



Universidad de los Andes  
Facultad de Ciencias  
Departamento de Física  
Laboratorio de Física Aplicada

**Búsqueda de agrupaciones en data proveniente de electrocardiogramas (ECG), mediante el Análisis de Componentes Principales (PCA) y el uso de Redes Neuronales.**

*Trabajo especial de grado.*

**Br. Abrahan David Quintero Teran**

Tutor: Prof. Juan Villegas

Jurados:  
Prof. Marcos Rodríguez  
Prof. John Ferreira

**Mérida – Venezuela**  
2024

## Resumen

Este es el resumen de mi tesis.

# Índice general

<b>1</b>	<b>El problema</b>	<b>2</b>
1.1.	Justificación . . . . .	3
1.2.	Objetivos . . . . .	3
<b>2</b>	<b>Marco Teórico</b>	<b>5</b>
2.1.	Antecedentes . . . . .	5
2.2.	Conceptos básicos . . . . .	12
2.3.	Conceptos específicos . . . . .	14
<b>3</b>	<b>Marco experimental</b>	<b>17</b>
<b>4</b>	<b>Resultados</b>	<b>18</b>
<b>I</b>	<b>Conclusiones</b>	<b>19</b>
<b>5</b>	<b>Conclusiones</b>	<b>20</b>
	<b>Bibliografía</b>	<b>21</b>

# Introducción

# Capítulo 1

## El problema

El electrocardiograma (ECG) es una técnica no invasiva que permite registrar y medir las señales eléctricas generadas por el corazón. Consiste en la captación de la variación temporal del potencial bioeléctrico durante cada ciclo cardíaco, utilizando electrodos colocados en la superficie cutánea del paciente que registran dicha actividad eléctrica y producen un gráfico que muestra el ritmo y la fuerza de los latidos del corazón, este produce un patrón muy específico, y cualquier anomalía puede indicar un problema cardíaco [1]. El análisis del ECG proporciona datos sobre el sistema cardiovascular, en particular, el corazón, lo que permite detectar diversas enfermedades que pueden afectar su funcionamiento óptimo. Estas enfermedades incluyen arritmias cardíacas, obstrucción de arterias, insuficiencia cardíaca y ataques al corazón [2]. Según la organización mundial de la salud [3], las enfermedades cardiovasculares (ECV) son la principal causa de muerte en hombres y mujeres en el mundo, con alrededor de 17,9 millones de personas que mueren al año a causa de estas. Entre los numerosos factores que llevan a esta consecuencia, se encuentran los errores provenientes de la interpretación manual de los electrocardiogramas (ECGs), por lo general, los médicos emplean características heurísticas diseñadas manualmente o utilizan arquitecturas de aprendizaje de características superficiales, esto puede generar como consecuencia, variabilidad entre los diagnósticos de los observadores e identificación de anomalías incorrectas que pueden llevar a ocasionar diagnósticos imprecisos y en consecuencia, tratamientos inadecuados. Además estos métodos manuales que utilizan arquitecturas de aprendizaje de características superficiales, descartan información relevante del ECG inmersa dentro de características que no son superficiales, lo que provee una baja exactitud en el diagnóstico a partir de las señales por lo que siempre será necesario de la supervisión de un experto

con el fin de corregir estos errores, además algunas enfermedades cardíacas, como la enfermedad de las arterias coronarias en etapa temprana, pueden no ser detectables en un ECG, también se debe resaltar, que el ECG se puede ver afectado por factores externos, como el movimiento del paciente, las interferencias electromagnéticas, la incorrecta colocación de los electrodos, estados ansiosos de los pacientes al momento de aplicarles el examen o el uso de medicamentos.

## **1.1. Justificación**

La interpretación manual de los ECGs está sujeta a errores humanos por fatiga, sesgos o variabilidad interobservador, lo que puede generar diagnósticos incorrectos y por ende, tratamientos inadecuados. Además, el ECG contiene información oculta (como edad, sexo o incluso identidad del paciente) que el ojo humano no puede cuantificar o identificar sistemáticamente. Aunque existen herramientas digitales para análisis de ECG, la mayoría se limitan a detectar arritmias básicas y no aprovechan el potencial que tiene el ECG como biomarcador para extraer características clínicas no evidentes e incluso llegar a identificar a cada paciente, es por esto que resulta importante desarrollar herramientas que puedan procesar los ECG para reducir o eliminar la presencia de ruidos e interferencias y así minimizar los errores manuales mediante algoritmos robustos y adaptables, descubriendo patrones clínicos con modelos profundos y hasta siendo capaz de identificar pacientes usando el ECG como firma biológica

## **1.2. Objetivos**

### **Objetivo General**

Analizar los datos provenientes de electrocardiogramas mediante la aplicación de análisis de componentes principales (PCA) y redes neuronales, con el objetivo de identificar agrupaciones en dicha data.

### **Objetivos específicos**

1. Implementar técnicas avanzadas de filtrado y eliminación de artefactos para obtener señales de electrocardiogramas más limpias.
2. Desarrollar un marco metodológico sólido que combine el análisis de componentes principales (PCA) y redes neuronales perceptrónicas para la identificación de grupos de datos.

3. Evaluar el rendimiento de los modelos propuestos utilizando los datasets de electrocardiogramas de las bases de datos de Physionet.

## Capítulo 2

# Marco Teórico

### 2.1. Antecedentes

La electrocardiografía sigue siendo uno de los bastiones en lo que se basa la cardiología moderna para el diagnóstico de cardiopatías, el electrocardiograma (ECG) suele ser el primer examen que se le realiza a cada paciente cuando se examina a profundidad, ya que es capaz de arrojar información importante del funcionamiento del sistema circulatorio y podría rociar indicios de problemas que pudieran estar pasando en otros sistemas importantes del cuerpo humano. La ciencia siempre ha acompañado la innovación y creación de nuevas tecnologías y la física, como baluarte entre las ciencias naturales, se ha interesado en contribuir en el mejoramiento de estas tecnologías, aportando eficacia y rendimiento desde ángulos no antes vistos. Esto ha conducido a investigar referentes, que permiten ampliar el panorama científico que hay detrás de un ECG. Ahora se presentan una serie de investigaciones que continúan con esta línea de investigación, que consiste en aplicar métodos computacionales a la salud, creciente interés que ha incrementado desde finales del siglo pasado.

#### A Real-Time QRS Detection Algorithm [4]

**Autores:** Pan, Jiapu; Tompkins, Willis J.

**Publicación:** IEEE Transactions on Biomedical Engineering **DOI:** 10.1109/TBME.1985.325532

**Resumen:** Pan y Tompkins presentan el primer algoritmo para detectar en tiempo real el complejo QRS de las señales de ECG. Este algoritmo es capaz de reconocer con fiabilidad los complejos QRS basándose en análisis digitales de la pendiente, la amplitud y la anchura. Además plantan las bases de filtrado que se le debe hacer a una señal de un ECG para reducir las falsas detecciones



causadas por los distintos tipos de interferencias presentes en las señales de ECG, aumentando así la sensibilidad de la detección.

**Metodología:**

1. **Filtro pasa-banda:** Como primer paso, se aplica un filtro pasa-banda, a la señal del ECG, para incrementar la relación señal-ruido, se sugiere un filtro con un ancho de banda de 5-15 Hz para maximizar el QRS y también reducir el ruido producido por el movimiento de los músculos y el desplazamiento de la línea base.
2. **Filtro derivativo:** Se aplica un filtro derivativo para proveer información acerca de la pendiente del QRS.
3. **Cuadratura e integración:** La señal filtrada se eleva al cuadrado para realzar los picos dominantes (QRS) y reducir la posibilidad de reconocer erróneamente una onda T como pico R. A continuación, se aplica un filtro de media móvil para proporcionar información sobre la duración del complejo QRS.

**Hallazgos clave:**

- La serie de filtros aplicados resaltan el contenido frecuencial de la rápida despolarización cardíaca y elimina el ruido de fondo.
- Identifica los QRS en tiempo real con un costo computacional bajo.
- Se reportó que el porcentaje de acierto del QRS fue de 99.3 %

**A Wavelet-Based ECG Delineator: Evaluation on Standard Databases [5]**

**Autores:** Martínez, J. P.; Almeida, R.; Olmos, S.; Rocha, A. P., Laguna, P.

**Publicación:** IEEE Transactions on Biomedical Engineering **DOI:** 10.1109/TBME.2003.821031

**Resumen:** En este artículo, desarrollan y evalúan un robusto sistema de delineación de ECG de una sola derivación basado en la transformada de ondícula (WT). Primeramente, detectan los complejos QRS para luego delimitar cada QRS, detectando e identificando los picos de las ondas individuales, así como también, los inicios y finales de cada complejo QRS. Finalmente, se realiza la determinación de los picos, inicios y finales de las ondas P y T. Posteriormente, se evalúa este algoritmo en varias bases de datos anotadas manualmente, como MIT-BIH Arritmia [6], QT [7], ST-T Europea [8] y otras bases de datos, desarrolladas para propósitos de validación.

**Metodología:**

- Usando la transformada de ondícula discreta, pueden detectar las ondas presentes en un ECG que están compuestas por pendientes y máximos (o mínimos) locales a diferentes escalas, ocurriendo a diferentes instantes de tiempo dentro del ciclo cardíaco.
- Primero, se detecta el complejo QRS para luego detectar e identificar las ondas individuales presentes en este y posteriormente, determinar los límites del complejo QRS.
- Por último se detecta y delimita las ondas T y P respectivamente.

**Hallazgos clave:**

- El algoritmo propuesto por Martínez *et al.* es capaz de identificar y ubicar en función del tiempo las ondas individuales de cada latido.
- Este método es lo suficiente robusto como para permitir la aplicación directa de este sobre señales de ECG sin previo filtrado.
- Obtuvieron un porcentaje de detección superior al 99.86 % para la base de datos Arritmia del MIT-BIH [6] y del 99.88 % para la base de datos QT [7].

**Feature extraction for heartbeat classification using independent component analysis and matching pursuits [9]**

**Autores:**Herrero, G.G.; Gotchev, A.; Christov, I.; Egiazarian, K.

**Publicación:** IEEE DOI: 10.1109/ICASSP.2005.1416111

**Resumen:** En esta investigación, se presenta un método basado en el algoritmo de búsqueda de coincidencias para la extracción de características de tiempo-frecuencia que pueden ser usadas para la clasificación de varios tipos de latidos anormales. Luego de esto, investigan sobre la usabilidad del análisis de componentes independientes (ICA) para extraer características espaciales de grabaciones de electrocardiogramas de varias derivaciones. El rendimiento de estos diferentes conjuntos de características es evaluado usando las grabaciones de la base de datos arritmia del MIT-BIH [6].

**Metodología:**

- Habiendo detectado los latidos anotados en la base de datos y usando las dos derivaciones presentes en dicha base de datos, el bloque de ICA estima para cada latido, una matriz de proyección que minimiza la dependencia estadística entre las dimensiones proyectadas.

- Las componentes de la matriz antes mencionadas, son usadas como características para la etapa de clasificación.
- La extracción de características de tiempo-frecuencia proyecta cada latido en conjuntos diferentes de paquetes de onda que se seleccionan para que coincidan con las estructuras características de los distintos tipos de latidos que se intentan clasificar.
- Por ultimo, las características de tiempo-frecuencia e ICA se clasifican mediante redes neuronales.

**Hallazgos clave:**

- Introducen ICA como extractor de características para el procesamiento de los ECGs.
- El rendimiento del sistema tuvo una precisión superior al 95 % en la clasificación de 5 tipos de latidos anormales.
- La potencia computacional requerida para el sistema propuesto es bastante alta durante el entrenamiento del extractor de características

**Sistema de adquisición multicanal y análisis de la señal electrocardiográfica de alta resolución aplicado a pacientes chagásicos [10]**

**Autores:** Dugarte, N.; Cuadros, J.; Medina, R.; Rojas, R.; Jugo, D. Nuñez, T.

**Publicación:** Conferencia: XI Congreso Internacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas

**Resumen:** En este trabajo, se reporta el desarrollo de un sistema integral para adquisición y posterior análisis de la señal electrocardiográfica de alta resolución (ECGAR) de pacientes chagásicos, en el cual utilizan maquinas de soporte vectorial de mínimos cuadrados (LSSVM) para determinar el inicio del complejo QRS y el final de la onda T, para estimar los intervalos QT y QT corregido y comparan su efectividad usando algoritmos de procesamiento implementados en la aplicación Cardiosoft [11].

**Metodología:**

- Se les realizó un registro ECGAR a 20 pacientes chagásicos, excluyendo pacientes con otras patologías y 20 pacientes de control.

- Utilizan LSSVM para determinar el inicio del complejo QRS y el final de la onda T, entrenadas en base a atributos extraídos de la señal preprocesada y de señales obtenidas mediante descomposiciones con Wavelets.
- Se estiman tanto el intervalo QT y el QT corregidos (QTc) con el uso de las técnicas antes mencionadas.

**Hallazgos clave:**

- Estudian la efectividad de este análisis en el reconocimiento de pacientes chagásicos al procesar 20 pacientes chagásicos y 20 sujetos de control.
- Los resultados muestran diferencias estilísticamente significativas entre ambos grupos.
- Validan los algoritmos de procesamiento implementados utilizando como referencia una aplicación denominada Cardiosoft [11].

**A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification [12].**

**Autores:** Khorrami, H.; Moavenian Majid.

**Publicación:** Expert Systems with Applications

**DOI:**<https://doi.org/10.1016/j.eswa.2010.02.033>

**Resumen:** En este estudio, se sugiere y comparan el uso de transformadas ondículas discretas (CWT), transformadas ondículas discretas (DWT) y transformada discreta del coseno (DCT), que ya están en uso, con el fin de mejorar la capacidad de dos clasificadores de patrones en clasificación de arritmias en ECG. Se utilizan como clasificadores, redes neuronales perceptrónicas y máquinas de soporte vectorial. Las señales ECG usadas son tomadas de la base de datos Arritmia del MIT-BIH que son usadas para clasificar cuatro diferentes tipos de arritmias.

**Metodología:**

- Las técnicas de extracción de características CWT, DWT y DCT son aplicadas separadamente a la base de datos antes de su clasificación.
- Utilizan estas características extraídas para clasificar los tipos de arritmia de cada señal ECG, usando redes neuronales perceptrónicas y máquinas de soporte vectorial.
- Comparan el rendimiento del entrenamiento, el rendimiento de las pruebas y el tiempo de entrenamiento de cada clasificador.

**Hallazgos clave:**

- Los resultados muestran que el mejor método de extracción de características dependerá del valor sustancial que se considere para el tiempo de entrenamiento y rendimiento en el entrenamiento y las pruebas.
- El rendimiento de las pruebas usando sólo la derivación II de la señal ECG de la base de datos usada muestra superioridad sobre las máquinas de soporte vectorial.
- El modelo que utiliza como extractor de características las DWT y una red neuronal perceptrónica como clasificador es el que tiene mejor rendimiento tanto en entrenamiento y pruebas, como en tiempo de entrenamiento.

**Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network [13].**

**Autores:** Ceylan, R.; Özbay, Y.

**Publicación:** Expert Systems with Applications. DOI:<https://doi.org/10.1016/j.eswa.2006.05.014>

**Resumen:** Este artículo presenta un estudio comparativo de la eficacia en la clasificación de señales ECG usando 4 tipos de estructuras que incluyen poderosas técnicas de reducción de dimensionalidad, como lo son PCA, WT y agrupamiento c-medios difuso (FCM); también usan redes neuronales perceptrónicas (NN) para clasificar tipos de arritmias, usando [6], estas son entrenadas para clasificar 10 diferentes tipos de arritmias.

**Metodología:**

- Utilizando PCA y WT para reducción de características en adición de FCM; y NN para clasificar los tipos de arritmia en las señales de ECG, comparan el desempeño de 4 tipos de estructuras propuestas, FCM-NN, PCA-NN, FCM-PCA-NN y WT-NN.
- Prueban cada uno de los modelos propuestos con las señales ECG tomadas de [6].

**Hallazgos clave:**

- Proponen como nuevo método de clasificación, la estructura FCM-PCA-NN.

- Las pruebas realizadas sugieren que la estructura FCM-PCA-NN puede generalizar mejor que PCA-NN y mucho más rápido de las demás estructuras propuestas.
- Si bien la estructura FCM-PCA-NN presenta el menor error ( $5,05 \times 10^{-9} \%$ ) al clasificar los tipos de arritmias, es importante señalar que el modelo PCA-NN no se queda atrás en desempeño, teniendo un error de ( $4,98 \times 10^{-9} \%$ ). Esta mínima diferencia sugiere que ambas arquitecturas, podrían considerarse alternativas viables.

### **Baseline wander removal of ECG signals using Hilbert vibration decomposition [15].**

**Autores:** Sharma, H.; Sharma, K. K.

**Publicación:** Electronics Letters. **DOI:** 10.1016/j.eswa.2010.02.033

**Resumen:** En este estudio, se propone una técnica para remover el desplazamiento de la línea base (BW, por sus siglas en inglés) de las señales ECG utilizando la descomposición vibracional de Hilbert (HVD). Se plantea que la primera componente (la componente de más alta energía) usando HVD en la señal ECG corresponde al BW de la señal. La técnica propuesta es comparada con otro método basado en descomposición modal empírica (EMD) y matemática morfológica en términos de criterios de correlación y relación señal-ruido.

#### **Metodología:**

- Para lograr remover el desplazamiento de la línea base, primero se descompone la señal ECG usando HVD en sus diversas componentes, siendo la primera de estas, la de más alta energía y por ende, la correspondiente al BW.
- Luego corrigen la señal ECG, sustrayendo la primera componentes de la descomposición vibracional de Hilbert de la señal ECG original.
- Para evaluar el desempeño de la técnica propuesta, se usan las señales electrocardiográficas de la base de datos arritmia [6] y le introducen un BW artificial obtenido de un filtrado paso bajo de señales aleatoriamente generadas.

#### **Hallazgos clave:**

- La técnica propuesta en este artículo, tiene un mejor desempeño que la aproximación basada en EMD para remover el BW.

- También se observa que la HVD es computacionalmente eficiente.
- El método propuesto es capaz de desempeñarse mejor bajo condiciones de severa distorsión de la línea base sin afectar la morfología real del ECG.

## **Nombre del artículo**

**Autores:**

**Publicación:**

**DOI:**

**Resumen:**

**Metodología:**

**Hallazgos clave:**

## **2.2. Conceptos básicos**

### **El corazón**

#### **Introducción al Electrocardiograma (ECG)**

La humanidad siempre, en su continuo deseo de aprender y entender más, ha querido desentrañar los secretos del cuerpo humano entendiendo su funcionamiento interno. Al principio con técnicas invasivas acordes a la tecnología disponible a la época, pero evolucionando continuamente, creando así exámenes cada vez menos invasivos, con el fin de mejorar el diagnóstico, siendo más preciso y oportuno. Entre esos exámenes se destaca el electrocardiograma (ECG) el cual es una representación visual de la actividad eléctrica del corazón en función del tiempo, que se obtiene desde la superficie corporal, con un electrocardiógrafo, este es el instrumento principal de la electrofisiología cardíaca y tiene una función relevante en el cribado y diagnóstico de las enfermedades cardiovasculares, alteraciones metabólicas y demás utilidades.

El primer «electrograma» humano fue publicado en 1887 por el fisiólogo británico Augustus Desiré Waller, de la St. Mary's Medical School de Londres. Utilizó un electrómetro capilar de Lipmann con electrodos aplicados a la espalda y el tórax del sujeto. Demostró que la contracción ventricular precedía a la actividad eléctrica. En su primer informe sobre un registro de la electricidad cardíaca realizado en la superficie corporal, Waller utilizó el término «cardiógrafo».

Einthoven empezó a experimentar con el potencial del capilar para captar corrientes eléctricas diminutas. En 1895 demostró cinco deflexiones que denominó ABCDE en 1895. Creó un ajuste matemático para tener en cuenta la inercia del sistema capilar, lo que produjo las curvas de corriente que vemos hoy en día. Siguiendo la tradición matemática establecida por Descartes, utilizó la parte terminal de la serie parte terminal de la serie alfabética (PQRST) para denominar estas derivaciones.

El pionero de la electrocardiografía, Waller dijo a finales de 1911: «No creo que la electrocardiografía vaya a tener un uso extensivo en los hospitales. A lo sumo puede tener un uso raro y ocasional para proporcionar un registro de alguna anomalía de la actividad cardíaca». Sin embargo, diez años de los estudios clínicos de Einthoven con los galvanómetros de cuerda transformaron este curioso fenómeno fisiológico en un dispositivo de registro clínico indispensable. Las asociaciones de la inversión de la onda T con la angina de pecho y la arteriosclerosis. en 1910, junto con otras arritmias, como el bigeminismo, bloqueo cardíaco completo, hipertrofia ventricular derecha e izquierda, fibrilación y aleteo auricular y ejemplos de diversas cardiopatías. Con su nueva técnica, estandarizó los trazados y formuló el concepto de «triángulo de Einthoven» relacionando matemáticamente las 3 derivaciones (Derivación III = Derivación II - Derivación I). En 1924, el «Padre de la electrocardiografía» recibió el Premio Nobel de Medicina [21].

En 1957, el médico estadounidense Norman Jefferis Holter inventó el ECG dinámico (DCG), a menudo conocido como Holter, en uno de los primeros intentos de combinar monitorización clínica y movilidad. Creó una mochila que pesaba unos 38 kg y tenía un dispositivo que podía registrar la actividad cardíaca del participante. Este portátil permite la monitorización continua de actividad eléctrica del sistema cardiovascular durante 24 horas, lo que ayuda a estudiar las arritmias y a localizar el lugar de la isquemia miocárdica. Reconociendo los beneficios potenciales de un dispositivo de monitorización de este tipo, Holter consiguió convertir su idea en una valiosa herramienta de diagnóstico reduciendo el tamaño y el peso a 1 kg con ayuda de Del Mar Avionics, un conocido fabricante de equipos aeronáuticos [21].

Durante las tres primeras décadas del siglo 20, el ECG de tres derivaciones periféricas fue largamente usado, especialmente luego de mejoras que lo hicieron más portable. A pesar de que el ECG de tres derivaciones era una manera fiable de evaluar arritmias, pronto se reconoció que el corazón incluía «zonas silenciosas» en las que un infarto de miocardio podría no ser detectado. En



1942, Emanuel Goldberger construyó las derivaciones precordiales (unipolares) usando el promedio de las diferencias de potencial de las tres derivaciones periféricas como terminal de referencia, que inicialmente fue creado por Frank N. Wilson, al cual se le conoce como terminal central de Wilson, que ahora se denominan como derivaciones precordiales (V1-V6), donde en 1938, la Asociación Americana del Corazón (AHA) y la Sociedad Cardíaca de Gran Bretaña recomendaron la estandarización del posicionamiento de los electrodos en el pecho para dichas derivaciones. También Goldberger propuso una manera de obtener lo que ahora se llaman derivaciones aumentadas, conocidas por las siglas a-VL, a-VR, y a-VF. 8 años después la AHA recomendó la estandarización del ECG de 12 derivaciones. [21]

En la era digital, la tecnología del silicio y los circuitos impresos han hecho posible la miniaturización electrónicos. Desde hace algún tiempo, la tecnología ha ganado popularidad en el campo de la medicina y la necesidad de los clientes de controlar su salud ha sido el principal motor. La influencia de los "vestibles"(wearables) ha hecho inevitable la continua investigación y desarrollo de nuevas funciones que pueden evaluar y transmitir datos biométricos en tiempo real.

El corazón consta de cuatro partes, dos aurículas y dos ventrículos, el ECG registra los impulsos eléctricos que estimulan estas partes y producen su contracción. Las células cardíacas en reposo se encuentran cargadas o polarizadas; pero la estimulación eléctrica las "despolariza", y se contraen.

Como se observa en la figura 2.1 el ECG consta de varias ondas representativas de cada etapa de un latido cardíaco, estas son:

- **Onda P:** registra la despolarización auricular.
- **Complejo QRS:** Es la despolarización ventricular.
- **Onda T:** representa la repolarización ventricular.

## 2.3. Conceptos específicos

### Análisis de Componentes Principales (PCA).

Los grandes conjuntos de datos están cada vez más extendidos en muchas disciplinas. Para interpretarlos, se necesitan métodos que reduzcan drástica-

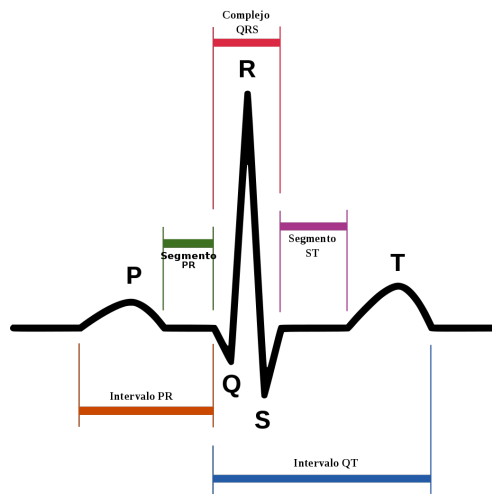


Figura 2.1: ECG del corazón con ritmo sinusal normal

mente su dimensionalidad de forma interpretable, de modo que se conserve la mayor parte de la información contenida en los datos. Se han desarrollado muchas técnicas con este fin, pero el análisis de componentes principales (PCA) es una de las más antiguas y utilizadas. Su idea es sencilla: reducir la dimensionalidad de un conjunto de datos conservando la mayor cantidad posible de «variabilidad» (es decir, de información estadística).

Esto significa que «preservar tanta variabilidad como sea posible» se traduce en encontrar nuevas variables que sean funciones lineales de las del conjunto de datos original, que maximicen sucesivamente la varianza y que no estén correlacionadas entre sí. Encontrar esas nuevas variables, los componentes principales (PC), se reduce a resolver un problema de auto-valores y auto-vectores.

Hasta que no se generalizó el uso de ordenadores electrónicos, que fue posible utilizarlo con conjuntos de datos que no fueran trivialmente pequeños. Desde entonces, su uso se ha multiplicado y se han desarrollado numerosas variantes en muchas disciplinas diferentes. Desde entonces, su uso se ha multiplicado y se han desarrollado numerosas variantes en muchas disciplinas diferentes.

La definición formal de PCA, en un contexto estándar, junto con una derivación que muestra que puede obtenerse como la solución a un problema de auto-valores y auto-vectores o, alternativamente, a partir de la descomposición del valor singular (SVD) de la matriz (centrada) de datos. El PCA puede

basarse en la matriz de covarianzas o en la matriz de correlaciones. Se discutirá la elección entre estos análisis. En ambos casos, las nuevas variables (las PC) dependen del conjunto de datos, en lugar de ser funciones de base predefinidas, por lo que son adaptativas en sentido amplio. Los principales usos del PCA son descriptivos y no inferenciales.[22]

## **Redes Neuronales**

Una red neuronal artificial es un grupo de neuronas artificiales interconectadas que interactúan entre sí de forma concertada. Se trata de hecho un procesador distribuido masivamente paralelo que tiene una propensión natural a almacenar el conocimiento experiencial y ponerlo disponible para su uso. Se parece al cerebro humano en dos aspectos: La red adquiere los conocimientos mediante un proceso de aprendizaje. El conocimiento lo adquiere la red mediante un proceso de aprendizaje, y para almacenarlo se utilizan las intensidades de conexión interneuronal, denominadas pesos. Los modelos de redes neuronales artificiales pueden utilizarse como método alternativo en análisis y predicciones. Funcionan como un modelo de «caja negra», que no requiere información detallada sobre el sistema. Imitan en cierto modo el proceso de aprendizaje de un cerebro humano porque aprenden la relación entre los parámetros de entrada, las variables controladas y no controladas estudiando datos registrados previamente. En este sentido, funcionan de forma similar a la regresión no lineal, pero son mucho más potentes que el análisis de regresión. Las redes neuronales son capaces de manejar sistemas grandes y complejos con muchos parámetros interrelacionados. Parece que simplemente ignoran el exceso de parámetros de entrada que tienen una importancia mínima y se concentran en los más importantes.

mas tema

## **K-means**

## Capítulo 3

# Marco experimental

...

## Capítulo 4

# Resultados

...

Parte I

Conclusiones

## Capítulo 5

## Conclusiones

...

# Bibliografía

- [1] L. Zhang, M. Karimzadeh, M. Welch, C. McIntosh, and B. Wang, “Chapter 7 - analytics methods and tools for integration of biomedical data in medicine,” in *Artificial Intelligence in Medicine* (L. Xing, M. L. Giger, and J. K. Min, eds.), pp. 113–129, Academic Press, 2021.
- [2] M. National Library of Medicine, “Electrocardiogram,” 2020.
- [3] O. Organizacion Mundial de la Salud, “Enfermedades cardiovasculares (cvds),” 2021.
- [4] J. Pan and W. J. Tompkins, “A real-time qrs detection algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.
- [5] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, “A wavelet-based ecg delineator: Evaluation on standard databases,” *IEEE transactions on bio-medical engineering*, vol. 51, pp. 570–81, 05 2004.
- [6] G. Moody and R. Mark, “The impact of the mit-bih arrhythmia database,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
- [7]
- [8] A. TADDEI, G. DISTANTE, M. EMDIN, P. PISANI, G. B. MOODY, C. ZEELLENBERG, and C. MARCHESI, “The european st-t database: standard for evaluating systems for the analysis of st-t changes in ambulatory electrocardiography,” *European Heart Journal*, vol. 13, pp. 1164–1172, 09 1992.
- [9] G. Herrero, A. Gotchev, I. Christov, and K. Egiazarian, “Feature extraction for heartbeat classification using independent component analysis



and matching pursuits,” in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 4, pp. iv/725–iv/728 Vol. 4, 2005.

- [10] N. Dugarte Jerez, R. Medina Molina, and R. Rojas Sulbarán, “Sistema para la adquisición de la señal electrocardiográfica de alta resolución,” *Universidad, Ciencia y Tecnología*, vol. 15, pp. 206 – 215, 12 2011.
- [11] Cardiosoft, “hfqrs-ecg - high frequency qrs electrocardiograph.”
- [12] H. Khorrami and M. Moavenian, “A comparative study of dwt, cwt and dct transformations in ecg arrhythmias classification,” *Expert Syst. Appl.*, vol. 37, pp. 5751–5757, 08 2010.
- [13] R. Ceylan and Y. Özbay, “Comparison of fcm, pca and wt techniques for classification ecg arrhythmias using artificial neural network,” *Expert Systems with Applications*, vol. 33, no. 2, pp. 286–295, 2007.
- [14] P. Laguna, R. Jané, S. Olmos, N. Thakor, H. Rix, and P. Caminal, “Adaptive estimation of qrs complex wave features of ecg signal by the hermite model,” *Medical and biological, engineering and computing*, vol. 34, pp. 58–68, 02 1996.
- [15] H. Sharma and K. Sharma, “Baseline wander removal of ecg signals using hilbert vibration decomposition,” *Electronics Letters*, vol. 51, pp. 447–449, 03 2015.
- [16] M. Song, J. Lee, S. Cho, K.-J. Lee, and S. Yoo, “Support vector machine based arrhythmia classification using reduced features,” *International Journal of Control, Automation and Systems*, vol. 3, 12 2005.
- [17] S.-N. Yu and C. Hsiang, “Electrocardiogram beat classification based wavelet and probabilistic neural network,” *Pattern Recognition Letters*, vol. 28, pp. 1142–1150, 07 2007.
- [18] C. Ye, M. Coimbra, and B. Kumar, “Arrhythmia detection and classification using morphological and dynamic features of ecg signals,” *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 2010, pp. 1918–21, 08 2010.

- [19] C.-J. Zhang, Yuan-Lu, F.-Q. Tang, H.-P. Cai, Y.-F. Qian, and Chao-Wang, “Heart failure classification using deep learning to extract spatiotemporal features from ecg,” *BMC Medical Informatics and Decision Making*, vol. 24, 01 2024.
- [20] V. Astudillo, D. Luna, and J. Muñoz Chaves, “Clasificación de arritmias cardiacas usando redes neuronales convolucionales en muestras de ecg,” *Revista EIA*, vol. 21, 01 2024.
- [21] R. Vincent, “From a laboratory to the wearables: a review on history and evolution of electrocardiogram,” *Iberoamerican Journal of Medicine*, vol. 4, no. 4, pp. 248–255, 2022.
- [22] I. T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments.,” *Philosophical transaction. Series A, Mathematical, physical, and engineering sciences*, 2016.