Desarrollar un modelo para predecir el nivel de obesidad basado en hábitos alimenticios y condición física. Para desarrollar el modelo deben comprender cómo los factores como la frecuencia del consumo de alimentos calóricos, la actividad física y la monitorización de calorías impactan el nivel de obesidad. Además, se deben explorar como las variables como la ingesta de alimentos calóricos, la frecuencia de consumo de vegetales, la actividad física y el consumo de agua pueden predecir el nivel de obesidad de un individuo.

Solucion:

Lo que se quiere con el modelo de aprendizaje es predecir el nivel de obesidad de un individuo. Esto es un problema de clasificación supervisada y vamos a trabajar con categorías como, por ejemplo: bajo peso, peso normal, sobre peso tipo I y II, el tipo de obesidad ya sea tipo I, II.

Trabajaremos con variables que sean predictoras

Como lo menciona en el enunciado del ejercicio monitorearemos, el consumo de alimentos de vegetales, consumo de alimentos calóricos, la actividad física, consumo de agua diaria, monitoreo de calorías, la frecuencia de consumos de alimentos de comida rápida, la calidad de sueño.

Lo siguente seria crear el modelo de obesidad que va leer el archivo ObesityDataSet.csv, con este modelo lo que va hacer codificar el texto como “MALE”, “YES”, a números.

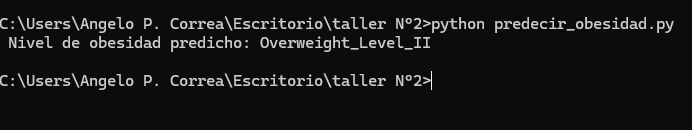


Con el modelo de obesidad ya creado y entrenado crearemos un script para hacer predicciones de nivel de obesidad con nuevos valores



* Básicamente cargar modelo entrenado el cual está guardado como joblid o pickle.
* También solicita y recibe nuevos datos como por ejemplo edad, altura, la actividad física del individuo.
* Transforma los datos al formato correcto
* Usando el modelo SVM para predecir nivel de obesidad.

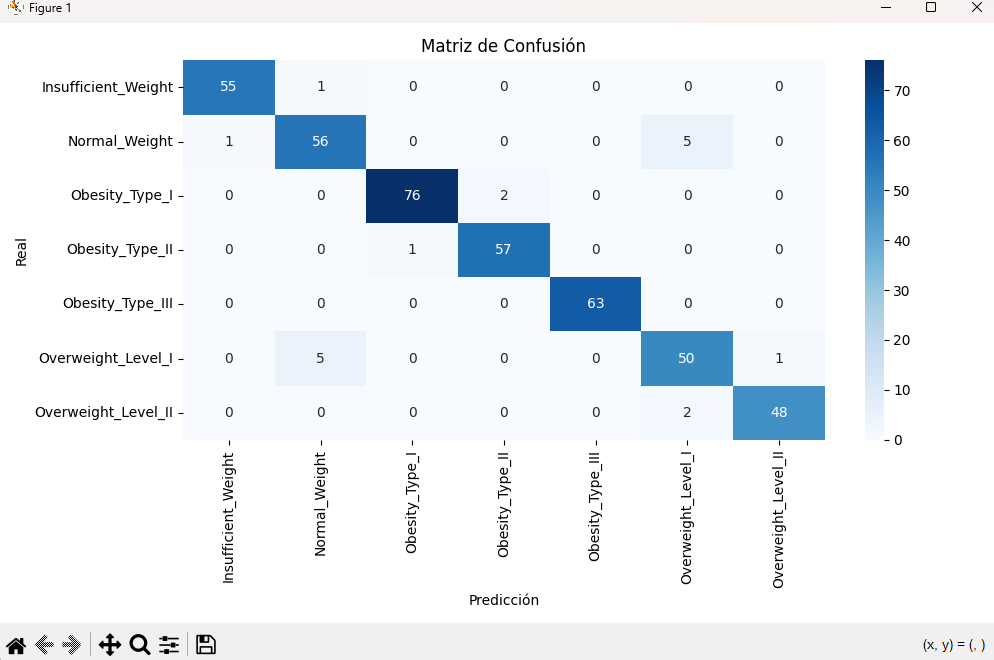
Ejecutamos en una terminal Python predecir\_obesidad.py y nos arroja lo siguente:

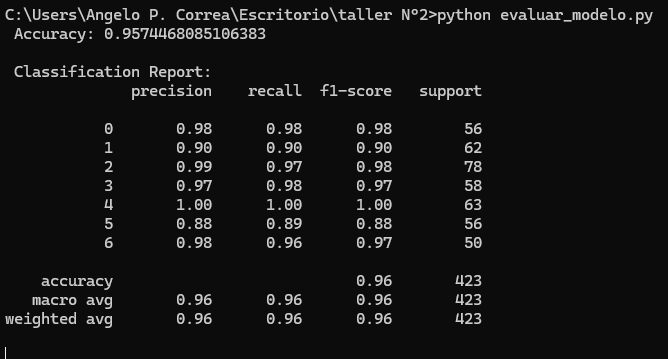


El siguiente paso sería evaluar el modelo



Aquí podemos observar que carga el modelo SVM previamente entrenado y lo evalúa usando datos de prueba. También nos permite clasificar el modelo métrico como precisión, matriz de confusión y reporte de clasificación nos ayuda a verificar el rendimiento.





Los numero del 0 al 6 esto representa la clase del nivel del tipo de obesidad si es

0= normal\_weight,

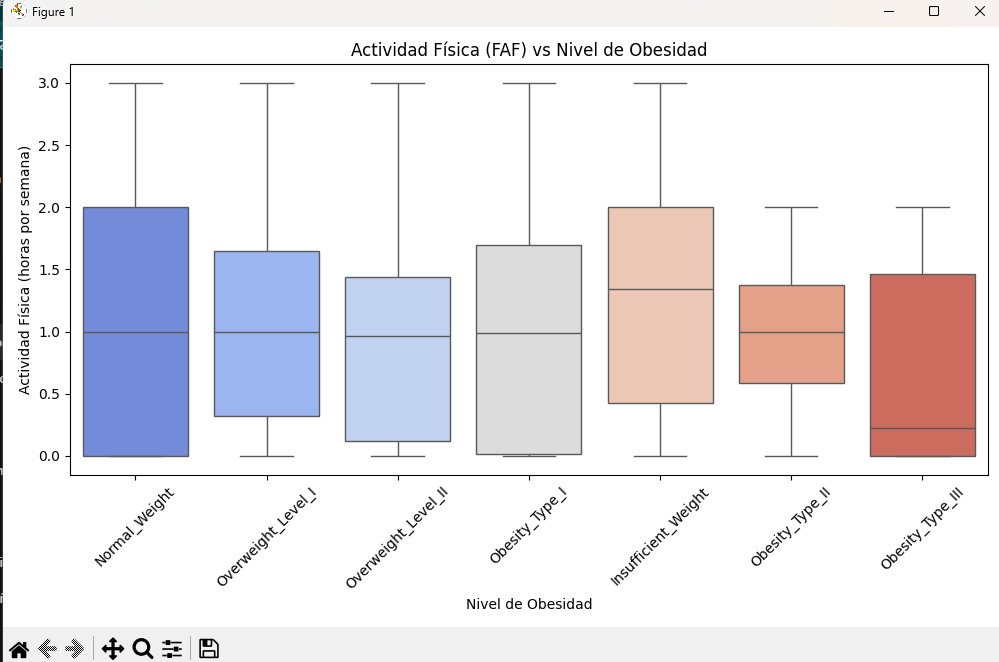
1 = obesity\_type\_I



Los siguiente seria mostrar en una gráfica como impacta algunas variables la actividad física vs la obesidad.



Con el script podemos visualizar la distribucion de cada variable según el nivel de obesidad.



Con este diagrama podemos visualizar el ObesityDataSet, para entender un poco la distribucion de las variables valores atípicos, correlaciones y patrones. Esto es muy importante al momento de entrenar el modelo, nos ayuda a decidir que posibles variables podemos transformar, eliminar o usar.

La importancia de las variables la usamos para poder visualizar que variables influyen más en la predicción del modelo. Esto nos ayuda a interpretar el comportamiento del modelo.



