

Clasificación de señales electrocardiográficas con python

García Espinosa David Alexis* Díaz Cuevas Bruno**
De la Torre Sánchez Mosses Arath***

* Universidad Nacional Autónoma México, Facultad de Ciencias
2019-2 (e-mail: davale@ciencias.unam.mx).

** Universidad Nacional Autónoma México, Facultad de Ciencias
2019-2 (e-mail: brunodiaz@ciencias.unam.mx).

*** Universidad Nacional Autónoma México, Facultad de Ciencias
2019-2 (e-mail: mosses_arath@ciencias.unam.mx).

Abstract: A partir de distintos métodos estadísticos se estudio una señal de un pulso cardiaco pudiendo clasificar la anomalía que presenta o si este es normal, los resultados obtenidos no son completamente precisos sin embargo se logró una clasificación.

Keywords: Electrocardiograma, Potenciales de acción, Corazón, Fisiología.

1. MOTIVACIÓN

Actualmente la tecnología permea cada una de las actividades cotidianas de la población, la creciente complejidad de las problemáticas que aquejan a la sociedad demandan una alta interdisciplinaridad de áreas, por lo que hemos decidido buscar un nicho de oportunidad tecnológico al intentar brindar al médico una herramienta que auxilie en su diagnóstico, en lo particular en el diagnóstico temprano de anomalías en el ciclo cardiaco, que de no atenderse pueden derivar en enfermedades mortales donde en México el 54 cardiopatías.

2. INTRODUCCIÓN

Anatómicamente hablando el corazón se puede dividir en cuatro cavidades, dos atrios, localizados en la parte superior del corazón y dos ventrículos en la parte inferior, citológicamente hablando el corazón posee 4 tipos de células que con el paso de un impulso eléctrico en distintos momentos del ciclo cardiaco producen una señal denominada electrocardiográfica la cual se puede subdividir en: Onda P: Corresponde a la creación del impulso eléctrico en el nodo sino atrial provocando la sístole auricular es decir una despolarización de los atrios.

Intervalo P-Q: Representa el tiempo de conducción del potencial desde las aurículas a los ventrículos.

Complejo QRS: Aquí es cuando ocurre la despolarización de los ventrículos (evento eléctrico) y consecuentemente la sístole ventricular (evento mecánico).

Segmento ST: Es el segmento comprendido entre el fin del complejo QRS y el inicio de la onda.

Onda T. Repolarización de los ventrículos (evento eléctrico) y generación de la diástole ventricular (evento mecánico). Existen parámetros de duración y amplitud asociados a

un ciclo cardiaco sano, sin embargo, existen alteraciones en este, particularmente nos centraremos en:

Latido supraventricular prematuro: sucede cuando la contracción atrial es creada por un impulso eléctrico proveniente de un foco ectópico, presentando disminuciones en la onda P y T, de aparecer de forma continua pueden degradarse en Complejos Atriales Prematuros (PAC).

Contracción ventricular prematura: sucede cuando el impulso eléctrico lo producen las fibras de Purkinje por lo que presentan un pico anormal distorsionado en el complejo QRS, de presentarse de forma regular son un signo importante de una cardiopatía no diagnosticada.

Fusión ventricular: sucede cuando existen 2 focos que producen la contracción eléctrica por lo que se aprecian picos sobresalientes en la onda P con respecto al resto del ciclo cardiaco.

El sistema cardiovascular como su nombre lo indica consiste principalmente en el corazón y los vasos sanguíneos que lo conforman como lo son arterias, venas y capilares, esté es un sistema complejo y se podría llegar a decir que es en parte autónomo puesto que una de sus funciones principales es poder bombear la sangre a todo el cuerpo, para lograr esto es necesario una contracción muscular conocida como sístole y una relajación muscular diástole estas 2 actividades pueden ser los mejores marcadores para padecimientos en el ritmo y funcionamiento por lo que su estudio individual es de suma importancia esta actividad muscular se registra gracias a el electromiograma el cual es un instrumento capaz de registrar los potenciales de acción asociados a cada etapa de la diástole y la sístole.

La conformación general del corazón se puede dividir en cuatro cavidades "simétricas" llamadas atrios(superiores) y ventrículos(inferiores) en cada bombeo debe haber un flujo de una cavidad a otra para que la sangre desoxigenada sea enviada a los pulmones se oxigene y regrese al corazón donde será bombeada nuevamente pero esta vez será envi-

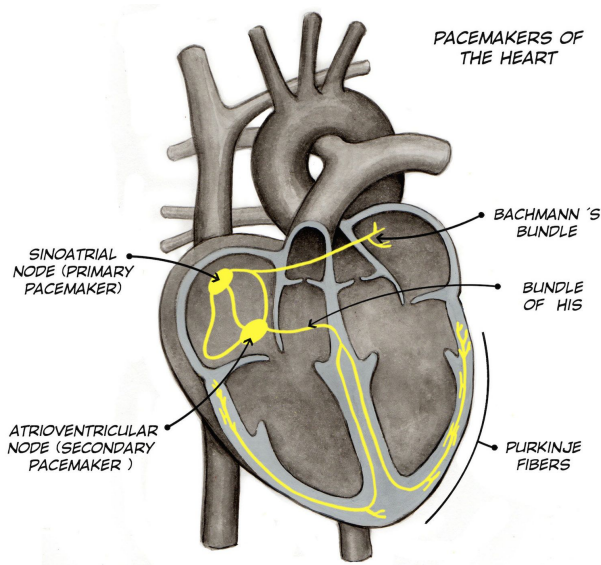


Fig. 1. Circuito eléctrico del corazón

ada a todo el circuito sistémico del cuerpo humano.

A continuación en el Figura 1 se muestra como es el circuito de la conducción eléctrica del corazón de donde la ruta de propagación de la señal comienza en el nodo sinusal bifurcándose en las ramificaciones de Bachman y al nodo auriculoventricular de ahí se propagará por el haz de His el cual enviará la señal a las ramificaciones posteriores y de Purkinje donde finalizará la propagación atenuándose y marcando el inicio de la relajación desde el haz de His y hasta el nodo auriculoventricular sección del ciclo correspondiente a la diástole ventricular.

2.1 Potenciales de acción

Para lograr hacer algún tipo de movimiento el músculo debe ser activado de alguna forma esto se logrará gracias a los potenciales de acción que se generan gracias a canales dependientes de voltaje encontrados de manera transversal en los túbulos T del músculo ($Ca_v1.1$) estos se encuentran enlazados mecánicamente por un receptor rianodina del retículo sarcoplásmico esto hace que la contracción muscular no sea totalmente dependiente de Ca finalmente la estimulación nerviosa es propagada por canales de Na como potenciales de acción a lo largo del tubo-T haciendo uso de las proteínas motoras y estructurales de la miosina, titina, filamentos de actina el movimiento de estos a lo largo del músculo de una manera organizada y estabalecida permiten ya sea contracción o relajación.

Cada uno de los pasos mencionados anteriormente son registrados gracias a la caracterización que logrará hacer el instrumento esto se debe a la distribución de las células y a la coordinación temporal que presentan ya que por ejemplo las células de marcapasos y de miocitos encontradas en distintas zonas del corazón, presentan sus potenciales de acción a distintos tiempos y la duración de los mismos también varía de acuerdo al tipo de célula, a pesar de ser un buen instrumento para identificar secciones del ciclo cardíaco su resolución no es tan potente como para distinguir entre una y otra célula por lo que el registro termina siendo una suma de señales asociadas e un conjunto de potenciales de

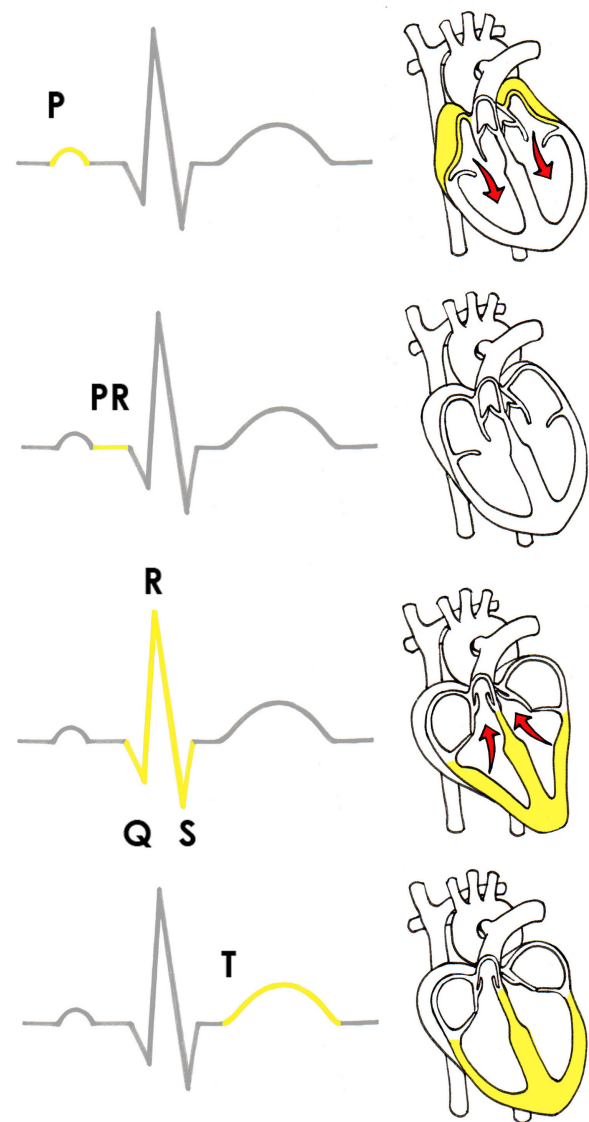


Fig. 2. Asociación de las ondas P,Q,R,S,T

acción individuales como se muestra en la figura 2 estos potenciales son clasificados en ondas P,Q,R,S,T,U.

Las cuales se les asigna uno de los pasos del ciclo cardiovascular siendo la onda P para identificar la sistole ventricular seguido del complejo Q,R,S donde se muestra la sistole ventricular y diástole atrial y finalmente la onda T y parte de la U donde se muestra la diástole ventricular.

2.2 Tipos de Derivaciones

Para poder obtener un registro del corazón por medio de un electrocardiograma es necesario escoger la forma correcta en la que se medirá el registro ya que al ser una medición indirecta debemos asegurarnos que lo que estamos midiendo realmente sea el corazón para tener una medición correcta existe diversas configuraciones para colocar electrodos esto consisten principalmente en 3 electrodos de superficie donde 2 estarán colocados como receptores de señal y el otro será el punto referencia conocido como tierra este se colocará comúnmente en un

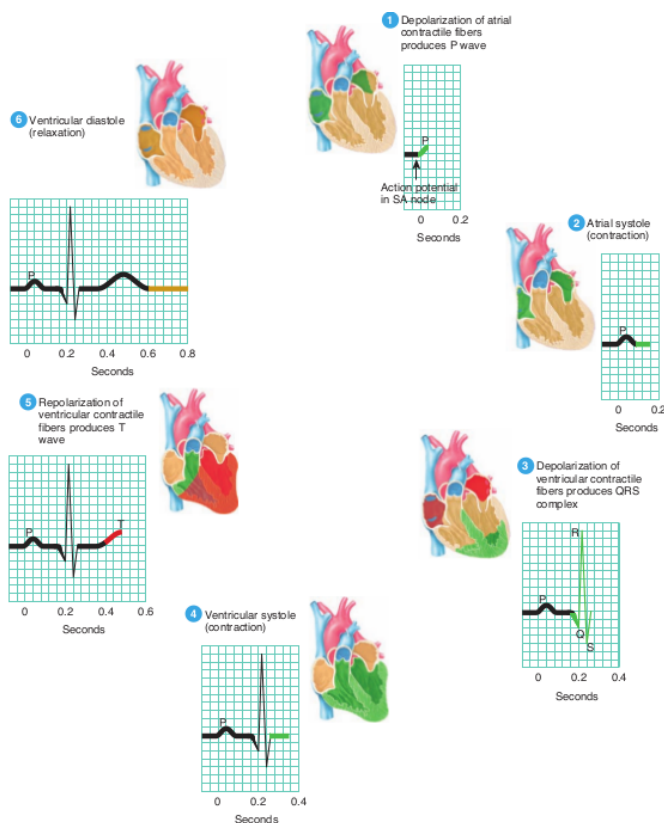


Fig. 3. Depolarización y repolarización atrioventricular

área donde el movimiento muscular sea mínimo ya que de no ser un punto estático podría haber ruido innecesario en la medición.

El electrocardiograma consta de 12 derivaciones, que son el resultado de la exploración indirecta del corazón desde distintos planos, tal como si una persona observara una montaña desde su base y su cima. De ese modo obtendría distintas panorámicas del accidente geográfico observado, pero la montaña no cambiaría; lo que cambia es el punto de vista del explorador, al situarse en lugares diferentes.

Las primeras derivaciones, que datan de principios del siglo XX aunque en realidad no fue entonces, sino mucho antes cuando se exploró la potencialidad del corazón como generador eléctrico fueron descritas por Einthoven, y desde entonces se les llama, derivaciones estándares o clásicas. En lo fundamental se basan en una concepción de bipolaridad (polo positivo menos polo negativo), y debido a ese hecho se les llama también derivaciones bipolares. Con posterioridad, surgieron las derivaciones unipolares de miembros, nacidas de los potenciales proyectados sobre ambos brazos y la pierna izquierda. Por cuestiones del idioma en que primero fueron estudiadas, se les conoce con los nombres de VR, VL y VF.

Más adelante las explicaremos en detalle. Por último, aparecieron las 6 derivaciones precordiales, también unipolares, que completan la exploración del corazón desde los planos anteriores, laterales y posteriores, y cuya utilidad en el diagnóstico "topográfico" es decisiva.

En términos concretos, Einthoven pensó que, siendo el corazón un generador de corriente y el cuerpo humano

un buen conductor, podría construirse imaginariamente un triángulo, formado por las raíces de los miembros, sobre cuyos lados se proyectarían las fuerzas eléctricas emanadas del músculo cardíaco. Dado que el corazón se inclina dentro del pecho hacia la izquierda, y como los brazos y piernas son prolongaciones de sus respectivas raíces, en la práctica empleamos los miembros superiores y el inferior izquierdo para construir el triángulo.

Esa región basal se proyecta sobre el brazo derecho, por lo que dicho brazo constituye el polo negativo de las derivaciones bipolares. Recordemos que la onda de excitación marcha de base a punta y, al aproximarse al brazo izquierdo y la pierna izquierda, los convierte en polos positivos. Teniendo ya constituidos los 2 polos del dipolo, las 3 derivaciones de Einthoven Figura 4

D1: Brazo izquierdo menos brazo derecho.

D2: Pierna izquierda menos brazo derecho.

D3: Pierna izquierda menos brazo izquierdo.

Con esta nomenclatura las derivaciones estándares quedan integradas como sigue:

D1 es igual a VL menos VR (brazo izquierdo menos brazo derecho).

D2 es igual a VF menos VR (pierna izquierda menos brazo derecho).

D3 es igual a VF menos VL (pierna izquierda menos brazo izquierdo).

Con esta nueva nominación vamos a desarrollar la ley de Einthoven. Si D2 es igual a D1 más D3, es lo mismo que si VF menos VR es igual a VL menos VR más VF menos VL. Suprimiendo los 2 puntos VL, uno positivo con el otro negativo, la ecuación queda:

$$D1 = VL - VR \quad (1)$$

$$D2 = VF - VR \quad (2)$$

$$D3 = VF - VL \quad (3)$$

Donde obtenemos:

$$D1 + D3 = VF - VR = D2 \quad (4)$$

Cada lado del triángulo de Einthoven está formado por 2 mitades: una negativa y otra positiva; ambas mitades están divididas en milímetros, que es la magnitud que emplearemos como unidad de medida en las ondas del electrocardiograma. Sin embargo para nuestros casos nos guiaremos en un caso particular de la derivación de Einthoven que se muestra en la figura 5

3. OBJETIVOS

Programar un algoritmo que utilizando una base de datos sea capaz de distinguir entre ECG's correspondientes a distintos latidos asociados a una persona sana o con alguna anomalía cardíaca

Particulares:

1. Realizar una función que sea capaz de identificar las ondas P QRS, T

2. Hacer un Ajuste polinomial de grado n

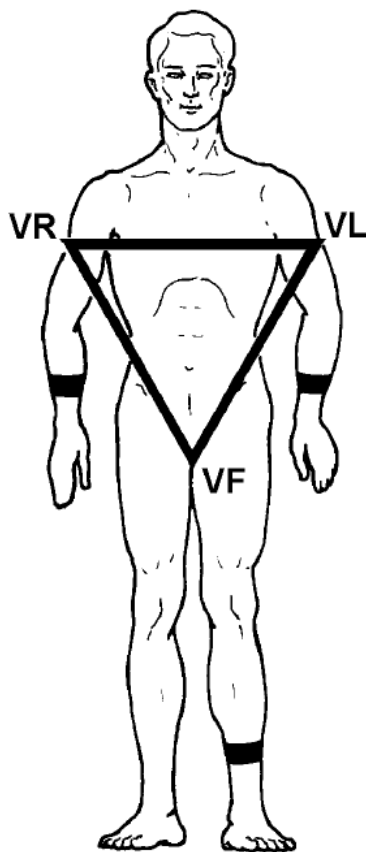


Fig. 4. Derivación del corazón

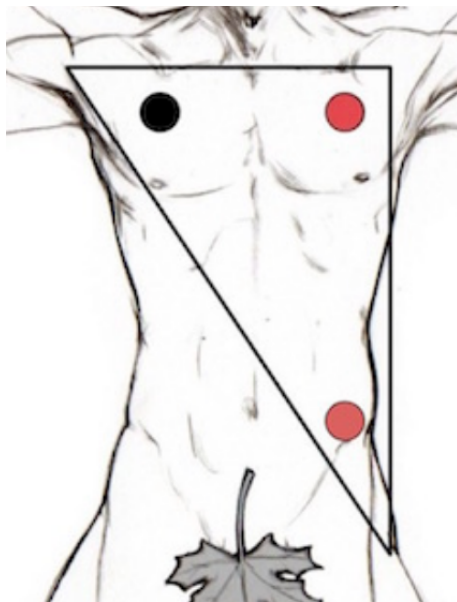


Fig. 5. Derivación del corazón utilizada comúnmente

3.Utilización de distintos métodos de sklearn

4.Realización de un algoritmo capaz de hacer Regresión Logística Multiclase

4. METODOLOGÍA Y DESARROLLO

Metodología

Para entrenar al algoritmo de clasificación se utilizó el archivo *mitbih.train.csv* y para probar su desempeño se utilizó *claseTest.csv* ambos encontrados en la base de datos que se adjunta en la Bibliografía. Se eligieron distintas variables para utilizar en la regresión logística sin embargo como se explica en los notebooks que se adjuntan en el github estos no funcionaron de manera eficiente.

Ya que los ECGs que contienen las bases de datos comienzan con el complejo QRS y terminan en el siguiente. Se creó una función que pretende identificar a los máximos de estas dos ondas en cada uno de los ECGs y regresa los valores anteriores al segundo QRS de cada uno. Posteriormente a cada señal le ajustó un polinomio de grado 15 con la finalidad de usar a los coeficientes de este como variables en el algoritmo de clasificación.

4.1 Análisis

Después de la construcción de tratamientos sobre los ECG explicados detalladamente en los notebooks hicimos uso de métodos estadísticos como la regresión logística y el ajuste polinomial para obtener información completa del ECG para su correcta clasificación de donde obtuvimos los siguientes resultados de puntuación:

Datos crudos sin modificar: puntuación: 0.98

Datos Cortados: puntuación:0.90

Datos de Betas: puntuación: 0.06

por lo tanto pudimos detectar y clasificar pulsos normales y algunos pulsos anormales sin embargo al tener puntos demasiado dispersos hay casos donde la clasificación es incorrecta, sugerimos revisar los notebooks mencionados.

5. CONCLUSION

Debido a como se presenta la información y la manera en que la obtenemos podemos asegurar tener un diagnóstico confiable por esta razón es importante perfeccionar los instrumentos de medición que se utilizan en la rama médica como pudimos registrar la señal de un corazón gracias a un electrocardiograma podemos obtener bastante información acerca de la fisiología de cada uno de los sujetos de prueba viendo cada una de las ondas P,Q,R,S,T,U que presentan sin embargo eso queda más allá de los propósitos iniciales sin embargo el analizar la frecuencia cardíaca nos da una buena aproximación a la salud y condición física de los sujetos.

Al usar como variables todos los datos de la señal para la función de regresión Logística Multiclase se obtuvo un mejor resultado que al utilizar como variables a los coeficientes de los polinomios que mejor se les ajustaron a la señal.

6. TRABAJO A FUTURO

Se pretende buscar más variables para el algoritmo de clasificación como la amplitud y el tiempo en que ocurren las ondas *P*, *QRS* y *T* el número de ondas encontradas.

También sería conveniente intentar variar el grado del polinomio que se le ajusta al polinomio para determinar si hay alguno que mejores la clasificación resultados.

7. BIBLIOGRAFÍA

Tortora, G. and Derrickson, B. (2013). Principios de anatomía y fisiología, 13a. edición. Buenos Aires [etc.]: Editorial Medica Panamericana.

Sepulveda Saavedra, J. (2012). Texto atlas de histología. Mexico, D.F.: McGraw-Hill Interamericana.

Stanfield, Cindy L.; Germann, William J. (2008), Principles of Human Physiology (3rd ed.), Pearson International Edition, p. 378,

Ataklte, F; Erqou, S; Laukkanen, J; Kaptoge, S (15 October 2013). "Meta-analysis of ventricular premature complexes and their relation to cardiac mortality in general populations". The American Journal of Cardiology. 112 (8): 1263–70.

Base de datos utilizada: (<https://www.kaggle.com/shayanfazeli/heartbeat>).

Repositorio Github donde se encuentran los programas:

<https://github.com/Ironsss/ProyectoFAlgoritmos20192.git>