

Sentiment Analysis on Demonetisation

DHEEMAN KUANER
BITS PILANI

<https://github.com/Dheeman1999/DeMonetisationAnalysis>

Problem Statement



- Find out the views of different people on the demonetization by analysing the tweets from twitter. Use the Twitter API to extract all tweets related to demonetisation in India by using appropriate filters. Divide each tweet text into words to calculate the sentiment of the whole tweet. Rate the word as per its meaning from +5 to -5 using the dictionary AFINN. The AFINN is a dictionary which consists of 2500 words which are rated from +5 to -5 depending on their meaning. You can download the dictionary from the following link: ****AFINN dictionary (<https://drive.google.com/file/d/1AjcJCNH2Oc9j-bJgjdELaUMIWZ3pXyfZ/view>)**** Now, you have to generate a sentiment rating for each tweet. After that, perform the sentiment analysis by filtering out positive sentiment tweets from negative sentiment tweets. Finally, provide a conclusion whether based on twitter data, the general sentiment is in favour of demonetisation or against, and by how much. You can provide insights on following parameters: Is the overall sentiment positive or negative? Is the latest sentiment positive or negative? When were the peak positive sentiment levels? When were the lowest negative sentiment levels? Who were the 5 biggest positive sentiment influencers? Who were the 5 biggest negative sentiment influencers? Ask more questions like these which can be answered by this dataset and then find answers to them. Finally present them using graphs and visualisations. ****Submission Guidelines**** The project must be made completely on command line. You should share the complete source code via Github only. And also share the screenshots of the complete application. Also submit a presentation consisting of graphs providing visualisation of insights and conclusions. Presentation should not have more than 6-7 slides.

Approach



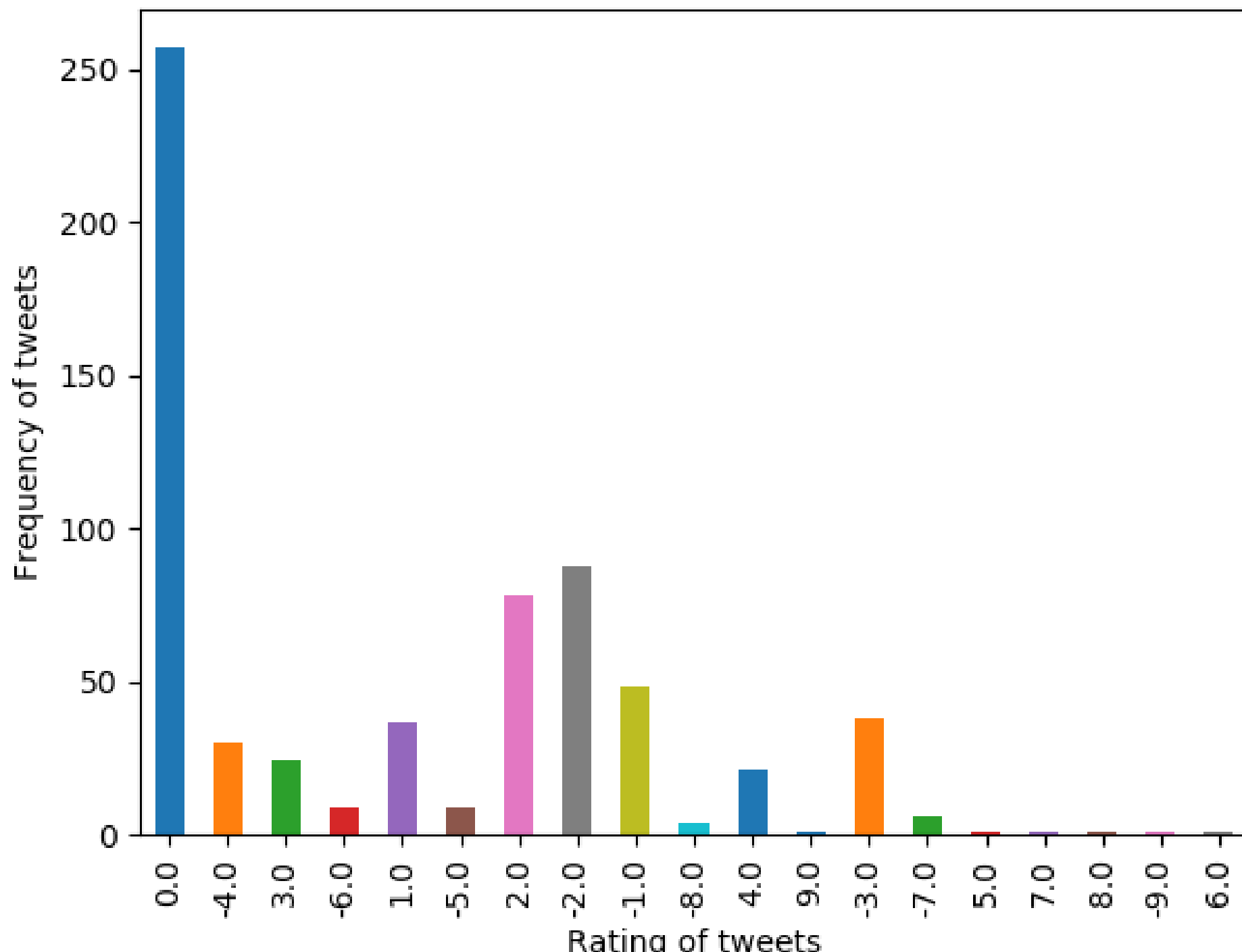
- The twitter API was used for the purpose.
- All the tweets with '#demonetization' in the last 7 days were scanned using twitter API and tweepy module of python.
- The data was cleaned by removing all the retweets, hyperlinks and abbreviations were substituted with non-abbreviated form so that a clear text was available for analysis.
- Initial analysis was done using AFINN dictionary and rating each tweet based on the dictionary using AFINN module in python.
- Secondary analysis was done using pre-trained TFIDF model.

Initial studies



- The initial studies using AFINN dictionary revealed the following data:
- Positive tweets(rating>0):25%
- Negative tweets(rating<0):36%
- Inconclusive tweets(rating=0):39%
- Overall sentiment:Negative
- Clearly the AFINN dictionary itself was incapable of comprehending all the data.
- Hence,TFIDF was used to analyse the inconclusive data.

Rating distribution of tweets

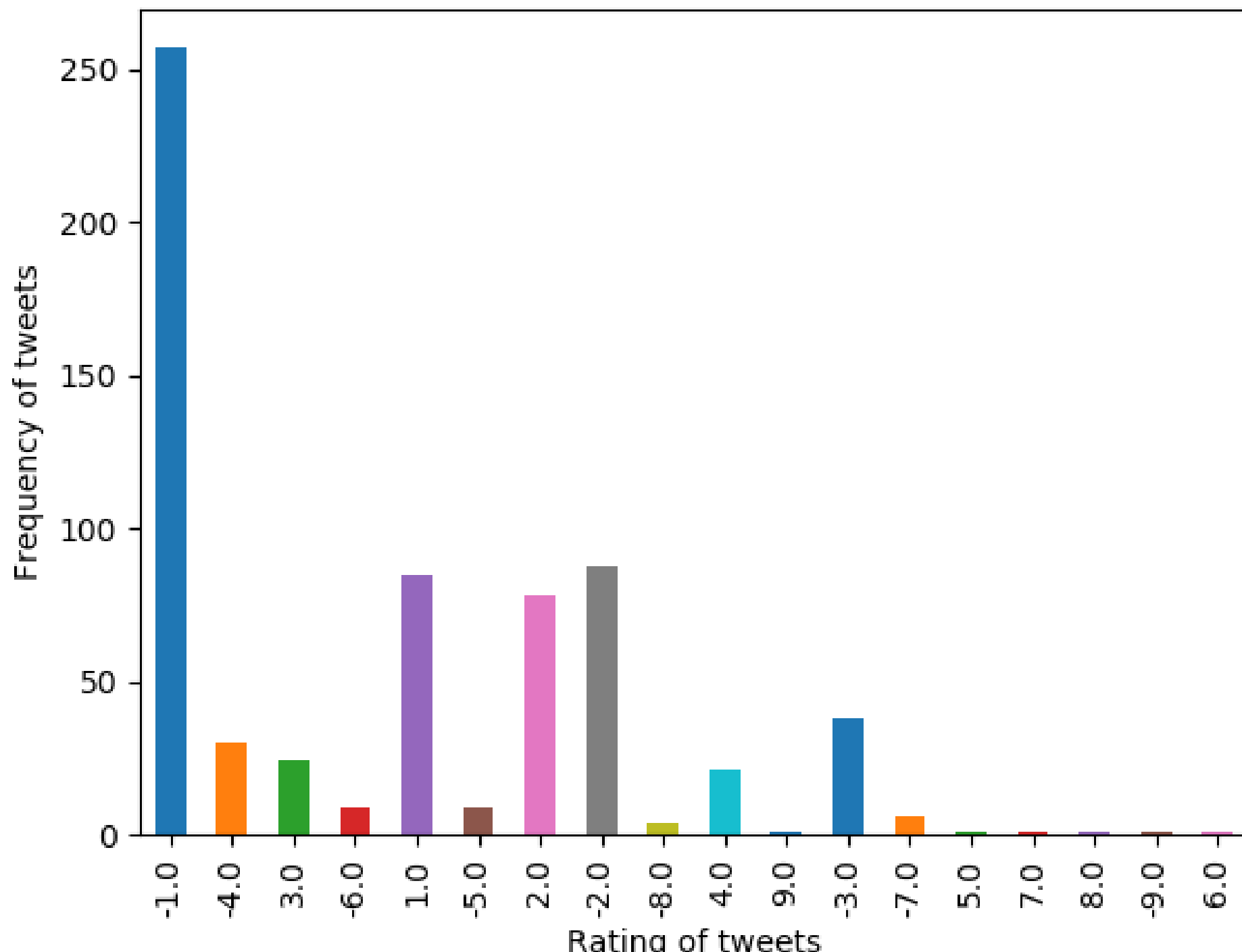


Post Secondary analysis results



- After applying TFIDF classifier on the inconclusive tweets, the following were the results:
- Positive tweets: 32.5%
- Negative tweets: 67.5%
- Overall Sentiment : Negative
- Overall Sentiment rating = $(\sum r_i f_i) / (\sum r_i) = -0.63$

Rating distribution of tweets



Other stats

- The other stats are accessible by the execution of the main module (`Demonetisation.py`) of the project.

