CS412 - HOMEWORK 1 – AHMET ALPEREN GÜNGÖR - 28847

https://colab.research.google.com/drive/1jxHgqOSsx_2WsSekAWaOWzF-qT5K19aL?usp=sharing

In this task, our aim is to implement a classifier for MNIST dataset using k-NN algorithm. While doing so, we need to train the model with training set, then tune the hyperparameters using validation set.

The dataset has 60000 training and 10000 test data, with 28x28 = 784 features. The training set will be further divided into 2 parts: %80 development (which will be used to initially train the data) and %20 validation (which will be used to tune hyperparameters – which is number of neighbors, K).

Parameters that we tried to tune number of K are [1,3,5,7,9,11,13] and the accuracies (using validation data) are given below in a table:

K parameter	1	3	5	7	9	11	13
Accuracy	0.9721	0.9712	0.9695	0.9692	0.9676	0.9665	0.9661

As it can be seen from the table, k=1 has the highest accuracy on validation set with an accuracy of 0.9721. Setting the number of neighbor parameter as 1 (the k=1 hyperparameter), k-NN classifier model is then trained again, but using combined training (development) and validation set. The accuracy of this model on the test set is 0.9691.