SENTIMENT ANALYSIS AND VISUALISATION OF TWITTER TRENDS

Dipanshu Nagpal, Shivani Suryawanshi, Priyanka Tikone

ABSTRACT

Twitter is an online microblogging and social networking platform, which allows users to write short status, updates of maximum length 280 characters. These tweets reflect public sentiment about various topics and events happening. Analysing the public sentiment can help, firms trying to find out the response of their products in the market, predicting political elections and predicting socioeconomic phenomena like stock exchange.

System does sentiment classification of tweets using machine-learning approach in python. A web application developed using flask in python displays these results of sentiment analysis in the form of various graphs like line graphs, maps, pie chart etc. This project deals with analysing sentiments behind the tweets be it positive or negative. Focus is not only on classifying the tweets, but also on making this task faster and more accurate by removing the parts of the tweets not contributing to the sentiment analysis, with the help of data pre-processing techniques. These pre-processing techniques include case conversion, punctuation removal, stopword removal, URL removal etc.

Classifier based on supervised machine learning algorithm classifies the sentiment present in a tweet. The model obtained from this classifier is applied on the tweets of currently trending topic, on twitter. The application will help in understanding people's views and emotions regarding the trending topics. It helps in understanding the effect of a particular topic on people, whether the topic garners majority of positive reviews or negative reviews.

INTRODUCTION

- Twitter is a widely popular micro-blogging platform for users to express their opinions about governmental issues, product items, sports and so forth. A tweet is a text-based post and has only 280 characters.
- Twitter is a "what's happening-right-now" social network and hence tweets are valuable sources for businesses, government and individuals to determine public's opinion or sentiment about an entity, product, people, topic, event etc.
- This process of analysing and summarizing the opinions expressed in these huge opinionated user generated data is usually called Sentiment Analysis or Opinion Mining.
- Sentiment analysis aims to determine the attitude of a speaker, writer, or other subject with respect to some topic or the overall contextual polarity or emotional reaction to a document, interaction, or event . The attitude may be a judgment, affective state or the intended emotional communication. The sentiment can be classified as either positive, negative or neutral. The objectives of these project are:
- To build a sentiment analysis model using machine learning approaches which, will automatically classify the political tweets acquired regarding a keyword, as positive, negative or neutral.
- To visualise the results obtained using the model according to various parameters like time, geographic location, age etc.
- To develop smooth, fast, efficient, reliable and easy to use web-based tool.
- Providing a user friendly menu and good entertainment visualization capabilities.

IMPLEMENTATION

1. Data Acquisition
2. Data Pre-Processing
   Pre-Processing methods- Case conversion
   Stop words removal
   Punctuation removal
   Stemming
   Lemmatization
   Spelling correction
3. Labelling
4. Feature Extraction
   Features- Term presence and frequency
   Parts of speech tagging
   Opinion words and phrases
5. Application of Machine Learning Algorithms
   Algorithms- Naïve Bayes Algorithm
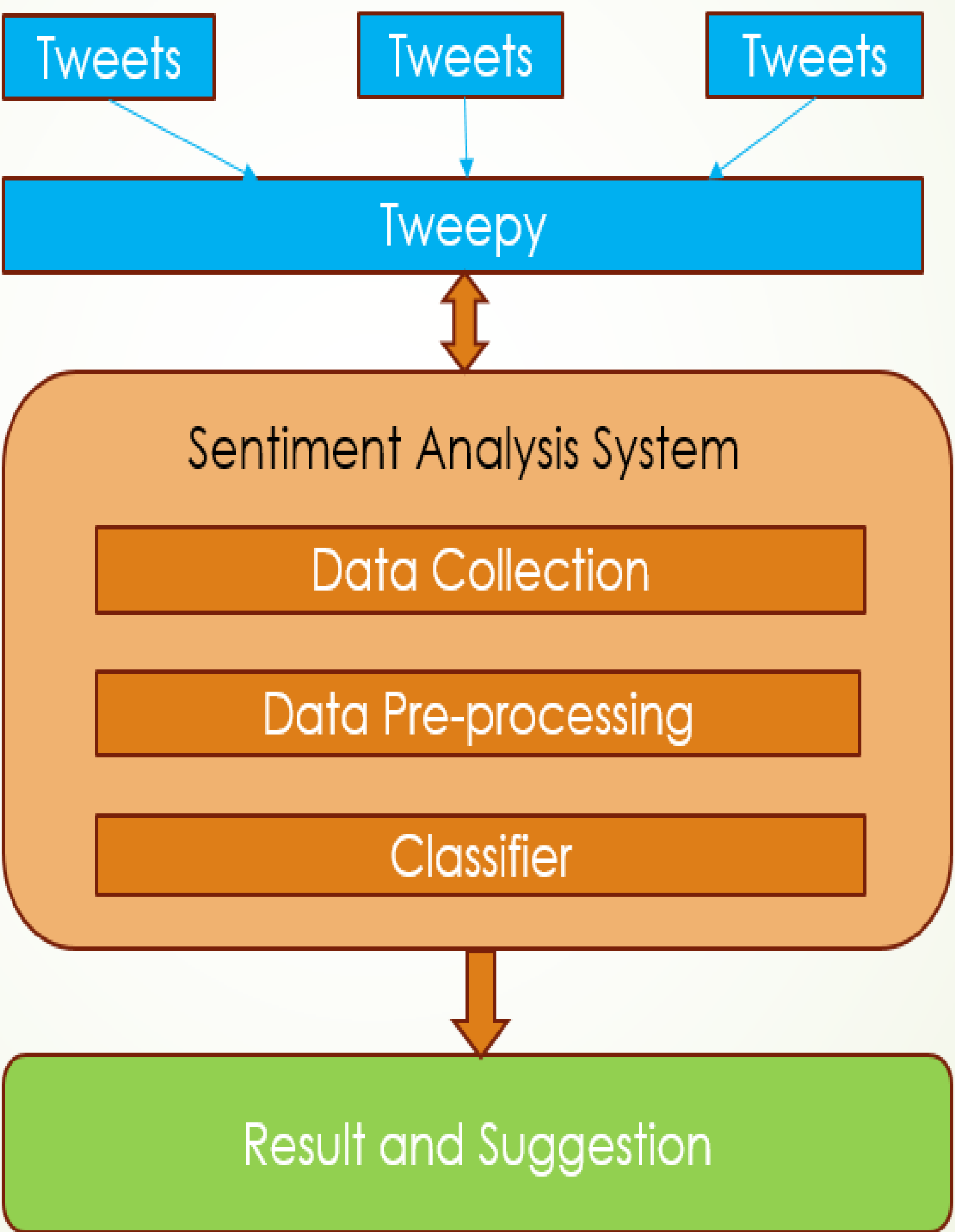   Support Vector Machine
   Decision Tree
6. Visualization
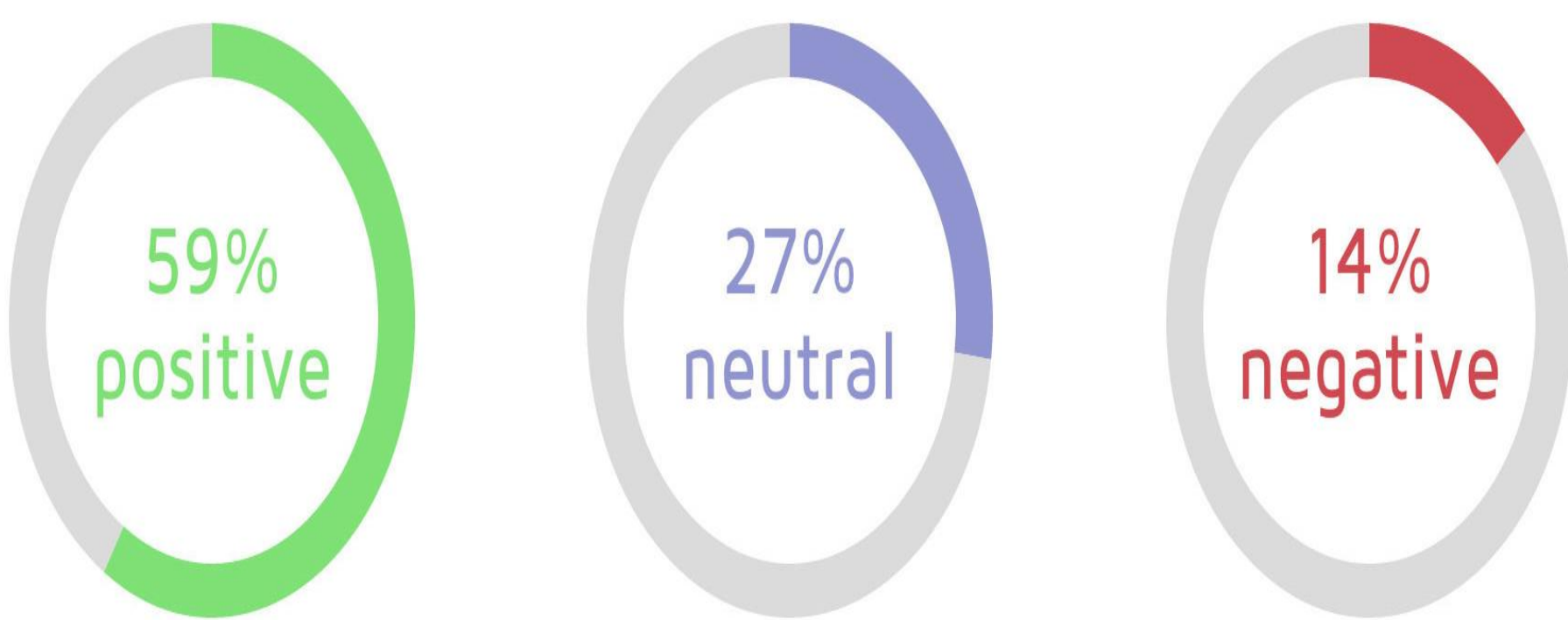
METHODOLOGY
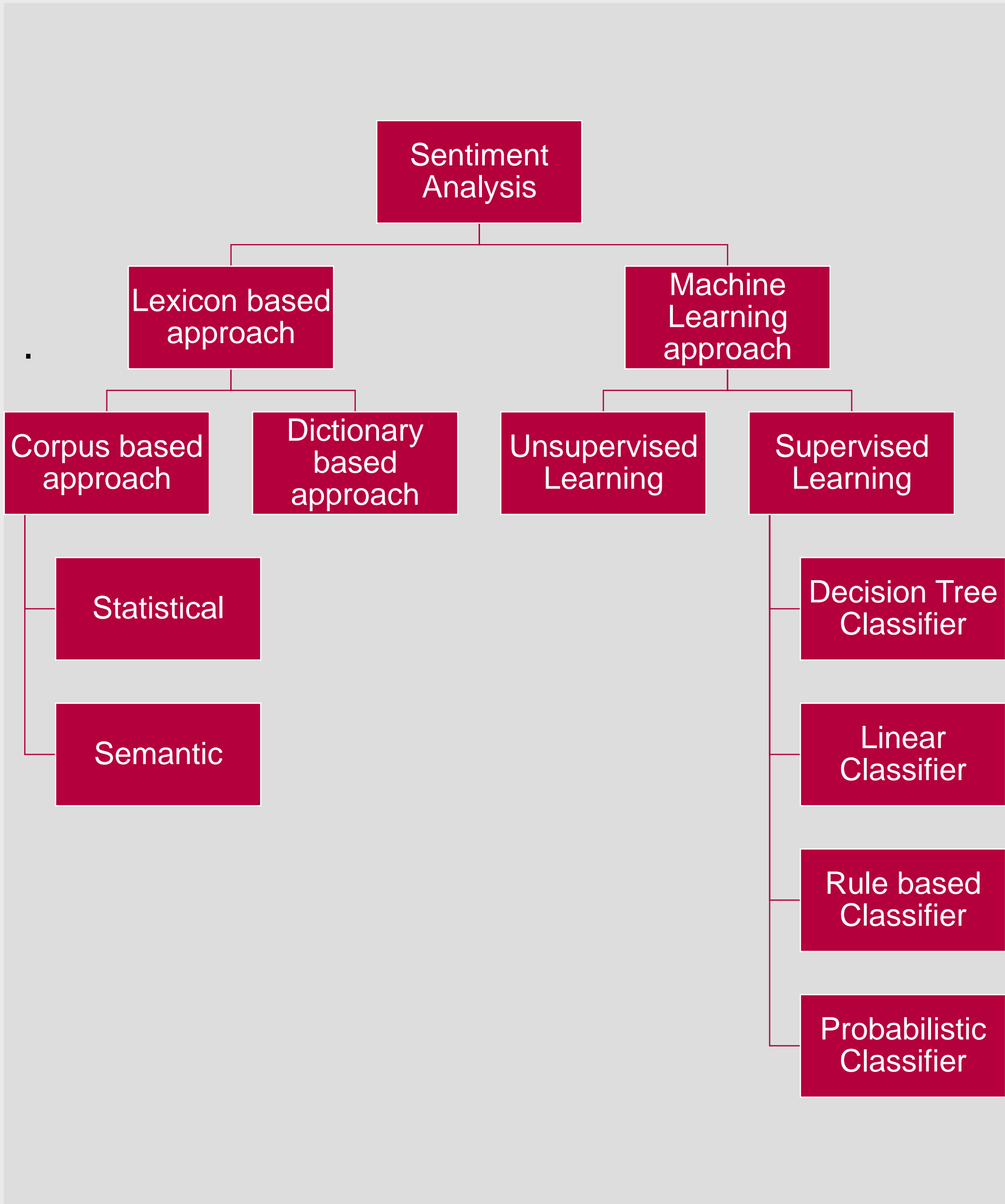
RESULTS
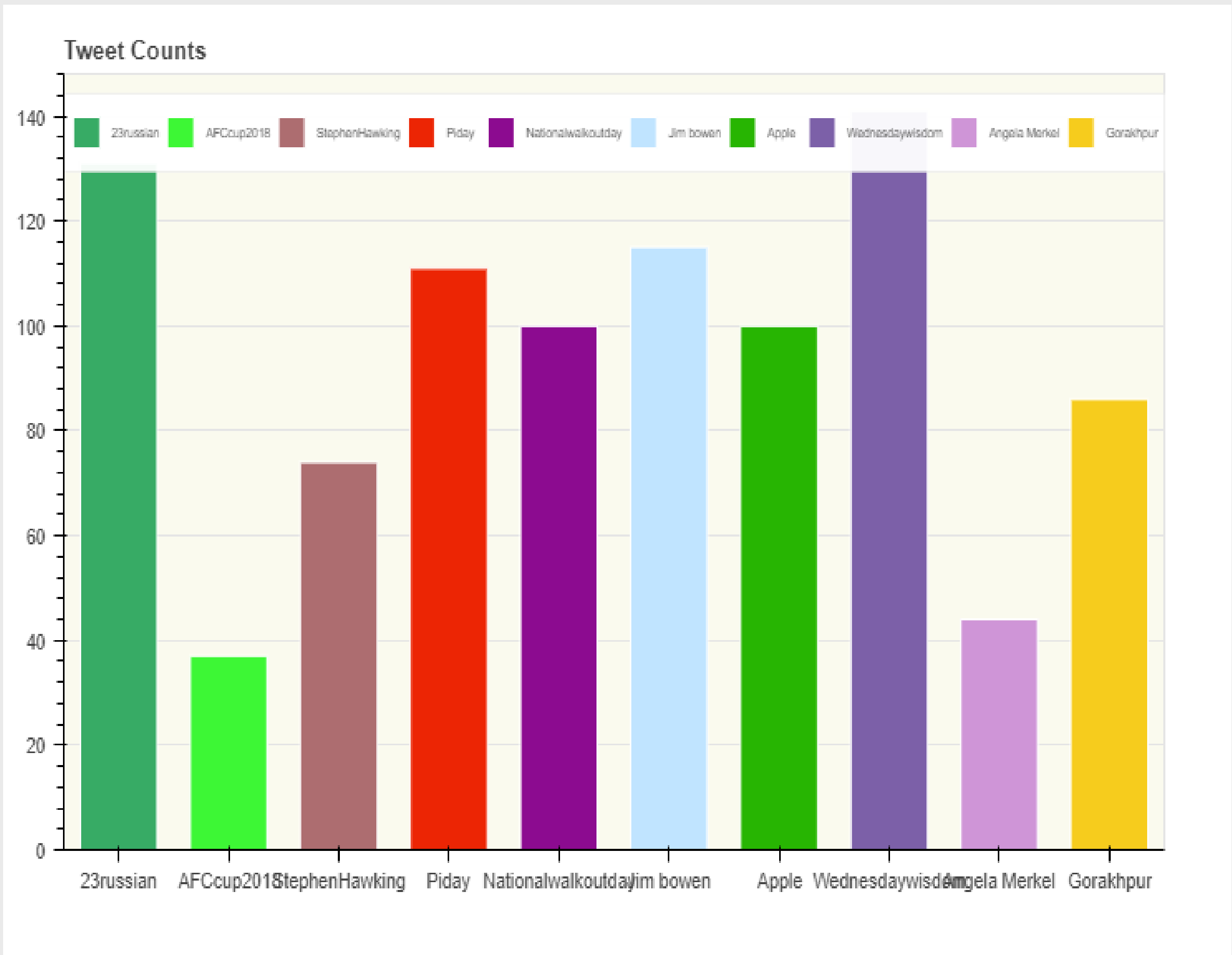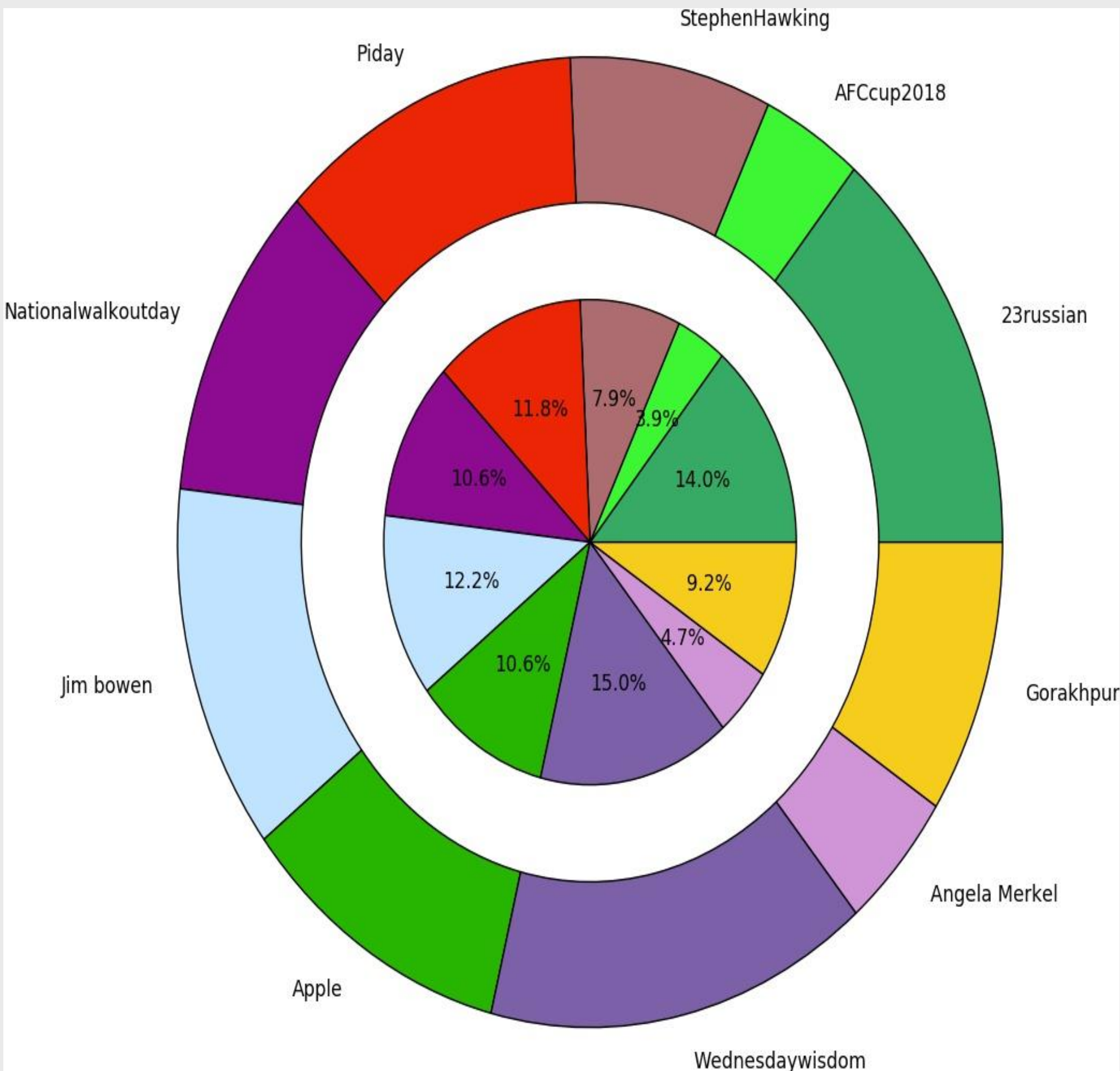
ARCHITECTURE

CONCLUSIONS

- In this project, we studied various methods and approaches of sentiment classification.
- After analysing all the methods, their advantages and disadvantages we concluded that machine learning approach is more efficient among all other methods.
- After experimentation, we found out that pre-processing tweet, to remove unwanted words, which do not express any sentiment, increases the accuracy.
- Unigram gives good result than bigram and trigram.

FUTURE SCOPE

- We conclude that the sentiment prediction accuracy will increase as the dataset increases.
- In this project, we focus on sentence level sentiment classification, but as a future work, it can be expanded to include feature/aspect level classification, which is useful in product review and recommendation system.
- The number of sentiment classes can be increased to get more refined sentiment prediction.

REFERENCES

- Afroze Ibrahim Baqapuri, "Twitter Sentiment Analysis",Department of Electrical Engineering, Anuja P Jain, Asst. Prof Padma Dandannavar.

- "Application of Machine Learning Techniques to Sentiment Analysis",*International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*