CSE 574
Introduction to Machine Learning
PA 3

Spring 2019
05/10/2019

Classification And Regression

Project Report

Prepared By:
Md Moniruzzaman Monir (50291708)
Archisman Patra (50291555)
Tushar Gupta (50291152)

Binary Logistic Regression(BLR):

Set	Accuracy	Error
Training	92.746%	7.254%
Validation	91.51%	8.49%
Testing	91.88%	8.12%

Above are the results after running Binary Logistic Regression on Training, Validation and Testing Data. We can see that the Training error is less than Testing error. So, we can infer that this Linear model performs better on the seen data, but when it gets the unseen dataset it gives a little more error(though there is not much of a difference). That's kind of natural for any Linear Model and there is not huge difference in error as the data is of similar type in all sets.

Multi-class Logistic Regression (MLR):

Set	Accuracy	Error
Training	93.11%	6.89%
Validation	92.38%	7.62%
Testing	92.53%	7.47%

Above are the results after running Multi-class Logistic Regression on Training, Validation and Testing Data. We can see that the Training error is slightly less than Testing error. So, we can infer that this Linear model performs better on the seen data, but when it gets the unseen dataset it gives a little more error. That's kind of natural for any Linear Model and the errors are almost equal as the pattern of data is same for all the sets.

Performance difference between multi-class strategy(MLR) with one-vs-all(BLR) strategy:

Set	MLR Accuracy	BLR Accuracy
Training	93.11%	92.746%
Validation	92.38%	91.51%
Testing	92.53%	91.88%

- In multiclass logistic regression we classify all the classes(total 10) of MNIST dataset at once, whether in one-vs-all(BLR) we only classify one class with respect to all other at a time, so multiclass has less time complexity and less chances of overlapping.

- We observed the accuracy of the multiclass was better than the BLR classification. That's because parameters are estimated independently in multiclass which helps to prevent wrong classification.

Support Vector Machine (SVM):

I. Using Linear Kernel:

Set	Accuracy
Training	92.574%
Validation	91.23%
Testing	91.19%

So, we can infer from the above results that Linear Kernel works like a linear model, as the results are almost same as the previous linear model we trained.

II. Radial Basis Function:

(a) Using Radial Basis Function (Gamma = 1)

Set	Accuracy
Training	100.0%
Validation	10.39%
Testing	11.73%

This setting gives very poor result on test data as this high value of gamma helps in overfitting the training data and we can infer that from the 100% Training accuracy.

(b) Using Radial Basis Function with value of gamma setting to default (all other parameters kept as default)

Set	Accuracy
Training	91.89%
Validation	92.05%
Testing	92.24%

(c) Using Radial Basis Function with value of gamma setting to default and varying value of C (1, 10, 20, 30, ...,100)

Now we iterate through the C values and record the optimum setting and then test the whole data on that setting. This C variable controls the importance we are giving to the Slack variable. So, here we observe a trade-off between the width of the margin and C value.

Below are the results for different values of C on Training, Testing and Validation

data:

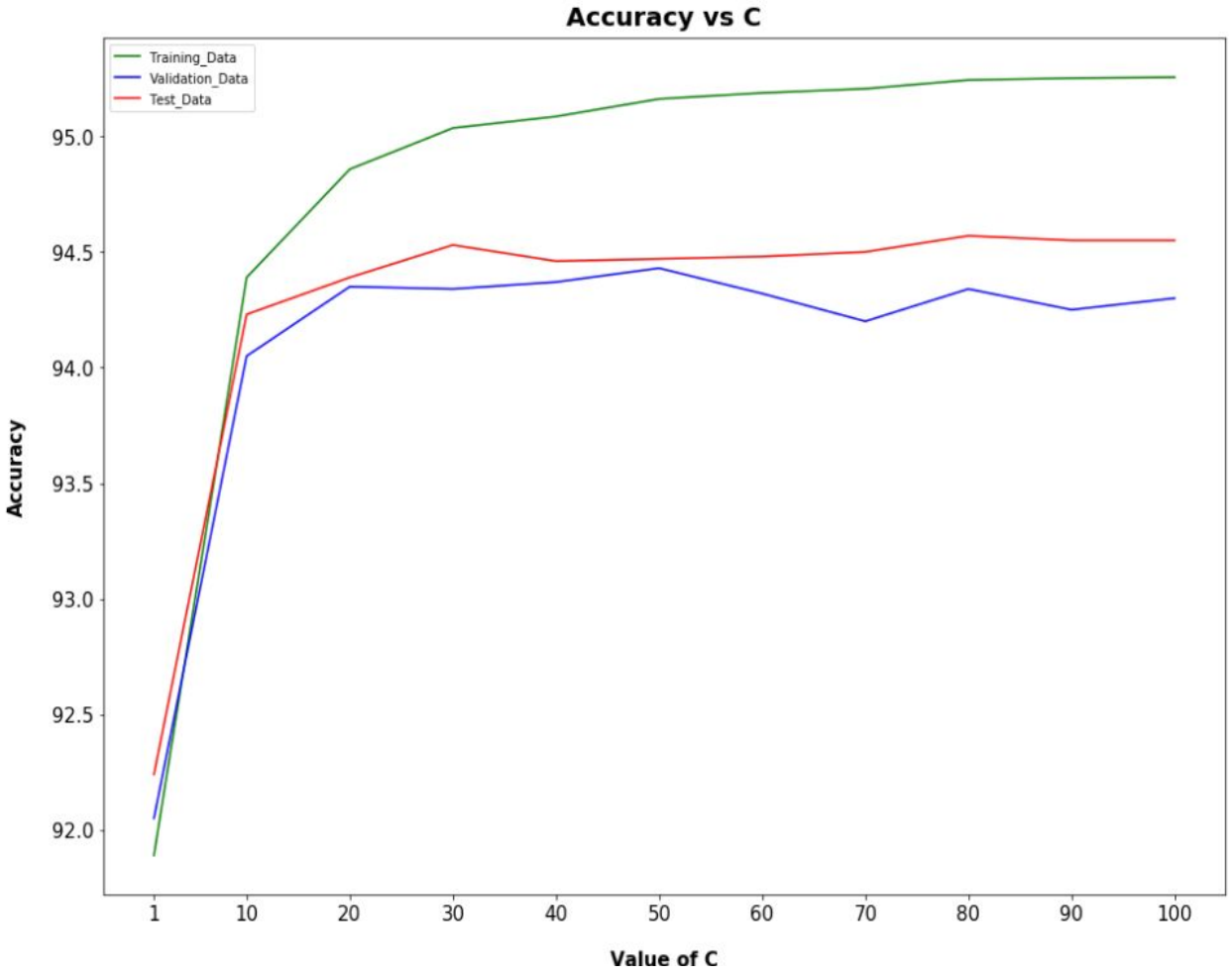
C	Training Accuracy	Validation Accuracy	Testing Accuracy
1	91.89%	92.05%	92.24%
10	94.39%	94.05%	94.23%
20	94.858%	94.35%	94.39%
30	95.036%	94.34%	94.53%
40	95.086%	94.37%	94.46%
50	95.162%	94.43%	94.47%
60	95.187%	94.32%	94.48%
70	95.206%	94.19%	94.5%
80	95.244%	94.34%	94.57%
90	95.252%	94.25%	94.55%
100	95.256%	94.3%	94.55%

So, we can conclude that we are getting the best result by setting gamma to default and taking C = 80

Results for the whole dataset using optimal parameters:

Kernel	C	Training Accuracy	Validation Accuracy	Testing Accuracy
RBF(Gamma=default)	80	99.34%	97.36%	97.26%

Plot of accuracy obtained on each of Training, Testing and Validation dataset with respect to various values of C:



So, we can also infer that our dataset is non-linear as it gives better result on this non-linear model.