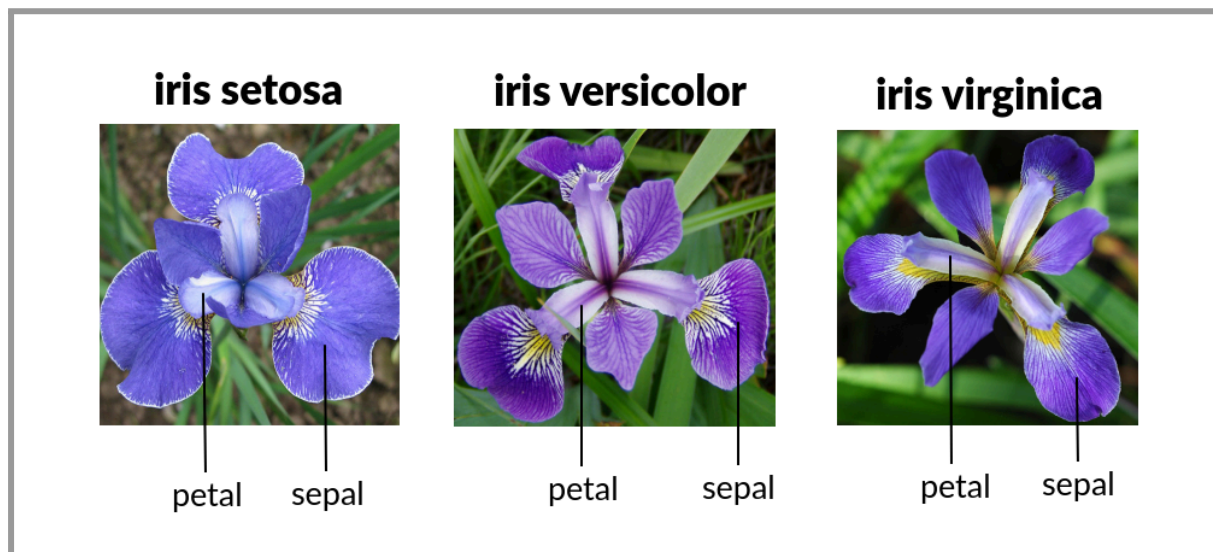


Guía de ejercicios de clasificación con KNN

Proponemos trabajar con el conjunto de datos Iris, que contiene 150 flores de 3 especies: setosa, versicolor y virginica. Cada flor tiene 4 atributos medidos en centímetros: largo del sépalo, ancho del sépalo, largo del pétalo, ancho del pétalo. La variable a predecir es la especie de la flor.



Primero deben descargar el archivo que está en el campus.

Ejercicios

1. Cargar el dataset y mostrar las primeras filas.
2. Hacer un scatterplot de dos atributos (por ejemplo, largo y ancho del pétalo), coloreando por especie.
3. Entrenar un clasificador k-NN para predecir la especie.
 - a. Probar distintos valores de k (por ejemplo: 1, 3, 5, 10, 20).
 - b. Probar entrenar usando solo dos atributos.
 - c. Para cada configuración, medir la exactitud del modelo (accuracy).
4. Elegir un valor de k que haya funcionado bien.
 - a. Visualizar la frontera de decisión correspondiente.
 - b. Predecir la especie de una flor con largo de pétalo 4.8 cm y ancho de pétalo 1.5 cm.