General Instructions  
1. In case the link to dataset is not useful, same dataset can be downloaded from any online resource.  
2. Each group is expected to submit jupyter notebook (.ipynb) with output for each cell.  
3. No extension on the deadline  
4. Submissions using other python IDEs will not be considered for grading

**Link to the Dataset**: [https://www.kaggle.com/mortena/fb-sentiment?select=fb\_sentiment.csv (Links to an external site.) (Links to an external site.)](https://www.kaggle.com/mortena/fb-sentiment?select=fb_sentiment.csv)

**Description of Data**: Dataset, consists of reviewText and Rating.

1. Download the file and set it as a Dataframe. (1 Marks)
2. Remove punctuations, special characters and stopwords from the text column. Convert the text to lower case. (3 Marks)
3. Create two objects **X** and **y**. **X** will be the ' Fbpost' column dataframe and **y**will be the “Label” column. create a **CountVectorizer** object and split the data into training and testing sets. Train a MultinomialNB model and Display the confusion Matrix (5 Marks)
4. Display the POS tagging on the first 4 rows of ‘Fbpost’ (4 Marks)
5. Build and display a dependency parser tree for the sentence:

“With the Night King riding on his back, Viserion breathes blue fire at the Wall, which then disintegrated.” (2 Marks)