

READ ME

Name: SUDHEERA MADDILA

G-Number: G01324503

This file contains an explanation of all the files in this submission:

Input File:

- fake_job_postings.csv (not included in the zip as mentioned in the instructions)

Output Files:

- Final output – It contains the output of every executed section (could not capture all the graphs)
- Graphs output – It contains the output of every executed section (captured all the graphs)

The entire code is in a single '.pynb' file named 'CS657 Assignment 2'. Almost every section of the code here has commented lines to give a brief explanation of what's happening in the code.

Conclusion:

The classifiers we evaluated produced results for the classification report that are good for both classes while initially due to the class imbalance the classifiers had some trouble in producing reliable results. We can also see that in specific cases bagging and boosting slightly improved the result of the classifier.