CSE 474 Machine Learning
Project Assignment 3
Group #45

By
Fatin Nazat
Ali Morshid

1. Logistic Regression

Logistic Regression with Gradient Descent

Training Set Accuracy: 85.544%

Validation Set Accuracy: 85.09%

Testing Set Accuracy: 85.14%

In your report, record and discuss classification results and accuracy.

Using logistic regression, our experiment resulted in 85% accuracy across all fields. Logistic regression considers all data points for determining a hyperplane that separates the data.

Direct Multi-class Logistic Regression (Extra Credit)

In your report, record and discuss classification results and accuracy.

Training set Accuracy:92.628%

Validation set Accuracy:92.26%

Testing set Accuracy:92.17%

Discuss the performance of multi-class logistic regression compared to the performance of logistic regression when using the one-vs-all strategy.

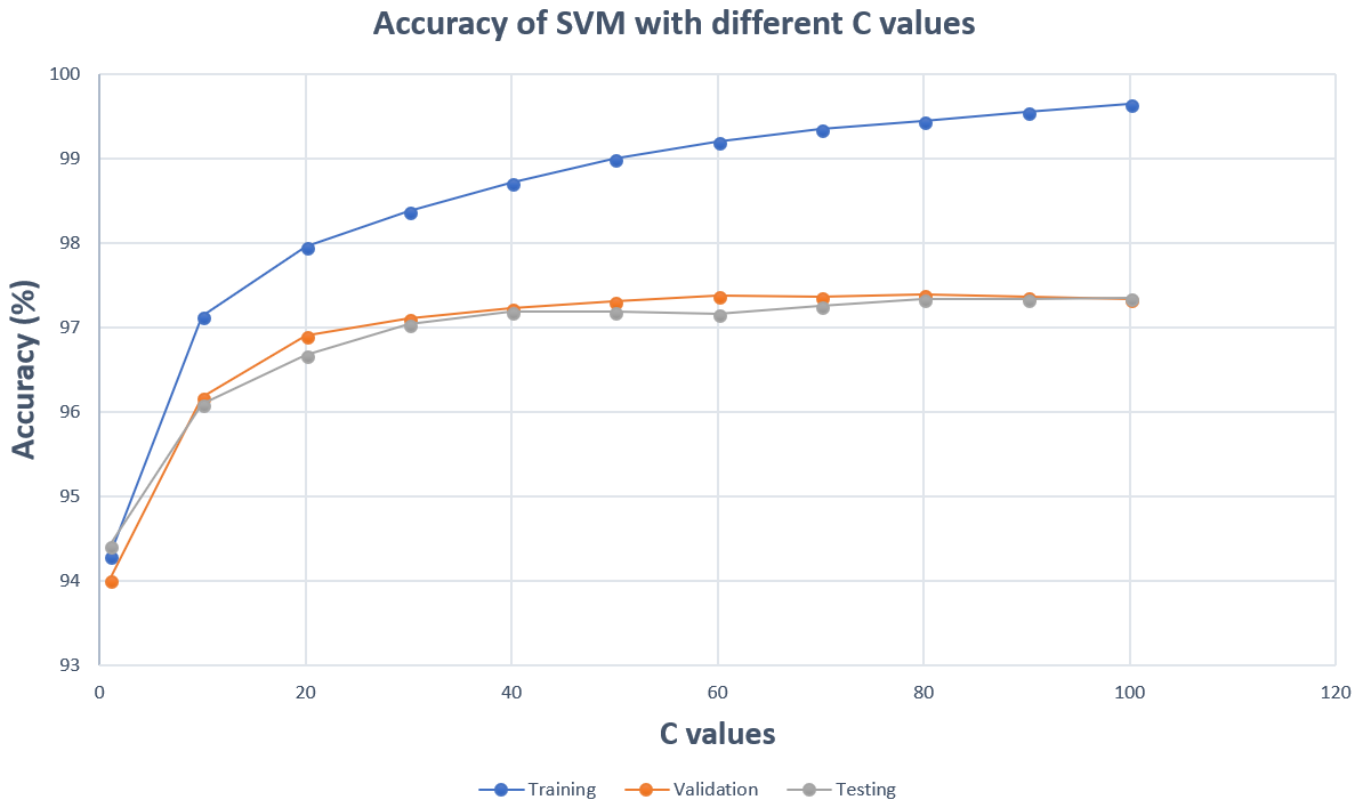
When using multi class logistic regression, we achieved 92% accuracy for training, validation, and testing. According to the results, multi class regression performs better than linear regression. Even when using the one vs all strategy and breaking it down to different categories it is more optimal to use multi classification instead.

2. Support Vector Machines

Kernel	**Training Accuracy (%)**	**Validation Accuracy (%)**	**Testing Accuracy (%)**
Linear	97.286	93.64	93.78
Radial Basis (Gamma = 0.1)	99.992	94.76	94.96
Radial Basis (Default)	94.294	94.02	94.42

Discuss results, comparing the selections of linear kernel and radial basis function kernel.

As we can see Radial Basis function kernel with gamma 0.1 performs the best in training, validation, and testing set. In comparison, default RBF has the lowest training accuracy even compared to Linear kernel. Therefore, gamma control does influence RDF's accuracy. In terms of validation and testing accuracy Linear kernel seems to perform the lowest.



In your report provide justification for the above selection of hyperparameters, as well as plots and discussion of your results.

By looking at the graph we can see that higher the C values greater the accuracy. This is because the value of C affects the weight of the error term. When the value of C increases a smaller margin hyperplane will be created and a higher margin hyperplane when C decreases. This directly affects the accuracy of testing, validation, and particularly training set. This may not be always the case because when C values are from 40-60, testing set decreases which might result in overfitting.