

Predicción de la apertura de los ojos en base a datos de un Electroencefalograma

Juan Miguel Torrez Manchego

jtorrezm@fcpn.edu.bo

Universidad Mayor de San Andres

RESUMEN:

INTRODUCCION: Durante mucho tiempo se tuvo un gran desarrollo de herramientas médicas para el estudio del cerebro y su comportamiento dentro del mismo, uno de ellos es el electroencefalograma (EEG) que su utilidad es monitorear el funcionamiento eléctrico cerebral, además de detectar ciertas anomalías tras realizar una serie de estímulos que requieran actividad cerebral.

Ante esta maravillosa herramienta se genera una gran duda, si se podría saber si una persona abre o cierra los ojos, solo con los datos tomados con el electroencefalograma.

Para ello se tomó una muestra a una persona durante 117 segundos, en los cuales se le pidió abrir y cerrar los ojos, mientras él era estudiado con un electroencefalograma y ver en qué lugares del cerebro surgen cambios al realizar dicha acción.

MATERIALES Y METODOS:

Dataset: los datos proceden de una página web llamada UCI, en el cual se observa 15 columnas, 14 columnas son los datos extraídos del estudio del EEG, el número de nodos en un aparato de EEG puede ser de 12 a 26 nodos, en el presente estudio se utilizó 16 nodos colocándolos la mitad en la parte izquierda y el resto en parte derecha del cráneo; solo 14 de los nodos fueron tomados en cuenta por la variación en estos y los 2 no tomados en cuenta no demostraron variabilidad durante el experimento.

la columna restante del conjunto de datos es un dato de dos valores de 1 y 2, indicando el valor de 1: ojos abiertos, y 2 ojos cerrados.

Durante el estudio que se realizó al individuo se obtuvo 14981 muestras durante 117 segundos.

Pre procesamiento:

Los datos del dataset, poseen una serie de ruido como se observa en el gráfico 1, para ello se aplicó una limpieza de datos, calculando la media, la desviación estándar para luego hallar los valores atípicos en el dataset. Una vez realizada esta técnica de pre procesamiento obtenemos una gráfica con menos ruido y con una tabla más corta con 14304 filas como se observa en la gráfica 2.

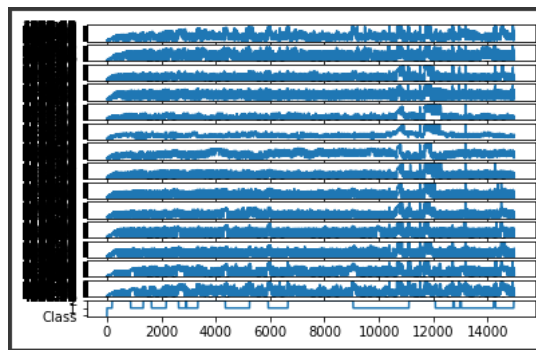


Gráfico 1: muestra el ruido existente en el dataset sin realizar pre procesamiento.

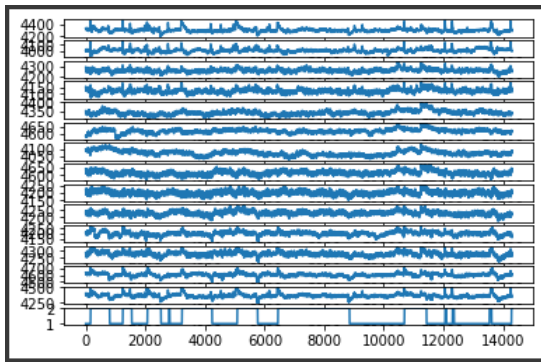


Grafico 2: Eliminacion de datos que excedían del valor atípico, generando menos ruido.

Para el dataset no se pudo aplicar la técnica de completado de datos, ya que no existían datos faltantes tras la eliminación de filas.

Ademas tratando de Normalizar la tabla por fila se observó una gran pérdida de información en los datos, siendo estos afectados para la predicción y el objetivo de la investigación.

Algoritmo de K-Nearest-Neighbor (KNN):

El algoritmo KNN es un método que busca las observaciones más cercanas a que se está tratando de predecir y clasifica el punto de interés basado en la mayoría de datos que le rodean, es un algoritmo supervisado, además es basado en instancia, esto nos indica que nuestro algoritmo no aprende explícitamente un modelo (como por ejemplo Regresión Logística o arboles de Decisión). En cambio, memoriza las instancias de entrenamiento que son usadas como “base de conocimiento” para la fase de predicción.