Assignment for week 13

Your given a training data set for some reviews (including labels). Your task is to use all your knowledge and skills you have gained so far in this course and try to predict the sentiment. You will need to predict the sentiments on the test data set. Test dataset is provided (without labels). Your predictions will then be evaluated live in the class. The group who achieves the highest ‘accuracy’ on the test data set will be the winner.

What you need to submit:

A csv file having the prediction for the test data. Please follow the naming conventions in the training data set. You will be evaluated on the following

1. data cleaning, 10%
2. exploratory data analysis 10%
3. feature engineering 10%
4. modeling process (you can use any modeling approach, custom, pretrained, NN, Sklearn ,pycaret, spark etc.) 10%
5. Accuracy on test: 60%

Please reach out to me if you have any questions