Readme.md 2024-11-05

Predicción de Riesgo de Evento Cardíaco con Señales de ECG Usando la Base de Datos PTB Diagnostic ECG y PTB-XL.

El proyecto se centra en la predicción del riesgo de eventos cardíacos utilizando registros de ECG (electrocardiogramas) de la base de datos PTB Diagnostic ECG y PTB-XL de PhysioNet. La finalidad es desarrollar un modelo de aprendizaje profundo que ayude a identificar pacientes en riesgo de sufrir eventos cardíacos, basándose en características extraídas de los datos electrocardiográficos.

Etapas contenidas en este proyecto

- I. Análisis exploratorio de datos.
- II. Selección de modelos
- III. Entrenamiento y evaluación de modelos
- IV. Comunicación de Resultados

Etapa I Análisis exploratorio de datos

Respecto a esta etapa, se realizaron las siguientes actividades:

Identificación de las propiedades de los registros ECG

Mediante el uso de la librería wfdb se identificaron propiedades relevantes de los registros ECG de la base de datos PTB Diagnostic ECG.

En este caso, tales registros tienen las siguientes propiedades:

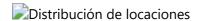
- record_name: Nombre del registro.
- n_sig: Número de señales en el registro.
- fs: Frecuencia de muestreo de las señales.
- **sig_name**: Nombre de las señales.
- **sig_len**: Longitud de las señales.
- comments: Comentarios sobre el registro.
- p_signal: Señales físicas.

Como se puede observar, en el atributo 'comments' existe una lista de información relevante sobre el paciente, en la cual se puede identificar que el valor de la llave 'Reason for admission' clasifica a al registro como 'Myocardal infarction' en este caso. También existe la clave 'Acute infarction (localization)' que indica la localización del infarto.

Debido a que esta base de datos tiene registros tanto de pacientes con Infarto al miocardio y de pacientes sanos, se procedió a realizar una busqueda de los registros que contienen la clave 'Reason for admission' con el valor 'Myocardal infarction' para identificar los registros de pacientes con infarto al miocardio.

Además se identificó los valores únicos relacionados a la clave 'Acute infarction (localization)' para identificar las localizaciones del infarto ya que en este apartado existe información faltante y ambigüa como se puede observar en la siguiente imagen:

Readme.md 2024-11-05



Para realizar la limpieza de estos elementos, se procedió a realizar un filtro en el cual los registros que contienen los valores 'no', 'n/a' y 'unknown' fueron eliminados para el análisis de datos dando como resultado la siguiente distribución de locaciones:

Distribución de locaciones filtradas

Identificación de la distribución de las clases

Se identifica la distribución de las clases STEMI y HC (Healthy Control) en la base de datos PTB Diagnostic ECG mediante el uso de un atributo "comment" propio de cada registro dando como resultado los las siguientes cantidades:

Distribución de las clases

Identificación de problematica y solución

Se identifica la problematica de la base de datos PTB Diagnostic ECG, la cual se refiere al desbalance de clases respecto a los registros STEMI y HC. La solución seleccionada respecto a la problematica cual se refiere al uso de otra base de datos (PTB-XL) para obtener y balancear los registros HC (Clase minoritaria).

Identificación de las propiedades de los registros ECG de la base de datos PTB-XL

Mediante el uso de la librería wfdb se identificaron las principales diferencias entre los registros ECG de la base de datos PTB Diagnostic ECG y PTB-XL.

En este caso, tales registros tienen las siguientes diferencias:

• **fs**: 500 Hz

• comments: Vacío

Los registros de la base de datos PTB-XL no tienen información en el atributo 'comments' por lo tanto su carga de datos para clasificarlos se debe de realizar de una manera distinta a la realizada con la base de datos PTB Diagnostic ECG.

En este caso, la información referida a los diagnosticos y etiquetas estan registradas en un .csv llamado ptbxl_database.csv.

Respecto a la frecuencia de muestreo, esta es diferente e inferior a los 1000hz de la base de datos PTB Diagnostic ECG.