
Predicción de Riesgo Cardiovascular utilizando
imágenes de electrocardiogramas
Prediction of Cardiovascular Risk Using
Electrocardiogram Images



Trabajo de Fin de Grado
Curso 2024–2025

Autor

Noelia Barranco Godoy

Directores

Belén Díaz Agudo

Juan A. Recio García

Doble Grado en Ingeniería Inforomática y Matemáticas

Facultad de Informática

Universidad Complutense de Madrid

Predicción de Riesgo Cardiovascular
utilizando imágenes de
electrocardiogramas
Prediction of Cardiovascular Risk Using
Electrocardiogram Images

Trabajo de Fin de Grado en Ingeniería Informática

Autor

Noelia Barranco Godoy

Director

Belén Díaz Agudo

Juan A. Recio García

Convocatoria: *Febrero 2025*

Doble Grado en Ingeniería Informática y Matemáticas

Facultad de Informática

Universidad Complutense de Madrid

28 de noviembre de 2024

Dedicatoria

{**TODO TODO TODO:** Hacer dedicatoria}

Agradecimientos

{**TODO TODO TODO:** Hacer agradecimientos}

Resumen

Predicción de Riesgo Cardiovascular utilizando imágenes de electrocardiogramas

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Palabras clave

{**TODO TODO TODO:** Máximo 10 palabras clave separadas por comas}

Abstract

Prediction of Cardiovascular Risk Using Electrocardiogram Images

{**TODO TODO TODO:** An abstract in English, half a page long, including the title in English. Below, a list with no more than 10 keywords.}

Keywords

{**TODO TODO TODO:** 10 keywords max., separated by commas.}

Capítulo 1

Introducción

RESUMEN: En este capítulo pretendemos introducir los objetivos de este trabajo.

1.1. Motivación

El auge de la inteligencia artificial ha hecho posible crear aplicaciones que hace unos años nos parecían imposibles. Como no podía ser de otro modo, uno de los campos que ha conseguido avances gracias a esto ha sido el de la medicina, ya que los avances en este campo permiten mejorar la calidad de vida de todo el mundo.

Una de las aplicaciones médicas de la inteligencia artificial que se han trabajado es el procesamiento automatizado de mediciones, creando algoritmos que permitan detectar si estas son normales o no, y quitar carga de trabajo a los médicos.

Este trabajo se centrará en estudiar las posibilidades de un algoritmo que, a partir de las mediciones que se toman en un electrocardiograma (ECG), detectar posibles anomalías.

Además de todo esto, la comunidad científica lleva un tiempo dándole importancia a la explicabilidad de los algoritmos, por lo que en este trabajo no solo veremos cómo de eficientes son las distintas aproximaciones, si no también cómo de explicables son nuestros modelos.

1.2. Aproximaciones al problema

El modelo más conocido es el de \mathcal{H} , que toma como entrada los datos del electrocardiograma tal como se recogen, que en esencia son doce mediciones de la actividad eléctrica en distintos puntos del cuerpo (expandiremos más en este tópico en la siguiente sección).

Nosotros exploraremos la posibilidad de convertir el electrocardiograma en diagramas de frecuencia-tiempo, y aplicarle diferentes transformadas antes de entrenar a los modelos, y estudiaremos si esto mejora su eficiencia, así como su explicabilidad.

1.3. Objetivos

El objetivo de este trabajo es modificar distintos modelos de predicción para que tomen como entrada imágenes en lugar de los datos del electrocardiograma, y ver como esto afecta a su rendimiento y explicabilidad.

1.4. Plan de trabajo

En el siguiente capítulo tendremos que familiarizarnos con los datos que proporciona un electrocardiograma, así como con los modelos ya existentes que tendremos que modificar y con los datos que hay disponibles de manera pública.

Luego modificaremos los modelos para que acepten como entrada imágenes y los volveremos a entrenar con los datos transformados de varias maneras diferentes.

Una vez entrenados los modelos modificados, evaluaremos su rendimiento y explicabilidad, y los compararemos con los modelos originales para ver si hemos conseguido una mejora.

Capítulo 2

Estado de la Cuestión

RESUMEN: En este capítulo estudiaremos las distintas bases de datos que podemos utilizar, así como los modelos que ya hay.

2.1. Electrocardiogramas

Un electrocardiograma es una prueba médica en la que se mide la diferencia de potencial entre varios puntos del cuerpo (cada diferencia de potencial recibe el nombre de derivación). Para hacer un ECG estándar, se ponen diez electrodos en puntos concretos del cuerpo y se generan doce ondas que representan la diferencia de potencial entre doce pares de estos electrodos.

Como es natural, no podemos tomar una cantidad continua de medidas del cuerpo, así que los electrocardiogramas funcionan a una frecuencia específica, lo que indica el número de muestras numéricas que toma el aparato medidor por segundo. Lo estándar es que los electrocardiogramas se tomen a 100Hz o a 500Hz dependiendo de la calidad del aparato medidor.

Con estas gráficas (podemos ver los datos de un electrocardiograma extraído de ? y dibujado mediante la librería de python *ecg_plot* en la Figura 2.1), un médico puede detectar alteraciones en el ritmo cardíaco que indiquen diferentes patologías cardíacas.

Por tanto, la información de un electrocardiograma puede almacenarse simplemente en doce vectores de tamaño 100 o 500 por segundo de duración del ECG. Estos son los datos con los que funcionan los modelos de predicción actuales. Nosotros utilizaremos la transformada de Fourier de tiempo reducido (o, por sus siglas en inglés, **STFT**) para crear una imagen (osea, una matriz) a partir de cada una de las doce derivaciones. Luego entrenaremos modelos ligeramente modificados para clasificar anomalías a partir de estas imágenes u otras transformaciones de las mismas.

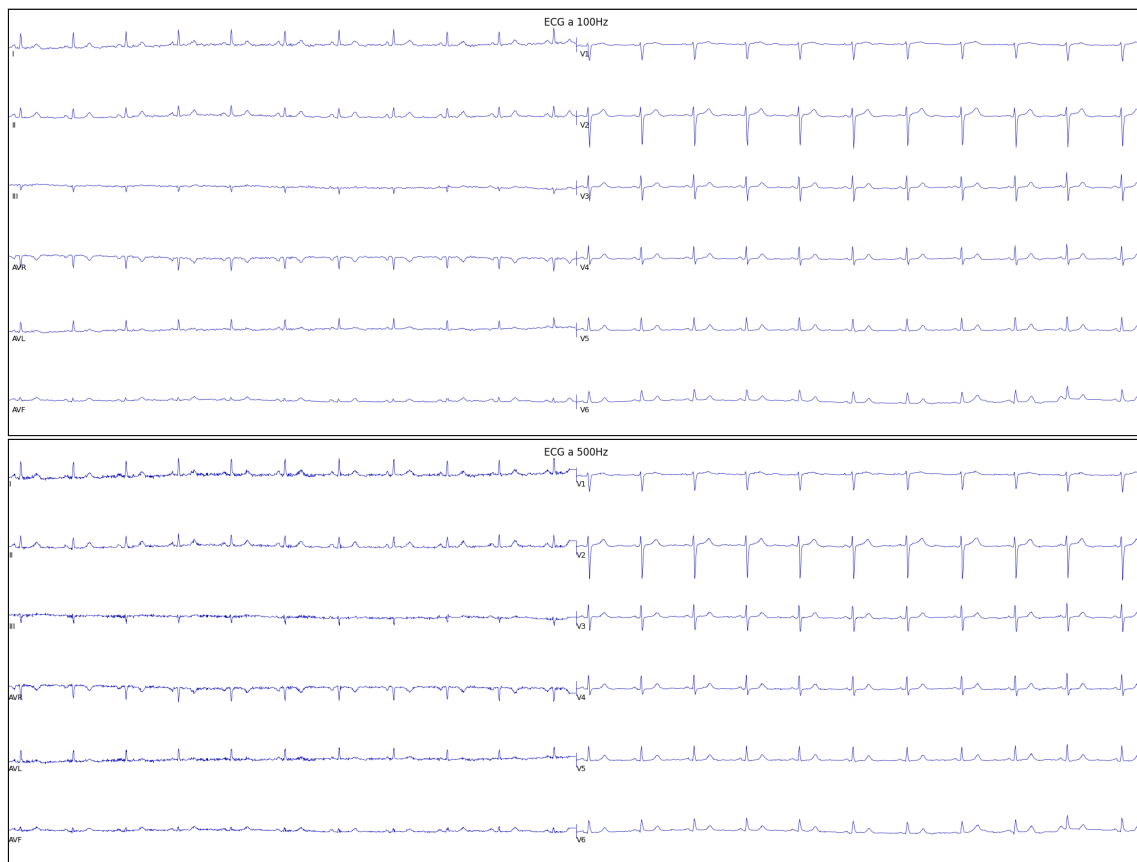


Figura 2.1: Mismo electrocardiograma de 12 derivaciones tomado a 100Hz (arriba) y 500Hz (abajo).

2.2. Bases de datos

Como queremos modificar y comparar el modelo de Ribeiro, lo ideal sería trabajar con la misma base de datos que este para entrenar y probar nuestros modelos, pero esta base de datos, llamada CODE (?), no es pública, por lo que esto no es una opción.

Existe una versión pública reducida de la base de datos llamada CODE-15 (?), que tiene al rededor de 350000 casos. No obstante, no podemos comparar un modelo entrenado con los datos de CODE-15 con otro entrenado con los datos originales, porque la ventaja en el conjunto de datos de entrenamiento haría imposible comparar la eficiencia y explicabilidad de manera equitativa. Por ello, vamos a volver a entrenar el modelo original de Ribeiro con una base de datos pública.

No obstante, en lugar de entrenar todos nuestros modelos con CODE-15, utilizaremos otra base de datos completamente pública, la PTB-XL (?). Esta base de datos cuenta solo con unos 22000 casos de prueba, pero todos estos casos vienen con el mismo formato, por lo que procesar los datos va a ser mucho más sencillo con esta base de datos. Además, dado que el objetivo es comparar si las modificaciones de los modelos mejoran el rendimiento, es mucho más importante que las bases de datos sean iguales y poder hacer varios entrenamientos que utilizar una base de datos enorme con la que nos lleve mucho más tiempo entrenar los modelos.

Un posible trabajo futuro sería escoger el modelo que mejor nos haya funcionado y volver a entrenarlo con una base de datos mayor (incluso con CODE si se tuviera acceso a ella), y luego compararla con el modelo de Ribeiro original, no obstante, eso queda fuera del alcance de este documento.

2.3. Modelos

Como hemos mencionado ya varias veces, el modelo que nos va a servir de referencia en este trabajo es el de Ribeiro, por ser el que mejores resultados ha obtenido hasta ahora. No obstante, también probaremos con otros modelos.

En concreto, existe una modificación del modelo de Ribeiro que, con tan solo 3 derivaciones, consigue unos resultados muy parecidos (?). Además, este modelo tiene la ventaja de estar ya entrenado con los datos de PTB-XL, por lo que no haría falta volver a entrenar el modelo original.

Capítulo 3

Preprocesamiento de los datos

Antes de poder trabajar

Capítulo 4

Conclusiones y Trabajo Futuro

Conclusiones del trabajo y líneas de trabajo futuro.

Antes de la entrega de actas de cada convocatoria, en el plazo que se indica en el calendario de los trabajos de fin de grado, el estudiante entregará en el Campus Virtual la versión final de la memoria en PDF.

Introduction

Introduction to the subject area. This chapter contains the translation of Chapter 1.

Conclusions and Future Work

Conclusions and future lines of work. This chapter contains the translation of Chapter 4.

Apéndice **A**

Título del Apéndice A

Los apéndices son secciones al final del documento en las que se agrega texto con el objetivo de ampliar los contenidos del documento principal.

Apéndice	B
----------	----------

Título del Apéndice B

Se pueden añadir los apéndices que se consideren oportunos.

Este texto se puede encontrar en el fichero Cascaras/fin.tex. Si deseas eliminarlo, basta con comentar la línea correspondiente al final del fichero TFGTeXiS.tex.

*–¿Qué te parece desto, Sancho? – Dijo Don Quijote –
Bien podrán los encantadores quitarme la ventura,
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

*–Buena está – dijo Sancho –; fírmela vuestra merced.
–No es menester firmarla – dijo Don Quijote–,
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

