En este trabajo trataremos un data set que consta de una serie de columnas que usaremos para predecir si os pacientes tienen insomnio, apnea del sueño, o no presentan ninguna enfermedad de este tipo.

Comenzaremos leyendo los datos y haciendo una pequeña visualización para ver el nombre de las variables, etc, y ver algunas cosas que debemos cambiar como podrían ser los tipos de algunas variables, e incluso modificar algunas columnas.

Eliminamos una variable que no necesitamos, cambiamos el tipo de otras varias a factor y posteriormente modificamos la variable de la presión sanguínea:

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

Pasamos a exportar los datos, y continuamos con algunas representaciones gráficas de variables que podríamos considerar de interés. Todas estas representaciones las hacemos separando la variable objetivo en sus distintas clases a ver si gracias a esa predictora podemos separarlo:

Gráfico, Gráfico de barras

El contenido generado por IA puede ser incorrecto.Nos damos cuenta de que no obtenemos una información clara con ninguna variable.

Pasamos a construir los modelos, comenzando con el upsampling de los datos puestos que tenemos pocos individuos que presenten apnea del sueño o insomnio. Una vez lo hacemos, para modificar los datos construimos nuestro pipeline con dos componentes principales para ver si de esta forma la representación es más sencilla que en los ejemplos anteriores.

A continuación, construimos una serie de modelos en los que entrenamos nuestros datos y posteriormente evaluamos el rendimiento del modelo en nuestro conjunto test, obteniendo algunas métricas y representaciones gráficas como la matriz de confusión y la región frontera.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.Y al final de cada modelo podemos observar que conclusión sacamos. Una vez analizamos todos los modelos pasamos a quedarnos con el mejor de ellos para usarlo en el siguiente apartado.

Como pestaña final tenemos un cuestionario en el que podremos introducir los datos de un nuevo paciente, y con el mejor modelo anterior obtenemos si el paciente tiene alguna enfermedad del sueño y la probabilidad de que la tenga.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.