Homework 5

Your Name: \_\_A20511552\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student ID:\_\_Vikas Sanil\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**(100 points) Python practice for classification**

**Use the Loans data, and run logistic regression and SVM techniques to find the best parameters and performance**

* Use Loans\_20K.csv data by using 10-fold cross validation
* Use Loans\_200K.csv data by using 75% as training, 25% as testing
* Loan term as label

Note:

* You need to change different/multiple parameters to find the best model.
* You can find data sets from “slide & data” on blackboard system

Submission

* The ipynb and saved html files

A comparison of different parameters and metrics (accuracy, F1)

Comparison of different parameters and metrics for Loans\_20K.csv data.



From the above result we can see "linear" kernal has more accuracy when margin is soft.

From the above result we can see "rbf" kernal has more accuracy when margin is hard.

Comparison of different parameters and metrics for Loans\_200K.csv data.



From the above result we can see "linear" kernal has more accuracy as iteration is increased.

From the above result we can not conclude which parameters have more impact to "rbf".