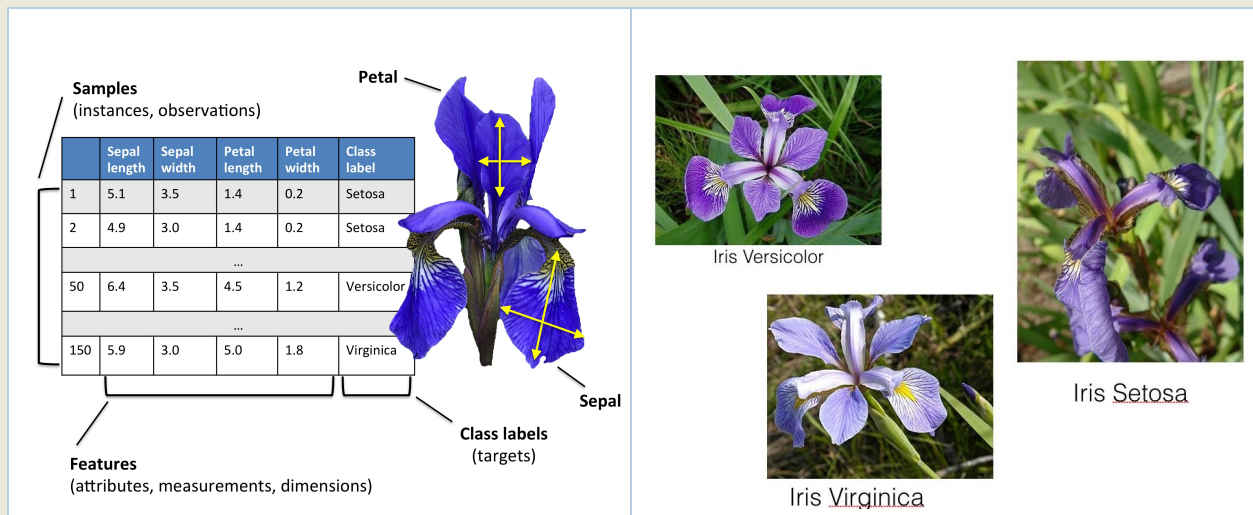


ML_CLASIFICACION_KNN_02

Modelo de Clasificación KNN para predecir el tipo de flor

ML

Un botánico está interesado en distinguir las especies de algunas flores (de iris) que ha encontrado. Para ello ha registrado algunas medidas asociadas con cada flor: la longitud y anchura de los pétalos (petal), y la longitud y anchura de los sépalos (sepal). Todas las medidas están en centímetros.



Además, dispone de algunas medidas de flores en las que previamente un experto botánico ha catalogado dentro de las especies ('setosa', 'versicolor' y 'virginica').

Construir un modelo de aprendizaje maquina (machine learning), basado en el algoritmo de KNN (k-Nearest Neighbors) que aprenda a clasificar el tipo de especie de la flor en función de las medidas conocidas.

Predecir el tipo de especie de flor si las dimensiones (longitud y anchura) de pétalo y sépalo son:

Longitud y anchura de sépalo: 7 y 3 cm

Longitud y anchura de pétalo: 5 y 1 cm

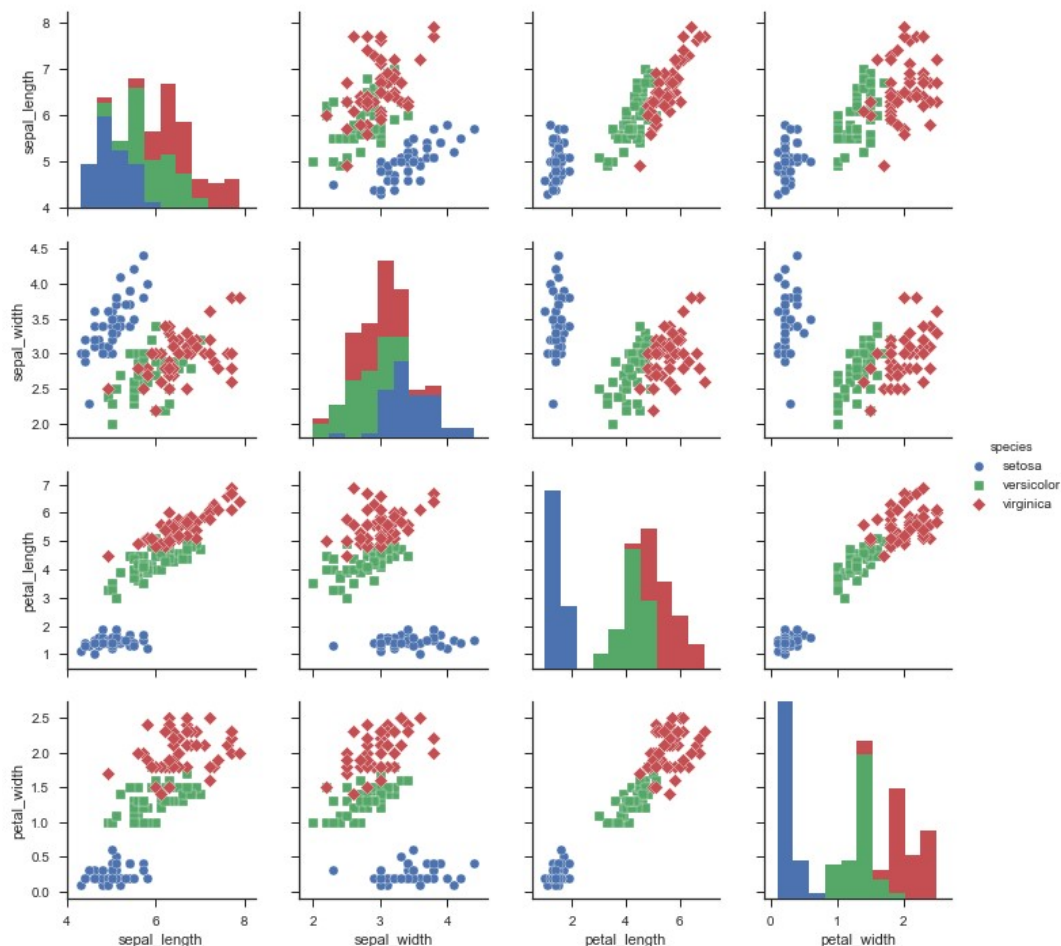
SOLUCIÓN

Importar las librerías necesarias para realizar la práctica.

```
# Librerías
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

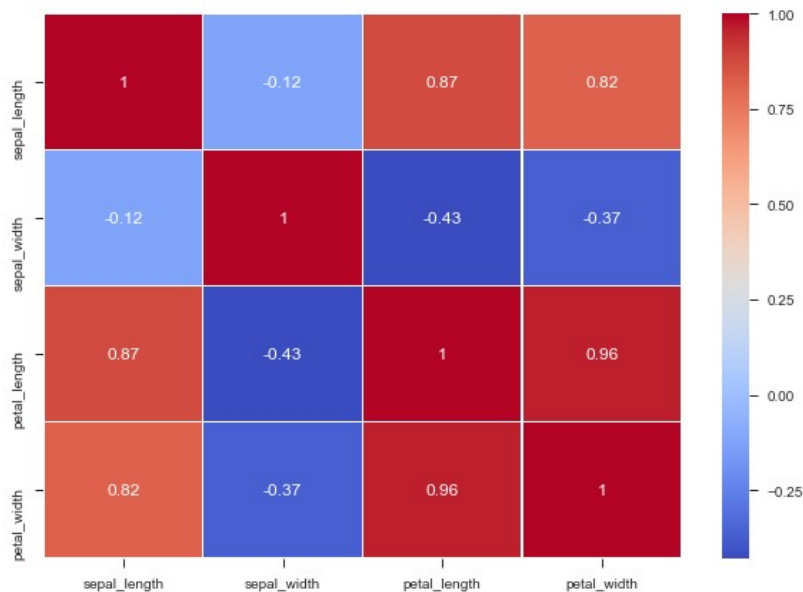
Cargar la base de datos (incluida ya en la librería 'seaborn'). Inspeccionar los datos:

```
# Cargar la base de datos 'Iris' y dar un primer vistazo con seaborn.
Inspeccionar los datos
df = sns.load_dataset("iris")
sns.pairplot(df, hue="species", markers=["o", "s", "D"])
```



Ver como se correlacionan.

```
# Ver como se correlacionan
fig = plt.figure(figsize=(10,7))
sns.heatmap(df.corr(),annot=True,cmap='coolwarm',linewidths=0.2)
plt.show()
```



Generar los datos X e y.

```
# Dividir el DataFrame 'df' en X (atributos) e y (etiqueta)
# X: contiene las 4 medidas (longitud y anchura de sépalo, y longitud y anchura de pétalo)
# y: contiene el valor de la clase o especie de flor (0-> setosa, 1-> versicolor, 2-> virginica)
X = df.drop('species',axis=1).values
y = df['species']
```

Dividirlos para entrenamiento y test.

```
# Dividir de forma aleatoria el conjunto de datos en entrenamiento (75 %) y test (25 %)

from sklearn.model_selection import train_test_split

X_train,X_test,y_train,y_test =
train_test_split(X,y,test_size=0.25,random_state=37, stratify=y)

print("X_train tamaño: {}".format(X_train.shape))
print("y_train tamaño: {}".format(y_train.shape))
```

```
print("X_test tamaño: {}".format(X_test.shape))
print("y_test tamaño: {}".format(y_test.shape))
```

```
X_train tamaño: (112, 4)
y_train tamaño: (112,)
X_test tamaño: (38, 4)
y_test tamaño: (38,)
```

Definir el clasificador basado en KNN y ajustar el modelo

```
# CLASIFICACIÓN CON KNN
# Crear el clasificador
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=1)

# Ajustar el modelo con los datos de entrenamiento
knn.fit(X_train, y_train)
```

Predecir para los datos de test.

```
# Predecir para los datos de test
y_pred = knn.predict(X_test)
print("Predicciones del Test:\n {}".format(y_pred))
```

Predicciones del Test:

```
['setosa' 'setosa' 'versicolor' 'setosa' 'setosa' 'virginica' 'versicolor'
'virginica' 'virginica' 'setosa' 'versicolor' 'setosa' 'virginica'
'virginica' 'versicolor' 'virginica' 'setosa' 'versicolor' 'versicolor'
'versicolor' 'setosa' 'versicolor' 'virginica' 'versicolor' 'setosa'
'versicolor' 'setosa' 'versicolor' 'setosa' 'versicolor' 'virginica'
'virginica' 'setosa' 'versicolor' 'setosa' 'virginica' 'virginica'
'versicolor']
```

Evaluar el modelo:

```
# Evaluar el modelo
print("Test score: {:.2f}".format(np.mean(y_pred == y_test)))
print("Test score: {:.2f}".format(knn.score(X_test, y_test)))
```

```
Test score: 0.97
Test score: 0.97
```

Predecir para los nuevos datos

```
# Predecir para nuevos datos
# Ejemplo:
# Longitud y anchura de sépalo: 7 y 3 cm
# Longitud y anchura de pétalo: 5 y 1 cm
y_pred_nuevo = knn.predict([[7,3,1,1]])

if y_pred_nuevo=="setosa":
    print("La especie es setosa")
```



```
elif y_pred_nuevo=="versicolor":  
    print("La especie es versicolor")  
else:  
    print("La especie es virginica")  
  
La especie es setosa
```