Predicción del tipo de especie de la flor Iris

pasos:

1. cargar datos

2.visualización de datos

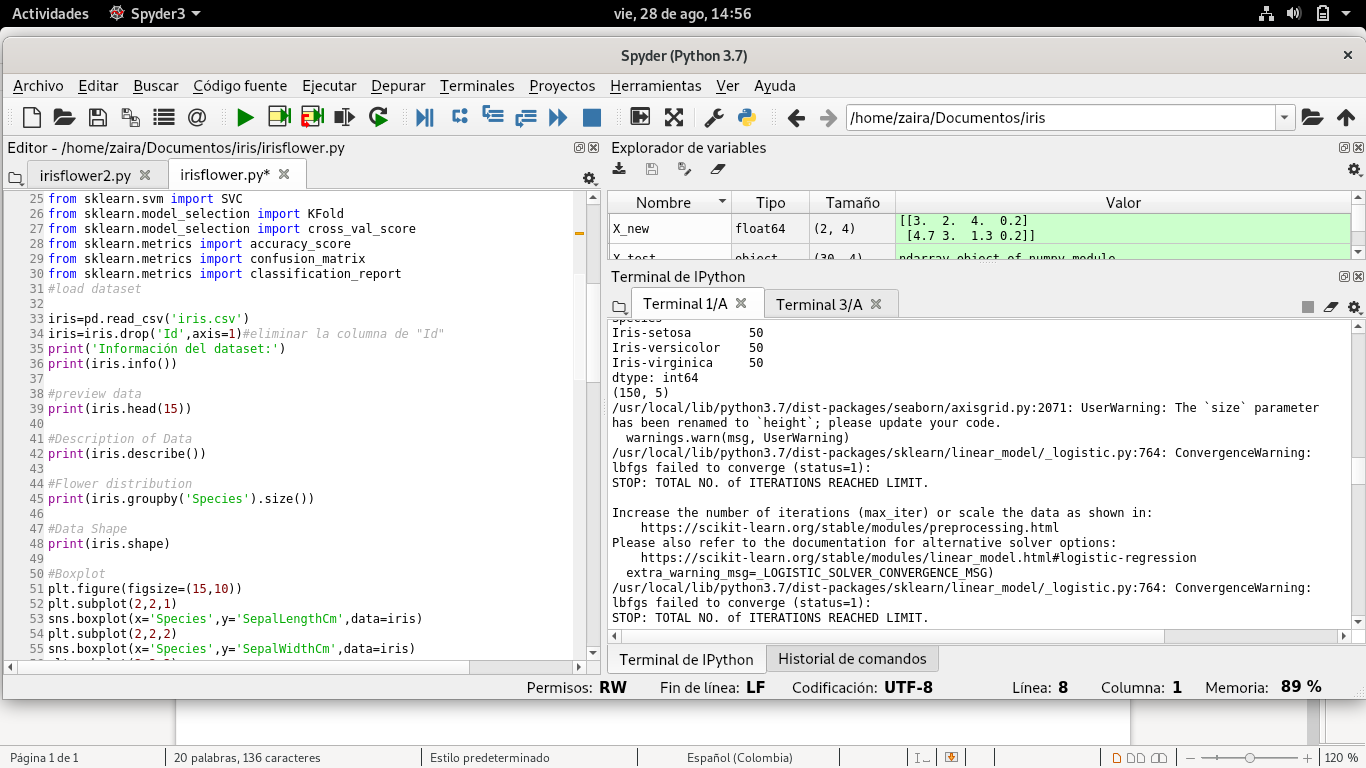
3.separar datos

4.entrenamiento

5.validación

1. Cargar datos

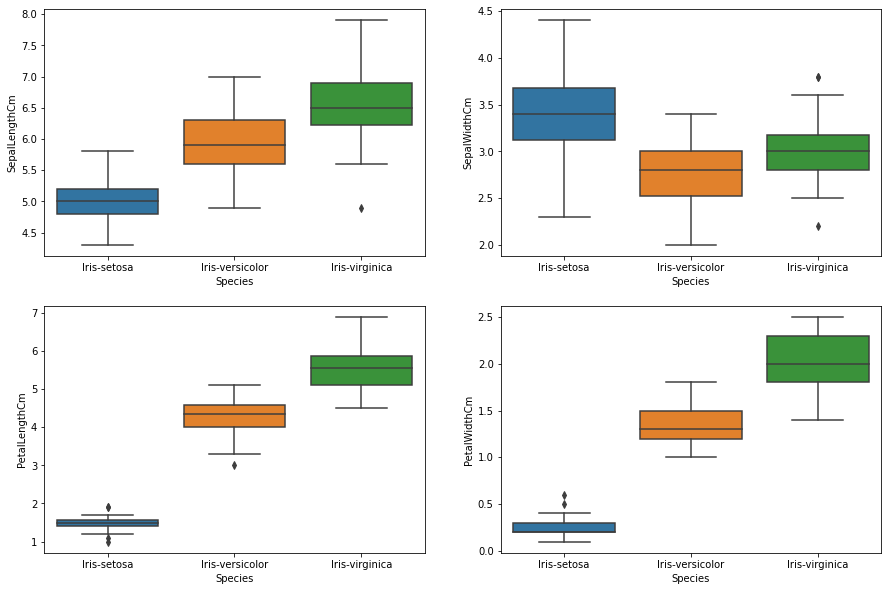
Es necesario tener en cuenta que se debe contar con el archivo del conjunto de datos que se van a usar.



2. Visualización de los datos

la visualización de los datos se realizará con ayuda de dos tipos de gráficas: Boxplot (diagrama de cajas) y diagrama de dispersión.

Boxplot:

La visualización de este tipo de gráfico nos permite realizar una comparación entre los datos de cada especie, pues nos indica el porcentaje de existencia de cada dato. Por ejemplo en la grafica de SepalWithCm se observa que el 50% de los datos de iris setosa se encuentran entre 3.1 y 3.7, los de iris Versicolor se encuentran entre 2.5 y 3.0, y los de iris Virginica se encuentran entre 2.8 y 3.1. El segmento que divide las cajas nos índica la media de cada conjunto de datos, esto ayuda a comprender si la distribución es simétrica o no.

En esta misma gráfica también se puede resaltar que los bigotes de cada caja son bastante pronunciados, lo que nos índica la existencia de datos atípicos

