**185C Project Instructions:**

Software Requirements:

1. Python v 3.5+
2. Jupyter notebook
3. Pyspark
4. Pandas
5. Numpy
6. Matplotlib.pyplot
7. Seaborn
8. wordcloud
9. Re
10. BeautifulSoup
11. Findspark
12. Warnings
13. Pyspark.sql

System Requirements:

8gb RAM

7 cores

Step1) Place the data set in the same folder where the code is present

Step2)Run the code

Step 3) Do not run one of the Model as it takes over 6 hours to execute (Warning is mentioned near the Model in the code)