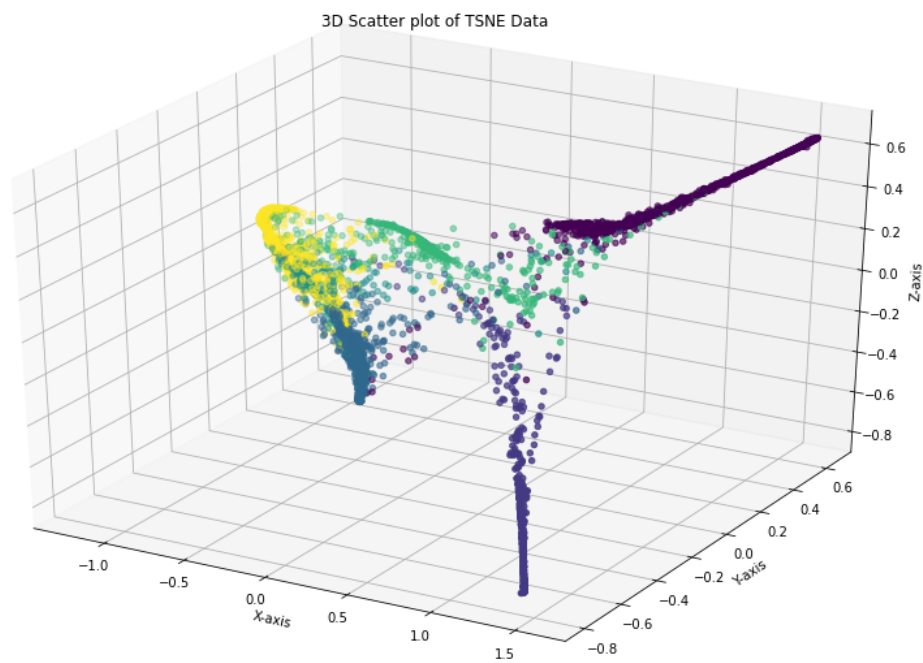
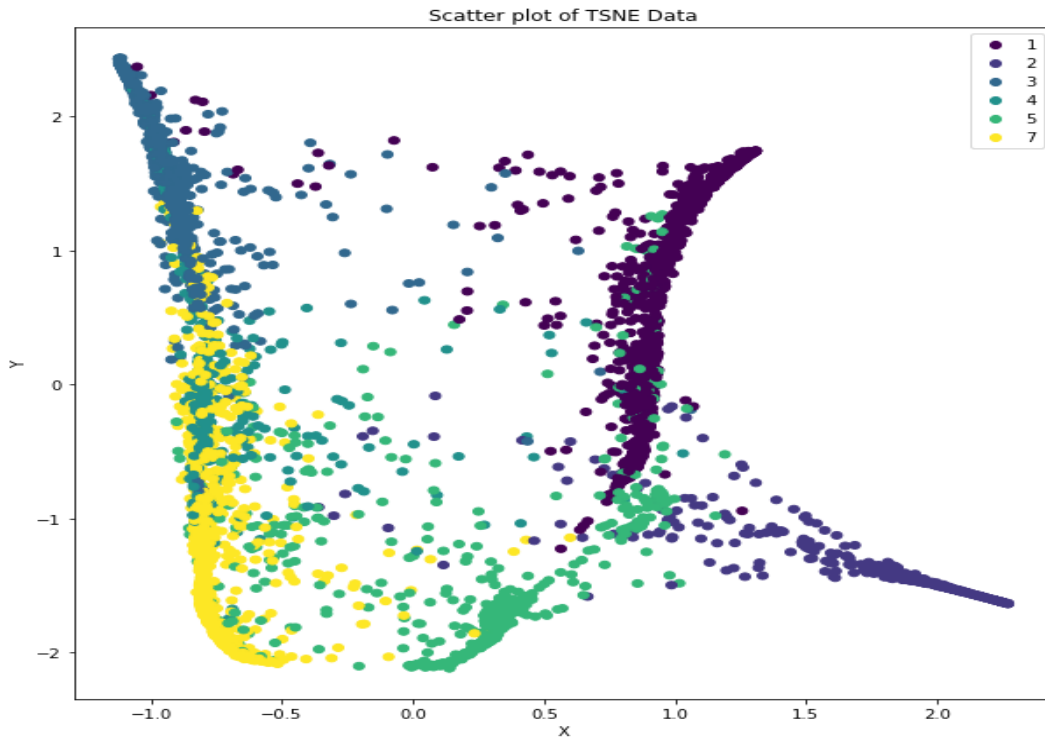
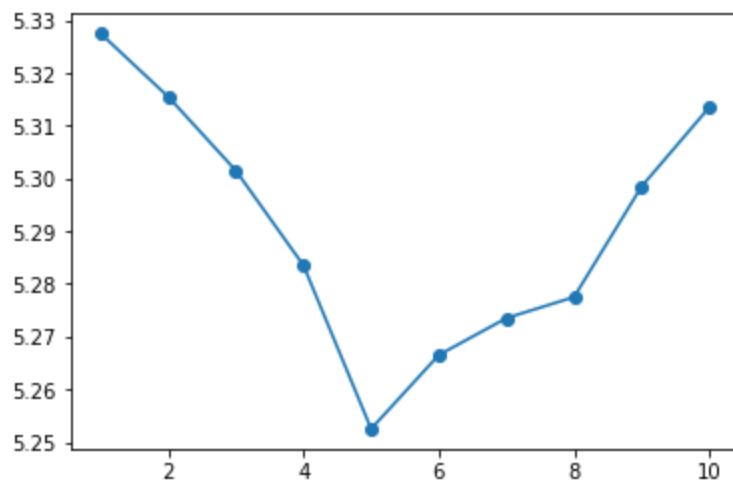


Answer 1:

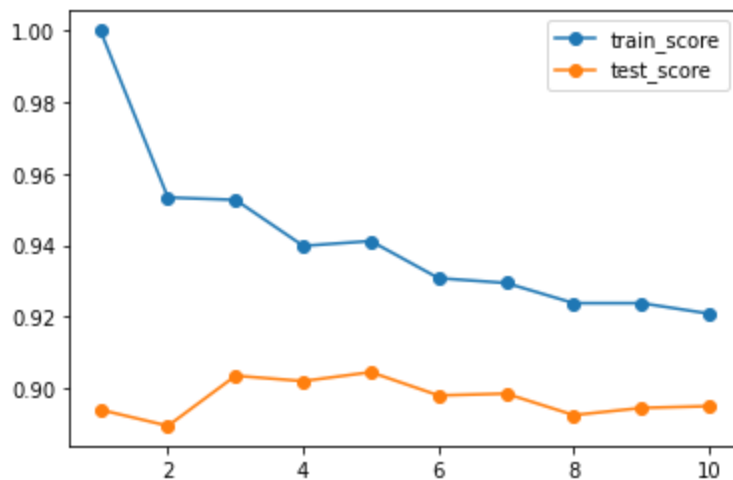
- (1) Loaded both the datasets (one for training called sat.trn and one for testing called sat.tst) by converting it to .xls files into the interface. Using TSNE plot on the dataset we get the following plots:



(2) By implementing grid search on self implemented KNN algorithm, we have plotted the error versus k plot as shown below:



Also as per accuracy of the k neighbors, we get the following graph as follows:



Hence, in these two graphs we see that the optimal number of $k = 5$ with training accuracy = 0.9411499436302142 and testing/validation accuracy = 0.9045

(3) We use k value as 5 for n_neighbors in sklearn implementation of KNN, to get the training and validation accuracies as shown below:

Training Accuracy = 0.9411499436302142

Validation Accuracy = 0.9045

Answer 2

- (1) Split the data into a train and test set with 80:20 using seed 42. Shown in code.
- (2) Implemented a NN architecture using sklearn with 3 hidden layers - [100, 50, 50].
Assumed a Sigmoid activation function in each layer. Accuracy and loss given below:

```
Train loss : 0.0006067526916367709
Test Loss  : 0.050758944335840495

Train accuracy : 1.0
Test accuracy  : 0.9898281304805332
```

- (3) Plotted the decision boundary for 3 different values of 'alpha' (2 extreme and 1 middle value) b/w the training samples.. Values taken for alpha - [0.0000001, 0.01, 1]

