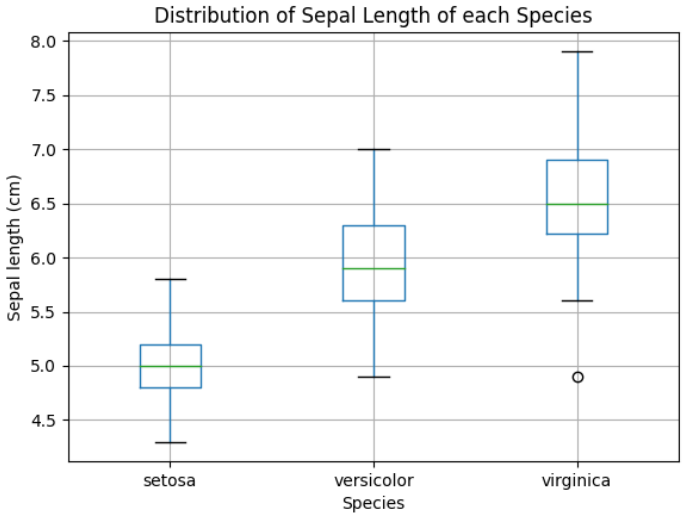
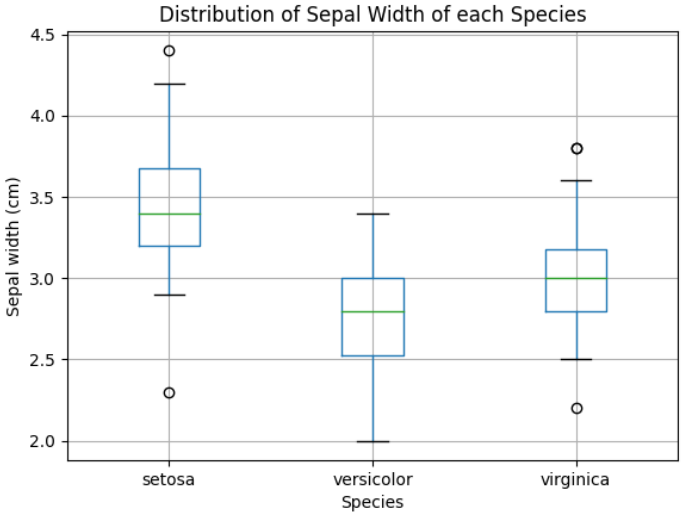
Q1:

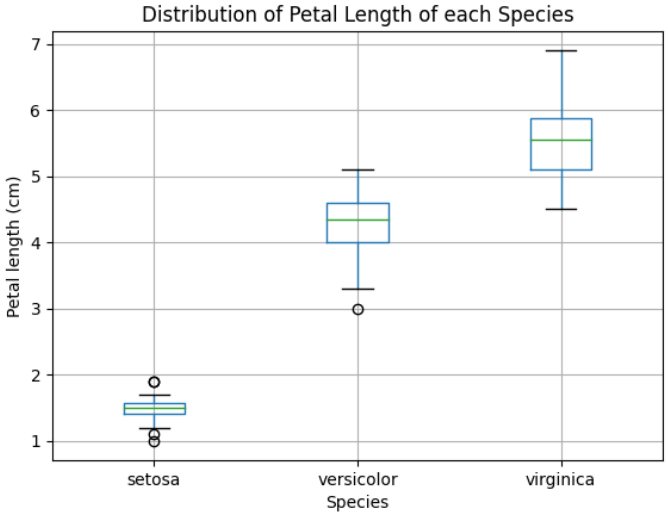
```
Python Console
Time [sec] (for loop): 0.38894033432006836
Time [sec] (np loop): 0.009993553161621094
np loop is  38.919123962210136  times faster than for loop.
```

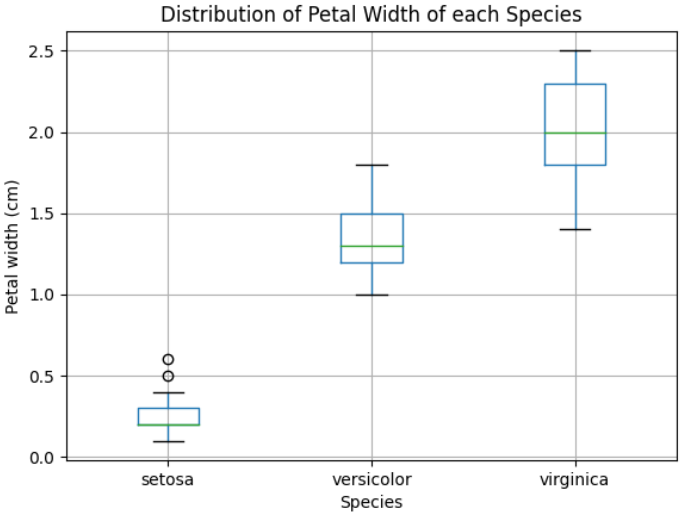
Q2:

b)

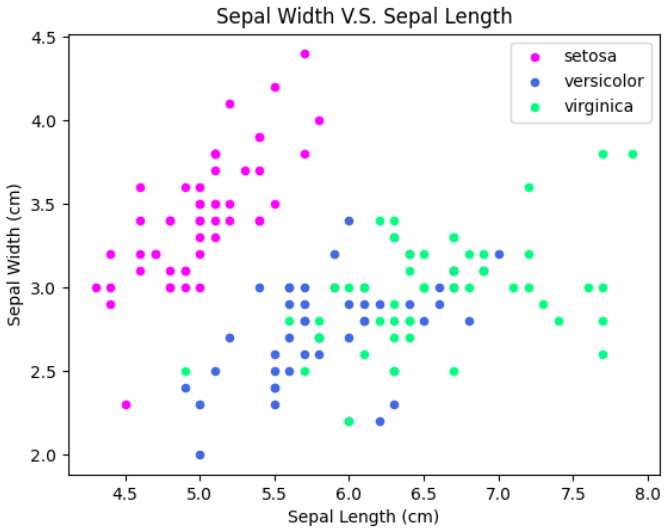


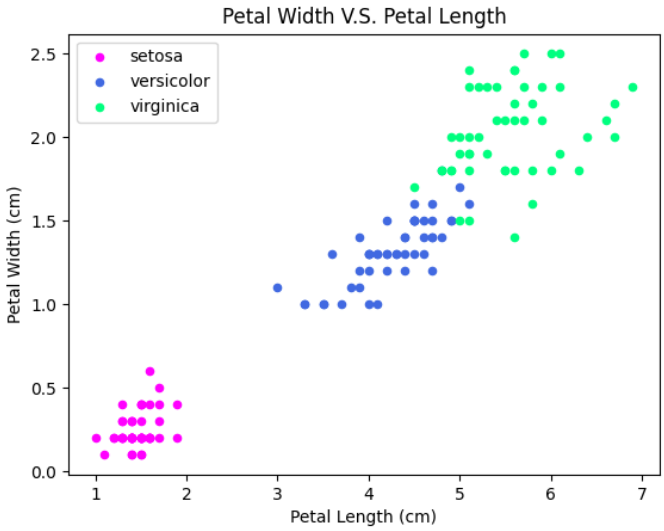






c)



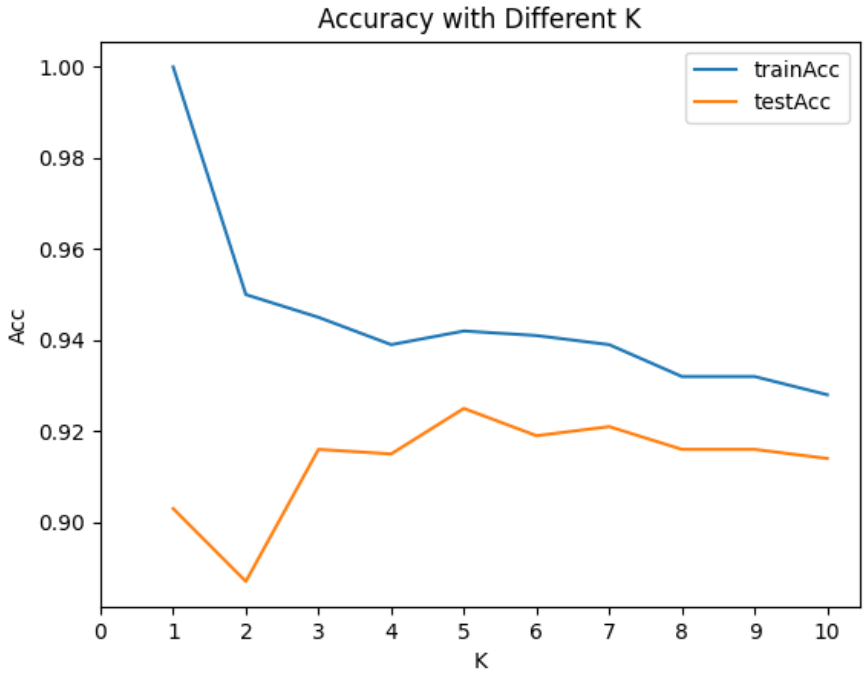


d) Rules to classify the species type:

1. If the petal length < 3cm and petal width < 1cm, the sample is setosa.
2. If 3cm <= petal length < 5cm and 1cm <= petal width < 1.75cm, the sample is versicolor
3. If petal length >= 5cm and petal width >=1.75cm, the sample is virginica.

Q3:

d)



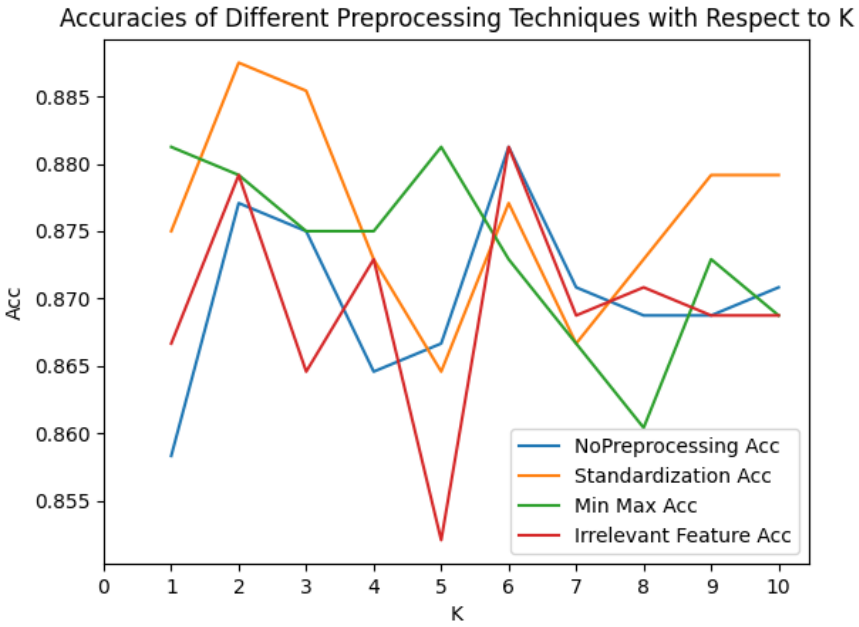
e)

The space complexity of the predict function is O(nd), since we need to store all the training data, which is an n*d array.

The time complexity of the predict function is O(n*(nd+nlogn+k)). Assume the testing set is also size n. We iterate through every instance in testing set, and in each iteration, we compute the distance between the instance and every training sample, which uses O(nd). Then, we run a quicksort on the distances which takes O(nlogn), and find the first k distances using O(k). Therefore, the time complexity is O(n*(nd+nlogn+k)).

Q4:

d)



Both standard scale and min max range scale improve the performance of KNN compared to non-preprocessed data, especially when k is in ranges of 1 - 4 and 9 - 10.

This dataset is not very sensitive to irrelevant features. Since most of the time, dataset with irrelevant features (red line) shows a similar accuracy with non-preprocessed dataset (blue line).