

MINI PROJECT (2021-22)

ON

Analysis of indian election using twitter

**NAME:**RANJANA SINGH

**ROLL NO:** 2014801

**SEMESTER:** 5th

**SUBMITTED TO:** DEEPAK UNIYAL

Abstract

**Background**

In this project system is going to study the general user tweets from the election point of view. Here the System will study the user view of Indian election. Based on the users tweets system analyses if there exist a pattern between the tweets and to analyse and draw meaningful inferences from the collection of these tweets collected over certain period; the proposed system identify the feasibility of development of a classification model to identify the political orientation of the twitter users based on the tweet content and other user based features. There are voting advice applications (VAAs) are online tools which are popularly used in deciding which party/candidate to vote for during an election in countries like Greece, Cyprus but still there in India there is no such application which focus on this . The proposed system plans to provide an online portal for recommendations and comparison of user’s political opinions.

INTRODUCTION

The main goal and challenge of the system is analysing twitter data for Indian election to see the impact of tweeter on Indian or particular state election. The proposed algorithm is consists of three main steps. Our proposed system is analysing system which is based on the mechanism that analyses User Tweets using Hash tags and Keywords.

The proposed system collects tweets using this Hash Tags which are nothing but the popular personalities/Parties. General public orientation toward these parties can be studied using the tweets the people have posted on the Tweeter.

PROBLEM DESCRIPTION

With all the real time data collected over period of time, the system will analyze and draw meaningful inferences from the collection of tweets.

By this , we are getting probability in percentage for people who can win in respective states by seeing their crime rates , popularity and by seeing their name hash tags on twitter.

ABOUT DATA

This Dataset is based on the Lok Sabha 2019 in India. There are a total of 2263 rows and 19 columns in this dataset. By using this dataset this data analysis project is created.Here we use JUPYTER NOTEBOOK to run these codes and analysis the dataset but you can use other platforms also to run the code.We have used a total of 3 datasets for this project.

REQUIREMENTS

**LANGUAGE USED:** Python

**SOFTWARE REQUIREMENTS:** Anaconda3 -> Jupyter

**HARDWARE REQUIREMENTS:** Windows Desktop

**REQUIRED INSTALLMENTS:**

* Python
* Anaconda3
* Jupyter

**Libraries:**

* Pandas
* Numpy
* Matplotlib
* Sklearn
* Seaborn

MOTIVATION

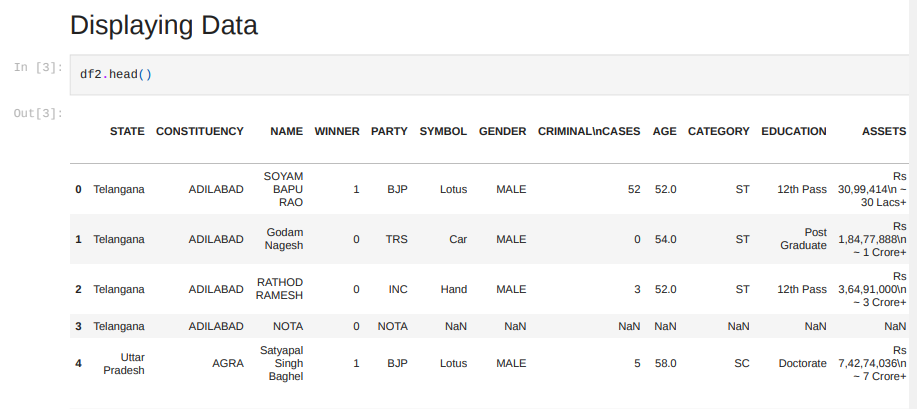
The 21st century is the century of Knowledge. For every answer people first use Google. The theme line is connecting people using Internet. When we talk about Internet along with knowledge we have Social sites are specially used for connecting world. People especially youth spend its lot of time on sites like Facebook, Twitter etc... Every current topic due to social media is supported with general public’s reviews and comment which makes a clear picture of that particular topic to world. The credit goes to popularity of social media.

METHODOLOGY

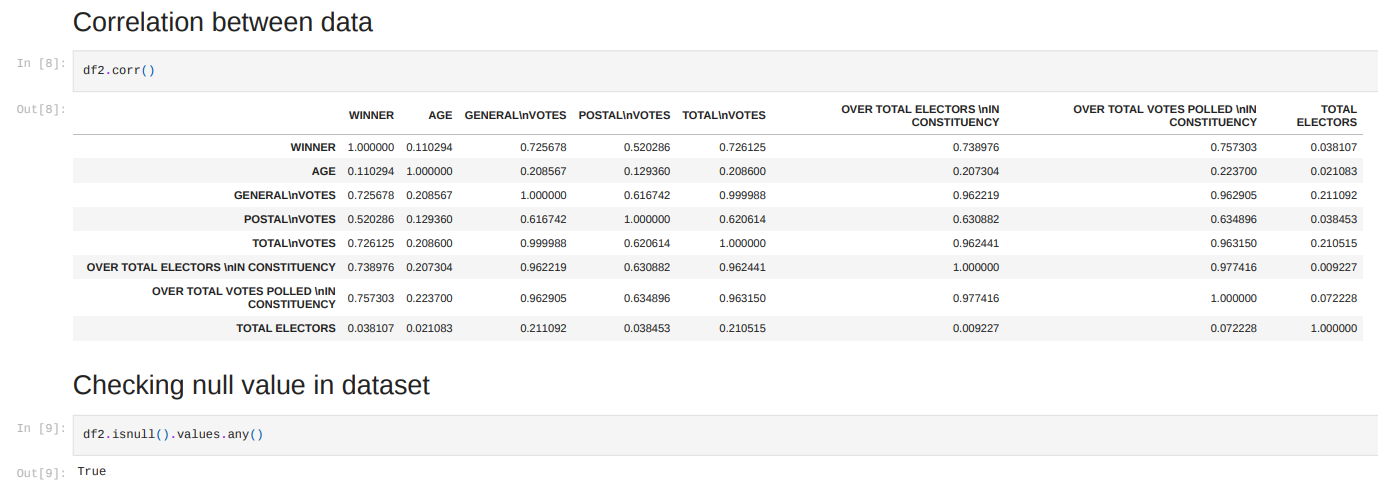
I used numpy and pandas for pre processing of data and sklearn for our machine learning task.I have used three datset and i’m analysing loksabha election of 2019 using twitter dataset.

1. Importing all required libraries and importing dataset and displaying:

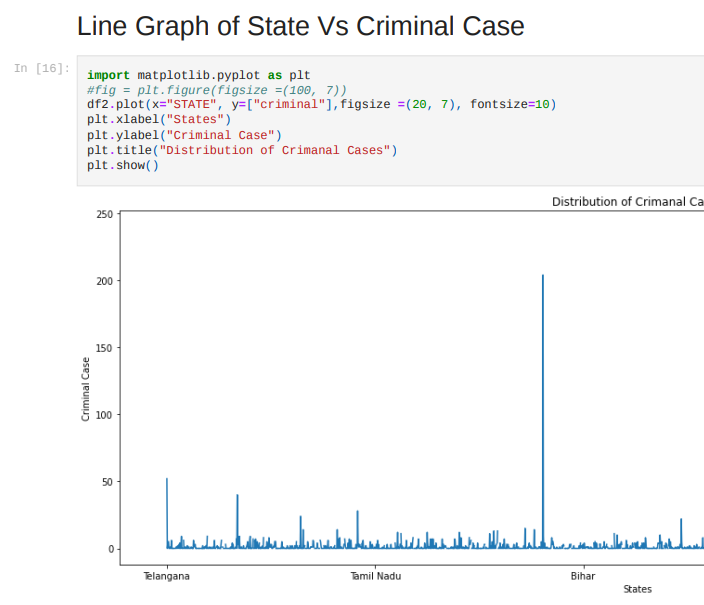
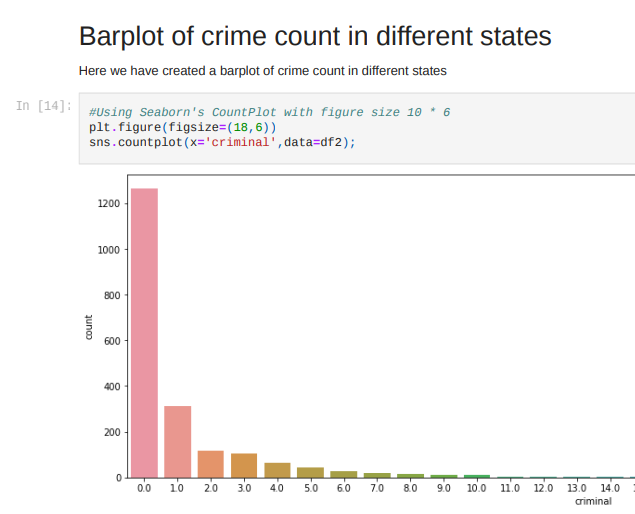




2.correlation of data:

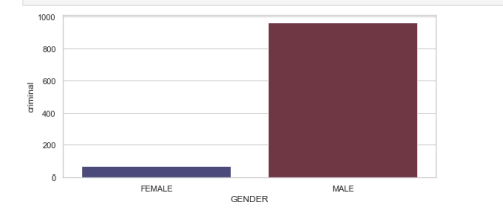


3.Counting crimes of candidates 

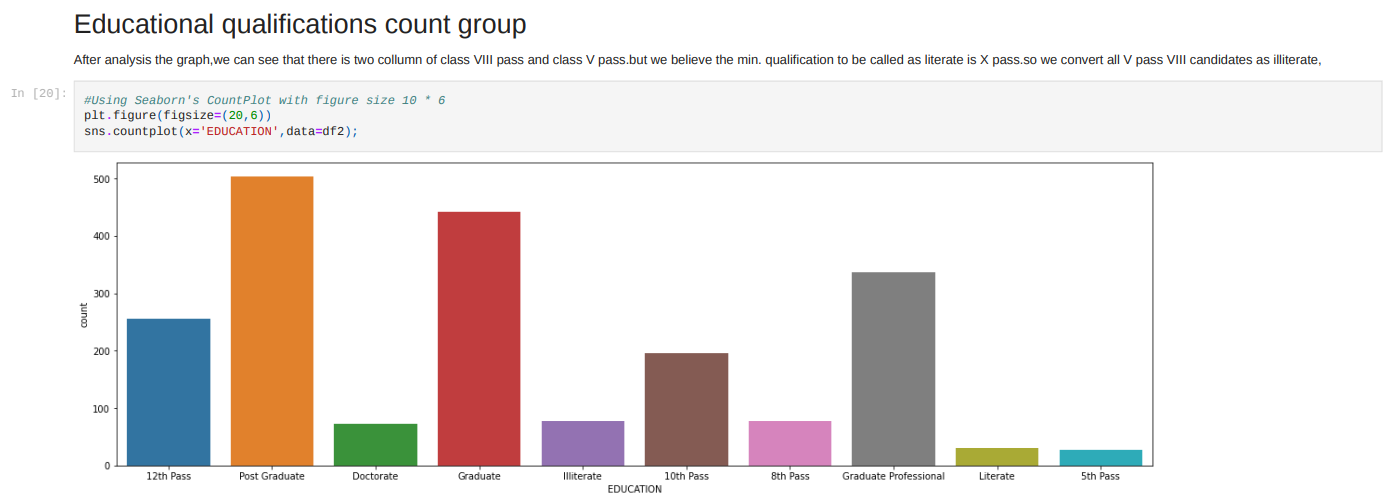




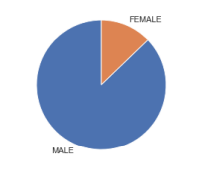
4. Barplot of Gender Vs Crime:



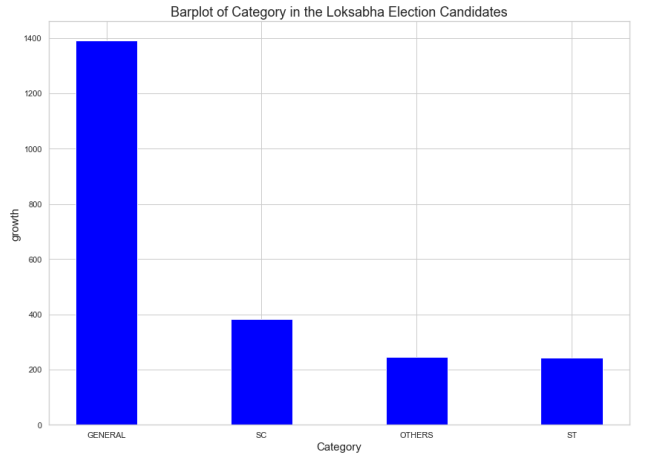
5.Seeing education qualification of candidates:

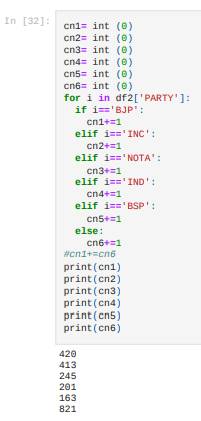


6.Pie chart of of Male Vs Female candidates:

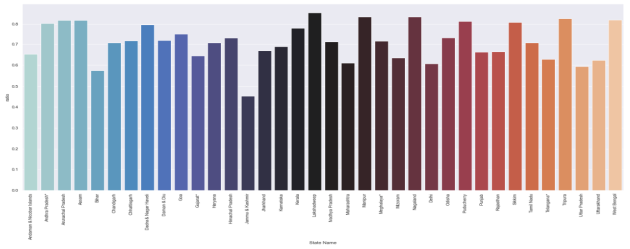
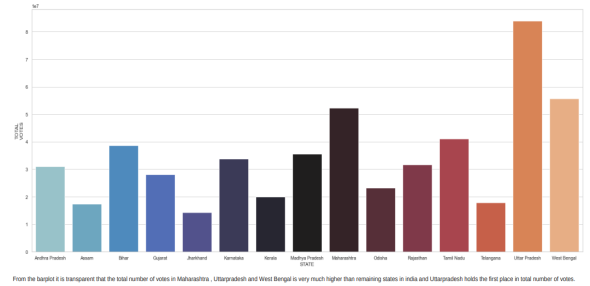


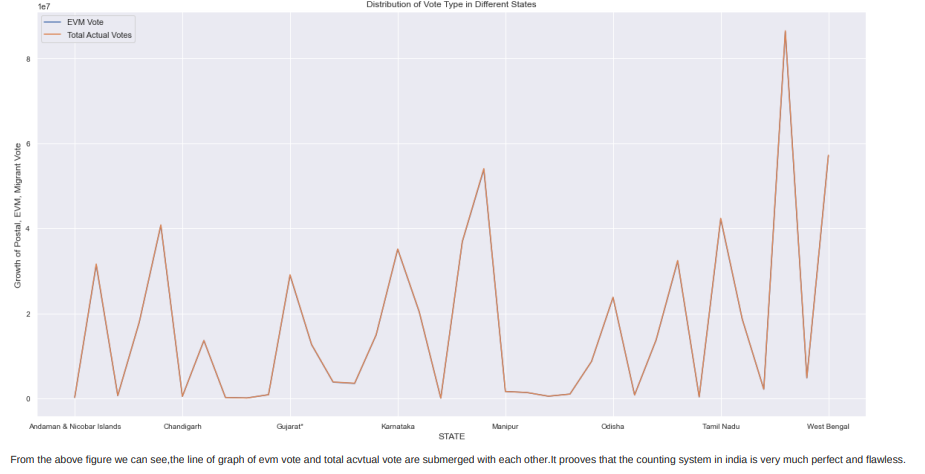
7.Barplot of cateogary growth for candidates:

8.Counting total numbers of allocation of candidates for different parties :

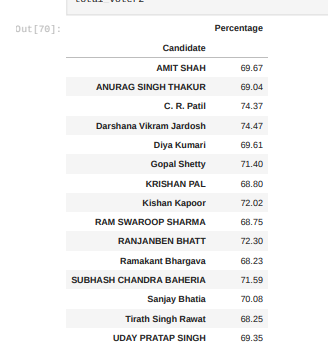


9.BARPLOT OF STATE Vs TOTAL VOTES & BARPLOT OF VOTE PERCENTAGE IN INDIA & LINE GRAPH OF STATE VS VOTE PERCENTAGE:





10. TOP WINNING CANDIDATE LIST WITH HIGHEST VOTE PERCENTAGE IN LOKSABHAELECTION 2019 INDIA:



CONCLUSIONS

From the above analysis, we can conclude the following points

From this analysis, we have found that Indians are very much aware of the voting system and political parties. The vote percentage is very good in Kerala, West Bengal, Manipur, Maharashtra, etc .The BJP candidates have participated more than the other parties in Lok Sabha 2019.It is transparent that the total number of votes in Maharastra, Uttar Pradesh, and West Bengal are very much higher than the remaining states in India and Uttar Pradesh holds first place in the total number of votes .The number of female candidates is very less than the number of male candidates .Unfortunately, we found that the candidates of Kerala have maximum criminal cases history, and West- Bengal, Uttar Pradesh, and Telangana are not far behind. The maximum no of criminal cases done by a single person is 240 and he is from Kerala . We have noticed that the history of the criminal case of candidates is maximum at the age of 49,37, and 51.

references

https://www.academia.edu/7486078/Use\_of\_New\_Media \_in\_Election\_Campaigning\_Lok\_Sabha\_Elections\_2014