CSE 474

Group #27

* Dakota Handzlik
* Danyang Luo
* Nicholas Phang

REPORT FOR ASSIGNMENT 1

Part 1: Linear Regression

1. Direct Minimization

The results of RMSE for training and test data for two cases are as follows:

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| With an Intercept | 46.77 | 60.89 |
| Without an Intercept | 138.20 | 326.67 |

As can be seen from the data, when the intercept is added to the linear model, the accuracy of the training and testing of the model is greatly improved.

1. Gradient Descent

The result of RMSE for training and test data after gradient descent based learning are as follows:

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| Gradient Descent | 48.10 | 54.98 |

It can be seen that when training the model. The results obtained will be slightly inferior to the results of direct minimization. But at the same time, the test results will be slightly better than the minimized results.

Part 2: Linear Classifiers

1. Perceptron Learning

The result of the minimized perceptron model are as follows:

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| Perceptron | 0.84 | 0.84 |

1. Logistics Regression Learning

The result of the minimized logistics regression model are as follows:

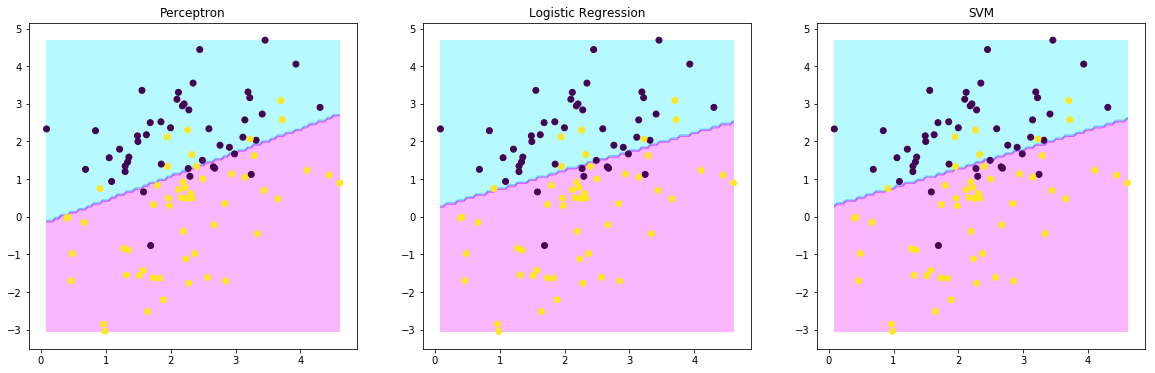
|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| Logistics Regression | 0.84 | 0.85 |

1. Support Vector Machine

The result of the linear support vector machine, using stochastic gradient descent method are as follows:

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| Logistics Regression | 0.86 | 0.86 |

Part 3: Comparing Linear Classifier

1. By comparing the results of the three linear classifiers, it can be seen that the support vector machine has the best classification performance.
2. 

As can be seen from the image, when the three classifiers are small for the amount of data, the results of the classification are roughly similar, but the results of the support vector machine will eventually be better. Through analysis, we can see that the first algorithm: the perceptron algorithm is mainly done by finding the RMSE minimum. The advantage of this method is that the algorithm is relatively straightforward, but the final performance is relatively low. The second algorithm: is to establish a relationship between the output index by the result of the function prediction. The advantage of this approach is that you can better learn the underlying laws and ultimately better describe the objects of modeling. The third algorithm: the support vector machine model that maximizes the soft interval is transformed into a linear support vector machine original optimization problem by the hint loss function and the addition of regularization terms. The hinge function is not only classified correctly but also when the confidence is high enough, the loss is zero. In other words, the hinge loss function has higher requirements for learning. However, by changing the value of T and the value of eta, it can be seen that the classification result of the support vector machine using the stochastic gradient descent is not stable.