera un total de 613 muestras, el resultado anterior se detalla en la figura 36, la aplicación de realizar diversas sublíneas en la primera línea de procesamiento es con el hecho de elegir la mejor técnica empleada para una correcta clasificación en su correspondiente etapa.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

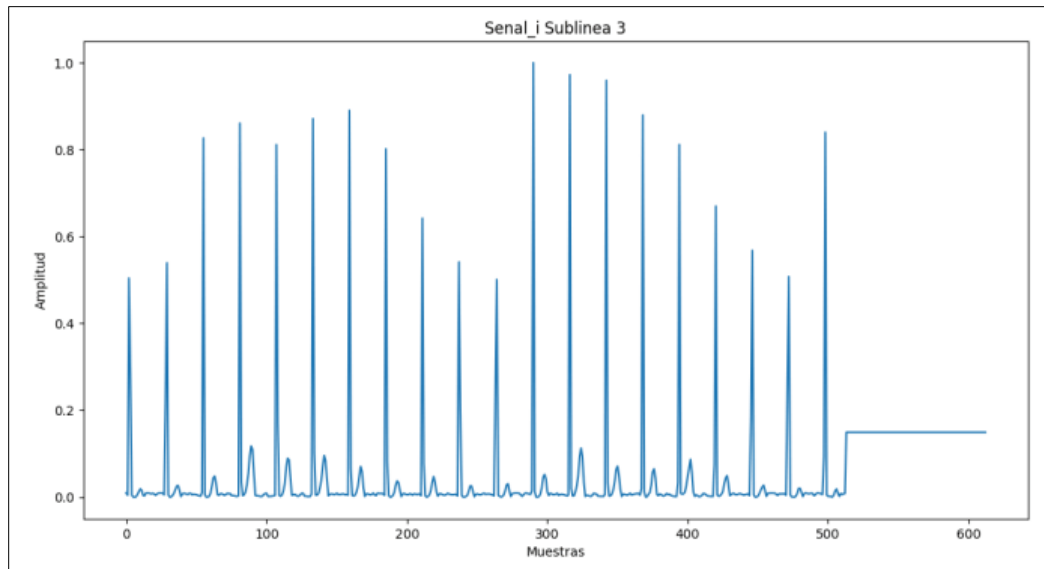


Figura 36 Entropía de la señal ECG con ventana de 30 puntos

Cabe resaltar que cada una de las sublíneas anteriores dadas a conocer se han normalizado una vez aplicada la técnica de procesamiento en base a la ecuación (7). Una vez finalizada la primera línea de procesamiento se procede a realizar la segunda línea de procesamiento en la cual se implementa la Transformada Discreta Wavelet, energía y normalización.

3.4.2. Segunda línea de procesamiento.

Para esta etapa se aplica la Transformada Discreta Wavelet (DWT), para ello se recurre al módulo *PyWavelets* de Python llamando a la función *dwt*, por tanto, por convención se escribe como *pywt.dwt*.

Uno de los mayores problemas al momento de trabajar señales bioeléctricas con Wavelet es la elección de la familia wavelet, por la razón en que las wavelets no están asociadas a las señales a analizar sino más bien a las limitantes matemáticas al momento de calcularse. Generalmente según investigaciones, la selección de la familia wavelet se halla en base a principios de observación sobre los resultados convenientes; por tanto, se elige la familia *Daubechies 4* ya que es la más semejante en la forma de



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



onda a las señales ECG y por el hecho de que se han implementado en diversos estudios y ha generado resultados óptimos.

Para la selección adecuada del nivel de descomposición a implementar, se aplican 2 niveles, por la razón de que al aplicar un nivel más alto que el seleccionado la señal pierde su forma y como consecuencia al cambiar la forma de la señal se está eliminando información importante donde podría estar la afección cardiovascular, por esa razón se toma el segundo nivel de descomposición para los coeficientes de aproximación ya que allí es donde se encuentra toda la información pertinente a tratar y a su vez porque genera una adecuada reducción de ruido presente en la onda sin alterar su forma y característica. En la figura 37 se observa la aplicación de wavelet con *db4*. La imagen inferior es un acercamiento a los tres primeros picos de la señal ECG.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

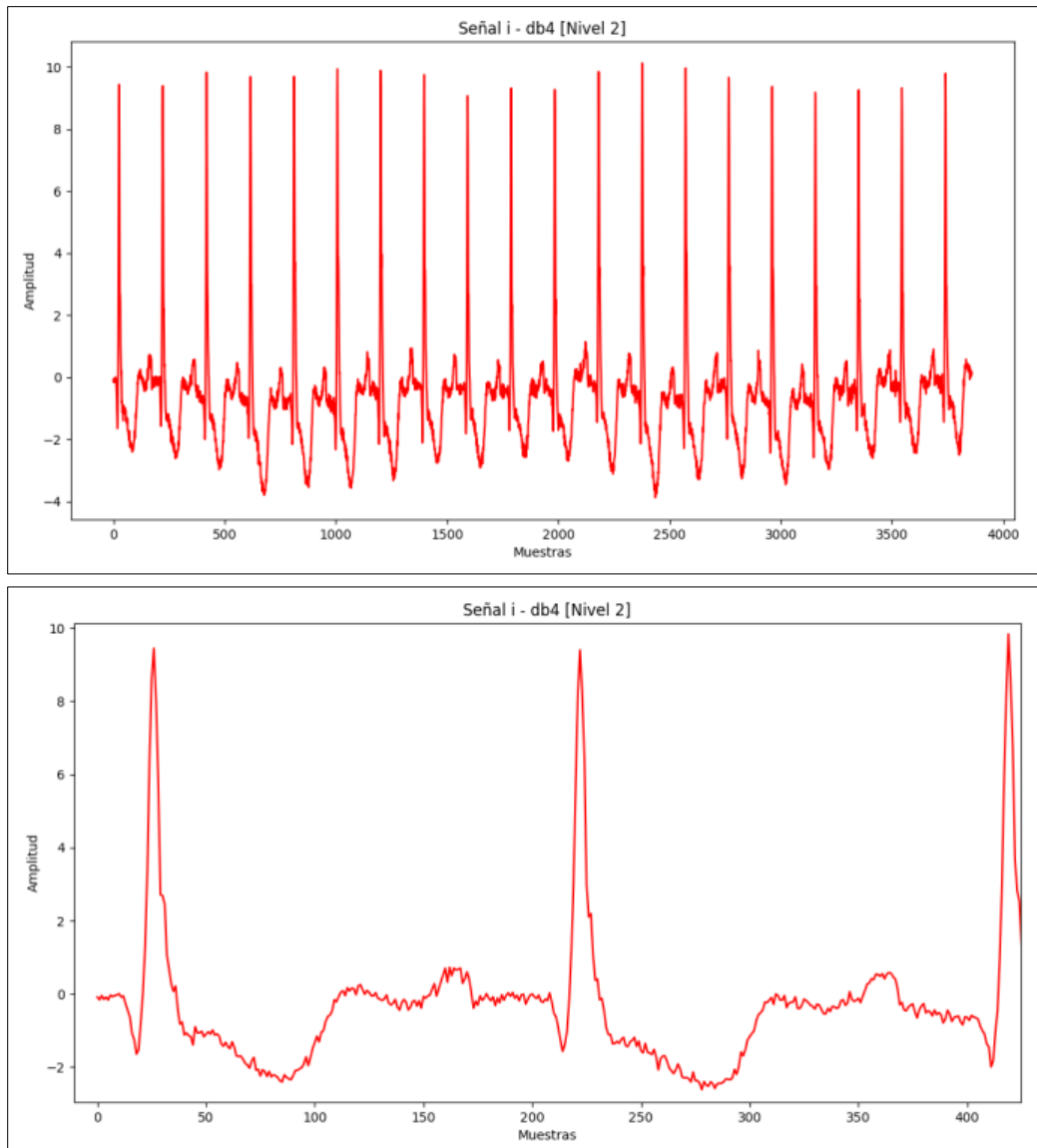


Figura 37 Wavelet con db4 a la señal ECG

El paso a seguir es aplicar energía a la señal ECG mediante la implementación de ventanas, también con los mismos valores de puntos que se ha aplicado en la primera línea de procesamiento. Para realizar esto, nos basamos en la ecuación (6).



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



En primer lugar, la ventana que se aplica será de 10 puntos, por tanto, los datos del vector señal se elevan al cuadrado y se suman mediante una sumatoria para estos primeros 10 puntos y luego el resultado generado se irá concatenando en otro vector mediante la función *append()* que trae integrada Python, una vez que el valor de ventana ha llegado al valor 10, esta incrementa su límite otros 10 puntos para así proceder a realizar el mismo paso con otros 10 puntos más hasta recorrer la señal por completo, el resultado de aplicar energía genera un total de 485 puntos, pues debido a que se ha aplicado Wavelet, esta reduce su valor a la mitad dependiendo de la cantidad de filtros que se aplique, la figura 34 muestra el proceso de aplicar dicha técnica.

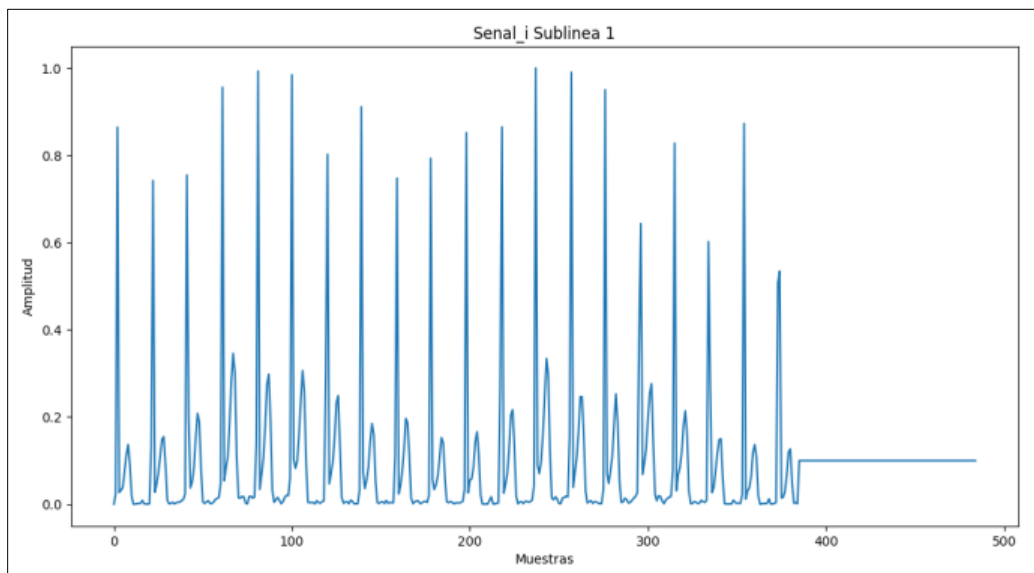


Figura 38 Energía de la señal ECG con ventana de 10 puntos

Del mismo modo la figura 39 muestra la señal ECG con energía por medio de ventana de 20 puntos.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

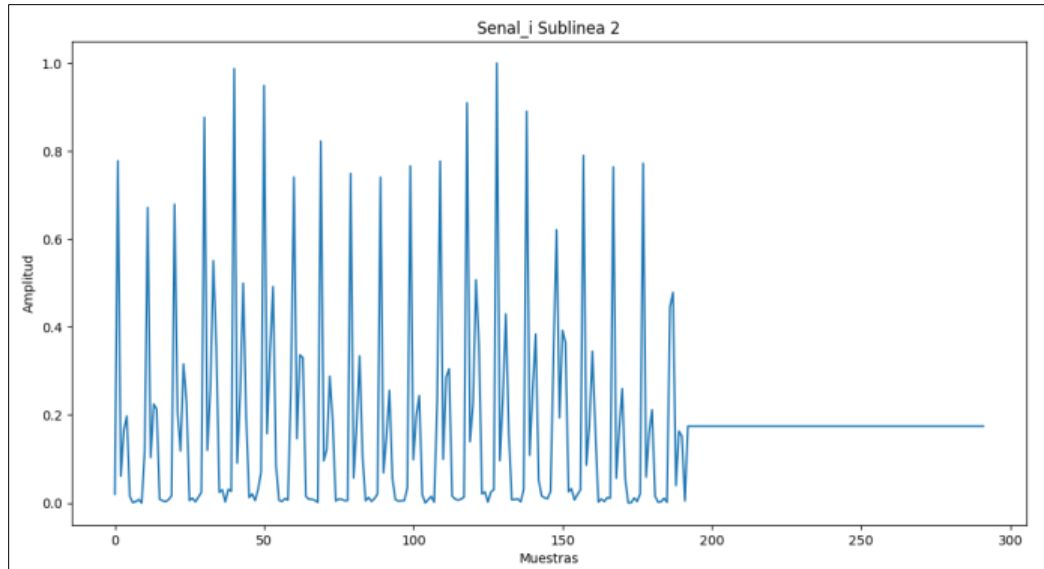


Figura 39 Energía de la señal ECG con ventana de 20 puntos

Para finalizar la segunda línea de procesamiento, en la figura 40 se aprecia la implementación del último proceso de aplicación de ventana con 30 puntos, cabe añadir que para la segunda línea de procesamiento después de haber calculado la energía en cada sublínea se ha normalizado la señal para suavizar la métrica.

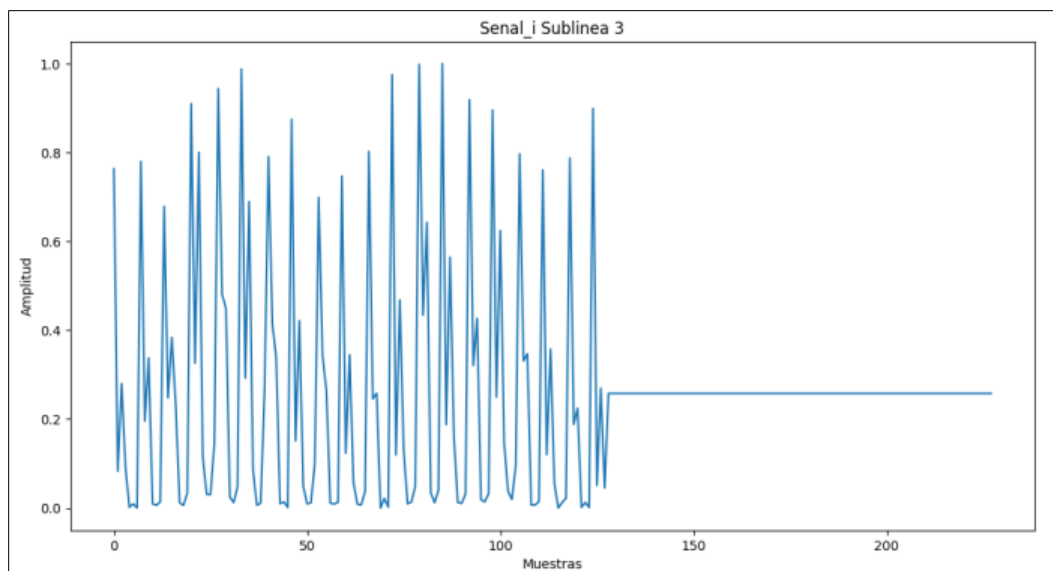


Figura 40 Energía de la señal ECG con ventana de 30 puntos

3.5. CLASIFICACIÓN APLICANDO MACHINE LEARNING.

Luego de obtener los resultados de la primera y segunda línea de procesamiento junto con las sublíneas en relación a las ventanas, cada uno de estos datos se forman en datagrama, es decir, se ingresa una señal, esta a su vez se debe realizar un ajuste de nivel (amplitud), posicionar en cero, no hay que olvidar que todas las señales que se han bajado desde la base de datos *Physiobank* trae consigo las 15 derivaciones del corazón, por tanto, al aplicar el preprocesamiento se transforma o agrupa como datagrama, teniendo como cabecera el nombre de cada una de las señales por columnas, cada columna contiene el total de cada derivada, lo anterior es un datagrama en Python, ahora, una vez que se tiene el datagrama como resultado del preprocesamiento, este datagrama (contiene las 15 señales) pasa por la etapa de procesamiento donde se aplican las técnicas de tratamiento y procesamiento de señales con las cuales se extraen los patrones o características que se requieren para la clasificación. En total se tienen 5 pacientes por cada patología y también se tienen 5 pacientes sanos, con base a lo anterior y a la aplicación de las etapas de preprocesamiento y procesamiento se construye una matriz de la siguiente forma:



- La matriz tendrá todos los pacientes de cada patología e inclusive los pacientes sanos, a su vez cada paciente (señal ECG) con tiene 15 derivaciones, así que el tamaño de la matriz será de $5_{\text{pacientes}} * 5_{\text{(Patologías + Sanos)}} * 15_{\text{Derivaciones}} = 375$ filas, en otras palabras, la matriz tendrá 375 filas contando todas las derivadas de cada patología y sanos por cada paciente con 15420 columnas en total.
- Se realiza la eliminación de la primera columna la cual es el conteo de los datos en total, lo cual no es necesario incluirlo mediante la función `iloc[]` en Python, pues cada una de las líneas de procesamiento se exportaran en formato CSV, y luego serán cargadas para la construcción de la matriz.
- Se aplica la transpuesta para reordenar los datos de la matriz debido a que están como datagrama y así poder realizar una clasificación de forma adecuada.
- Se exporta la matriz en formato `.txt`.

Los pasos anteriores se realizan de igual manera para la agrupación de señales que presenten patologías y para aquellas que son de pacientes sanos y así tener por separado la matriz de cada conjunto para la creación del modelo con respecto a las técnicas de clasificación a implementar.

La matriz de entrenamiento y matriz de clasificación o prueba se generan mediante la ley de Pareto, donde, el 80% de la matriz descrita anteriormente es para entrenamiento y generación del modelo y el 20% de los datos sobrantes de la matriz principal se usan para la clasificación del modelo y saber que tan preciso es. Cada una de las técnicas implementadas en este proyecto se basa en la ley de Pareto 80/20, para este caso la clasificación de VSM se realiza mediante la implementación de dos kernel, el primero para Kernel Lineal y el segundo para Kernel RBF.

3.5.1. Support Vector Machine (VSM).

VSM es una técnica muy compleja y útil para la clasificación, por tanto, nos permite clasificar los datos mediante la etiquetación de los datos ya que su gran ventaja permite etiquetar una o varias clases, pues esta técnica de Machine Learning es de aprendizaje supervisado, así que se deben etiquetar los datos a clasificar y se realizar de la siguiente forma a través de la tabla 3.





ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Tabla 3 Etiquetado de datos

Dato	Etiqueta
Afección valvular	1
Bloqueo de rama	2
Disritmia cardíaca	3
Hipertrofia	4
Paciente sano	5

En base a los datos de entrenamiento (datos históricos) y los datos etiquetados se entrena la máquina de soporte vectorial para construir un modelo donde nos prediga la clase de cualquier nueva muestra (matriz de clasificación o prueba) que se suministra a dicho modelo. Una VSM es un modelo el cual representa los puntos de muestra en un espacio vectorial de tantas dimensiones como columnas de datos que se tengan en la matriz separando las clases de dos o más categorías diferentes mediante la creación de un hiperplano que el mismo algoritmo establece donde el modelo previamente entrenado entregara como respuesta una clase de etiquetación de correspondiente a las ya etiquetadas.

3.5.2. Naïve Bayes (NB).

En Naïve Bayes el clasificador asume que el efecto característico generado por parte de una característica en particular en una clase es independiente de otras características. Para generar el modelo y realizar el respectivo entrenamiento, dado que NB es de tipo lenguaje supervisado se dispone de un etiquetado previo de una forma distinta a como se plantea en VSM, por tanto, las etiquetas se realizan de forma binaria y se comparan solamente entre dos datos, es decir, cada una de las patologías se pasa a clasificar con señales de pacientes sanos, para entender mejor el etiquetado, se tiene la table así que los datos se etiquetan mediante las 4, 5, 6 y 7.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Tabla 4 Etiqueta de datos entre paciente con afección valvular y paciente sano

Dato	Etiqueta
Afección Valvular	1
Paciente sano	0

A diferencia del etiquetado con VSM, en Bayes es distinto, por tanto, el resultado generado por el modelo NB entregara como resultado un 1 si la persona padece de dicha enfermedad y 0 si la persona no presente ninguna afección cardiovascular, en otras palabras, es un paciente sano.

Tabla 5 Etiqueta de datos entre paciente con bloqueo de rama y paciente sano

Dato	Etiqueta
Bloqueo de rama	1
Paciente sano	0

Del mismo modo aplica para la patología de disritmia y se compara con una de paciente sano, pues el interés de aplicar NB es clasificar si la persona padece solamente esa patología o no. La tabla 6 muestra la descripción anterior.

Tabla 6 Etiqueta de datos entre paciente con disritmia y paciente sano

Dato	Etiqueta
Disritmia	1
Paciente sano	0

De igual manera, aplica para la patología de hipertrofia ventricular con paciente sano. Esto se refleja en la tabla 7.

Tabla 7 Etiqueta de datos entre paciente con hipertrofia y paciente sano

Dato	Etiqueta
Hipertrofia	1
Paciente sano	0



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



3.5.3. Árboles de decisión.

La ventaja de esta técnica en comparación de las demás, es que el ser humano es capaz de ver la interpretación de lo que está sucediendo. Su estructura se define como un diagrama de flujo donde cada nivel o nodo es una pregunta con una respuesta de *si* o *no*. Esta técnica a diferencia de VSM permite etiquetas de una sola clase, por tanto, la etiquetación previa se realiza del mismo modo como se muestra con las tablas 4, 5, 6 y 7. Se selecciona como criterio de división óptima de atributos *entropy* (entropía), la cual se encarga de medir la información que indica el desorden de las características, esta división óptima es elegida por la característica que tenga la menor entropía, obtiene su valor máximo cuando la probabilidad de dos clases es la misma.

3.6. PYTHONANYWHERE (CLOUD).

La solución del quinto objetivo se plantea en este apartado, trabajar y ejecutar código Python desde servidores externos a través de un navegador web se está convirtiendo en una alternativa bastante atractiva. Debido al avance que presenta Python a un ritmo vertiginoso se presentan demasiadas complicaciones al momento de alojar código en la nube y desde allí ejecutar el procesamiento mediante un aplicativo web.

Una solución a esto es *Pythonanywhere*, ya que permite crear y ejecutar códigos de Python y aplicaciones completas en la nube, este Cloud no implica mantenimiento ni configuración del servidor web para iniciar, solo es cuestión de escribir código o subirlo al servidor. Los programas pueden ser escritos en un editor de código basado en la web y con espacio de almacenamiento dedicado a cada plan. [41]

Este servicio se clasifica como modelo SaaS, donde claramente se puede acceder a las aplicaciones desde diversos dispositivos por medio de un navegador web y a su vez es de tipo Nube pública.

La pregunta que puede generar controversia sería ¿Cómo puedo publicar un código de Python en una aplicación Web?, la respuesta a esta pregunta sería mediante la implementación de *Flask*.

Flask es un micro-framework escrito en Python que permite crear de forma sencilla aplicaciones web bajo el patrón MVC. Este patrón es una forma de trabajar el cual



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



permite diferenciar y separar lo que es el modelo de base de datos (datos de la App que se guardan en BD), la vista general del aplicativo web (página HTML) y el controlador donde se gestionarán las peticiones. Al momento de instalar Flask en el dispositivo se tienen herramientas necesarias para crear aplicaciones web funcionales. [42]

Para realizar este proceso se accede a la página principal de *Pythonanywhere*, donde se es necesario registrarse y así poder obtener los recursos gratuitos limitados que ofrece este servicio.

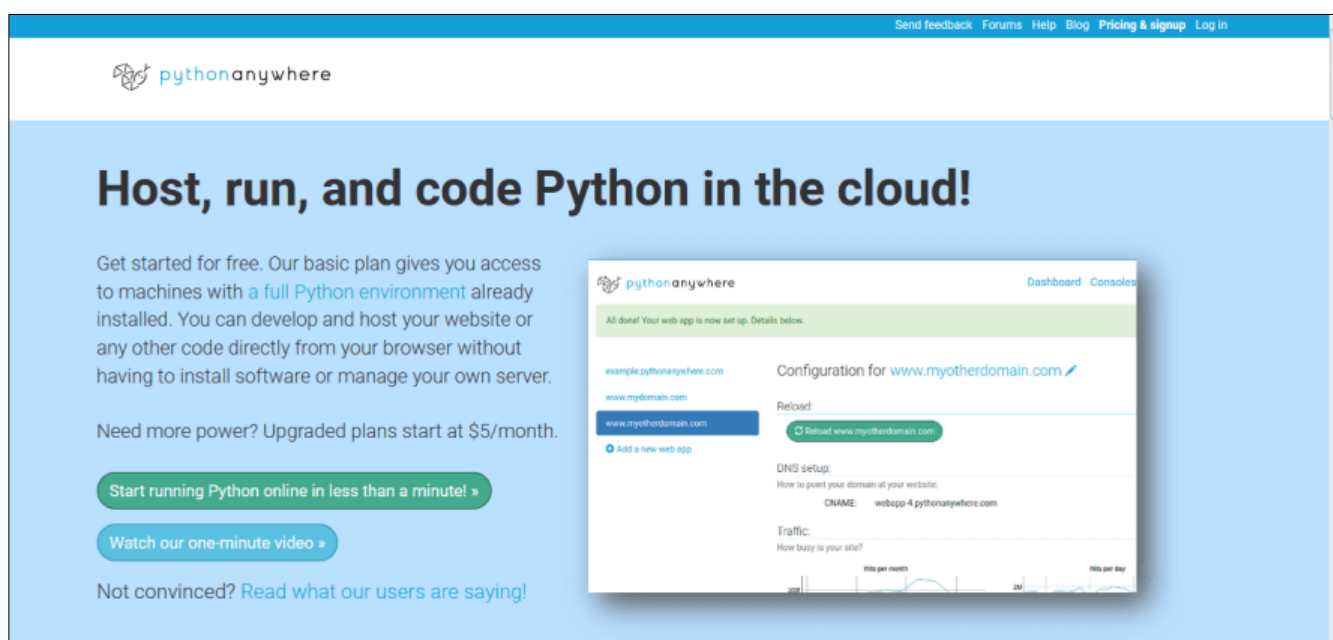


Figura 41 Página principal de Pythonanywhere

Al crear la cuenta gratuita en *Pythonanywhere* esta tiene varias limitaciones, como la capacidad de almacenamiento y los recursos computacionales que nos presta el servicio Cloud, con respecto a esto en *Pythonanywhere* el uso de la CPU se presta por segundos, es decir, se obtiene una cantidad de segundos de CPU incluidos cada día, al finalizar las 24 horas este recurso limitante se reinicia nuevamente. Una vez creada la cuenta gratuita, desde la Dashboard de *Pythonanywhere* se accede a las aplicaciones dando clic en Open Web tab.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750

SC-CER96940



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



The screenshot shows the PythonAnywhere dashboard for user 'Nandoathlete'. It includes navigation links for Dashboard, Consoles, Files, Web, Tasks, and Databases. The dashboard displays system usage (CPU and storage) and provides options to manage recent consoles, files, and notebooks. A red box highlights the 'Open Web tab' button under the 'All Web apps' section.

Figura 42 Creando nueva App web

Cabe aclarar que ya se tiene configurada una app web, por tanto, los pasos siguientes es el proceso de como configurar una nueva App web con Flask, subir código y ejecutar la aplicación web. Damos clic en Add a new web app.

The screenshot shows the 'Web' configuration page for 'Nandoathlete.pythonanywhere.com'. It features a red box around the 'Add a new web app' button, indicating the next step in the process.

Figura 43 Agregando nueva web app

Como ya se tiene agregado un dominio el cual es *Nandoathlete.pythonanywhere.com* no se dispone de poder crear, esa es otra limitante de *Pythonanywhere*, por tanto, se ha pagado solo 5.25 dólares para obtener el beneficio de más segundos en la Pc y con un almacenamiento de 2GB en vez de 1GB lo cual es lo que ofrece la cuenta gratuita. Por tanto, se dispone de figura con otra cuenta. Una vez que se agrega la nueva app web, se debe elegir el framework de la app web, el cual da a elegir entre Django, Flask, Botton, web2py, para este caso se selecciona Flask, luego de forma predeterminada se



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



estable la ruta de la aplicación, después de esto el entorno está configurado, la figura 44 muestra la app web creada.

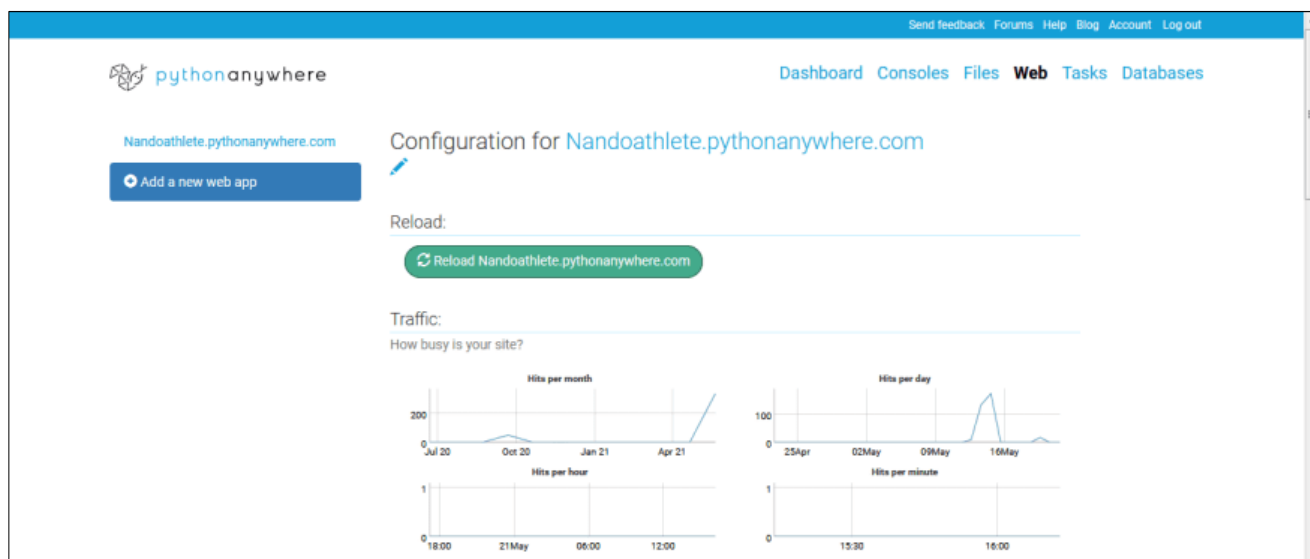


Figura 44 App web flask creada

Para terminar la configuración del servicio Cloud, desde la consola de comandos que ofrece *Pythonanywhere* se crea un entorno virtual y se instalan los requerimientos necesarios, como la versión de Python con la que se ha trabajado todo el proyecto y los módulos que se requiere para que se ejecutan las etapas de procesamiento, carga de los datos y demás. Otra limitante de cargar archivos a este entorno es que solo permite cargar archivos con 100MB, por tanto, se recurre a usar Git Hub para cargar el código y poder clonarlo desde la consola que ofrece el servicio cloud, una vez que se bajado el archivo, ya se tiene prácticamente todo el aplicativo web listo para ser usado, la figura 45 muestra el resultado de la ejecución del aplicativo web en la red, por tanto para acceder desde cualquier lugar del mundo basta con escribir el nombre de dominio en un navegador web.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 45 Aplicativo web



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



CAPÍTULO IV: RESULTADOS

En el presente capítulo se expondrán todos los resultados y evidencias del trabajo de grado donde se explica con detalle cada uno de los pasos realizados para la solución de cada objetivo planteado para su respectivo desarrollo.

4.1. BASE DE DATOS PHYSIOBANK.

Las señales ECG se adquieren por medio de una base datos en línea la cual es de acceso gratuito para los investigadores y demás personas. Las descargas de estas señales se realizan a través de las bases de datos que contiene el aplicativo web. Cada base de datos dentro de Physiobank consta de un conjunto de grabaciones identificados por un nombre de registro. En la mayoría de los casos un registro que se ha cargado consta de tres archivos, la mayoría de estos registros incluyen un archivo binario *.dat* (señal) que con tiene las muestras digitalizadas de una o más señales, un archivo *.head* (encabezado) que describe los nombres, URL del archivo, formato de almacenamiento, número, tipo de señal, entre otros, algunos de estos registros incluyen un archivo *.atr* el cual denota las anotaciones donde se describen las características de las señales. La visualización de los registros se realiza desde la sección *PhysioBank Databases* y la descarga de los mismos se realiza mediante la sección *PhysioBank ATM*.

4.2. TÉCNICAS DE TRATAMIENTO Y PROCESAMIENTO DE SEÑALES.

Para el desarrollo de esta etapa, se clasifica en dos líneas de procesamiento, para la primera línea de procesamiento se aplica la técnica de Transformada Discreta de Fourier (DFT) mediante el algoritmo de desarrollo empleado en Python denominado Transformada Rápida de Fourier (FFT), una vez que se aplica a cada una de las derivaciones del corazón (12 derivadas) se procede a realizar el filtro de forma redundante por la razón de que la BD no proporciona la frecuencia de corte de la señal



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



para realizar un mejor filtro con respecto al aplicado, por tanto, se realizan diversos filtros hasta encontrar el adecuado en donde no degrade la señal (pérdida de forma), obteniendo como resultado una señal más limpia y sin pérdida de las características, al finalizar este paso se procede a transformar dicha señal(es) al dominio del tiempo, empleando la Inversa de la Transformada Rápida de Fourier mediante el módulo *numpy* llamando a su clase *np.fft.ifft*, para lo cual se debe recuperar la parte real de la señal con la función *real* que trae implementada Python por sí mismo, una vez aplicada la *ifft* se procede a concatenar cada una de las derivadas del corazón para después proceder a convertirlas en un array, después de ello se procede a aplicar *entropía* por medio de un ciclo *for*, se recorre cada fila (derivada del corazón) para calcular el valor elevado al cuadrado de cada uno de los puntos de la señal, estos resultados se vuelven a concatenar en la misma variable y luego se procede a calcular el *logaritmo en base dos* de cada uno de los datos elevados al cuadrado que se ha realizado anteriormente para generar las sublíneas, las cuales son el cálculo de *entropía* por medio de ventanas de 10 puntos, esta corresponde a la primera sublínea de la línea 1 de procesamiento, por tanto, las demás sublíneas (en total son 3 para la primera línea de procesamiento) constan de ventanas de 20 y 30 puntos, cabe resaltar que al calcular cada una de las sublíneas estas se normalizan mediante la implementación de la ecuación (7), con lo anterior se concluye la primera línea de procesamiento.

De otro modo, la segunda línea de procesamiento consta de la aplicación de la Transformada Discreta de Wavelet y energía. El primer paso consiste en aplicar DWT mediante el módulo *PyWavelets* de Python llamando a su clase *dwt*, para esta técnica de procesamiento se ha aplicado dos niveles de descomposición aplicando la familia *Daubechies 4* seleccionando el coeficiente de aproximación, por la razón de que ha generado una mejor eliminación de ruido causado por otras frecuencias y la no degradación de la misma, pues al implementar otro nivel más de descomposición la señal pierde su forma y no es óptimo para el proceso.

Una vez que se ha aplicado la DWT se procede con la primera sublínea de la segunda línea de procesamiento, la cual consiste en la aplicación de *energía*; esto se realiza mediante un ciclo *for* donde se va acumulando cada uno de los puntos de la señal elevado al cuadrado mediante un contador el cual va indicando la cantidad de puntos de la señal, dicha acumulación se concatena en otro vector y se incrementa 10 puntos más al contador una vez que éste llega a su límite de conteo mediante un condicional para seguir nuevamente con el cálculo de energía, lo anterior aplica para las demás sublíneas las cuales conllevan al cálculo de la energía con ventanas de 20 y 30 puntos,



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



del mismo modo, una vez que se ha realizado el proceso de cálculo de energía a cada derivada, dicha señal se normaliza.

Las ventanas se usan frecuentemente en el análisis y procesamiento de señales para evitar discontinuidades al inicio y final de los bloques que se analizarán, en otras palabras, permiten escoger un margen de prueba para extraer las características de forma fraccionada y poder observar el comportamiento de la señal. Una pregunta que podría surgir es ¿Por qué se toman ventanas de 10, 20 y 30 puntos?, la respuesta a esta cuestión es que los puntos que se han elegido permiten visualizar a la señal de 3 distintas formas y así optar por la de un mejor resultado, no necesariamente deben ser esa cantidad de puntos, pero para este caso por tener una señal de poca longitud de puntos se seleccionan estos valores para que al momento de calcular las técnicas de procesamiento implementando ventaneo la cantidad de puntos sea adecuada para tener una señal conforme y se puedan visualizar los patrones característicos. Al momento de aplicar ventanas cambia el espectro de la señal.

Continuando con esta etapa, otra posible pregunta sería ¿Por qué seleccionó estas técnicas de procesamiento?, según varios artículos de investigación, tesis, publicaciones en revistas y libros, estas han generado diversos resultados óptimos en el área de procesamiento de señales y son las más usadas, cabe resaltar que no solamente existen estas técnicas, también hay otras. Una de las ventajas al momento de aplicarlas es que ya se tiene el conocimiento y la experiencia en el momento de haber trabajado con estas técnicas, a su vez debido a las limitaciones de tiempo y diversas dificultades lamentablemente no se pudo estudiar otras técnicas que no se hayan trabajado dentro del programa, pero si se llegaron a mencionar.

4.3. CLASIFICACIÓN.

Para la etapa de clasificación, una vez finalizada las respectivas líneas de procesamiento descritas anteriormente, se procede a obtener la matriz completa, pero antes de continuar es necesario dar a conocer las características de las señales en cuanto a longitud, cantidad de información, variables de entrenamiento y clasificación, entre otros.

Las señales que se han descargado desde la base de datos Physiobank tienen aproximadamente una longitud de 60000 puntos, es necesario dar a conocer que



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



algunas de las señales que pertenecen a las cardiopatías a trabajar no tienen esta misma longitud, por tanto, por simetría se realiza un recorte de hasta 15420 puntos para todas las señales a procesar.

Realizado lo anterior, se pasan las señales por cada una de las líneas de procesamiento, como resultado en cuanto a longitud de datos se aprecia en la tabla 8.

Una vez que las señales se han pasado por la etapa de preprocesamiento, estas son agrupadas en un DataFrame y allí se recortan las señales hasta el valor dado anteriormente, por tanto, dicho DataFrame hasta el momento tendrá una sola señal con 15420, todo este proceso se aplica para todas las señales correspondientes a cada patología,

Para una mejor explicación a lo escrito anteriormente, se tienen 5 patologías a tratar, las cuales cada una de estas patologías tiene 5 pacientes (señales) con los cuales se van a trabajar y estas señales/pacientes (tipo array) a su vez traen consigo 15 señales demás (derivadas), en la tabla 9 se evidencia la cantidad de filas de cada patología.

En base a lo anterior se recurre a la construcción de la matriz completa, por consiguiente, se realizará la respectiva división de datos de entrenamiento y clasificación.

Tabla 8 Longitud de datos de las líneas de procesamiento

Técnica de procesamiento	Línea de procesamiento	Sublínea de procesamiento	Fila	Columna
DFT – Entropía - Normalización		-----	375	15420
	Línea 1	Sublínea 1	375	1641
		Sublínea 2	375	870
		Sublínea 3	375	613
DWT – Energía - Normalización		-----	375	3860
	Línea 2	Sublínea 1	375	485
		Sublínea 2	375	292
		Sublínea 3	375	228



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



Cuando se habla de matriz completa, hace referencia a que contiene todas las señales que han pasado por cada línea de procesamiento, se describe de la siguiente forma:

- ✓ Se selecciona 1 señal (paciente) de cualquier patología, esta se pasa por la etapa de preprocesamiento y procesamiento, se tendrá como resultado:
 - Línea 1 – Sublínea 1 (se genera un archivo .csv)
 - Línea 1 – Sublínea 2 (se genera un archivo .csv)
 - Línea 1 – Sublínea 3 (se genera un archivo .csv)
 - Línea 2 – Sublínea 1 (se genera un archivo .csv)
 - Línea 2 – Sublínea 2 (se genera un archivo .csv)
 - Línea 2 – Sublínea 3 (se genera un archivo .csv)

Todos esos datos son solamente para el primer paciente, se repite esto para todos los pacientes que tiene cada patología, en otras palabras, se obtiene:

$$6_{archivos\ csv} * 5_{pacientes} * 5_{patologías} = 150\ archivos\ csv$$

- ✓ En el script donde se crea la matriz completa como tal, se cargan cada uno de estos archivos .csv, al mismo tiempo se agrupan en sus patologías, es decir, se cargan todos los archivos .csv de pacientes que pertenezcan a cierta patología, luego se procede a cargar los demás pacientes de las otras patologías. Dentro del script se realiza un ajuste a los DataFrame (los archivos .csv se han exportado como DataFrame) el cual consiste en la eliminación de la primera columna la cual contiene los índices (posiciones) de las filas de dicho archivo, luego se convierten estos DataFrame en array por la razón de que cuando se exportan como archivo .csv estos se definen como (Columna, Fila), por tanto, para trabajar con los puntos de la señal se requiere aplicar transpuesta en Python para que se convierta en array y quede como (Fila, Columna).



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



En la tabla 9 se describe como está conformada la matriz en sí.

Tabla 9 Matriz completa

Matriz	
Afección valvular	(5 pacientes * 15 señales (derivadas))
Bloqueo de rama	(5 pacientes * 15 señales (derivadas))
Disritmia	(5 pacientes * 15 señales (derivadas))
Hipertrofia	(5 pacientes * 15 señales (derivadas))
Pacientes sanos	(5 pacientes * 15 señales (derivadas))
Total	375 filas

- ✓ Con base a la tabla anterior, se construye en el mismo script donde se crea la matriz completa, valga la redundancia, una matriz de ceros con las siguientes características:
 - $Matriz_Completa = np.zeros(375, Valvular_P1.shape[1])$
Donde tendrá 375 filas por la cantidad de columnas que contiene una de las señales que se han cargado (este valor varía dependiendo de la señal que se cargue con respecto a la línea de procesamiento por la cual ha pasado).
- ✓ La matriz de ceros se irá llenando con los datos que se han cargado con el siguiente orden:
 - Afección valvular (15 señales “Derivadas”)
 - Bloqueo de rama (15 señales “Derivadas”)
 - Disritmia (15 señales “Derivadas”)
 - Hipertrofia (15 señales “Derivadas”)
 - Pacientes sanos (15 señales “Derivadas”)

Finalizada la creación de la matriz completa, cabe resaltar que al momento en que se aplica DFT + Entropía + Normalización (primera línea de procesamiento) y DWT + Energía + Normalización (segunda línea de procesamiento) sin haber aplicado las sublíneas (ventaneo) contiene una cantidad de puntos que se aprecian en la tabla 8, luego una vez que se aplican las sublíneas como consecuencia la cantidad de puntos



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



cambia, por consiguiente, en base a esos resultado se selecciona la mejor técnica de procesamiento para finalizar dicha etapa.

De otro modo, la cantidad de filas y columnas que se genera en cada línea de procesamiento (incluyendo sublíneas) será siempre la misma al momento de seleccionar otra señal que no está dentro de la matriz para proceder a realizar la respectiva clasificación, puesto que, estas sublíneas en si son las características que se necesitan al momento de realizar la predicción con el modelo generado por cada técnica IA.

El paso a seguir es crear el modelo de clasificación mediante las técnicas IA mencionadas anteriormente, por tanto, como las técnicas pertenecen a Machine Learning de tipo lenguaje supervisado se deben etiquetar los datos históricos (se le denomina datos históricos al conjunto de datos que se tiene para la creación del modelo) para tener como referencia a que clase pertenece cada patología.

Se aplican técnicas de VSM con kernel lineal y RBF, Naïve Bayes y árboles de decisión, para estas dos últimas técnicas el etiquetado se realiza de forma binaria (se tiene en cuenta las tablas 4, 5, 6 y 7) donde solo se podrá comparar cada una de las patologías con pacientes sanos de forma independiente, pues algunas de las técnicas de Machine Learning no son multiclase.

La creación del modelo implica dividir la matriz completa en una matriz de entrenamiento con la cual se entrena el modelo y una matriz de clasificación para medir la precisión de la predicción resultante mediante las técnicas empleadas. La división de los datos (matriz completa) para entrenamiento y clasificación se aplica teniendo en cuenta la ley de Pareto 80/20 mediante la función del módulo *sklearn.model_selection* llamando a su clase *train_test_split*, con respecto a lo anterior se tendrá el 80% de los datos de la matriz completa para el entrenamiento y generación del modelo y el 20% (*test_size = 0.2*) restante se emplea para la clasificación/testeo y así calcular la estimación, dicho cálculo se realiza mediante el módulo *sklearn.metrics* llamando a la función *precision_score*



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



La aplicación de diversas técnicas para la clasificación se emplea como motivo para la selección de la mejor técnica teniendo en cuenta el cálculo de la precisión. En la figura 46 se aprecia los porcentajes de precisión/clasificación empleando la técnica VSM con kernel lineal y con kernel RBF.

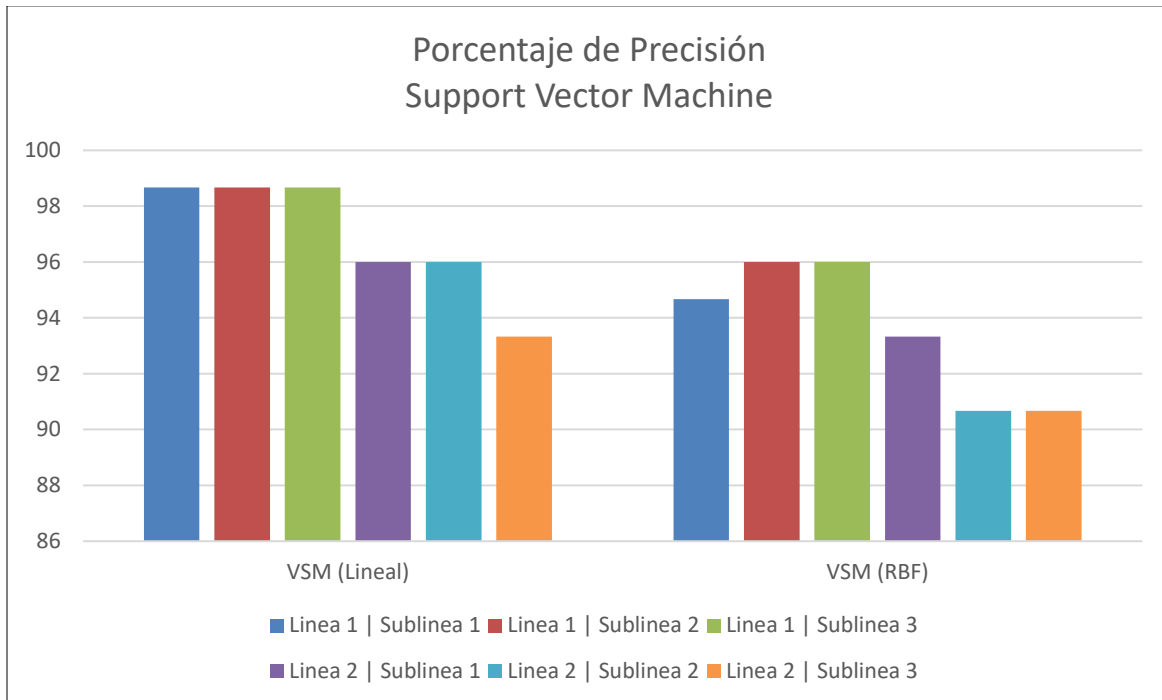


Figura 46 Porcentaje de estimación empleando VSM

Donde,

Línea 1 | Sublínea 1, es la aplicación de DFT, entropía por medio de ventanas de 10 puntos junto con la normalización.

Línea 1 | Sublínea 2, es la aplicación de DFT, entropía por medio de ventanas de 20 puntos junto con la normalización.

Línea 1 | Sublínea 3, es la aplicación de DFT, entropía por medio de ventanas de 30 puntos junto con la normalización.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Línea 2 | Sublínea 1, es la aplicación de DWT, energía por medio de ventanas de 10 puntos junto con la normalización.

Línea 2 | Sublínea 2, es la aplicación de DWT, entropía por medio de ventanas de 20 puntos junto con la normalización.

Línea 2 | Sublínea 3, es la aplicación de DWT, entropía por medio de ventanas de 30 puntos junto con la normalización.

La implementación de SVM para cada señal se realiza mediante la etiquetación de datos según el siguiente orden:

- Etiqueta 1 -> Afección Valvular (75 Señales)
- Etiqueta 2 -> Bloqueo de Rama (75 Señales)
- Etiqueta 3 -> Disritmia Cardíaca (75 Señales)
- Etiqueta 4 -> Hipertrofia Ventricular (75 Señales)
- Etiqueta 5 -> Paciente Sano (75 Señales)

Una vez que se han etiquetado los datos, se tienen las variables necesarias para crear el modelo con VSM, por tanto, se tendría la variable X la cual es la misma matriz completa y la variable y que contiene las etiquetas (matriz de etiquetas), seguidamente se procede a realizar la división de los datos en matriz de entrenamiento y matriz de clasificación de la siguiente manera:

$X_{train}, X_{test}, y_{train}, y_{test} = \text{traint_test_split}(X, y, \text{test_size} = 0.2, \text{random_state} = 0)$

Donde,

X_{train} , es la matriz de datos de entrenamiento

X_{test} , es la matriz de datos de prueba

y_{train} , es la matriz de etiquetas para el entrenamiento

y_{test} , es la matriz de etiquetas para las pruebas

una vez realizado lo anterior, se procede a generar el modelo mediante el módulo SVC, tal y como se muestra de la siguiente forma:

$\text{Kernel} = \text{SVC}(\text{kernel} = \text{'linear'})$



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



Por último, se necesita realizar un ajuste a los datos, para después proceder a clasificar. Se generaron 60 modelos en función de cada línea y sublínea de procesamiento, a partir de estos modelos se calculan los porcentajes de estimación/predicción para la elección de la mejor técnica de clasificación, lo anterior aplica a las demás técnicas con respecto a la carga de la matriz, división de los datos en matriz de entrenamiento y clasificación y el porcentaje de estimación de la predicción, con respecto a las demás técnicas de Machine Learning para la generación del modelo se recurre a los módulos que se instalan en Python por medio de PIP.

El etiquetado de Naïve Bayes se realiza de una manera donde no sea multiclase, puesto que ha generado resultados de precisión bastante buenos cuando se clasifica solamente entre una patología y un paciente saludable. La figura 47 muestra los porcentajes de clasificación al emplearse etiquetas de una sola clase.

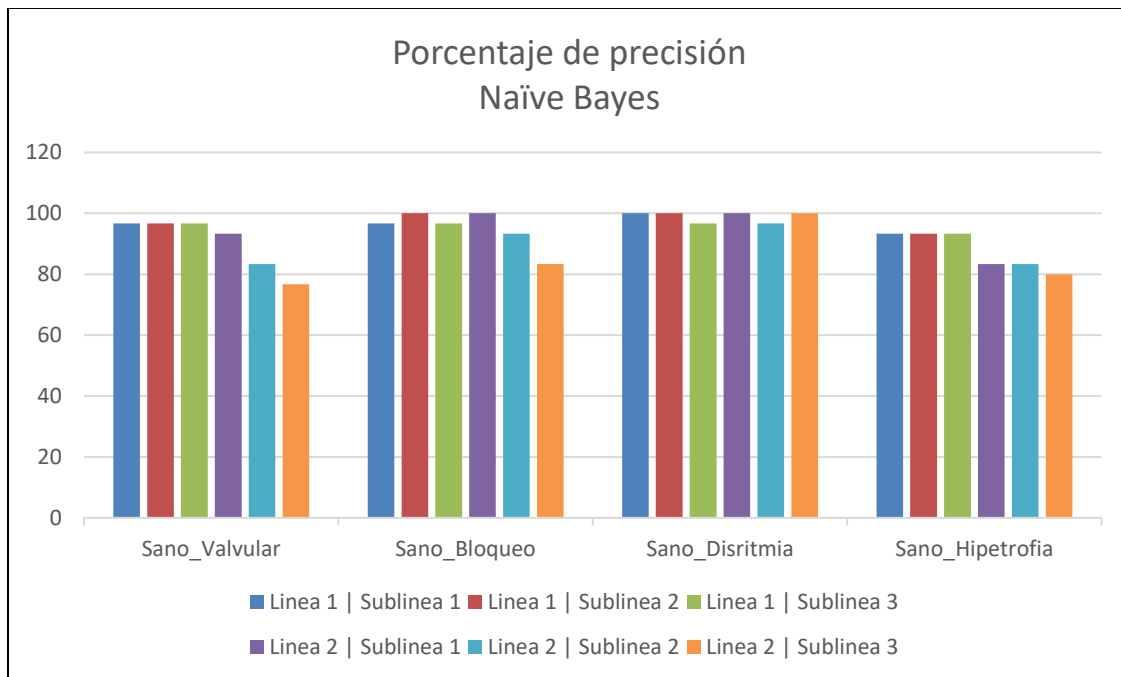


Figura 47 Porcentaje de estimación empleando Naïve Bayes

Con respecto a la figura 47, se observan porcentajes bastante satisfactorios, el problema de seleccionar esta técnica de clasificación implica que solo hará la predicción para una única patología o sí el paciente no contiene ninguna patología cardíaca.



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Para la generación del modelo de Naïve Bayes se realiza de la siguiente forma:

$$\text{Kernel} = \text{GaussianNB}()$$

De igual modo, la división de los datos en entrenamiento y clasificación a partir de la matriz completa, se realiza de la misma manera con que se realiza en SVM.

La técnica de Machine Learning denominada árboles de decisión no emplea multiclases, por tanto, para la etiquetación de los datos se realiza el mismo modo que Naïve Bayes. La figura 48 muestra el resultado del porcentaje de clasificación/estimación al emplearse árboles de decisión.

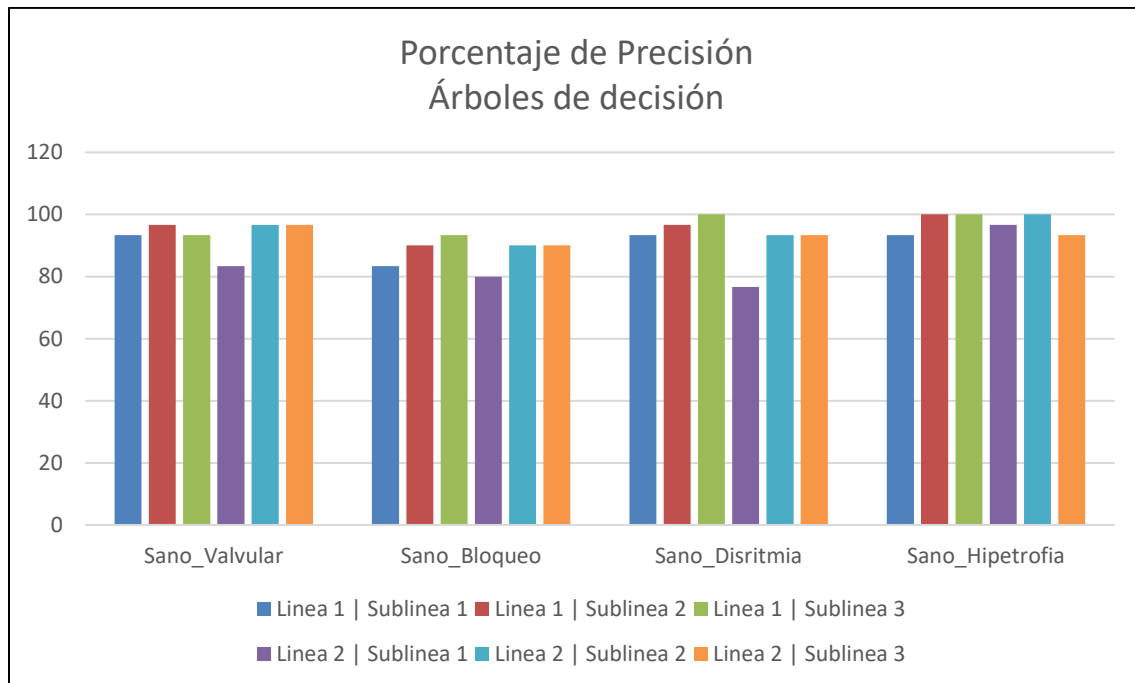


Figura 48 Porcentaje de estimación empleando Árboles de decisión

Para la generación del modelo de Naïve Bayes se realiza de la siguiente forma:

$$\text{Kernel} = \text{DecisionTreeClassifier}(\text{criterion} = \text{"entropy"})$$



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



Se selecciona VSM con kernel lineal como la técnica para la clasificación de los datos, por el hecho de ser multiclase ya que devuelve como resultado en la predicción las etiquetas multiclases a usar y por tanto, genera un mejor porcentaje de clasificación, a su vez, la técnica empleada para el sistema clasificación de señales ECG se basa en la línea de procesamiento con la sublínea 3.

Se eligieron estas tres técnicas de Machine Learning (ML), valga la redundancia, por ser una de las técnicas de ML que se usan con mayor frecuencia para la predicción de datos, a su vez por la razón de haber realizado varios cursos en plataformas educativas para obtener un mejor conocimiento de la temática, de igual modo el dominio de estos temas no se logra con totalidad por la complejidad que estas presentan en cuanto a su parte matemática y conceptual debido al poco tiempo que se tiene para lograr este trabajo y la gran capacidad de temas y recursos educativos que se tiene para comprender con exactitud todo el proceso de generación de modelos, clasificación, visualización y demás.

4.4. INTERFAZ GRÁFICA CON TKINTER

La implementación de una interfaz gráfica de usuario es fundamental para el desarrollo completo del sistema. Mediante la implementación de *tkinter* con Python se diseña una GUI de fácil uso y comprensión, la interfaz consta de dos ventanas, para la primera se le denomina ventana principal o ventana de ejecución/presentación, esta dispone de dos botones, el primer botón tiene la acción de ejecutar la segunda ventana, y el segundo botón es para cerrar/finalizar la GUI.

La segunda ventana es donde el usuario digita los parámetros que se requieren junto con la carga de las datos (señal ECG) y así generar el resultado de la etapa de preprocesamiento, posteriormente el usuario puede generar el procesado de las señales y por último clasificar la señal dependiendo de la derivada que ha seleccionado.

La ejecución de la GUI como aplicación es primordial, por tanto, para lograr esto, la figura 49 se aprecia como se logra lo anterior. Linux permite crear y ejecutar aplicaciones mediante la configuración de archivos *.desktop*, en la ventana de comando de Ubuntu se observa la ruta donde nos encontramos ubicados la cual es */usr/share/applications*, la ventana de archivos muestra todos los archivos que están listos para ser ejecutados como aplicaciones, para convertir el sistema clasificador en



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



una aplicación de Ubuntu, se crear un archivo con nombre y extensión *UI.desktop* dentro de este se escribe el nombre que llevará por título la ventana de ejecución, la versión, el ejecutable (script *.py*), el ícono que tendrá la aplicación al momento de buscarse en el menú de aplicaciones de Ubuntu, y por último el tipo que se le otorgará a este archivo, el cual será de tipo *application*.

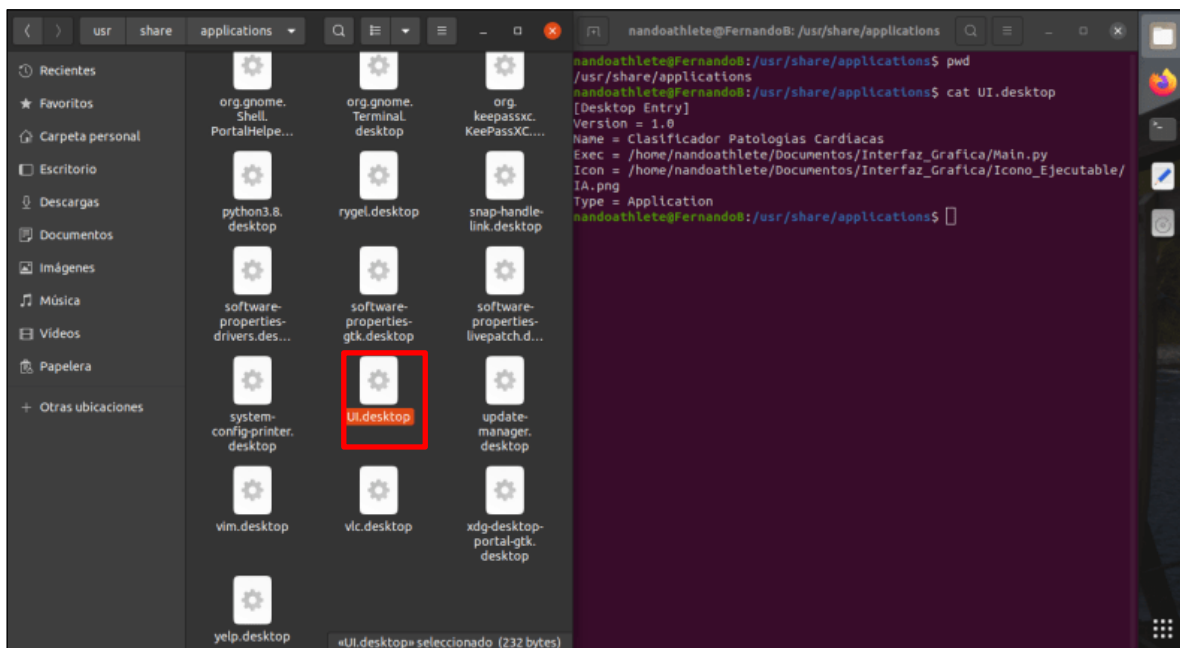


Figura 49 Configurando archivo .desktop

Una vez que se ha configurado dicho archivo dentro de la ruta específica, se procede a buscar la aplicación desde el menú de Ubuntu tal y como se aprecia en la figura 50.



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!

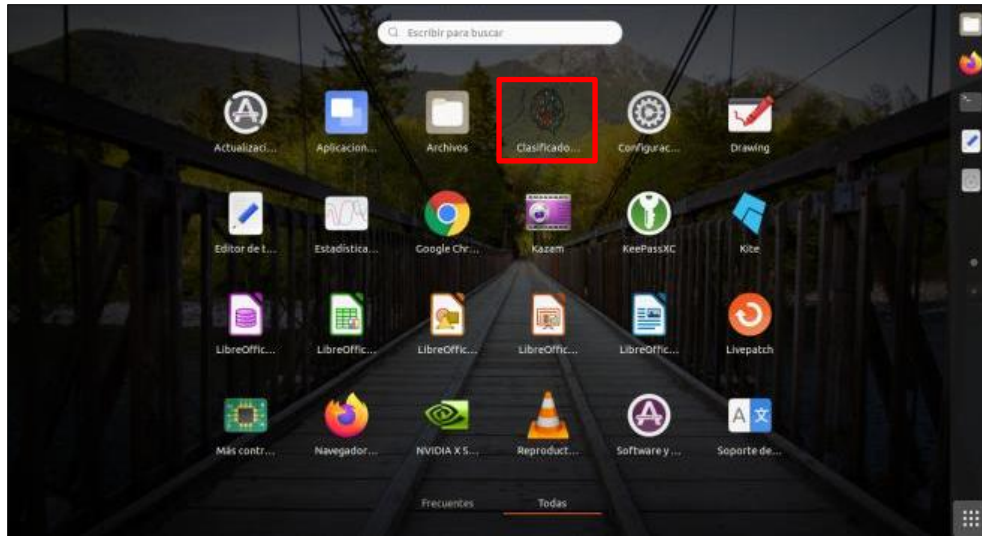


Figura 50 Aplicación Sistema clasificador de patologías cardíacas

Al ejecutarse, aparecerá la primera ventana descrita anteriormente, esto se logra ver en la figura 51.



Figura 51 Ventana principal GUI

Para comenzar a realizar la carga de la señal, preprocesamiento, procesamiento y demás, se da clic en el botón *Comenzar* para ejecutar la segunda ventana, en la figura 52 se puede ver el contenido de la misma, donde se dan unas instrucciones cortas para que el usuario entienda como es el funcionamiento de la GUI.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



Figura 52 Ventana de procesos

Esta GUI se puede dividir en tres secciones, la primer sección es la de preprocesamiento. Allí es donde se realiza la carga de la señal, los datos que ingresa el usuario se usan para ajustar la señal y demás. En la figura 53 se muestran señales ECG que no hacen parte del modelo generado mediante la técnica de clasificación seleccionada, se escoge alguna señal al azar que se requiera de procesar y clasificar, para este caso se selecciona la señal que lleva por nombre *s0032_rem.mat*, para saber cuál es la ganancia y frecuencia con la que se han grabado y cargado las señales ECG en la base de datos, también se ha descargado la información de nada una.

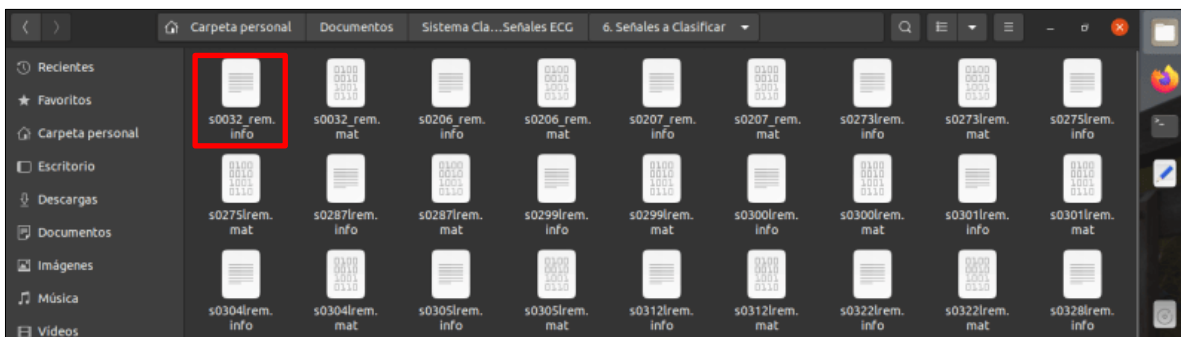


Figura 53 Señales ECG para clasificación



La figura 54 muestra el contenido de los archivos `.info` que traen consigo cada una de las señales una vez descargadas, se tienen en cuenta los datos que se piden en los campos de texto de entrada de la interfaz GUI.

```

1 Source: record ptbdb/patient168/s0032_re
2 val has 15 rows (signals) and 38400 columns (samples/signal)
3 Duration: 0:38
4 Sampling frequency: 1000 Hz Sampling interval: 0.001 sec
5 Row Signal Gain Base Units
6 1 i 2000 0 mV
7 2 ii 2000 0 mV
8 3 iit 2000 0 mV
9 4 avr 2000 0 mV
10 5 avl 2000 0 mV
11 6 avf 2000 0 mV
12 7 v1 2000 0 mV
13 8 v2 2000 0 mV
14 9 v3 2000 0 mV
15 10 v4 2000 0 mV
16 11 v5 2000 0 mV
17 12 v6 2000 0 mV
18 13 vx 2000 0 mV
19 14 vy 2000 0 mV
20 15 vz 2000 0 mV
21
22 To convert from raw units to the physical units shown
23 above, call the 'rdmat.m' function from the wfdb-matlab
24 toolbox: https://physionet.org/physiotools/matlab/wfdb-app-matlab/
  
```

Figura 54 Información de las señales ECG

El paso a seguir es digitar la información que se solicita y cargar la señal a procesar y clasificar, en la figura 55 se muestra la ventana de carga de archivos a través de tkinter, cabe resaltar que los archivos a cargar serán `.mat`.

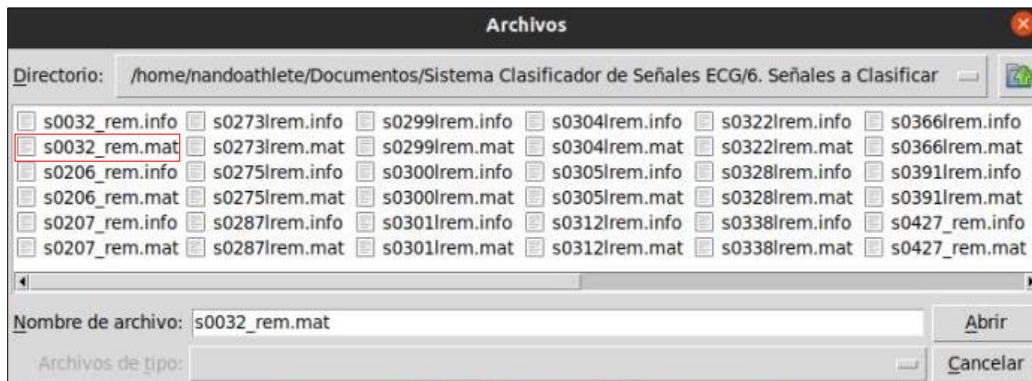


Figura 55 Carga de señales a la GUI

Una vez que se ha cargado la señal junto con los datos que el usuario debe ingresar, el sistema da como resultado el preprocesamiento de la señal en una ventana independiente, la figura 56 muestra cada una de las derivadas del corazón una vez ya preprocesadas, esto da como finalizado la primera sección.

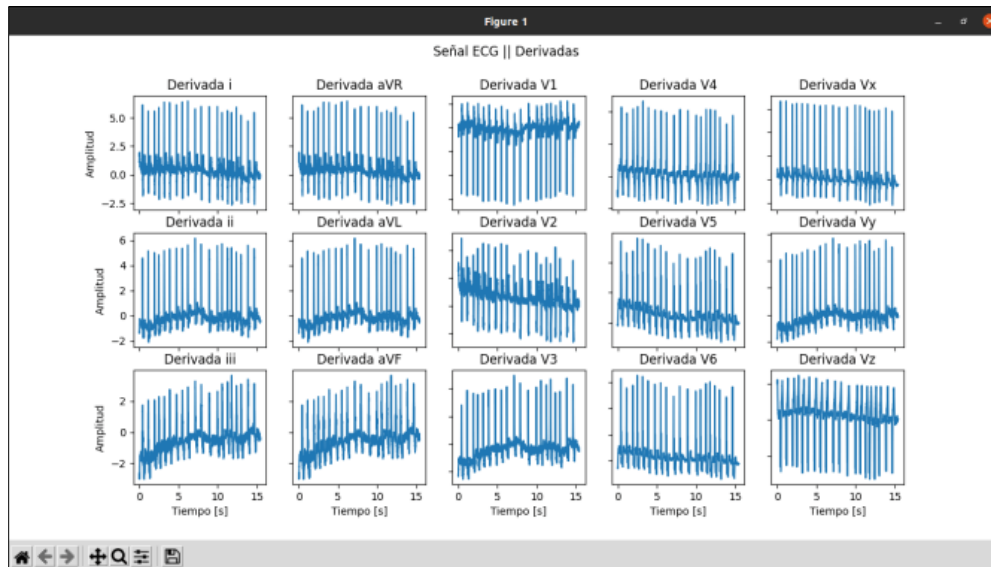


Figura 56 Preprocesamiento de las derivadas del corazón

La sección dos se basa en el procesamiento de la señal, allí es donde se aplican las técnicas de procesamiento de señales para la extracción de los patrones característicos, el usuario solamente debe dar clic en el botón el botón *Procesar*, una vez que éste de clic se mostrará una ventana con los resultados de las características de cada una de las señales cargadas, esto se puede observar en la figura 57, con esto se concluye la sección dos.

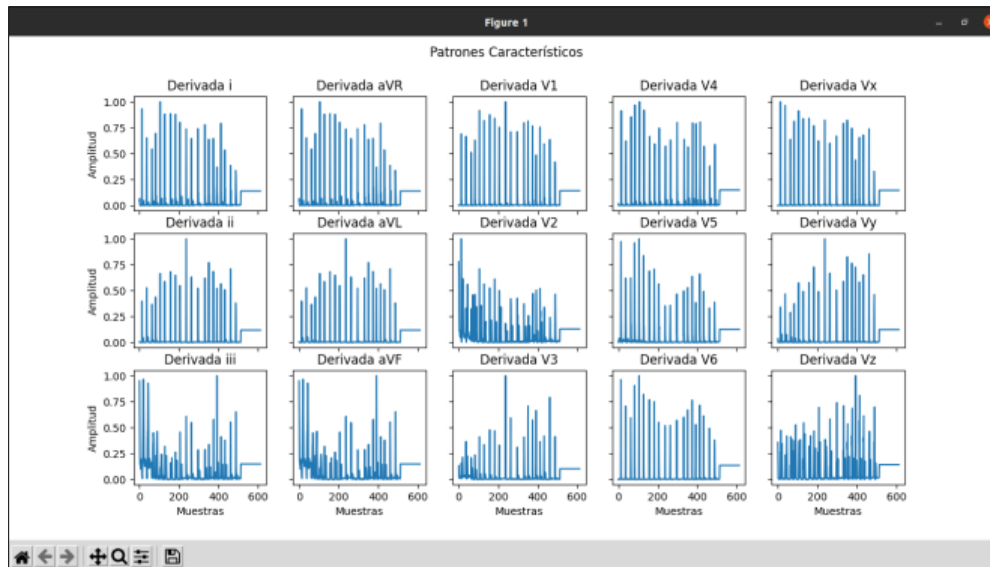


Figura 57 Patrones característicos

Para la sección tres el usuario (especialista en cardiología o alguna entidad de salud), deberá seleccionar alguna de las derivadas del corazón, cabe recordar que las derivadas no es más que la dirección en la que se observa el corazón y analizar si se presentan o no afecciones cardíacas, esta selección se realiza por medio de un *combobox* (lista desplegable) que contiene *tkinter*, una vez que selecciona la derivada el paso a seguir es dar clic en el botón *Clasificar*, la figura 58 muestra la selección de la derivada a seleccionar.



Figura 58 Proceso de Clasificación

Una vez que ha sido seleccionada se procede a clasificar, el resultado de la predicción de la señal que se ha cargado al sistema clasificador de patologías cardiovasculares se muestra en la figura 59.



Figura 59 Resultado de la predicción



4.5. IMPLEMENTACIÓN DE CLOUD COMPUTING

La implementación de Pythonanywhere para la inclusión de Cloud Computing permite ejecutar código Python por medio de un micro-framework denominado Flask, por tanto, al momento de cargar una señal al servidor Cloud ésta se almacena, se procesa, y se clasifica en la misma nube. Para trabajar con Flask en el Cloud, se debe crear un entorno virtual e instalar los paquetes necesarios para el funcionamiento del mismo y del sistema clasificador como una aplicación web, la figura 60 muestra los archivos que conforman todo lo relacionado a flask y el entorno virtual.

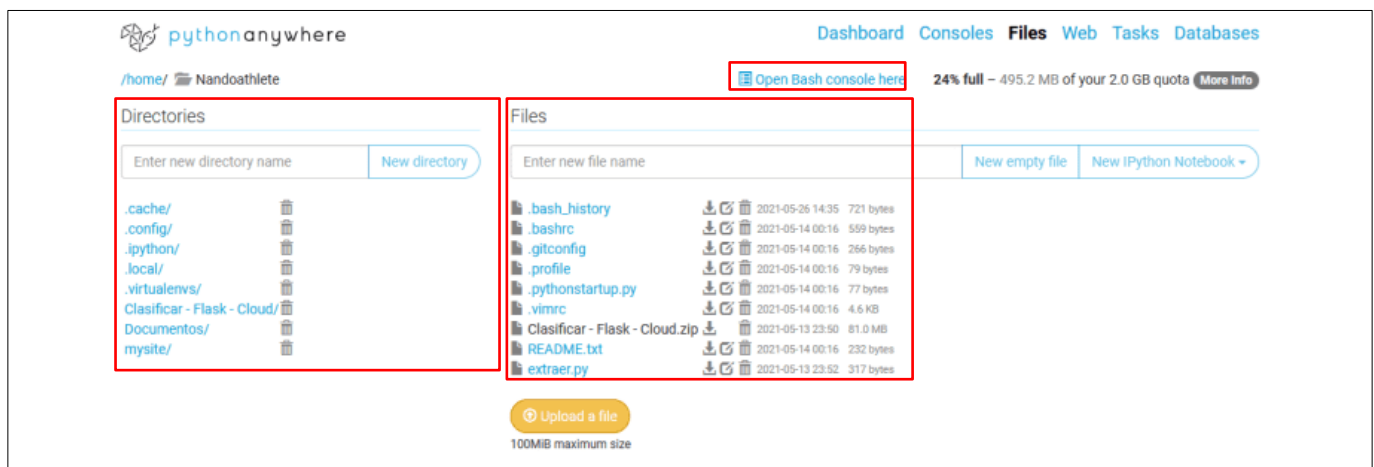


Figura 60 Directorios y archivos en Pythonanywhere

La carpeta denominada **.virtualenvs** contiene el entorno virtual el cual se ha llamado Flask, allí dentro están todos los módulos, submódulos y funciones que se requieren al momento de ejecutar el aplicativo web, por tanto, en caso de alguna actualización del aplicativo y requiere de otros módulos o eliminación de algunos, se debe activar el entorno virtual por medio de la consola la cual se puede acceder desde la opción **Open Bash Console Here** que ofrece *Pythonanywhere* y desde allí se realiza dicho paso. La carpeta que lleva por nombre **Clasificar – Flask – Cloud** es la que contiene todos lo funcional a flask esto se aprecia en la figura 61. Para el funcionamiento de los archivos **.html** los cuales dan estructura a las páginas que se ejecutan, estos se almacenen en la carpeta **templates** y se usan por medio de la función *render_template* que ofrece Flask, la carpeta **Model** contiene el modelo que se ha generado a partir de la implementación de la técnica VSM con kernel lineal, este modelo se llama desde el



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



código principal denominado *Main.py*. La carpeta llamada **static** contiene todas las imágenes ya sea en formato *.jpg*, *.png*, *.jpeg* e inclusive si se requiere de algún script *.css* este se almacena allí.

Por último la carpeta **uploads** es donde se almacena únicamente archivos *.mat*, es decir, las señales que se han cargado al Sistema Cloud para ser procesadas y clasificadas, esto requiere de almacenamiento y en caso de que haya llegado al límite, estas se pueden eliminar.

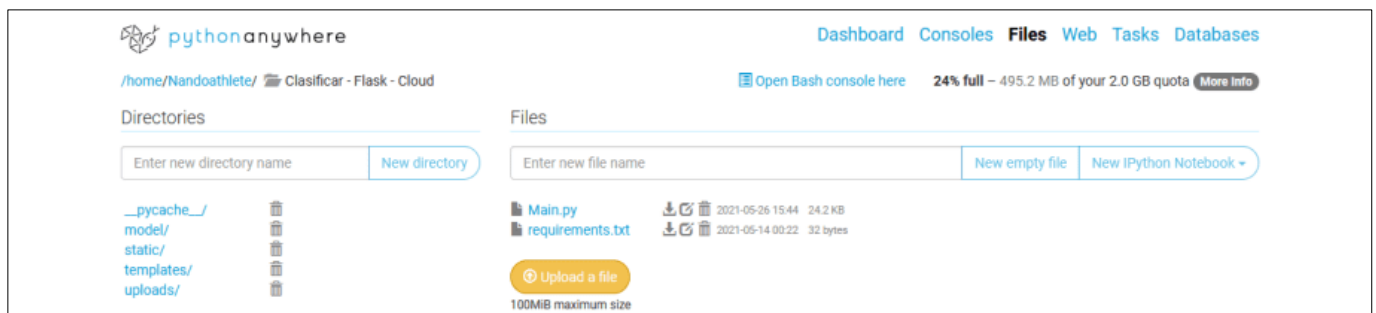


Figura 61 Archivos que conforma Flask

La correcta configuración y resolución de problemas que se pueden presentar en Pythonanywhere debido a errores de sintaxis, módulos no incorporados o incompatibilidad de Python con ciertas opciones permite que el sistema se ejecute de forma correcta, con lo anterior se presentaron problemas graves al momento de la ejecución, uno de ellos es la variable de entorno, por tanto, se debe agregar el valor de dicha variable en la configuración del archivo **postactivate**. Esta variable en la configuración del archivo principal (*Main.py*) tiene como valor la ruta donde se van a guardar los archivos que se cargan al sistema por tanto se debe exportar el mismo valor e *postactivate*, la figura 62 muestra la configuración de los parámetros de cada archivo.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



```
Keyboard shortcuts: Normal Share Save Save as... >>> Run
1 from flask import Flask, render_template, request, redirect, url_for
2 import os
3 import joblib
4 import pandas as pd
5 import numpy as np
6 import scipy.io as sio
7 from werkzeug.utils import secure_filename
8
9 #Definiendo aplicación
10 app = Flask(__name__)
11 app.config["UPLOAD_FOLDER"] = "/home/Nandoathlete/Clasificar - Flask - Cloud/uploads/"
12

/home/Nandoathlete/.virtualenvs/postactivate
Keyboard shortcuts: Normal Share Save Save as...
1 #!/bin/sh
2 # This hook is sourced after every virtualenv is activated.
3 export UPLOAD_FOLDER="/home/Nandoathlete/Clasificar - Flask - Cloud/uploads"
4 #export ALLOWED_EXTENSIONS="mat"
```

Figura 62 Configuración de variable de entorno

Otro error que se debe corregir es la configuración de WSGI (Web Server Gateway Interface) el cual es un estándar que permite la comunicación a través de protocolo HTTP, la ventaja que ofrece Pythonanywhere es que ya lo trae incluido y lo que se debe hacer es la configuración del archivo WSGI desde la *Web App* damos clic al enlace con la etiqueta *WSGI configuration file* tal y como se observa en la figura 63.

```
Code:
What your site is running.
Source code: /home/Nandoathlete/mysite Go to directory
Working directory: /home/Nandoathlete/ Go to directory
WSGI configuration file: /var/www/nandoathlete_pythonanywhere_com_wsgi.py
Python version: 3.8

Virtualenv:
Use a virtualenv to get different versions of flask, django etc from our default system ones. More info here. You need to Reload your web app to activate it; NB - will do nothing if the virtualenv does not exist.
/home/Nandoathlete/.virtualenvs/flask
Start a console in this virtualenv
```

Figura 63 Archivo WSGI

Una vez estando dentro, se debe configurar la ruta del directorio donde se encuentra nuestro proyecto, por tanto, el valor de la variable *project_name* se establece como la ruta donde se encuentran todos los archivos que conforman la Web App Flask, una vez hecho esto, se debe importar el nombre del script *Main.py* (script principal) como módulo y desde ese módulo se llama a la función *app* (aplicación web definida dentro





ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



del script *Main.py*), pues está es la que contiene todo el código flask, para validaciones, seguridad de los datos, llamado de archivos html y demás.

```
#!/usr/bin/env python
# This file contains the WSGI configuration required to serve up your
# web application at http://your-username.pythonanywhere.com/
# It works by setting the variable 'application' to a WSGI handler of some
# description.
#
# The below has been auto-generated for your Flask project
import sys

# add your project directory to the sys.path
project_home = '/home/Nandoathlete/Clasificar - Flask - Cloud'
if project_home not in sys.path:
    sys.path = [project_home] + sys.path

# import flask app but need to call it "application" for WSGI to work
from Main import app as application # noqa
```

Figura 64 Configuración del archivo WSGI

Finalizado los ajustes, se procede abrir el enlace siguiente:
<http://nandoathlete.pythonanywhere.com>

La figura 65 muestra el aplicativo web una vez se ha ingresado al enlace. Se observa una aplicación web sencilla y fácil de usar, esta cuenta con el ingreso de datos que el usuario debe ingresar en base a la señal, se tiene la opción de elegir la derivada la cual desea clasificar para saber si existe algún problema cardiovascular ya mencionados anteriormente, se cuenta con validaciones con validaciones y retornos a la página principal.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 65 Aplicativo web

En caso de que el usuario no ingrese ningún dato en los campos de textos, y haya seleccionado la señal a cargar, se retorna a una página web donde se le indica que los campos están vacíos y éste podrá retornar a la página principal dando clic en el vínculo *Regresar*, la figura 66 muestra el resultado al realizar lo anteriormente descrito.



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 66 Campos de texto vacías

Otro caso es en el que el usuario ha ingresado los datos en los campos de texto, pero no ha seleccionado ningún archivo, retornará a otra página web indicando al usuario que no se ha cargado ningún archivo, esto se ve reflejado en la figura 67.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!

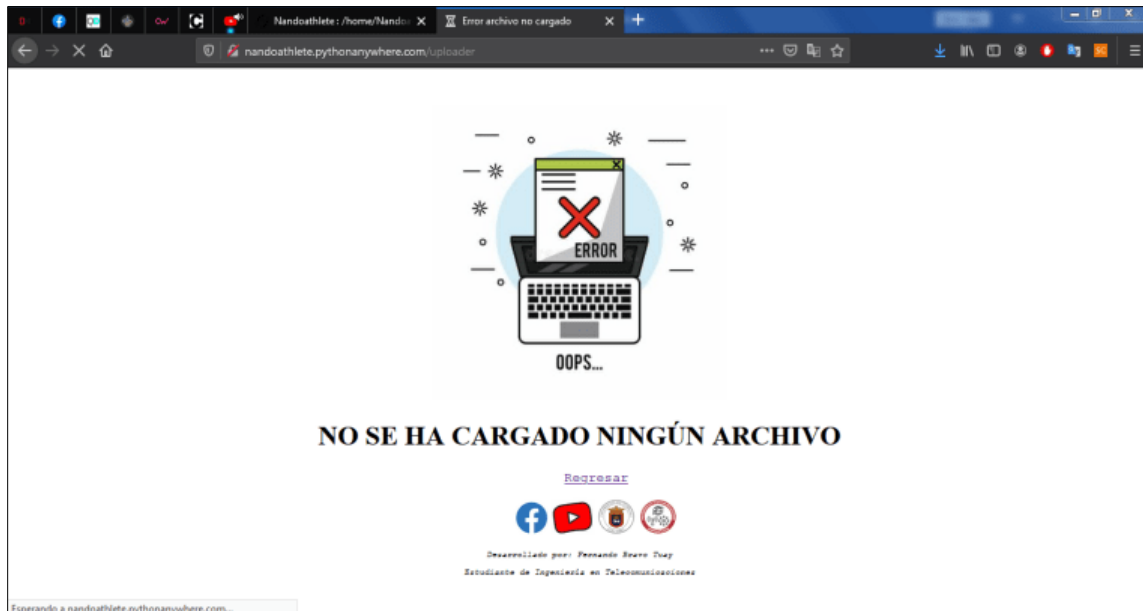


Figura 67 Archivo no cargado

De igual modo, se ha aplicado seguridad en manejo de datos y validación del mismo mediante la implementación del módulo *werkzeug* que trae consigo *Flask* usando la función *secure_filename*, por tanto, el aplicativo web solo permite cargar archivos que sean *.mat*, en caso de que no cargue archivos con dicha extensión se retornará a una página indicando al usuario de que no es un archivo válido. El módulo *werkzeug* verifica el nombre del archivo, es decir, se pasa el nombre del archivo y éste devuelve una versión segura de tipo ASCII para una máxima portabilidad. La figura 68 muestra la carga de un archivo *.jpg* (archivo que no acepta el aplicativo web).



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 68 Cargando archivo que no es .mat

Al realizar lo anterior, se retorna a otra página web donde le indica al usuario que el archivo que ha cargado no es válido, lo cual deberá volver a la página principal y cargar el archivo correcto, dicho resultado se muestra en la figura 69.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 69 Validación de archivos .mat

Una vez concluidas las validaciones, se procede a cargar una señal *.mat* para ser procesada en el sistema y esperar el resultado de la predicción. En la figura 70 se observa la señal que se ha cargado y los parámetros de los campos de entrada, se debe tener en cuenta que, si se tiene una señal con una longitud menor a la establecida en la configuración de todo el sistema, no podrá realizarse dicha predicción, por tanto, lo recomendable es que se tenga una señal por encima de los 20.000 puntos de muestras. Una vez que se ha cargado la señal, se procede a dar clic en **Enviar**, cabe aclarar que si no se ha seleccionado una derivada, está por defección toma el valor de la que se presenta una vez se inicia la aplicación web.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 70 Comprobación del sistema

El resultado se puede apreciar en la figura 71, donde se muestra por medio de una imagen gráfica la afección que puede presentar el paciente.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



Figura 71 Resultado de predicción

Para dar finalidad a esta etapa, se hace un comparativo en cuanto a carga de datos, velocidad de procesamiento en menor tiempo, visualización de los resultados, y las ventajas que podría ofrecer tanto la interfaz gráfica de escritorio como el aplicativo web.

La interfaz gráfica de escritorio ofrece mayor velocidad de carga de datos, visualización, procesamiento (dependiendo de las características de la máquina anfitriona), el inconveniente que presenta es que se ha desarrollado para el sistema operativo (SO) Linux, por tanto, si se requiere de ejecutar dicha interfaz en otro SO, se debe realizar la instalación de los respectivos módulos y mismas versiones de estos para que se ejecute y a su vez bajo la misma versión de Python. En caso de que alguna entidad requiera uso del sistema clasificador de tendrá que hacer el respectivo traslado de la GUI al equipo de dicha entidad para que esta pueda ser ejecutada.

Por otro lado, la implementación de Cloud Computing mejora a comparación de la GUI, en cuanto a de carga de las señales a la nube depende del tamaño del archivo a subir y velocidad de subida que tenga el usuario hacia la red, el procesamiento se realiza bajo la misma nube, por lo tanto, no requiere de una máquinas con excelentes especificaciones, sino más bien depende de la capacidad y servicios que ofrezca la



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL

Avanzamos... ¡Es nuestro objetivo!



nube, para este caso Pythonanywhere, algunas veces sacrificar ciertos aspectos por otros, es recurrente y necesario, pues al no tener una interfaz gráfica de usuario, solamente se usan los servicios que ofrece la plataforma donde se ha diseñado, por tanto, debido a que se está procesando en la misma nube, se requiere de más consumo “computacional” y para lograr esto se realiza mediante la suscripción por un mejor servicio en cuanto a velocidad de procesamiento y almacenamiento, por tanto, como ventaja cualquier persona podría acceder a este medio y solamente requeriría de una buena conexión a la red, gracias a estas nuevas tecnologías permiten que los sistemas se cargados en la nube para ofrecer un servicio a quien lo necesite.

Cloud Computing es bastante novedoso en cuanto a prestaciones virtuales a cambio de un pago para optar por un mejor servicio, gracias a esto no se requiere de un mayor consumo computacional y es prácticamente flexible para el uso en línea desde cualquier lugar del mundo.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



CAPÍTULO V: CONCLUSIONES

- Debido a que las señales no sean linealmente separables y no son estacionarias, implica una mayor complejidad con respecto al análisis para la extracción de los patrones característicos.
- Las afecciones cardiovasculares se trabajaron de forma global, a causa de que la base de datos que proporciona las señales ECG no especifica la clasificación de las afecciones, es decir, para cada una de estas se dispone de más patologías, por tanto, el analizar una señal con antecedentes de afección valvular no es lo mismo que analizar una señal que presente cardiopatía por hipertrofia, para la primera con lleva más afecciones tales como estenosis cardíaca, estenosis mitral, entre otras y para la segunda, están las enfermedades por hipertrofia ventricular izquierda y derecha, hipertrofia ventricular izquierda y derecha, entre otros, por tanto, realizar un análisis con mayor profundidad requiere de mucho más tiempo, más análisis y más conocimiento por parte de la medicina en el área cardiovascular.
- La selección de la primera línea de procesamiento con ventanas de 30 puntos se realiza por elección propia y por el hecho de otorgar un mejor porcentaje de clasificación de tipo multiclase a comparación de las demás técnicas empleadas.
- La implementación de interfaz gráfica de usuario es importante para realizar los pasos de preprocesamiento, procesamiento y obtener como resultado una predicción de forma visual a comparación de ejecutar código por medio de la terminal de comandos del sistema operativo Linux, por lo cual se emplea tkinter permitiendo diseñar ventanas agradables y de fácil uso para el usuario.
- Se selecciona como servicio Cloud de almacenamiento Pythonanywhere por la experiencia que se ha obtenido años atrás realizando pruebas de aplicaciones con API REST mediante la implementación de Flask, lo cual es de gran ayuda por el hecho de permitir ejecutar código Python en aplicativos webs de forma más rápida, fácil y manejable.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



REFERENCIAS

- [1] Pozas G., “El electrocardiograma y su tecnología,” *Educ. Medica*, vol. 8, no. 24, pp. 27–31, 2008, [Online]. Available: http://www.cmzh.com.mx/media/108934/rev_24_electricardiogramasutecnologia.pdf.
- [2] C. V. Silva and V. G. Rojas, “Diseño e implementación de un sistema electrocardiográfico digital,” *Rev. Fac. Ing.*, no. 55, pp. 99–107, 2010.
- [3] “Minimanual de ECG Ilustrado,” 2005, [Online]. Available: <https://muymedico.com>.
- [4] J. Garcia Niebla, “Errores y artefactos más comunes en la obtención del electrocardiograma,” *Siacardio*, vol. 1, no. 8, pp. 1–29, 15AD, [Online]. Available: <http://www.siacardio.com/wp-content/uploads/2015/01/ECG-Capitulo-10-Errores-y-artefactos-comunes-en-ECG.pdf>.
- [5] R. Sastre, Iván;Pérez, “ANATOMÍA Y FISIOLOGÍA DEL CORAZÓN,” *Anatomía y Fisiol. del corazón*, pp. 1–30, 2014, [Online]. Available: <https://www.faeditorial.es/capitulos/infarto-miocardio.pdf>.
- [6] J. Ramirez, “Fisiología cardiaca,” *Revista Médica MD*, vol. 1, no. 3, pp. 3–6, 2009.
- [7] P. Acosta, A. Álvarez, J. Betancourt, J. Cabrales, and N. Ascencio, *Electrocardiografía*, Primera Ed. Bogotá.
- [8] D. E. Mohrman and L. J. Heller, *Fisiología Cardiovascular*. 2007.
- [9] B. E. V. Angarita *et al.*, “Arritmias.”
- [10] D. S, “Transformada Discreta de Fourier (DFT).” [Online]. Available: <https://sites.google.com/site/dson978/signals-systems/discrete-fourier-transform>.
- [11] G. Pazmiño Castillo and R. M. García Oquendo, “La transformada de Fourier y su aplicación en el procesamiento de imágenes,” *Tsafiqui*, no. 1, p. 7, 2008, doi: 10.29019/tsafiqui.v0i1.23.
- [12] K. Najarian and R. Splinter, *Biomedical Signal and Image Processing*. CRC Press, Taylor & Francis Group, 2006.



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



- [13] J. Semmlow, *Biosignal and Medical Image Processing*. CRC Press, Taylor & Francis Group, 2009.
- [14] M. Larrota, "Aplicación de la transformada wavelet discreta en el filtrado de señales bioeléctricas," *Umbral Científico*, no. 5, pp. 92–98, 2004.
- [15] "La Transformada Wavelet." [Online]. Available: <http://bibing.us.es/proyectos/abreproy/11511/fichero/PFC+Silvia+Blasco+Vadillo%252FCapítulo+9+--+Anexo+2.pdf+>.
- [16] B. Medina, J. E. Sierra, and A. B. Ulloa, "Técnicas de extracción de características de señales EEG en la imaginación de movimiento para sistemas BCI," *Espacios*, vol. 39, no. 22, pp. 36–48, 2018, [Online]. Available: <https://www.revistaespacios.com/a18v39n22/18392236.html>.
- [17] G. Garg, V. Singh, and A. P. Mittal, "Wrapper Based Wavelet Feature Optimization for EEG Signals." doi: 10.1007/s13534-012-0044-0.
- [18] B. A. Medina and R. Alvarez López, "Caracterización de Señales EEG mediante Wavelet Packet y Entropía Difusa para Tareas de Imaginación Motora," *Ingeniería*, vol. 22, no. 2, p. 226, 2017, doi: 10.14483/udistrital.jour.reving.2017.2.a04.
- [19] E. Neira, "MODELO DE IDENTIFICACIÓN DE LOCUTOR EN ENTORNOS GSM, APLICACIÓN EN COLOMBIA," Universidad Nacional de Colombia, 2010.
- [20] A. Ramiro and V. Alvarado, "Machine Learning para Todos Aldo Valdez Alvarado," *ResearchGate*, no. January 2019, 2020, doi: 10.13140/RG.2.2.13786.70086.
- [21] G. Betancourt, "LAS MÁQUINAS DE SOPORTE VECTORIAL (SVMs)," *Sci. Tech.*, vol. XI, pp. 67–72, 2005, [Online]. Available: <https://www.redalyc.org/pdf/849/84911698014.pdf>.
- [22] L. Gonzalez, "Naïve Bayes." <https://aprendeia.com/naive-bayes-teoria-machine-learning/>.
- [23] M. Quintero and A. Edgar, *Alternativas de Uso de la Tierra para los Llanos Orientales de Colombia*. pp. 1–46.
- [24] L. Gonzalez, "Árboles de Decisión - Clasificación." <https://aprendeia.com/arboles-de-decision-clasificacion-teoria-machine-learning/>.
- [25] A. Youssef, "Exploring Cloud Computing Services and Applications," *J. Emerg.*



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



- Trends Comput. ...*, vol. 3, no. 6, pp. 838–847, 2012, [Online]. Available: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Exploring+Cloud+Computing+Services+and+Applications#0>.
- [26] N. Quintero and A. Fuentes, “Computacion En La Nube,” pp. 46–51, 2014.
- [27] “Cloud Computing La revolución tecnológica para las empresas.” [Online]. Available: <https://systemsgroup.es/noticias-it/cloud-computing/32212/>.
- [28] T. Cardiovasc, “Listado de Aplicacion IA en cardiología,” vol. 70, no. December 2016. pp. 2016–2017, 2017.
- [29] S. Sahoo, M. Mohanty, S. Behera, and S. K. Sabut, “ECG beat classification using empirical mode decomposition and mixture of features,” *J. Med. Eng. Technol.*, vol. 41, no. 8, pp. 652–661, 2017, doi: 10.1080/03091902.2017.1394386.
- [30] Q. Li, C. Rajagopalan, and G. D. Clifford, “A machine learning approach to multi-level ECG signal quality classification,” *Comput. Methods Programs Biomed.*, vol. 117, no. 3, pp. 435–447, 2014, doi: 10.1016/j.cmpb.2014.09.002.
- [31] U. R. Acharya *et al.*, “A deep convolutional neural network model to classify heartbeats,” *Comput. Biol. Med.*, vol. 89, no. August, pp. 389–396, 2017, doi: 10.1016/j.combiomed.2017.08.022.
- [32] Ö. Yıldırım, P. Pławiak, R. S. Tan, and U. R. Acharya, “Arrhythmia detection using deep convolutional neural network with long duration ECG signals,” *Comput. Biol. Med.*, vol. 102, no. August, pp. 411–420, 2018, doi: 10.1016/j.combiomed.2018.09.009.
- [33] B. R. de Oliveira, C. C. E. de Abreu, M. A. Q. Duarte, and J. Vieira Filho, “Geometrical features for premature ventricular contraction recognition with analytic hierarchy process based machine learning algorithms selection,” *Comput. Methods Programs Biomed.*, vol. 169, pp. 59–69, 2019, doi: 10.1016/j.cmpb.2018.12.028.
- [34] J. Díaz, J. Fernández, and E. Guerrero, “Diagnóstico automático del síndrome coronario agudo utilizando un sistema multiagente basado en redes neuronales,” *Rev. Colomb. Cardiol.*, vol. 24, no. 3, pp. 1–6, 2017, doi: <http://dx.doi.org/10.1016/j.rccar.2016.11.010>.
- [35] C. D. E. L. Adulto, L. Morales Muñoz, G. Quintana, and L. F. Niño, “Modelo computacional para la identificación de endofenotipos y clasificación de pacientes con artritis reumatoide a partir de datos genéticos, serológicos y clínicos, utilizando técnicas de inteligencia computacional,” *Rev. Colomb. Reumatol.*, vol.



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750



ACREDITACIÓN INSTITUCIONAL
Avanzamos... ¡Es nuestro objetivo!



22, no. 2, pp. 90–103, 2015, doi: 10.1016/j.rcreu.2015.05.005.

- [36] J. J. Sprockela, J. J. Diaztagleb, W. Alzatec, and E. González, “Redes neuronales en el diagnóstico del infarto agudo de miocardio,” vol. 21, no. 4, 2014.
- [37] GERMAN ARLEY PORTILLA GONZÁLEZ, “DESARROLLO DE UN SISTEMA PARA CLASIFICACIÓN DE CARACTERES OBTENIDOS DE SEÑALES ELECTROENCEFALOGRAFÍCAS (EEG) APLICADO EN PACIENTES CON AFASIA MOTRIZ.,” Universidad de Pamplona, 2017.
- [38] A. Lizcano-Portilla, L. Mendoza, and Z. Nieto-Sánchez, “Procesamiento de señales cerebrales provenientes de estímulos visuales y auditivos utilizando análisis wavelet y redes neuronales artificiales,” *Rev. UIS Ing.*, vol. 19, no. 2, pp. 119–125, 2020, doi: 10.18273/revuin.v19n2-2020013.
- [39] E. Arias, L. García, O. Gualdrón González, and R. Corrêa, “Redes neuronales artificiales en conducción de calor multidimensional transitorio,” *Rev. UIS Ing.*, vol. 2, no. 1, pp. 51–61, 2003.
- [40] “Base datos physiobank.” [Online]. Available: <https://archive.physionet.org/physiobank/>.
- [41] “Pythonanywhere.” 2019, [Online]. Available: <https://opensofty.com/es/2019/11/8/los-mejores-ides-basados-en-la-nube-para-python/>.
- [42] J. Muñoz, “Que es Flask.” 2017, [Online]. Available: <https://openwebinars.net/blog/que-es-flask/>.



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750