

Assignment 2

Write jupyter notebook scripts for all of the questions.

1. K Nearest Neighbor classifier (50 points)

Apply K-Nearest Neighbor classifier to the iris dataset. Select any TWO of the four variables (sepal length, sepal width, petal length, petal width) for this problem so that the KNN prediction results could be easily displayed in a 2-D plane. Use the `train_test_split` method in sklearn to split the iris data for training and testing. Use 20% of the samples for testing. Display the prediction results in a figure similar to Figure 1

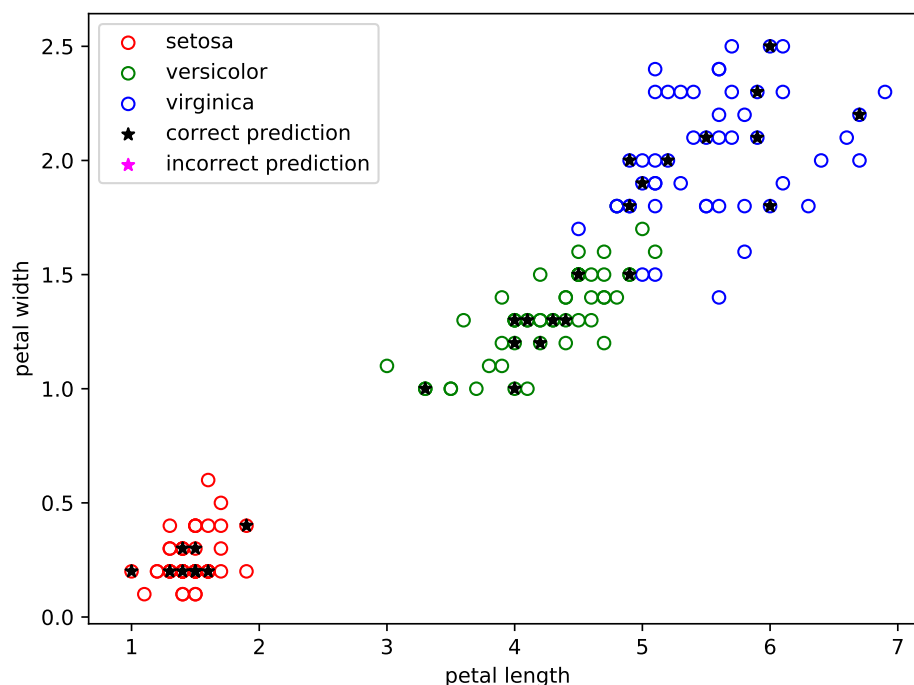


Figure 1: Example figure to display results from KNN classifier.

Information about how to use sklearn KNN can be found here: [sklearn K-Nearest Neighbor](#)

Information about how to use sklearn train-test-split can be found here: [train test split](#)


```
In[3]: prediction_accuracy
Out[3]:
array([0.96666667, 0.93333333, 0.96666667, 0.96666667, 1.
        0.96666667, 0.93333333, 1.
        , 0.9
        , 0.9
        ])
In[4]: print(np.mean(prediction_accuracy))
0.9533333333333334
```

Figure 3: Example prediction accuracy to display.

Information about how to use sklearn decision tree can be found here: [decision tree classifier](#).

Information about how to plot a decision tree can be found here: [plot a decision tree](#).

Information about how to compute prediction accuracy can be found here: [sklearn model evaluation](#). Look for `metrics.accuracy_score`