CSE 574: Introduction to Machine Learning
Programming Assignment 3

Group Members:

Name: Karan Rajendra Shah
UB Number: 50354506

Name: Pranav Moreshwar Sorte
UB Number: 50413353

Name: Thelma Melissa Gomes
UB Number: 50416894

Q1. Logistic Regression

Task → To implement Logistic Regression to classify hand-written digit images into correct corresponding labels.

1. Record the total error (the error returned from the blrObjFunction()) with respect to each category (the 10 categories in MNIST) in both training data and test data

Category/Class	Total error in training data	Total error in testing data
0	0.020149624265758993	0.0019206466531900303
1	0.021309752958694077	0.0015799673798645329
2	0.06200516976608354	0.042991523158585544
3	0.07541099797081077	0.042991523158585544
4	0.04432670984757564	0.0210129054863936
5	0.08306663204896231	0.056517851333542694
6	0.03413941560035174	0.012581380304563667
7	0.04320632115433112	0.02074643931589003
8	0.110407760048935	0.08357236686358212
9	0.09773073457744491	0.07404120465980162

2. **The reason why there is difference between training error and test error:**
In Logistic Regression we consider all data points in training data while creating the hyperplane that separates the data as per their classes. It learns a hyperplane which tends to overfit on the learning data. That's why we observe the difference in training error and testing error

3. **Accuracy**
Training set accuracy: 92.712%
Validation set accuracy: 91.491%
Test set accuracy: 91.94%

Q2. Support Vector Machine

Task → Perform Classification using Linear and RBF SVM by using scikit-learn.

Kernel	Gamma	Training Accuracy	Validation Accuracy	Test Accuracy
Linear	0	99.83%	91.43%	91.99000000000001%
RBF	0	98.79%	96.05%	96.16%
RBF	1	100.0%	10.02%	11.360000000000001%

Analysis of the above observations –
If we compare the result of RBF kernel with gamma as default and linear kernel, we can observe that RBF kernel SVM with gamma as default gives a much better performance as compared to linear model. If we try to increase the gamma of RBF kernel to 1, we can observe that the model tends to overfit the data. The generalizability of the RBF SVM model with gamma = 1 is very poor as it gives 100% accuracy on training data and fairs very poorly on testing and validation data

Task → Perform RBF SVM with C values ranging from 1 to 100

C	Training Accuracy %	Validation Accuracy %	Testing Accuracy %
1	98.83%	96.08%	96.21%
10	100%	96.92%	96.3%
20	100%	96.89%	96.92%
30	100%	96.89%	96.92%
40	100%	96.89%	96.92%
50	100%	96.86%	96.93%
60	100%	96.865%	96.931%
70	100%	96.84%	96.92%
80	100%	96.87%	96.91%
90	100%	96.91%	96.90%
100	100%	96.83%	96.92%

Analysis of the above observations –
After performing the iteration over C on the SVM RBF kernel, we observe that after a certain range of C values, there is no such difference in the training, testing and validation accuracies. The first value of C i.e. 1 tends to give a better model in terms of generalizability. After the 1st value of C, the model overfits the data which is treated as poor performance.

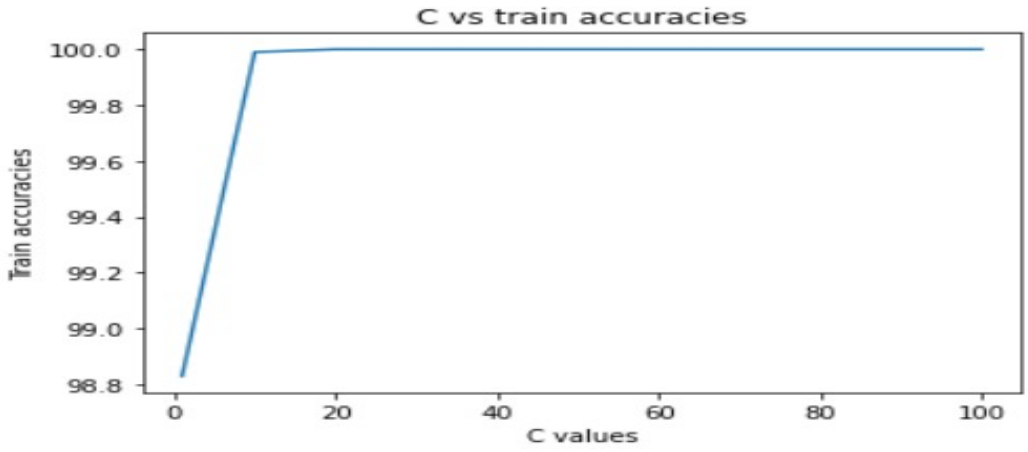
After careful observations in the above experiments using SVM, we choose the below hyper parameters

Kernel – RBF
Gamma – default
C – 1

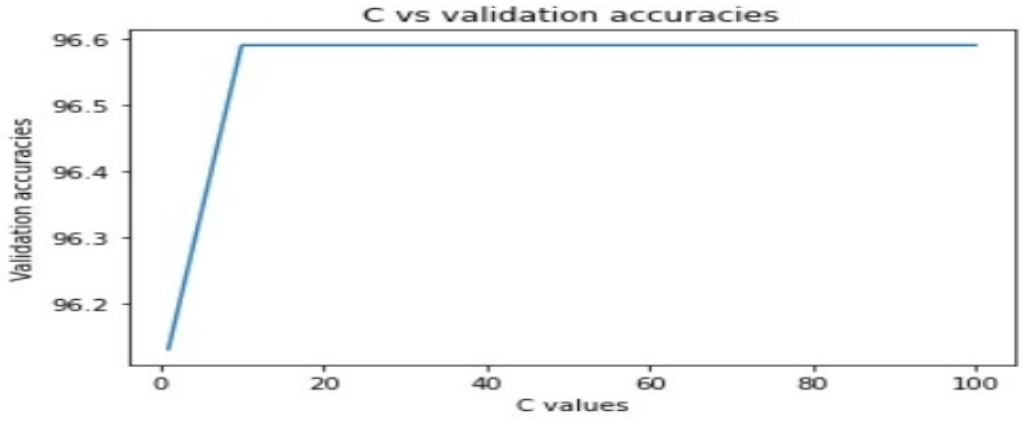
Performing SVM on entire data using best hyperparameters -

Kernel	Gamma	C	Training accuracy	Validation accuracy	Test accuracy
RBF	0	1	98.982%	97.89%	97.87%

Train accuracy:



Validation accuracy:



Test accuracy: