# Monitoring change in popularity

# of political parties using sentiment analysis

Animesh Yadav1, Vivek Rao2, Yash Tibrewal3

Department of Information Technology,

Mukesh Patel School of Technology Management and Engineering, NMIMS Deemed-to-be University,

Mumbai, India,

1 [animeshy@gmail.com](mailto:animeshy@gmail.com), 8655725330,

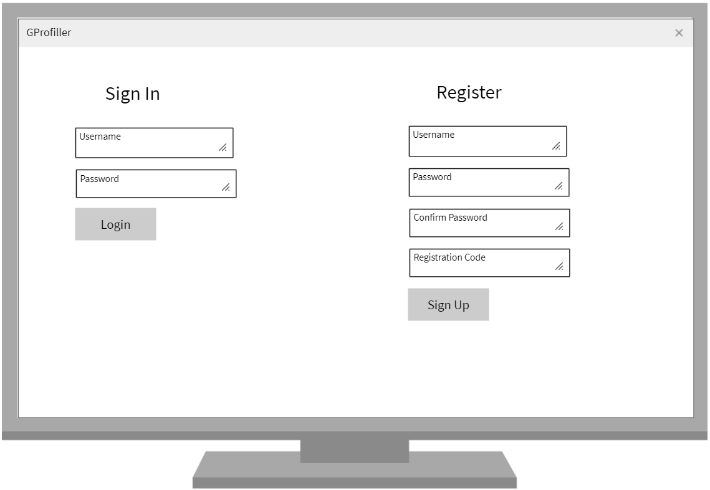
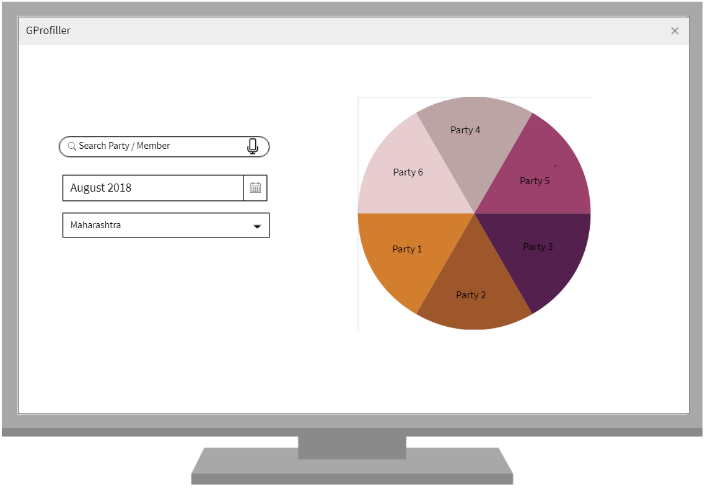
2 [v\_rao97@yahoo.com](mailto:v_rao97@yahoo.com), 9699130765,

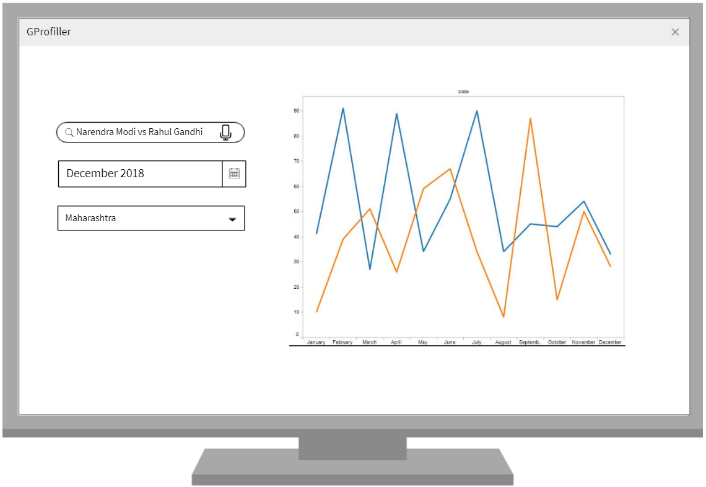
yashkush.tibrewal@gmail.com, 8879034882

1. SCOPE

**Project Scope Objective**- To develop an application that utilizes the techniques of Natural Language Processing (NLP) which involves text analysis of user generated tweets and application of machine learning algorithms to gauge the popularity of prospective candidates along with their respective parties contesting in the Lok Sabha General Elections-2019.

**Project Scope Deliverables**- The final deliverable consists of an application along with an appropriate UI using Python language. The predictions will be graphically represented/visualized on the screen with the popularity ratings of one party compared to other parties. Intermediate deliverables include software engineering reports and code(s) implemented with incremental requirements/tasks/functions.

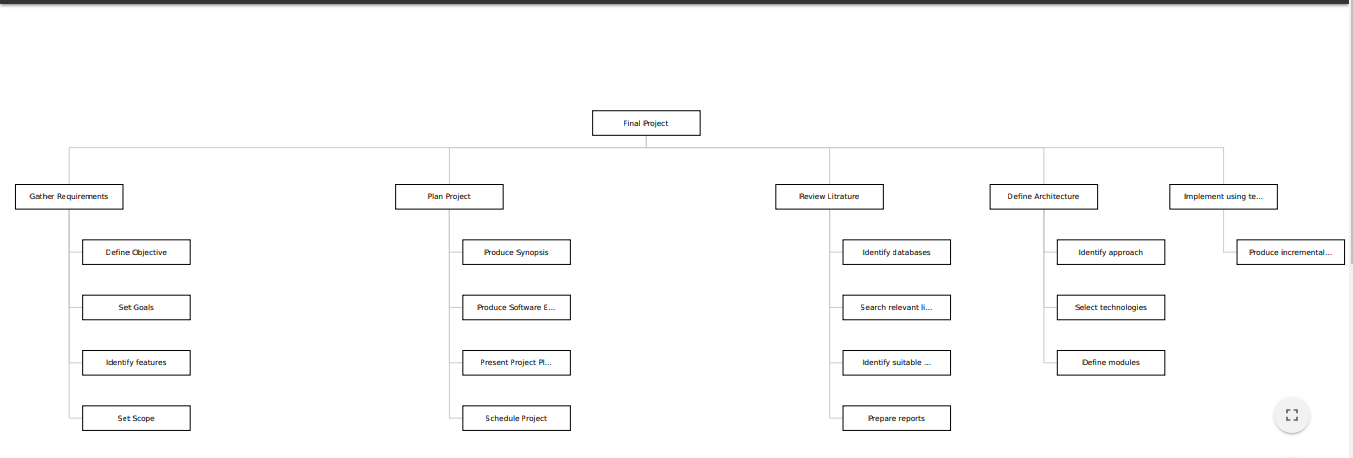


**Project Scope Milestones**- Key events that mark significant progress in the project include feasibility study, analysis, logical design, incremental implementations as per schedule and final delivery of the project.

**Scope**: Tweets in English, Images, Emoticons.

**Out of Scope**: Fake news (tweet)

**WBS-**



**Description**-The main objective of the project is to classify the user generated opinions about the political parties contesting in the elections from their official social media handles as well as from their leaders. Then from the extracted data set we will execute the algorithms on the processed data obtained from the social media platform. Using these responses, we will observe the trends in popularity of a political party and/or the candidates in the election. We will also be able to gain an insight into the “Consumer Confidence” during the time a particular party is ruling. Here, consumer confidence refers to the public opinion regarding the health of the country’s economy in terms of job, healthcare and employment opportunities. Consumer confidence about the economic health will also be analysed.

Based on the positive or the negative sentiment mined, the political parties shall be allotted some positive or negative points, based on which the popularity will be concluded.

1. LITERATURE REVIEW

Our estimated approach in solving the problem is influenced by the established work in this particular domain. Sentiment analysis, also known as opinion mining, utilizes various approaches as presented in [1]. Typical methodologies include Lexicon based and Machine Learning based. The authors [1] have stated that the context of the word is not taken care of when data is classified based on word dictionaries. On the other hand, the algorithmic approach builds a model from the trained data set to predict/classify the remaining dataset. We intend to incorporate stages in the project which will perform the respective defined functions. As suggested in [1], we intend to incorporate stages of computations. Namely- building an initial dictionary which will act as “training” data, classification and prediction based on modified dictionary. The data will be extracted based on a particular topic and initial classes/labels will be defined.

Next what follows is the task of pre-processing the raw data. As stated in [2], the unstructured data fetched from twitter needs to undergo cleaning and transformation process. After the computation is performed, the data needs to be stored in a database. Various tasks will be involved in this step as the algorithmic implementation requires formatted data to execute upon.

In order to classify a tweet into one of the defined classes, opinion classifiers will be employed as suggested by [2]. This particular project will be employing machine learning algorithms which fall under the category to supervised learning. Based on metrics, the performance of the algorithms will also be compared along with achieving the stated project objective. Each of the independent work conducted faces a few limitations in achieving the accuracy/performance. Our estimated approach will incorporate the unique features of each work which will complement the others and the overall project as a whole.

1. REFERENCES

[1] El Alaoui, I., Gahi, Y., Messoussi, R., Chaabi, Y., Todoskoff, A. and Kobi,, “*A. (2018). A novel adaptable approach for sentiment analysis on big social data*”, *Journal of Big Data*, 5(1).

[2] Anjaria, M. and Guddeti, R. (2014)., “*A novel sentiment analysis of social networks using supervised learning*”,  *Social Network Analysis and Mining*, 4(1).

[3] Ratab G, Umar S, Saba R, Washma A, Beenish Z,. “*Pre Processing of Twitter’s Data for Opinion Mining in Political Context*”, 20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2016, 5-7 September 2016, York, United Kingdom