A Review: Prediction of Election Using Twitter Sentiment Analysis

Pritee Salunkhe, Avinash Surnar, Sunil Sonawane

Abstract— **Twitter is a micro-blogging service that provides short messages on a daily basis for events, products, entity etc. Now a day's researchers dealing with utilizing Twitter to monitor people reactions in political activities like debates and campaigns. On the basis of that prediction an election can be made. Analysis of the prediction of election results using messages of either political parties or politician. The use of tweet content considered as a valid indicator of political sentiment. Sentiment analysis is used for analysis as well as predicting the emotions from the text patterns regarding political issues, products, entity, election etc. Sentiment analysis consists an analyzing texts to extract information. Basic sentiment analysis allows to determine or measuring the polarity (negative or positive) of sentiment.**

Index Terms— **Ideological Learning, Opinion Mining, Sentiment Analysis, TwitterSphere.**

I. INTRODUCTION

Election may be a vital half within the democracy. It's the most instrument of democracy wherever the voters communicate with the representatives. One vital component in associate degree election is that the election polls/survey. The polls did not accurately predict the election outcomes. thus attempting to resolve the accuracy and high value problems, we have a tendency to study the chance of exploitation knowledge from social media because the knowledge supply to predict the result of associate degree election. . Social media has become the foremost widespread communication tool on the net. many scores of messages area unit being denote each day within the widespread social media sites like Twitter and Facebook .[1] explicit in their paper that social media websites become valuable sources for opinion mining as a result of folks post everything, from the main points of their existence, like the product and services they use, to opinions concerning current problems like their political and non secular views. The social media suppliers change the users to precise their feelings or opinions the maximum amount as potential to extend the interaction between the users and their sites.

Twitter is a microblogging service launched in 2006 . On Twitter, every user can read and write short messages with up to 140 characters, it's called "tweets", which are visible on a publicly. There are more than one million messages per hour on twitter. The main purpose behind microblogging was to provide personal status updates. some political analysts are use "Twittersphere" as an indicator of political opinion, others have suggested "pointless babble". Twitter is used for sentiment analysis and opinion mining to predict election results.

II. RELATED WORK

In this section, we discuss related works about predicting the result of an election using Twitter. Researchers use a different approach, there are researchers who try to discover the political preference of a user, then relate it to the election and there are others who use selected tweet related to the upcoming election and figure out vote preference of the user using that data.

Various strategies used for inferring political leaning such as profile information, user behavior, user graph, Twitter specific feature (reply/re-tweet), and sentiment from tweet content. For example, In[2], the authors used tweet containing parties' name in several political events to assign a political/ideological leaning of the user who posted the tweets. Similar to the previous method, [3] used the tweets and retweets of a user regarding a political party to infer the political leaning. [4]assigned a score to every member which a Twitter user is following, then a political preference is assigned based on that score. In [5], the authors compared several features such as user's posting behavior, linguistic content, follower, reply and retweet. They found out that the combination between user profile and linguistic outperform other feature.

The second approach is by using selected data just days or weeks prior to the election. The prediction could be derived by comparing the number of tweets mentioning each candidate or by comparing the number of tweets that has positive sentiments towards each candidate. The earliest research stated that the number of tweets mentioning a party reflects the election result was shown in [6] where they found out that the prediction result from Twitter were only slightly worse than offline election polls. While [7] is the first

Pritee Salunkhe, *Department of Computer Science and Information Technology , Dr. Babasaheb Ambedkar Marathwada Universisty, Aurangabad, India, +91 8888033060.*
Avinash Surnar, *Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada Universisty, Aurangabad, India, +91 9921615326.*

Sunil Sonawane, *Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada Universisty, Aurangabad, India, +91 8552082329.*

research in which argued that sentiment detection approach from Twitter can replace the expensive and time intensive polling.

Researches have tried to compare these two methods, for example, [8] that tried to predict congress and senate election in several states of the US. They showed that though the method is the same, the prediction error can vary greatly. The research also showed that lexicon based sentiment analysis improve the prediction result, but the improvement also vary in different states. Same result was shown in [9] where they predict the result of Irish general election using both methods and [10] which predict the Italian primary election. All of the research showed that sentiment detection do reduce the error of the prediction result. Because of that, several researchers focused on improving the sentiment analysis, such as [10] and [11]who used more sophisticated sentiment analysis than lexicon based in the US presidential election, France legislative election, and Italy primary election.

Other than using sentiment analysis, the prediction result from Twitter can be improved by using user normalization. This is based on the fact that in an election, one person only have one vote. [12]implemented this method and showed that the prediction result of 2011 Dutch senate election was improved. [13] takes further step by adding census correction on the user normalization. [14] also implemented this method in several south American countries. He collected more than 400 million of tweets, and got a very good result (low difference with the election result) predicting Venezuela presidential election. But when applying in Ecuador and Paraguay presidential election that has much less dataset, the error of the prediction increase significantly.

Other methods proposed by researchers are by utilizing interaction information between potential voter and the candidates and creating trend line from the changes in follower of the candidates. [15]used interaction information such as the number of interaction, the frequency of interaction, the number of positive and negative terms in the interactions in the Canadian legislative election. The candidates were grouped into four parties, and based on their result, they argued that that the generated content and the behavior of users during the campaign contain useful knowledge that can be used for predicting the user's preference. [16]tried to utilize the size of candidates' network (follower in Twitter and friend in Facebook), but the result showed that it was not a good predictor of election results. One interesting result from their research is that despite the huge size of social media, it has small effect on the election results. Therefore, it only make a difference in a closely contested elections.

However, there are several researchers arguing that research in this area is still premature and requires a lot of development before it can give satisfying prediction result.[17] argued that prediction model using Twitter only able to predict the result from the top candidates/parties and slight variable changes in the model did impact the prediction result. In [18], the authors listed several drawback of the research in this topic such as, most predictions are actually a post-hoc analysis, no commonly accepted way

exists for "counting votes", the sentiment analysis methods are not reliable, no data cleansing step, demography and self-selection bias has not been addressed. In [19], in addition to previously stated drawbacks, gave several suggestions such as the importance of geographical and demographical bias, the noise in the social media, the reproducibility of proposed methods, and MAE should be use rather than only winner prediction.

III. CONCLUSION

As the number of tweets mentioning a party imitate the election result leaning towards party. In this paper, various strategies are discussed regarding to political leaning like user graph, twitter specific features, user behaviour for the appropriate prediction of the elections. Here we provide a view related to several features as comparison using user's posting behaviour, linguistic content, reply. The combination between user profile and linguistic outperforms other features. The different approaches are discussed for analysis like selecting data just days or weeks prior to the election. The prediction could be made by comparing the number of tweets that has positive sentiments towards each political candidate. Other methods like utilizing interaction information between potential voter and the candidates could be important for analysis of election prediction.

REFERENCES

[1]Pak, A. &. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *LREC*.

[2]Wong, F. M. (2013). Quantifying Political Leaning from Tweets and Retweets. *ICWSM*.

[3]Boutet, A. K. (2012). What's in your Tweets? I know who you supported in the UK 2010 general election. *Proceedings of the International AAAI Conference on Weblogs and Social Media*.

[4] Golbeck, J. &. (2011). Computing political preference among twitter followers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

[5] Pennacchiotti, M. &. (2011). Democrats, republicans and starbucks afficionados: user classification in twitter. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* , 430-438.

[6] Tumasjan, A. S. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *ICWSM, 10*, 178-185.

[7] O'Connor, B. B. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM, 11*, 122-129

[8] Gayo-Avello, D. M. (2011). Limits of electoral predictions using twitter. *ICWSM*.

[9] Bermingham, A. &. (2011). On using Twitter to monitor political sentiment and predict election results.

[10] Ceron, A. C. (2014). Using Sentiment Analysis to Monitor Electoral Campaigns: Method

Matters—Evidence From the United States and Italy. *Social Science Computer Review.*

[11] Ceron, A. C. (2013). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society, 16(2)*, 340-358.

[12] Sang, E. T. (2012). Predicting the 2011 dutch senate election results with twitter. *the Workshop on Semantic Analysis in Social Media* (pp. 53-60). Association for Computational Linguistics.

[13] Choy, M. C. (2012). US Presidential Election 2012 Prediction using Census Corrected Twitter Model. *arXiv preprint arXiv:1211.0938.*

[14] Gaurav, M. S. (2013). Leveraging candidate popularity on Twitter to predict election outcome. *Proceedings of the 7th Workshop on Social Network Mining and Analysis. ACM.*, 7.

[15] Makazhanov, A. R. (2014). Predicting political preference of Twitter users. *Social Network Analysis and Mining*, 1-15

[16] Cameron, M. P. (2013). Can Social Media Predict Election Results? Evidence from New Zealand. *No. 13/08.*

[17] Jungherr, A. J. (2012). Why the pirate party won the german election of 2009 or the trouble with predictions: A response to tumasjan, a., sprenger, to, sander, pg, & welpe, im "predicting elections with twitter: What 140 characters reveal about political sentiment". *Social Science Computer Review, 30(2),*, 229-234.

[18] Gayo-Avello, D. (2012). No, you cannot predict elections with twitter. *Internet Computing, IEEE, 16(6)*, 91-94.

[19] Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review.*