CSE 574:Programming Assignment 3

Instructor: Dr. Mingchen Gao

Classification and Regression

Submitted by:
Muthuvel Palanisamy
Person # : 50246815

Due Date : May 10, 2018

University at Buffalo
The State University of New York

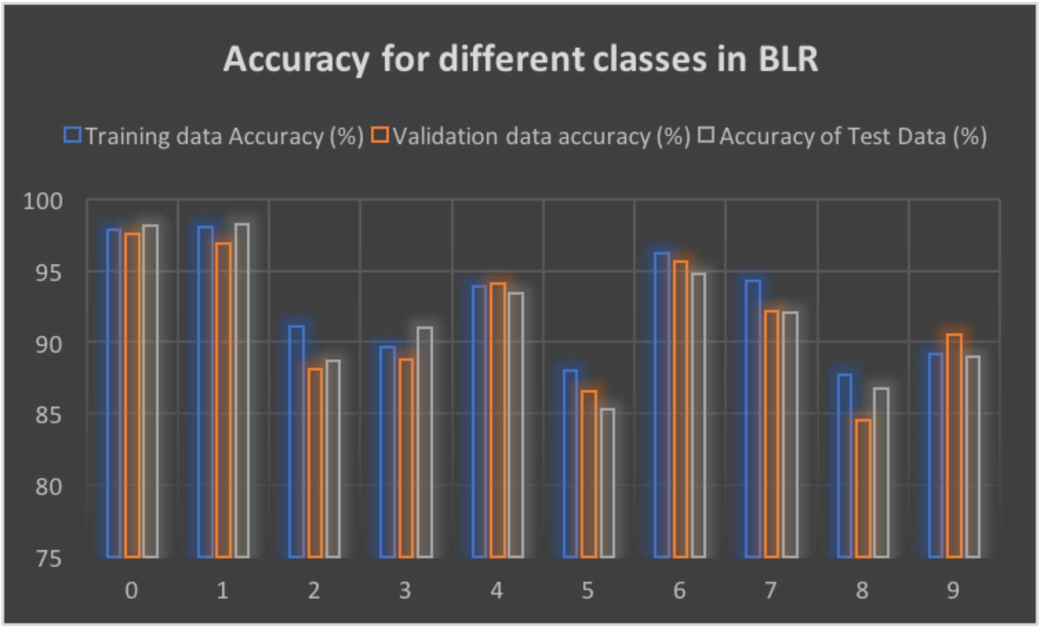

Figure 1: Comparison of for different classes in Binomial Logistic Regression

1 Logistic Regression

1.1 Binomial Logistic Regression(40 code + 15 report = 55 points)

Final Error from Gradient Descent = **0.09734**

Training Data Accuracy = **92.75**%

Validation Data Accuracy = **91.49**%

Test Data Accuracy = **91.87**%

0 and 1 are among the classes with the highest accuracy while digits like 8 and 5 have lower accuracy in Binary Logistic Regression. Digit 8 has the lowest accuracy in validation data.

For many classes, the training error is slight lower than the test error. See Figure 1 for Comparison of Accuracy of Different classes

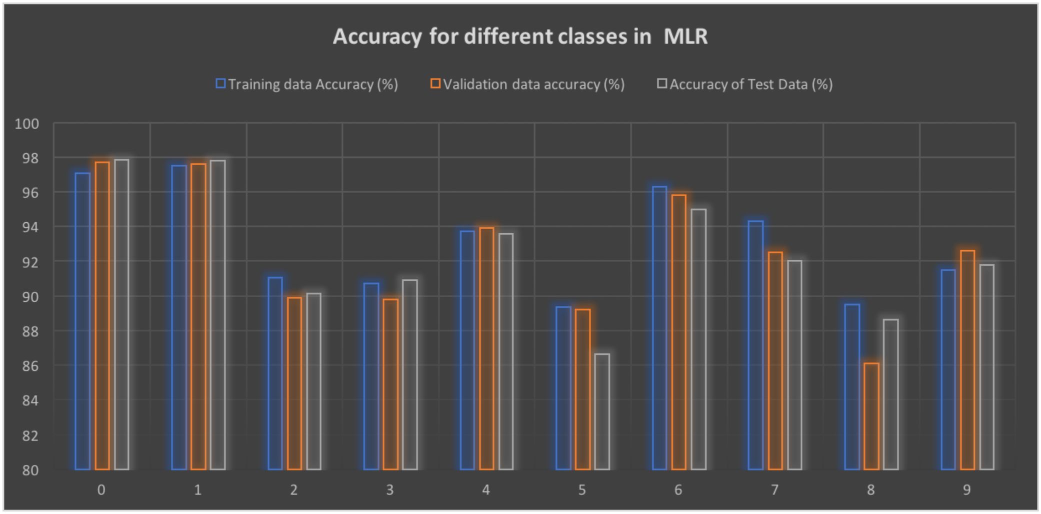

Figure 2: Comparison of for different classes in Multinomial Logistic Regression

1.2 Multinomial Logistic Regression(10 code + 10 report = 20 points)

Final Error from Gradient Descent = **0.02458**

Training Data Accuracy = **91.86**%

Validation Data Accuracy = **93.37**%

Test Data Accuracy = **93.95**%

Similar to Binomial Logistic Regression, 0 and 1 are among the classes with the highest accuracy while digits like 8 and 5 have lower accuracy in Multinomial Logistic Regression. But Multinomial logistic regression tends to produces more similar accuracies in training as well as testing comparatively for data from a same class. See Figure 2 for Comparison of Accuracy of Different classes

1.3 Comparison and Analysis

* On comparing Binomial and Multinomial Logistic Regression, the latter produces slightly better accuracy on the test data.

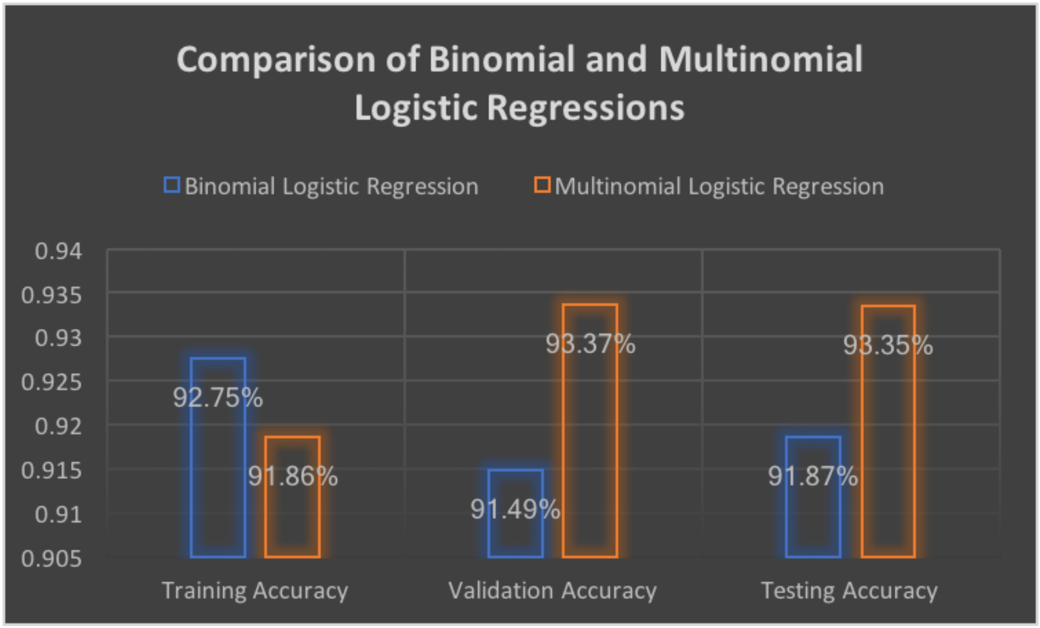

Figure 3: Comparison of the Accuracy of Regression Models

* Both perform well and give out an accuracy greater than 90%. See Figure 3 for comparison chart.
* Though not applicable to this case, a very low training error generally causes over fitting which may need regularization.
* Multinomial Logistic Regression is clearly faster that its binary counterpart

2 Support Vector Machines(20 code + 25 report = 45 points)

2.1 Using Linear Kernel (with default parameters)

Training Data Accuracy = **97.29**%

Validation Data Accuracy = **93.64**%

Test Data Accuracy = **93.78**%

2.2 Using Radial Bias Function

2.2.1 When $\gamma = 1$

In Radial Bias Function, when $\gamma = 1$, the accuracy is very low for validation and test data data: over fitting occurs when $\gamma = 1$. Hence it is ideal to avoid this setting in SVM with Radial Kernel

Training Data Accuracy = **98.15**%

Validation Data Accuracy = **12.12**%

Test Data Accuracy = **13.76**%

2.2.2 When γ = default

Training Data Accuracy = **94.30**%

Validation Data Accuracy = **94.02**%

Test Data Accuracy = **94.42**%

2.2.3 When C = [1, 10, 20,..., 100]

As the value of C increases from 0 to 100, the accuracy increases gradually and converges to an optimum vale. We can infer that C needs to sufficiently larger, but too large values of C does not have any much impact on the accuracy after a point. Table 1 shows detailed output accuracies when C value is varied from 1 to 100. Figure 4 shows how accuracy varies for different values of C.

2.3 Comparison and Analysis

* $\gamma = 1$ causes overfitting and hence produced very low test accuracy - this might because large gamma leads to high bias and low variance models

* Linear SVM is comparatively faster than than radial. The latter with gamma = 1 took around 3 hours to run for the MNIST dataset

* Radial SVM provides better accuracy in comparison with Linear SVM. This might be because Linear SVM might under fit the data

C	Training Accuracy	Validation Accuracy	Test Accuracy
1	94.30%	94.02%	94.42%
10	97.13%	96.18%	96.10%
20	98.95%	96.90%	96.67%
30	98.37%	97.10%	97.04%
40	99.10%	97.23%	97.19%
50	99.00%	97.31%	97.19%
60	99.29%	97.38%	97.16%
70	99.34%	97.36%	97.26%
80	99.43%	97.39%	97.33%
90	99.54%	97.36%	97.34%
100	99.61%	97.41%	97.40

Table 1: SVM with Radial Bias Function

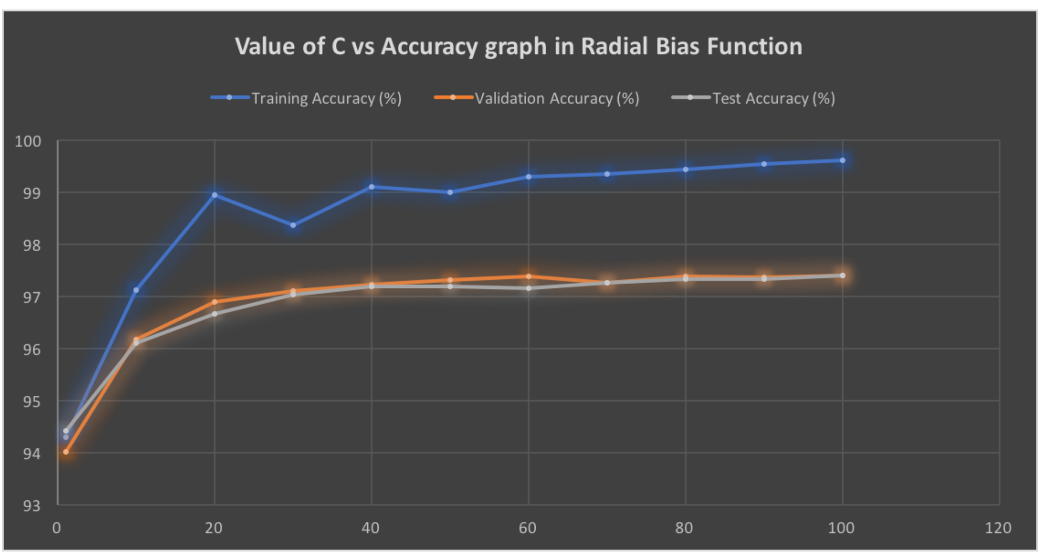

Figure 4: C vs Accuracy in SVM with Radial Bias Function

Model	Training Accuracy	Validation Accuracy	Test Accuracy	Time (s)
BLR	92.75%	91.59%	91.87%	706
MLR	91.86%	93.37%	93.95%	150
Linear SVM	97.29%	93.64%	93.78%	1018
Radial SVM	94.30%	94.02%	94.42%	1270

Table 2: Comparison of Different Classifiers

* Radial SVM with properly tuned parameters will always be better than Linear SVM if not equal

* It is clear that the accuracy of Training data is always greater than that of testing data from Figure 4

3 Comparison of SVM and Logistic Regression

* Radial SVM provides the best accuracy but takes more time to train. Check table 2.

* Multinomial Logistic Regression takes the least time with good efficiency. Refer Figure 5 and 6

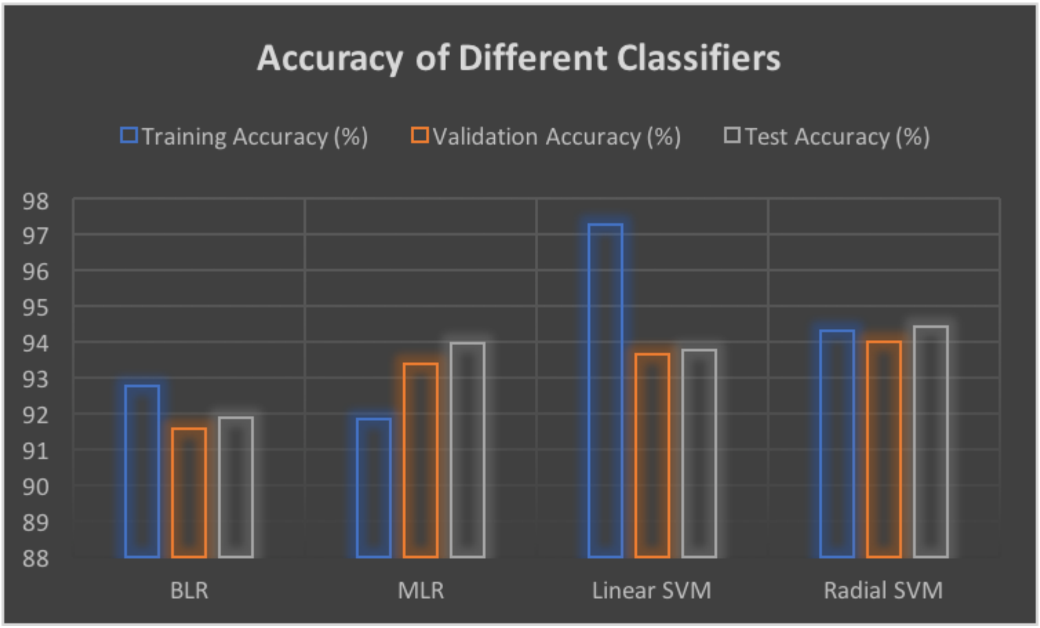

Figure 5: Accuracy of Different Classifiers

Figure 6: Time Taken to train different Classifiers