CSE 574 Introduction to Machine Learning

Programming Assignment 1 Report

(Group 48- Pavithra Harindranath , Vikas Deep Shukla, Prakul)

Problem 1: Linear Regression with Direct Minimization

Results:

RMSE without intercept on train data - 138.20
RMSE with intercept on train data - 46.77
RMSE without intercept on test data - 326.76
RMSE with intercept on test data - 60.89

We can see that RMSE is considerably lower when we use an intercept for both train and test data. A low RMSE indicates a better fit to the data. Therefore, the model fits better with an intercept for this scenario.

Problem 2: Using Gradient Descent for Linear Regression Learning

Results:

Gradient Descent Linear Regression RMSE on train data - 48.20
Gradient Descent Linear Regression RMSE on test data - 54.90

While comparing RMSE on train data, we can see that Gradient Descent Linear Regression RMSE is lower than RMSE without an intercept after direct minimization, but RMSE with intercept is slightly lower in direct minimization when compared to Gradient Descent Linear Regression RMSE.

Whereas, on the test data, Gradient Descent Linear Regression RMSE is better than RMSE after direct minimization(both with and without intercept), which means that Gradient Descent learning is the better choice between the two as it has a lower RMSE.

Problem 3: Using Gradient Descent for Perceptron Learning

Results:

Perceptron Accuracy on train data - 0.84
Perceptron Accuracy on test data - 0.84

Problem 4: Using Newton's Method for Logistic Regression Learning

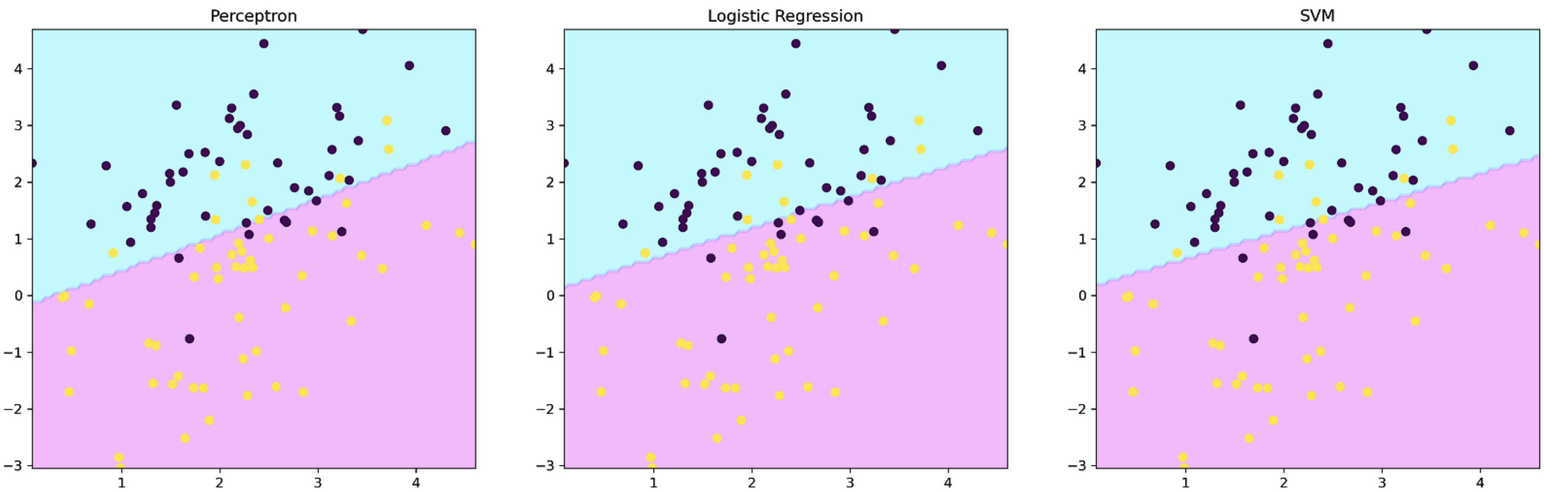
Results:

Logistic Regression Accuracy on train data - 0.84
Logistic Regression Accuracy on test data - 0.86

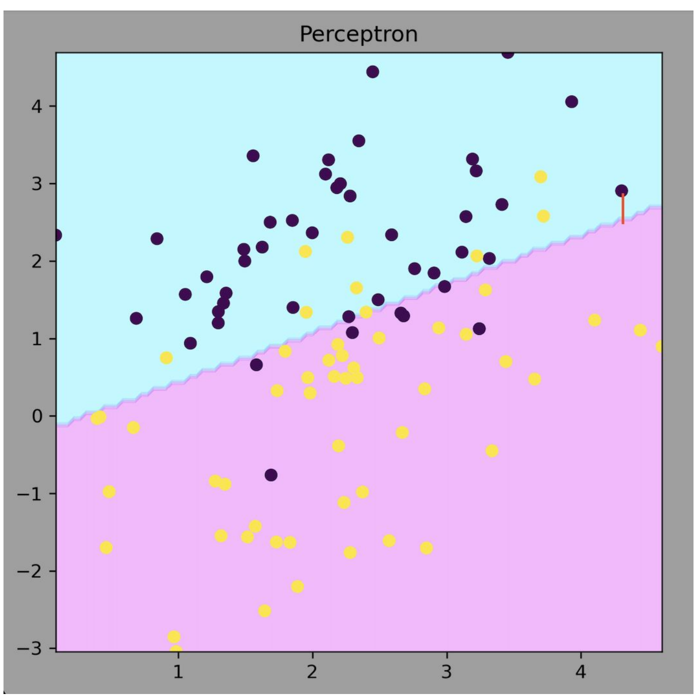
Problem 5: Using Stochastic Gradient Descent Method for Training Linear Support Vector Machine

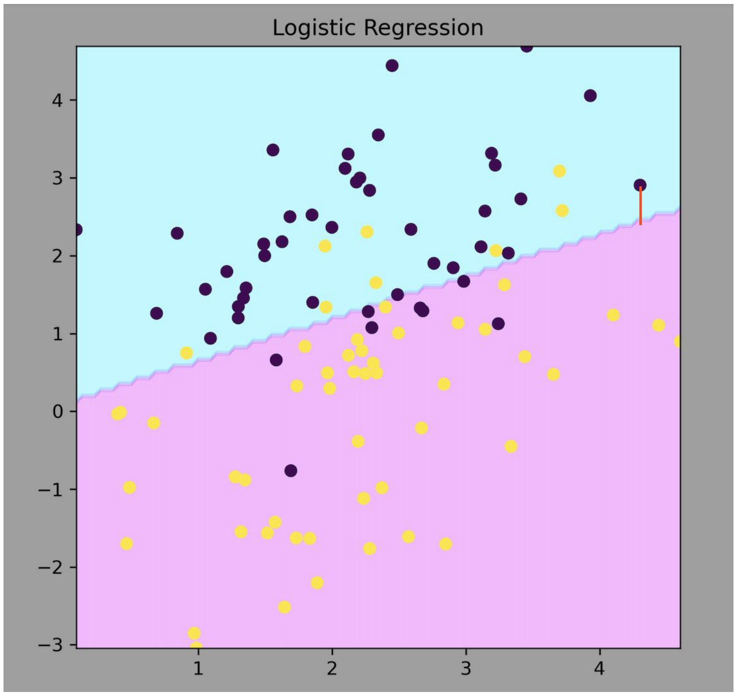
Results:
SVM Accuracy on train data - 0.85
SVM Accuracy on test data - 0.87

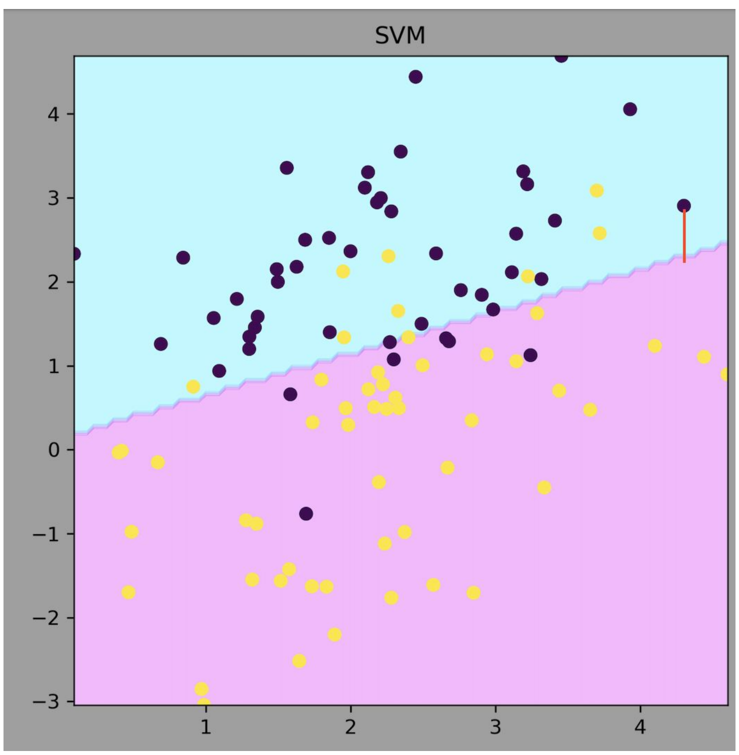
Problem 6: Comparing Linear Classifiers



<u>Observation</u>:







‘SVM’ model shows higher accuracy among all the models.

On comparing the decision boundaries for all three models, SVM model shows the best decision boundary(Marked by distance in red.)

Upon closer inspection, we observe that the decision boundary is fairly straight and has reduced unevenness when compared to other models.

Perceptron having the most jittery decision boundary and SVM having the least.
If the decision boundary is too jittery--->it could lead to overfitting of the model.