Assignment 3 Report

For Logistic Regression, these are the accuracy values that we have obtained for train\_data, validation\_data and test\_data.

|  |  |
| --- | --- |
| Training Set Accuracy | 92.328% |
| Validation Set Accuracy | 91.46% |
| Testing Set Accuracy | 91.92 |

Onto SVM, we will break down each part into different charts to make it easier to see.

Linear Kernel

|  |  |
| --- | --- |
| Training Set Accuracy | 97.286% |
| Validation Set Accuracy | 93.64% |
| Testing Set Accuracy | 93.78% |

Radial Basis Function when Gamma equals 1

|  |  |
| --- | --- |
| Training Set Accuracy | 100% |
| Validation Set Accuracy | 15.48% |
| Testing Set Accuracy | 17.14% |

Radial Basis Function when Gamma is default

|  |  |
| --- | --- |
| Training Set Accuracy | 94.294% |
| Validation Set Accuracy | 94.02% |
| Testing Set Accuracy | 94.42% |

Radial Basis Function when Gamma is default and C equals to 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100

We will separate each accuracy into its own chart for the 100 values

Training Set Accuracy

|  |  |
| --- | --- |
| C = 1 | 94.294% |
| C = 10 | 97.132% |
| C = 20 | 97.952% |
| C = 30 | 98.372% |
| C = 40 | 98.706% |
| C = 50 | 99.002% |
| C = 60 | 99.196% |
| C = 70 | 99.34% |
| C = 80 | 99.438% |
| C = 90 | 99.542% |
| C = 100 | 99.612% |

Validation Set Accuracy

|  |  |
| --- | --- |
| C = 1 | 94.02% |
| C = 10 | 96.18% |
| C = 20 | 96.9% |
| C = 30 | 97.1% |
| C = 40 | 97.23% |
| C = 50 | 97.31% |
| C = 60 | 97.38% |
| C = 70 | 97.36% |
| C = 80 | 97.39% |
| C = 90 | 97.36% |
| C = 100 | 97.41% |

Testing Set Accuracy

|  |  |
| --- | --- |
| C = 1 | 94.42% |
| C = 10 | 96.1% |
| C = 20 | 96.67% |
| C = 30 | 97.04% |
| C = 40 | 97.19% |
| C = 50 | 97.19% |
| C = 60 | 97.16% |
| C = 70 | 97.26% |
| C = 80 | 97.33% |
| C = 90 | 97.34% |
| C = 100 | 97.4% |

Here are the graphs for training, validation and testing accuracy

Here is a graph with all 3 accuracies side by side.

From the chart, you can tell that, as the value of C increases, the accuracies tend to increase as well. The test\_data and validation\_data accuracy slightly decreases at a certain C values, but it jumps right back up when it hits the next C value.