Election Prediction On Social Media

Praveen Kumar Singh[1*], Anurag Seetha[2]

[1,2] Department of Computer Science & Application, MCU, Bhopal, India

[*]*Corresponding Author: Praveen_sin2000@yahoo.com, Tel.: +91-9827393494*

Abstract— Social media today is the most popular medium of communication, due to its immediacy. According to Statista, the number of social media users in India is 226 million (2018) and this is expected to go up to 336 million by 2021. The 2014 Lok Sabha elections witnessed a significant usage of social media by political parties and leaders, especially the BJP and their then PM designate Narendra Modi to disseminate their ideology, policies and programmes and highlight the shortcomings / corruption-related scandals of the previous regime. All this helped in creating what is called the 'Modi wave', and led to BJP sweeping the 2014 polls. After 2014, most political parties realised the importance of social media and registered their presence on platforms like Facebook, Twitter, Instagram. About 65 percent of India's population is within the age group 18-35. This group spends almost 4 hours on the internet. Political parties are therefore targeting this group of voters for mobilisation, as most of them use Twitter / Facebook to consume news. This paper represents various issues, methodologies, techniques and research work carried out for election prediction.

Keywords—Social media, performance indicators, sentiment analysis, prediction

I. INTRODUCTION

Social media which allows people to participate in online activities and shatters the barrier for online users to create and share information in any place at any time generates massive data in an unprecedented rate. With such a large amount of social media user activity records, it makes the analysis of online social media user possible. Mining the user online activities patterns will greatly improve the Internet based services and enable many real world applications such as content/item recommendation, personalized information retrieval, event prediction and etc. From application perspective, social media provides online signals that can sense people physical world activities. Social media can be defined as a collection of internet-based applications that expand the ideological and technological foundations of Web 2.0 and that permit the creation and exchange of user-generated content [1]. Social networking sites are the interface between people and social media, and for many the "Internet" is synonymous with social networking sites[2]. One of the most interesting characteristic of social media is represented by the term "user-generated content", which refers to different forms of media content, publicly available and created by end users [3]. Therefore, people use social media not only to consume online information, but also to produce unique content themselves, transforming from content "consumers" to content "producers".

Today youth is increasingly using numbers of social sites such as Facebook, Twitter, WhatsApp, Blog and LinkedIn as platforms for communicating with friends, family and work colleagues. Social Media's quick development shows its influence on society and is a crucial part of the advancement of information and communication technologies. Given its popularity, social media could be used as an influential channel for opinion leading, including agenda-setting and public opinion formation.

Social Media has rapidly grown in importance as a platform for political activities in its different forms. Social media platforms, such as Twitter, Facebook and YouTube provide new ways to encourage citizen get involved in political life, where elections and electoral campaigns have a central role. India, a largest democracy in the world, is secular and socialist country. Here, election of representatives of people involves management of the mega event in the world.

During 2014 general elections, there were around 0.9 million polling stations across the country. The electoral roll is a list of all people in the constituency who are registered to vote. Only those people with their names on the voting list are allowed to vote. For a party to form the government, it must win at least 272 seats or prove its majority with the help of alliances. The responsibility of conducting elections in a fair manner lies with the Election Commission of India (ECI). The elections of 2014 had a battle between few major national parties, namely the Indian National Congress (INC)

or commonly known as Congress, the Bhartiya Janta Party (BJP), the Aam Aadmi Party (AAP) and some other national parties like the Samajwadi Party (SP), the Bahujan Samaj Party (BSP) etc. forming the third front.

There was a significant change in the General Elections 2014 from the General Elections 2009; this was the change in the role played by the social media during the elections. It has been observed worldwide that the democracies have been engaging in dialogues with the public over the social media [4]

Almost all the major parties made their presence felt on the social media with the official accounts and verified pages of their leaders and the parties. As per the media reports, professional help was taken by the parties to improve the image of their party and leaders on the social networking sites. General Elections in India is the biggest democratic exercise in 5 years. And this time with social media as an added factor, it generated a bulk of data everywhere. The differences between parties could be a result of differences in target audience, content strategy and other factors, but these factors are not yet thoroughly explored. Additionally, the higher scores for certain candidates could reflect that a party candidate has an already large social network in the real world. Nevertheless, there seems to be a powerful influence from Social Media on reaching and influencing voters. The increasing dependence on technology for basic communication also highlights the importance of analyzing how SNS are affecting daily processes. Sites like Face book, Twitter, Youtube and LinkedIn are influencing the way users establish, maintain and cultivate a range of social relationships [22].

With all the real time data with us, we aimed to relate some semantics with the data. We propose to develop a framework for analyze the social media data extracted from different applications and find out predictions for election by using efficient and improved approach.. We wanted to see what the popularity on different parameters for different politicians and political parties.

Section II contain details about social media and election. Section III describe related work that had been done by various researchers, Section IV represents research gap found. Section V contain conclusion and future work.

II. SOCIAL MEDIA AND ELECTION

Types of social media platforms: The most popular Social Networking Sites in the United States in March 2015 (based on market share of visits) were Facebook as number one, followed by Youtube, Google Plus and then Twitter [13]. Table below represents various social media platforms:

Table 1. Social media platforms

Types	Examples
Social Networks	Facebook, LinkedIn, MySpace, Googleplus
Blogs	Blogger, WordPress
Microblogs	Twitter, Tumblr
Social News	Digg, Reddit
Social Bookmarking	Delicious, StumbleUpon
Media Sharing	Instagram, Youtube
Wikies	Wikipedia
Review Sites	Yelp, Tripadvisor

Proposed work will mainly focus on Facebook , Twitter and Youtube as these Social Networking Sites are most applicable for the purpose of this research due to the valuable and 'chatty' data which can be gathered from these communication channels.

Social media Key Performance Indicators:
Following are social media indicators used for analysis of social media platforms[14]:
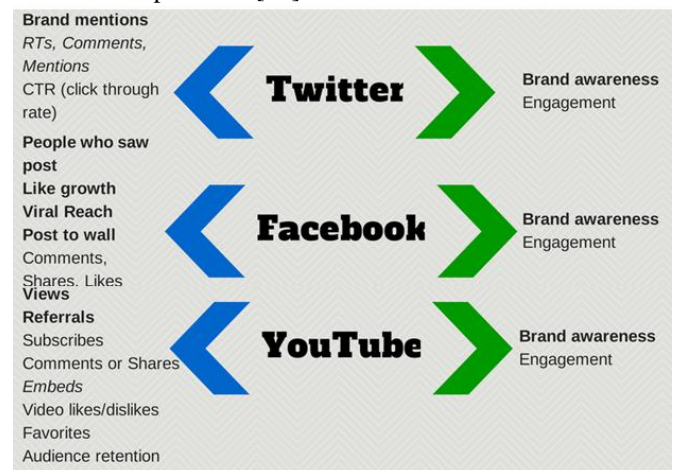

Figure 1: Social Media KPIs

Sentiment Analysis: Sentiment analysis is the process of detecting the contextual polarity of the text. It determines whether given text is positive, negative or neutral. It is otherwise called as opinion mining too, since it derives the opinion or attitude of the speaker.

In general, sentiment analysis has been investigated mainly at three levels [15]. In document level the main task is to classify whether a whole opinion document expresses a positive or negative sentiment. This level of analysis assumes that each document expresses opinions on a single entity. In sentence level the main task is to check whether each sentence expressed a positive, negative, or neutral opinion. This level of analysis is closely related to subjectivity classification, which distinguishes objective sentences that express factual information from subjective sentences that express subjective views and opinion. Document level and

the sentence level analyses do not discover what exactly people liked and did not like. Aspect level performs finer-grained analysis. Instead of looking at language constructs (documents, paragraphs, sentences, clauses or phrases), aspect level directly looks at the opinion itself.

Comparison of different approaches of sentiment analysis [16] is shown below:

Table 2. Sentiment Analysis Approaches

Approaches	Findings
Machine Learning	Supervised & unsupervised both, Dictionary is not necessary, High accuracy
Rule Based	Supervised & unsupervised both, Sentence level works better than word level, Accuracy depends on rules
Lexicon Based	Unsupervised, Learning is not required, Powerful resources are not available

Impact of social media on Indian politics: For politician, the coming loksabha election is not about getting elected; it is about getting socially elected. Social media has changed the way people think, write and react but political pundits believe it has also influence the way people vote. Most of the political parties are aware of the pulse and the impulse of the public on social media timelines, which is why more and more politician are realizing the importance of social media as an electoral tool. As a result, for the general election, digital strategies have become central to planning political rallies and party manifesto are no longer conspiracies of a coterie but are laid bare in the public eye, if not crowd-source to voters at large. Canvassing demand a party's interactive presence on social platform like twitter, Facebook, You tube, Google plus.

But do social media really bridge the divide between candidates and voters? Does digital democracy have the power to change political fortunes? What makes for a winning political strategy and just how is social media at the helm of a social churn in our country? So the analysis and interpretation of the above question will be carried out by this research , by conducting survey among the young voters. This election will be a litmus test to put the spotlight on the question we started with; can social media influence these voters? The answer may not be an unequivocal yes, but it certainly isn't a no. It's an answer in progress.

The relationship between social media and politics is rooted in the desire for change. Today the public at large is using the online information and networking access to find solution. Civil society with its rising frustration over political apathy, is trying to use social media for a change. This online revolution may be less about technology and more about changing human behaviour.

Social media connects people and gets them talking and sharing is one thing. But it become more meaningful when it allow campaigner to know the voters, target a specific audience, splice demographics, mobiles support, and urge people to participate. When some of these people actively engage in political debates, they become a great tool in spreading the word and influencing opinion. More and younger Indians are getting on the social media bandwagon, and this will increase with time, According to Internet and Mobile Association of India(IAMAI) and IMRB International, the number of Internet users in India had reached 205 million in October this year. And by June 2014, India will have 243 million internet users, at which point of time, it is expected to overtake the U.S as the second largest Internet base in the world [13].

Social networking sites can be described as networks of friends for social or professional interactions [17]. Indeed, online social networks have completely changed the spreading of information by making it easy to share and digest information on the internet [18].

The gradual switch from traditional media to social media is something that must be taken into account by all political campaigns because if taken advantage of, social media opens up the possibility for politicians to structure their campaign to address constituents with different beliefs on a more personal base. With political discussions occurring on these sites, a relevant question is arose: could activity on sites like Twitter be a predictor of election results? [19] discovered that the relative volume of tweets closely mirrored the results of the German federal elections. The researchers concluded that Twitter was being used as a platform for political deliberation, and that the number of tweets reflected voter presence, which closely resembled the live political debate. Additionally, research continues to point out that just searching for political information is more popular than participating in the discussion. Getting information and being active in politics are used more by well educated people. The more the public relies on Facebook and Twitter as not only necessary lines of communication, but as rewarding parts of the communicative process, the more value and influence Facebook and Twitter will have aside from campaigning, politicians have found social media useful for connecting with their constituency. UK Members of Parliament (MPs) have been using Twitter to enhance communication with their constituency since 2007. The influence of media on political participation has been studied extensively in the United States. In taking advantage of new communication platforms during the 2008 presidential election, Barack Obama proved the benefit of using social media for political campaigns. The study found that young people tend to get political information from social media more than any other age group. It is becoming clear that online tools play a significant role in shaping public opinion and setting political agendas [20]

Politicians, citizens and journalists increasingly adopt new social media like Twitter, Facebook and YouTube to support their political purposes, be it to engage with other stakeholders in the political public sphere, campaign, disseminate or retrieve information, or contribute to rational-critical debate[21]. An overview of the Indian social media literature suggests that there has been very little research on the impact of the social media on political campaign in general and voting behaviour in particular.

III. RELATED WORK

Lei Wang and John Q Gan [6], presents a new method for election prediction based on Twitter data analysis is proposed and applied to predict the 2017 French Election. Researcher states that in Twitter based election prediction it is critical to extract informative keywords or features reflecting true sentiment of voters. In addition, traditional prediction models may not be suitable for the data from social networks. To calculate the popularity of a candidate in the election, this paper proposes the following formula:

$$popularity(a) = \left[\frac{pos(a)}{pos(a)+neg(a)} \right] \left[\frac{N(a)}{N(a)+N(b)} \right]$$

where N(a) and N(b) are the number of tweets that are related to candidate a and candidate(s) b respectively, pos(a) and neg(a) are the number of positive and negative tweets for candidate a respectively. As neutral tweets are also considered in the proposed method, the sum of pos(a) and neg(a) is not necessarily equal to N(a). If there are more than two candidates, b represents all the candidates but candidate a. To make the sum of the popularities of all the candidates equal to 100%, the popularities are scaled if needed.

The proposed method considers neutral tweets related to specific candidates, which has been proved to increase prediction accuracy in our case study of predicting the 2017 French election result.

Munaza Ramzan, E. Annapoorna & Shikha Mehta [7] in their paper focuses on predicting the winning party in UP elections with the help of public opinion on Twitter. The paper intends to understand whether tweets can be used as useful method in predicting the election results or is it just a social hype. Twitter API is used to extract tweets about the UP elections. Public opinion is then ascertained with the help of sentiment analysis on the tweets. After sufficient amount of tweets are collected for analysis, we are simply expanding on positive tweets and getting 45.13% for BJP, the highest percentage as compared to other parties. On comparing this result with the exit polls and the actual election results, the prediction of winning party by the twitter data analysis is correct. .In this work, a prediction model is built for elections that would be based on sentiment analysis (positive or neutral or negative) done on the tweets. It can be seen that doing sentiment analysis on Twitter data for opinion mining

can be used for prediction of election results. Further accuracy can be increased if the data is collected just before the commencement of the elections.

Onur Varol et al [8] works on early detection. Early detection, occurring immediately at trending time, is a more challenging problem due to the minimal volume of activity data that is available prior to trending. Our supervised learning framework exploits hundreds of time-varying features to capture changing network and diffusion patterns, content and sentiment information, timing signals, and user meta-data. We explore different methods for encoding feature time series. Using millions of tweets containing trending hash tags, we achieve 75% AUC score for early detection, increasing to above 95% after trending. They uses KNN classifier. They evaluate the robustness of the algorithms by introducing random temporal shifts on the trend time series. Feature selection analysis reveals that content cues provide consistently useful signals; user features are more informative for early detection, while network and timing features are more helpful once more data is available. This work represents an important step toward the automatic detection of campaigns. The problem is of paramount importance, since social media shape the opinions of millions of users in everyday life. Further work is needed to study whether different classes of campaigns (say, legitimate advertising vs. terrorist propaganda) may exhibit characteristics captured by distinct features.

Ben Ltaifa Ibtihela et al [9], they study the problem of Tweet Categorization (TC), which aims to automatically classify tweets based on their topic. The accurate TC, however, is a challenging task within the 140-character limit imposed by Twitter. The majority of TC approaches use lexical features such as Bag of Words(BoW) and Bag of Entities(BoE) extracted from a Tweet content. In this paper, they propose a semantic approach of improving the accuracy of TC based on feature expansion from external Knowledge Bases(KBs) and the use of eXtended WordNet Domain as a classifer.

Parama Fadli Kurnia & Suharjito [10] analyzes social media information. Social media is a platform to share information that is very liked by everyone nowadays because some of the facilities that make it easier for us to communicate with each other, share documents, chat and even create a community. In addition, we can also analyze the content of social media by using several methods in data mining, so that we can get new the information to support decision making that can bring benefits to individuals and companies. The purpose of this research, to create a business intelligence dashboard to observe the performance of each Topic or channel of news posted to social media accounts such as Facebook and Twitter. Topical performance in social media is the number of Topics in articles posted to social

media getting like, share, comment etc. To be able to know the Topic of a news post in social media, used some text classification techniques such as Naive Bayes, SVM and Decision Tree. The comparative results of the algorithms are taken which has the best accuracy of SVM for subsequent implementation in the data warehouse. Meanwhile, the business intelligence dashboard data source will be sourced from the data warehouses that have been made before.

IV.   RESEARCH GAP AND ISSUES

A study says nearly 45% of Indian web users connect on social media to discuss politics. With elections around the corner, there is a lot of chatter and speculation about the impact of social networks on the outcome, specifically with the way it has fuelled some recent movements. Indian parties are becoming increasingly aware of the value of Internet platforms to attract tech-savvy voters. So political parties in urban areas are increasingly becoming tech-savvy, realizing this is the only way to reach out to the articulate young as this could change the democratic political dialogue. While political parties still rely on the traditional and old fashioned ways of campaigning such as posters, rallies, cardboard cutouts as well as house-to-house canvassing to win voters, many politicians realize that social media is set to influence the upcoming vote. An overview of the Indian social media literature suggests that there has been very little research on the impact of the social media on political campaign in general and voting behaviour in particular.

The use of social media in Indian politics is growing and it shows sign of future upsurge. Furthermore, there is abundant research to show the importance of social media in political activities in India. However, what is less studied is the use of social media in influencing an individual Indian voter and getting him/her to cast a final vote. It seems that it makes a difference if politicians use Social Media with or without strategy. Use of Social Media does not always result in a more effective political campaign. It heavily depends on how its use is designed, which emphasizes the need for further research. This research therefore will investigate the use of the three social media tools – Facebook, Twitter and YouTube. The task of election trend prediction on single social media platform like twitter or facebook is not complicated. In literature, we observe that mostly work had been done on twitter. But for finding more accurate outcome it is important to take more common social media platform also. This will increase criticality in achieving result due to different type of their structure and indicators. So it is a problem to get an efficient and accurate prediction by integrating all these factors of different platforms along with sentiment analysis.

V.   CONCLUSION AND FUTURE SCOPE

In this paper we have studied different aspect related with election prediction system using social media. There has been little research on this and that will mostly focuses on twitter. There exists a need for considering many other platforms like facebook, whatsapp, youtube etc. More work on sentiment analysis and key performance indicators on social media need to be carryout. In future development of a system is required, which focuses on all social media platforms with optimized approach in prediction.

REFERENCES

[1]Kaplan, A.M and Haenlein, M. "Users *of the world, unite! The challenges and opportunities of Social Media.*" Business Horizons 53, 59-68, **2010.**

[2]Hinton, S. and Hjorth, L. "*Understanding Social Media*" 1st edition, SAGE Publications, UK, London, **2013.**

[3] Gil de Zúñiga, H., Molyneux, L. & Zheng, P. "*Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships*" Journal of Communication 64   612–634 International Communication Association, **2014.**

[4] Hong, S., and Nadler, D." *Does the early bird move the polls?: the use of the social media tool'twitter'by us politicians and its impact on public opinion."* In Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, ACM, pp. 182–186, **2011.**

[5] R. Effing, J. van Hillegersberg, and T. Huibers , "*Social Media and Political Participation: Are Facebook, Twitter and YouTube Democratizing Our Political Systems?. ",*ePart 2011, LNCS 6847, pp. 25–35, **2011.**

[6] Lei Wang and John Q Gan, "*Prediction of the 2017 French Election Based on Twitter Data Analysis*" , IEEE **2017.**

[7] Munaza Ramzan, E. Annapoorna, Shikha Mehta, "*Are Tweets The Real Estimators Of Election Results?*", Proceedings of 2017 Tenth International Conference on Contemporary Computing IEEE **2017.**

[8] Onur Varol, Emilio Ferrara, Filippo Menczer, Alessandro Flammini, "*Early Detection of Promoted Campaigns on Social Media*", arXiv:1703.07518v1 [cs.SI] 22 Mar **2017.**

[9] Ben Ltaifa Ibtihela,, Hlaoua Lobnab, Ben Jemaa Maher, "*A Semantic Approach for Tweet Categorization*", 22nd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Elsevier **2018.**

[10] Parama Fadli Kurnia, Suharjito, "*Business Intelligence Model to Analyze Social Media Information*", 3rd International Conference on Computer Science and Computational Intelligence **2018.**

[11] Kaplan, A.M., Haenlein, M. "*Users of the world, unite! The challenges and oppurtunities of Social Media*". Business Horizons, 53, 59-68. **2010.**

[12] Constantinides, E., Fountain, S.J. . "*Web 2.0: Conceptual foundations and marketing issues*". Journal of Direct, Data and Digital Marketing Practice, 9(3), 231244. **2008.**

[13] Statista , "*U.S. Social Media visit share. Retrieved from Statista*", **2015.**

[14] "*Social Media KPIs*", measuring your social media and a kpi-cheat sheet, **2014.**

[15] Neha S. Joshi, Suhasini A. Itkat, " *A Survey on Feature Level Sentiment Analysis*" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , **2014,**

[16] Devika M D, Sunitha C, Amal Ganesha, "*Sentiment Analysis:A Comparative Study On Different Approaches*", ICRTCSE Elsevier **2016**

[17] Trusov, M., Bucklin, R.E., & Pauwels, K., "*Monetary Value of word-of-mouth marketing in online communities. Quantifying the Effectiveness of Word-Of-Mouth Marketing*", **2009**

[18] Akrimi, Y., & Khemakhem, R. "*What Drive Consumers to Spread the Word in Social Media? Journal of Marketing Research & Case Studies, 1-14.* ", http://dx.doi.org/10.5171/ **2012.**

[19] Tumasjan, A., T. Sprenger, P. Sandner, & I. Welpe, "*Predicting elections with Twitter: What 140 characters reveal about political sentiment*". In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 178–185, **2010.**

[20] Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., & Malone, T.W., "*Evidence for a Collective Intelligence Factor in the Performance of Human Groups.*" Science, 330, 686, **2010**

[21] Sauter,T, & Bruns, A., "*Exploring Emotions on Polarity, Conservativism and Public Performance in the Twitter Debate on Australian Politics*". Selected Papers of Internet Research, **2013.**

[22] Aindrila Biswas, Nikhil Ingle and Mousumi Roy, "*Influence of Social Media on Voting Behavior*", Journal of Power, Politics & Governance , Vol. 2, No. 2, pp. 127-155, June **2014.**

Authors Profile

Mr. Praveen Kumar Singh pursuing Ph. D. from Makhanlal Chaturvedi Rashtriya Patrakarita Evam Sanchar University, Bhopal.
Pursed Master of Computer Application from MP Bhoj University. He is having more than 15 years of teaching experience. His research area is social media and data mining.