In this homework, I implemented the one-versus-all support vector classification algorithm

**Part 2)**

Read Data from

hw06\_images.csv

hw06\_labels.csv

**Part 3)**

Data is divided into train and test set. first 1000 images to the training set and the remaining 4000 images to the test set.

**Part 4)**

I used lab08 code to implement one-versus-all support vector classification algorithm. First took Gaussian Kernel Function from lab08.

Text

Description automatically generated

Then calculated Gaussian Kernels for train and Test data.

Text

Description automatically generated

I defined OVA(c, s=10) function to find predictions for training and test sets. This code works similarly with lab08 code. However, to make it one-vs-all I iterated K times. At each step I set one label to 1, and others as -1. Then trained one- versus-all support vector classification algorithm to calculate train and test labels.

Text

Description automatically generated

get label scores using given formula in the class. Which is used also in lab08.

Text, letter

Description automatically generated

Text

Description automatically generated with medium confidence

**Part 5-6)**

After getting predict labels scores as train=(5,1000) test=(5,4000) array. Get predicted label using argmax. Finally printed confusion matrix using pd.crosstab.

A picture containing table

Description automatically generated

**Part 7)**

I this part, I draw the classification accuracy for training and set data points as a function of 𝐶(0.1, 1, 10, 100 and 1000) with the kernel width parameter 𝑠 to 10. For each parameter I run Ova function then calculated accuracy errors from these predictions for train and test sets.

Graphical user interface, text, application, email

Description automatically generated

Then plotted accuracy values of train and test sets for different Regularization C parameters. Increasing regularization parameter(C), gives better results in terms of accuracy for train set. Increasing C, we increase penalization of mispredicted labels. That’s why it is going to be overfitting. That causes decrease in test set after parameter C = 10.

Chart, line chart

Description automatically generated