Survey on Election Prediction Using Machine Learning Technique

Prof. Sangeeta Alagi[1], Mr. Tejas Kolambe[2], Mr. Vishal Bibe[3], Mr. Karan Gite[4],

Mr. Sanket Chachar[5]

[1]Guide, Computer Engineering Department, Genba Sopanrao Moze College of Engineering, Balewadi, Pune, India

[2, 3,4,5] Students, Computer Engineering Department, Genba Sopanrao Moze College of Engineering,

Balewadi, Pune, India

ABSTRACT

In today's era educational organization strongly needs devices which are ready to access and use and also various operating system platforms are required for different purpose. Another important thing is many educational organizations still used Virtual machine method to used different OS platforms for various operations. This project come up with the tweets data from the and analysing the results from the tweets. Newspaper is the crucial part of the human life. Because newspaper is a source of information. Every type of information is present there on the newspaper. But now a day's life becomes very easy because of the social media. People can use social media according to their convenience anywhere anytime. Social media is one of the important parts of the human life. Social media keeps every generation updated. There are many social sites with people are connected. Twitter and face book are the famous social media among the generation. People get influence with the social media. Whatever is happening over the world, social media keeps you in connect. So, the people are more incline toward the social media. When about elections, if any data related on any elections party. People started reading and thinking on that. In the form of comments, they will share their opinion that what are their thoughts about the elections and the political parties. Some people are in support with or some are in oppose of. So, the above scenario, from the social media which political party is on top. Predictions of elections from the social media from the twitter data we are predicting the election outcomes.

Keywords: - *Election, social media, Twitter, prediction, twitter data.*

1. INTRODUCTION

Social impact plays vital portion in a human life. Social impact implies when an individual comments on a few social tweets. In the event that that individual responds on it whether it is positive or negative which may alter the circumstances [1]. Social media interface individuals all over the world with in few moments. In this extend by utilizing twitter information or Facebook information. We are attempting to figure out the decision result. When a individuals responded on something which calculate the extremity of the circumstance in favour or in oppose. In this paper we are cantering on the Indian race expectation. For illustration, Consider any Indian political party and the different clients that's dynamic client and dormant user [1]. When any party related news is there on the social media and the individuals will begin responding on it. These responses may be positive, impartial and negative. Utilizing testing method, we are getting different points and aspects.

What social media precisely implies, which helps human in knowing that what happens within the corner of the world. The social media parading genuine as well as sometimes fake news from which the human feelings get influenced. Politics is the thought-provoking topic. Decisions are the foremost anticipated minute for the people. Individuals are exceptionally much inquisitive about, who is reaching to win the decisions or who will be the administering party to run the show. This paper is likely to anticipate the decision results utilizing twitter or

Facebook information. We'll collecting tweets from the twitter and after that perform pre-processing on that information.
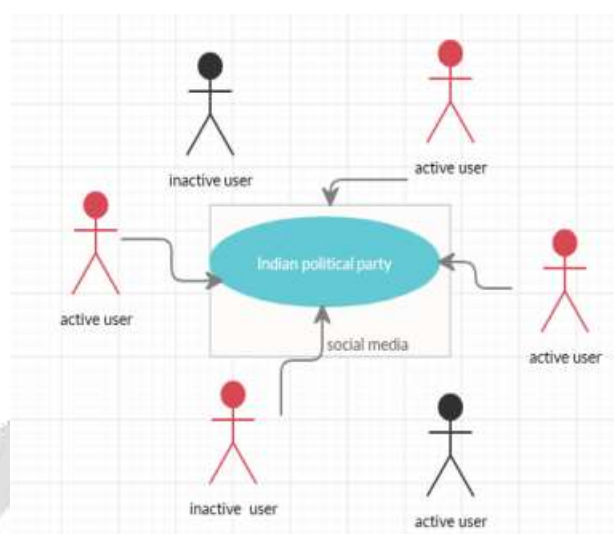

Fig. 1 Local example of social influence

2. LITERATURE SURVEY

Jiezhongqiu [1] proposed system on Social Influence Prediction with Deep learning, here in this paper described the study of the social influence using deep learning technique and obtained the final result in the graphical format. This framework basically worked on the OAG, Digg, twitter, weibo. The main idea behind the paper is to work on the local network. Using sampling techniques and the learning techniques simultaneously.

Parul Sharma and Teng-Sheng Moh [2] research on Sentiment Analysis on Hindi Twitter data, in this paper using Twitter Archiver tool obtained the Hindi language tweets. Then completed the data mining on 42,235 tweets. By using Naïve Baye's, Dictionary Approach and Support Vector machine algorithms tried to calculate the result of the general elections. Among the three algorithms the accuracy of the Support vector Machine is higher. So from that the prediction is clear that which party had chances of winning. They have also computed the precision and recall.

Steven T. Smith [3] studied on Influence Estimation on Social Media Networks Using Causal Inference this research is about the social media influence estimation. During French presidential elections 2017 find out the influence by using causal inference. Context are used for the graph sampling, Graph Filtering and for Graph influence.

Quanzeng [4] You Approach to Social Prediction of elections. The approach of this paper is the connection between image and the social media. Competitive Vector Auto-Regression model joins an image and the text in multimedia network. The proposed system is simply about when an election candidate post image on the social media and the viewer comments on that candidate related image. CVAR model calculates the accurate prediction of the elections.

Jason Anastasopoulos [5] studied on Political image analysis with deep neural network according to this proposed system posting text and images on social media is like daily routine though Political Candidate can interact with the voters directly. Images are the medium of communication between the candidate and the voters. Deep learning, neural network, machine learning and image analysis methodology are used in this proposed system. In studies we found that what methodologies or techniques used early to achieve specific target. We required high power system to perform this project related work or we can use cloud instances which gives better power for processing and storage space for data. You can develop and use your own cloud and perform NLP and MTL operation on that. [9],[18].

3. PROPOSED SYSTEM

In this model, the social media i.e., twitter or Facebook is the centre of the system. There were the number of the user are handling many accounts on the social media, among them twitter and Facebook are the trending social media accounts. As any Indian political party tweets on the twitter, user starts reacting on that topic whether they are agreeing or not or they have any other opinion on it. Firstly, from the tweets, we are going to remove the unnecessary words. Then correctly identify the meaning of the word from the given tweet called as lemmatization process.
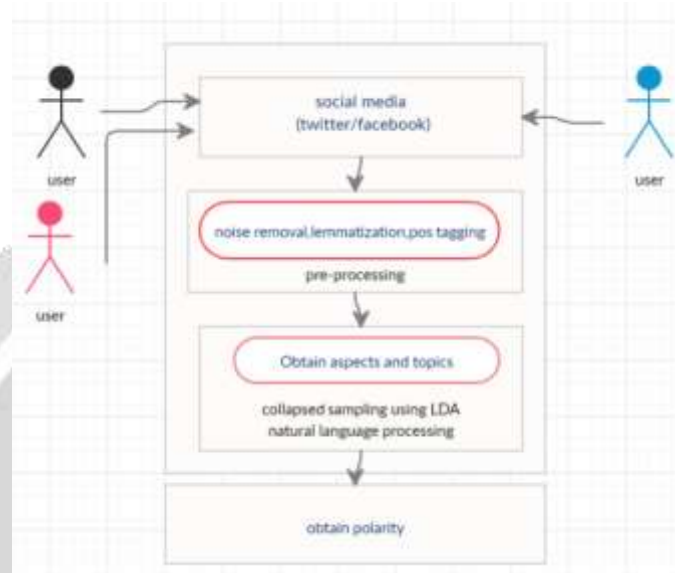

Fig. 2 Proposed System

This example of part of speech tagging shows the exact working of the process. It will separate all the part of speech as pronoun, verb, adjective, adverb, preposition, conjunction, noun etc. [16] This is called as the part of speech tagging.
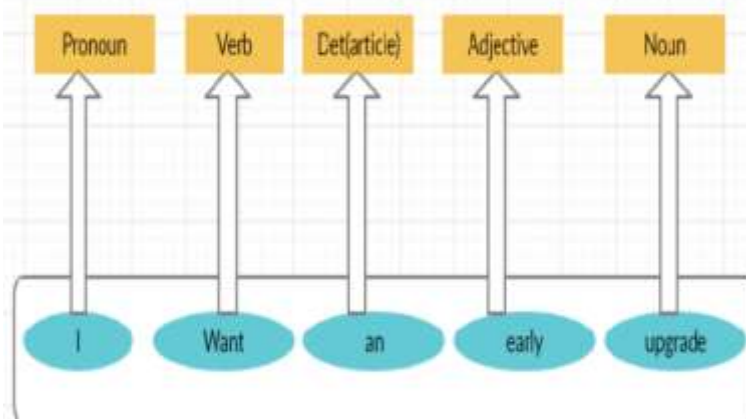

Fig. 3 Example of Parts of speech tagging

Collapsed sampling using Latent Dirichlet Allocation (LDA). The below diagram is the o diagram of the LDA system. This diagram gives detailed information about the proposed system. This diagram clarifies the actual working of the system [15]. Here is the actual working of the system is that firstly how the topics and aspects are obtain from the twitter data. Natural language processing helps to extract the main topic from the twitter data set with the help of the Latent Dirichlet Allocation (LDA).
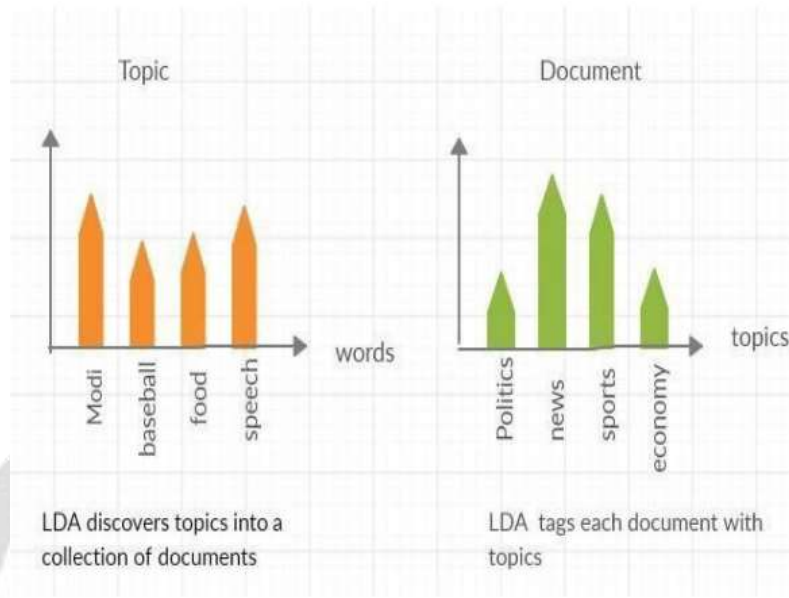

Fig. 4 Latent Dirichlet Allocation (LDA)

4. COMPARISION IN METHODOLOGY

Sr. No.	Parameter	Natural language Processing (NLP)	Machine Learning Technique (MLT)	Decision Tree Learning (DTL)
I	Parent	1-Computer science 2- Artificial Intelligence 3-Information Engineering 4-Linguistics	1-Artificial Intelligence	1- Data Mining 2- Machine Learning
II	Purpose	Learn Specific pattern from huge data	Perform Intelligent Things/automatically learn without any help of programmer	Based on different inputs and obtained the desired output
III	Technical Algorithm	1- Hidden MarkorvModel 2- Continuous bag of words 3-N gram Algorithm	1-Supervised i)Support Vector Machine ii) Naïve Bayes 2-Unsupervised i)Neural network ii) Hidden Markov iii) ICA, PCA	1-C 4.5 Algorithm 2-ID3 Algorithm 3-CART Algorithm

IV	**Used For**	NLP basically used to understand the language	Machine learning working on Learning things from Experience	Decision tree learning basically working on classification
V	**Processing Time**	Based on the algorithm data and structure of the program. It takes less time to perform operation	Based on the supervised and Unsupervised. On an average it takes less time in unsupervised learning but in supervised it consume more time to perform operation	Based on the algorithm of size of data, it takes more time compared to others
VI	**Learning Frequency**	We have to feed first and solution may or may not be useful to other problem	Self-learning and learn each and every cycle of the solution which is further useful	We have to feed first and part of solution is useful
VII	**Working On**	To understand the structure and meaning of the complex problem	To learn the data complex problem by observation or by experience	Observed the data or the specific data of complex problem
VIII	**Problem Solving Capacity**	Solve complex problem based on the data size	Solve complex problem based on technique. Once setup then large dataset also processes in less time	Solve problem based on data representation in the form of tree
IX	**Application**	1- Information retrieval 2- Sentiment analysis 3- Information	1-sentiment analysis 2-fraud prediction 3-election prediction	1-Rapid miner 2- mat lab

Table 1: Comparative study

This comparison shows why we combine two Methodologies Natural language processing and Machine Learning.

If we see the comparisons, we found that the natural language processing is better to understand the language but every time we not feed data to the NLP. We required the machine that learn or understand the data by itself and those parts comes under unsupervised machine learning. It gives better output, better execution time and accurate results rather than single methodology.

This combination of methodologies handles the complex problem very easily and takes less time to perform operation. Decision tree learning is not enough to fulfil requirements of this project.

5. FUTURE SCOPE

In future we will try to solve this problem by using much better algorithms or techniques than these and implement stable and better sentiment analysis model which is useful in real time environment for prediction purpose.

6. CONCLUSION

In this paper, we saw that how social media influence the people to take their decision. After seen some literature survey's we decided our proposed system which can be modify as per the requirement of this project.

The important part of this project is which methodology we are using for election prediction. There are many methodologies present for prediction or for sentiment analysis but only few of them work efficiently. Such methodologies are NLP, MLT and DTL. NLP and MLT are mostly used techniques or methodologies.

7. REFERENCES

 [1]JiezhongQiu,JianTang,HaoMa,YuxiaoDong,Kuansan Wang and Jie Tang,"Deeplnf:Social Influence Prediction with Deep Learning",The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data mining,August19- 23,2020,London,United Kingdom.ACM,NewYork,NY,USA

[2]Mark Carmen,markKoerber,JiuyongLi,Kim-Kwang Raymond Choo and Helen Ashman,"Manupulating Visibility Of Political and Apolitical Threads On Reddit Via Score Boosting", 2018 17th IEEE International Conference On Trust,Security and Privacy in Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering.

[3]Steven T. Smith,Edward K. Kao,Danelle C. Shah,OlgaSimek and Donald B. Rubin,"Influence Estimation On Social Media Networks Using Inference",2021 IEEE Workshop on Statistical Signal Processing(SSP)

[4]ZhenhuaTan,(Member,IEEE),DankeWu,GuangmingYang,And   Zhang   Bin,"APMSID:Activated Probability for Multi-Score Information Diffusion in Online Social Networks,Volume6,2020"

[5]QuanzenYou,LiangliangCao,YangCong,Member,IEEE,XianchaoZhang,andJieboLuo,Fellow,IEEE,"A Multifaceted Approach to Social Multimedia-based Prediction of Elections ",IEEE Transactions on Multimedia (Volume: 17 Issue:12, Dec. 2020) [6]Parul Sharma and Teng-Sheng Moh,"Prediction of Indian Election Using Sentiment Analysis on Hindi Twitter",2021 IEEE International Conference On Big Data (BigData)

[7]RohiniAshokPandit,Dr.GulveAvinashKushal,Dr.PraveenShetiye,"Implementing   Feature   Selection Technique to estimate influence from voting records",International Journal of Research and Analytical Reviews(IJRAR) 2021 IJRAR June 2021,Volume 6,Issue 2

[8]InternetLink:https://en.m.wikipedia.org/wiki/Natural_language_processing

[9]InternetLink:https://en.m.wikipedia.org/wiki/Machine_learning

[10]InternetLink:https://www.quora.com/Whats-the-difference-between-machine-learning-AI-and-NLP

[11]Rekha S. Thorat,Dr. Praveen C. Shetiye,Dr. Avinash K. Gulve,"Prediction Using Satellite Imagery And Machine Learning ",International Journal of Scientific And Technology Research Volume 9,Issue2,FEB2020

[12]InternetLink:https://rutumulkar.com/blog/2016/NLP-ML/

[13]InternetLink:https://www.google.com/search?q=latent+dirichlet+allocation+explained&sxsrf=ALeKk0 0eOH0epoqKh37vw
XnWdyXwNs0VQQ:1583255113047&tbm=isch&source=iu&ictx=1&fir=G0tsIyD6o71fM%253A%252Cv NBqmfQ2m8O4vM
%252C%252Fm%252F0cc726&vet=1&usg=AI4_kTeCAul3I7dOcz0STMpU6tdbajiLQ&sa=X&ved=2ahU KEwicOf05P7nAh Vf7HMBHa8bCEwQ_B0wHHoECAsQAw#imgrc=G0tsIyD6o-71fM

[14]InternetLink:https://www.google.com/search?q=part+of+speech+tagging&sxsrf=ALeKk01uMmVBoIv n0hn6b4s2tymdbc1

kAA:1583255227189&tbm=isch&source=iu&ictx=1&fir=Km4FsjGqh56Y6M%253A%252C1EX4mmNB
zdWDSM%252C%2
52Fm%252F037xvs&vet=1&usg=AI4_kSCIrquHcKPJYcwQ6mpsRP_DOrB8w&sa=X&ved=2ahUKEwjf
5J6r5f7nAhUU6nM BHTDcBZEQ_B0wGXoECAsQAw#imgrc=Km4FsjGqh56Y6M:

[15]SupriyaB.Moralwar ,Sachin N.Deshmukh,"Different Approches of Sentiment Analysis",JCSE
International Journal of Computer Science AndEngineering,Volume-3,Issue-3