Computer RA Task

a) Implement a K-Nearest Neighbor (KNN) classifier that uses the leave-one-out cross validation approach for determining the best K value. Apply the classifier to the training data of the scanned images of the 10 digits (0 to 9) provided in the file "Task Dataset.zip". The zip file contains two folders: "Train" and "Test". The "Train" folder contains 240 images for each digit while the "Test" folder contains 20 images for each digit. The folder contains a filed named "Training Labels.txt" which includes the labels of the 2400 images in order. The images in the "Train" folder should be used in the leave-one-out cross validation. Use maximum K of 100.

Deliverables:

• Your code.

• A plot of the classification error obtained for the training data during the validation process versus the choice of K. Name your file "KNN.jpg".

b) Use the test data to test your classifier. Apply your KNN classifier with the best value of K as obtained from part (a). The folder also contains a text file named "Test Labels.txt" which include the labels of the 200 images in order.

Deliverables:

• Your code.

• A confusion matrix showing the number of images of the Test folder of each digit that were classified to belong to different digits (For example: Number of images of 0 that were classified as 0, 1, 2, …, 9, and so on for other digits). Convert the confusion matrix to an image and save it as "Confusion.jpg".

Important Notes:

- Do not use Python functions for KNN classifier. You have to implement your own version of all needed functions. However, you are allowed to use the function that computes the norm of a vector or its equivalent and the sorting function.
- To speed up the process of your function, in part (a), you should first compute the distance between each image and all other images and store such values in some data structure. You can then start changing K and get the nearest neighbors of each image from the values you stored instead of re-computing the distances with every change of K.

- You should write a comprehensive **scientific report** introducing "KNN", discuss whether if it is appropriate for this task or not. Moreover, you should present your results in the best form through necessary figures and graphs (if any) from your point of view, and you should end it with an appropriate conclusion. This report is extremely important, and it is the deliverable with the highest weight.