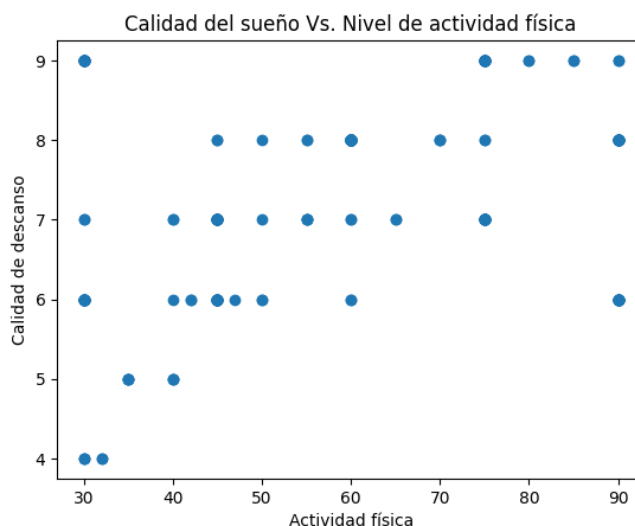


“Herramienta predictiva sobre la calidad del descanso en la población en general”

Machine Learning

Descripción

Este es un proyecto de machine learning con Python a través del cual queremos predecir como influirán x cambios de los hábitos de comportamiento en la calidad del descanso de las personas. Para esto se utilizan variables fijas como pueden ser la edad y el género; como variables móviles como pueden ser el nivel de actividad física, la ocupación ó las horas de descanso. Para este proyecto se ha usado un set de datos de kaggle.com sobre “La salud del sueño y el estilo de vida”.



El trabajo se ha dividido en tres partes: la primera correspondiente al tratamiento de los datos. Se corrigieron, eliminaron y/o asignaron valores nuevos a los inexistentes o nulos.

Dividimos los datos entre datos de entrenamiento y datos de test para posteriormente entrenar a nuestro modelo. Y eliminamos la variable predictora “calidad de descanso”.

```
datos_entrenamiento = df.sample(frac = 0.8, random_state = 0)
datos_test = df.drop(datos_entrenamiento.index)
```

Las librerías usadas han sido: pandas, matplotlib, numpy.

La segunda etapa corresponde a la de entrenamiento.

Utilizamos una función lineal (regresión) como modelo para entrenar por considerarlo más sencillo. En esta fase advertimos que nuestro modelo presenta underfitting, no está aprendiendo. Actualmente se está trabajando limitando las variables con la intención de facilitar a nuestro modelo el patrón.

Conclusiones

De no dar resultados, el siguiente paso será buscar otro modelo más complejo ya que se considera que hay datos suficientes para intentarlo.

Quedan pendientes las pruebas de test y la aplicación de nuestro algoritmo como herramienta predictiva añadiendo nuevos valores.

Author: Lourdes Espinola.