DS201_Lab9 – Classification

Use sklearn to train the classifier model using Iris data:

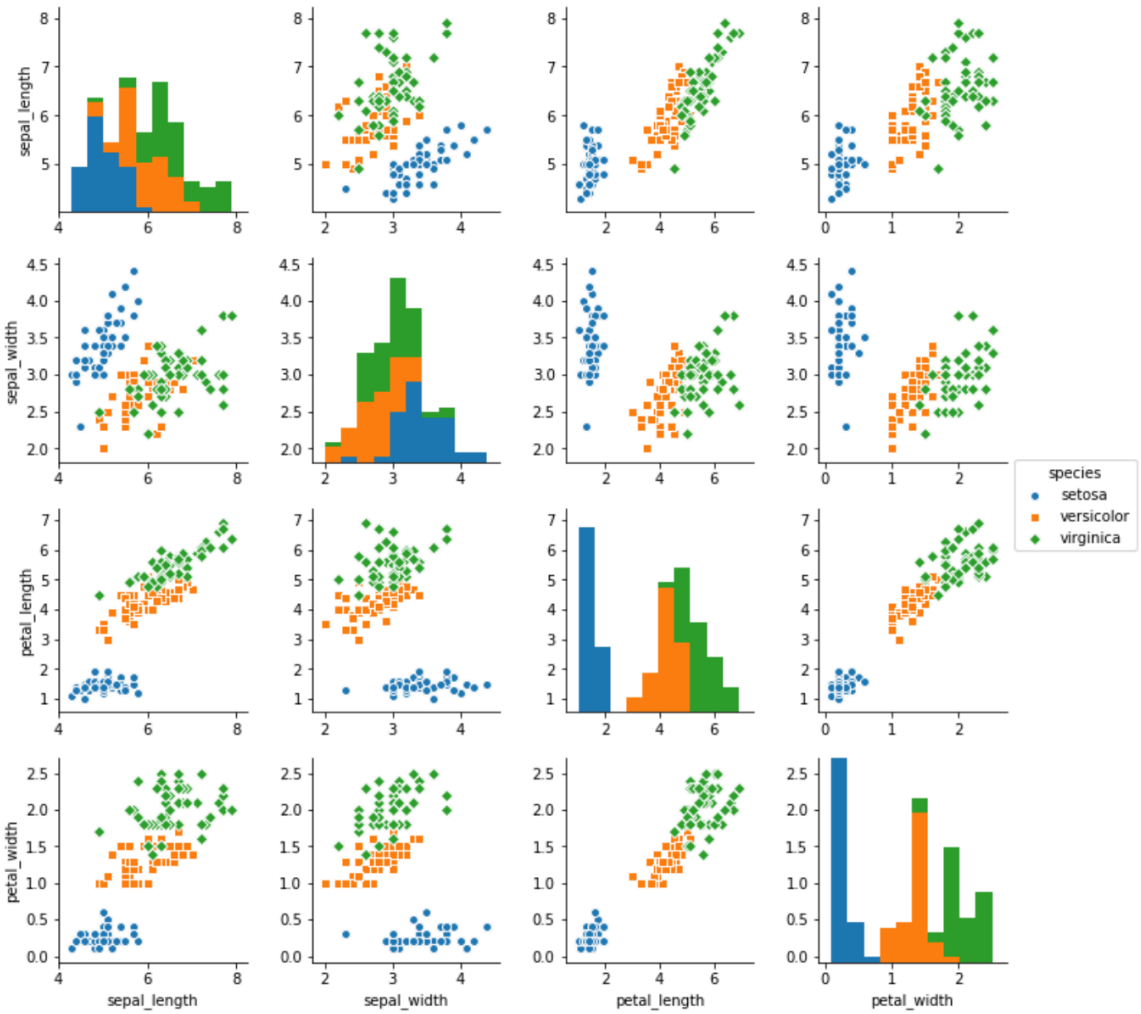
```
from sklearn.datasets import load_iris
iris = load_iris()
```

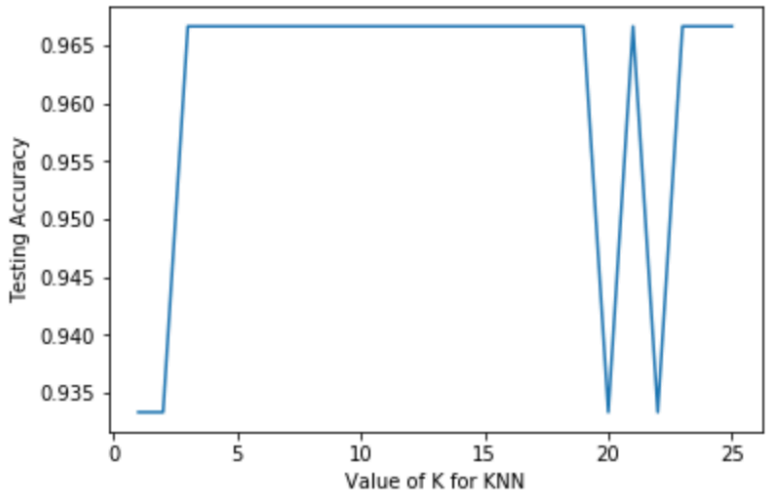
```
iris.head(150)
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
5	5.4	3.9	1.7	0.4	setosa
6	4.6	3.4	1.4	0.3	setosa
7	5.0	3.4	1.5	0.2	setosa
8	4.4	2.9	1.4	0.2	setosa
9	4.9	3.1	1.5	0.1	setosa
10	5.4	3.7	1.5	0.2	setosa

1. Create a seaborn pairplot with above data with species as hues and add markers. (Figure at the end)
2. Use the Iris data to train multiple classifiers:
   - KNeighborsClassifier(n_neighbors=1)
   - KNeighborsClassifier(n_neighbors=5)
   - LogisticRegression
3. Predict the type of two new iris flowers:
   [3, 4, 4, 2], [4, 4, 3, 2]
   By using above three classifier models:
4. Evaluate KNN model using different neighbor size
   - Create a train/test split of the iris data (70/30)
   - Train a model with neighbor size = 5 and predict the species in test set
   - Plot a line chart of KNN accuracy of k=1 to k=25 then pick the best k value.
5. Answer in the .pynb file code block:

Submit as DS201XLab09_Classifier_LastFirst.ipynb on Canvas


sepal_length
sepal_width
petal_length
petal_width
species
setosa
versicolor
virginica


Testing Accuracy
Value of K for KNN