PREDICTING ELECTIONS: SOCIAL MEDIA DATA AND TECHNIQUES

Muhammad Bilal
[a]*School of Computing and IT,*
[b]*Centre for Data Science and Analytics,*
Taylor's University,
Subang Jaya, Malaysia
mbilal.csit@gmail.com

Abdullah Gani (SMIEEE)
[a]*School of Computing and IT,*
[b]*Centre for Data Science and Analytics,*
Taylor's University,
Subang Jaya, Malaysia
abdullah.gani@taylors.edu.my

Mohsen Marjani
[a]*School of Computing and IT,*
[b]*Centre for Data Science and Analytics,*
Taylor's University,
Subang Jaya, Malaysia
mohsen.marjani@taylors.edu.my

Nadia Malik
Department of Management Sciences,
COMSATS University Islamabad,
Islamabad, Pakistan
nadia_malik_16@yahoo.com

***Abstract* — During the last decade, the use social media has provided a virtual community where the users express their intention, opinion and communicate with others. Organizations and researchers are very much interested to investigate the polarity of the user's opinions for making predictions and planning of future plans. The use and influence of social media in politics have been recognized by many researchers and political parties in making electoral predictions and devising future campaign strategies. As the literature on making electoral predictions is increasing day by day, this study aims to give an overview of the current state of approaches being utilized, social media used for data collection and the outcomes of the proposed approaches, either succeeded in making accurate electoral predictions using social media data and approaches used or not. This study gives a quick overview and guides researchers in understanding the approaches previously adopted and data sources utilized in predicting electoral results.**

Keywords—Election Prediction, Social Media Data, Sentiment Analysis, Politics, Twitter, Facebook

I. INTRODUCTION

Social media is revolutionizing the ways of social life of individuals by creating a virtual community for self-expression, connecting and collaborating with others. The sharing of information on social media results into a huge amount of user-generated content which discloses personal information and interests that can be used in making predictions inconspicuously. The data collection and analysis from social media data is time and cost-effective when compared with the traditional approaches. The researcher reported that the predictive power for measuring individual attributes from social media data does not affect by the representative nature of the sample to be a true representative of the population [1]. Twitter is one of the most frequently used social networking platforms being for data analysis and has approximately 100 million active users on daily basis. The online website's data can also be utilized for predicting and reporting job opportunities [2,3]. The Twitter data has been analyzed by researchers in healthcare [4], classification and comparative analysis [5] and predicting elections [6 – 9]. In addition to making predictions, many researchers also investigated the uncertainty of results [7].

Presently, political parties increasingly rely on social media platforms i.e. Twitter and Facebook for political communication, interacting with voters and promotions. This increased use of social media by political candidates for attracting potential voters has been reflected by the 2011 general elections in New Zealand. In a modern democracy, the problem of predicting electoral results is very popular and has attracted the attention of researchers from computing domain to predict election results based on data collected from social media platforms [10,11]. Prior studies i.e. Skoric et al., 2015 [12], on predicting electoral results reported accurate outcomes using data collected from social media networks and techniques such as sentiment analysis, volumetric analysis and social media analysis for countries including United States [8], United Kingdom [6], Ireland [13] and Germany [9].

The electoral studies are very different from other studies being conducted in political science domain, as their goal is not to explain the election results but to forecast outcomes. The number of studies investigating the association of social media and electoral results is rising quickly [14-17]. The influence of social media on electoral results has also been explored by previous studies. As the literature on prediction of electoral results using social media data is expanding, this paper aims to explore, summarize and present a mini review of existing literature related to election predictions using social media data. The paper is organized as section II provides a quick overview of related work and reports data sources used along with techniques, section III, discusses the studies presented in related work and section IV, presents the conclusions.

II. RELATED WORK

Previously, the polling data has been considered as the best estimator for forecasting electoral results. But, recently the polling data has been considered partial and erroneous. Therefore, a research, investigated the accuracy of polls in comparison with sentiment analysis results performed on Twitter tweets. Surprisingly, the results reported that Twitter was 3.5% more biased in popular votes and 2.5% more biased in state results when compared with the polls. The study concluded that the predictions based on Twitter data are worse when compared with polling data [18]. To examine the effectiveness of the previously proposed methods for predicting electoral results i.e. sentiment analysis, by applying on a new dataset of Twitter tweets related to politics. The dataset of 234,697 tweets was collected using the Twitter streaming Twitter API. The data preprocessing was performed on the collected tweets by removing the hashtags, links and names of accounts. Emotions and similes were replaced by full form i.e. "☺" to <happy>. The study reported that there have been many limitations attached to the previous methods due to which they were inadequate in predicting electoral results using social media data. It has been suggested that to improve the accuracy of predictions the researchers should not only rely on the polarity of words alone. The approaches of POS Tagging, Sense disambiguation should be adopted in preprocessing along with considering the contextual and lexical features of words [19].

The traditional techniques of predicting election results i.e. polling have become unreliable with the frequently changing technology. Due to the increased use of smartphone and easily available internet facility, the social and digital media has become the platform for presenting political views. Numerical comparisons have been done for the slogans used in Twitter tweets for US elections 2016 and visualized using WordCloud. The results reported by the study were inconsistent then the actual outcomes of the elections and it was suggested by the researchers to further consider and evaluate qualitative aspects for making electoral predictions. The Trump win was not predicted by the Twitter for the states of Michigan and Wisconsin by any of the approaches being utilized by the study [20]. A study explored the relationship between the size of the social network of candidates and the chances of winning the elections using regression analysis and data collected from Facebook and Twitter. Three models were proposed in which the number of votes was taken as a dependent variable and number of Facebook connections, along with other factors were taken as independent variables. The results reported the size of the network and the chances of a win are significantly correlated to each other. Hence, the election results were reported to be inferred by looking at the size of social network. However, the magnitude of the effect is very small and the social media data was reported to be predictive for elections with close competition. The proposed model is elaborated in Figure 1 [21].

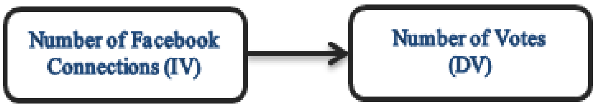

Figure 1: Regression model representation IV & DV

Social network techniques i.e. volumetric analysis, sentiment analysis, has been utilized by authors to evaluate the predictive power of Twitter data for inferring electoral results for three countries, Pakistan, India, and Malaysia. The data preprocessing was performed on approximately 3.4 million Tweets collected using Twitter streaming API. To separate the tweets in English language a natural langue toolkit of python was used. 90% of the Tweets from Pakistan and India are English, but on the other hand only 23% of the tweets were in English. The equation (1) shows volumetric analysis approach, where $\mathrm{Vol_x}$ for any party x represents the volume of tweets and $\mathrm{c_x}$ represents tweets count. Equation (2) shows sentiment analysis whereas equation (3) represents social network analysis approach. In equation (2) $\mathrm{Sent_t}$ is for sentiment of tweets $\mathrm{pos_t}$ represents positive tweets where as $\mathrm{neg_t}$ represents negative tweets. In equation (3) the centrality score of a party x is represented by $\mathrm{Net_x}$ out of n number of parties and the raw score is represented by $\mathrm{s_i}$ .The performance was measured using Mean Absolute Error (MAE) and is represented by equation (4). The results reported that the Twitter data was not effective for making election predictions for Malaysia, but in the case of Pakistan and India, it appeared as an effective and efficient for electoral predictions. By combining multiple techniques the proposed model for predicting electoral outcomes was also effective for candidates and parties having small vote count [22].

$$\mathrm{Vol_x} = \frac{\mathrm{c_x}}{\sum_{j=1}^{n} \mathrm{c_j}} \% \qquad (1)$$

$$\mathrm{Sent_t} = \{\, 1, \mathrm{pos_t} > |\mathrm{neg_t}| - 1, \mathrm{pos_t} < |\mathrm{neg_t}|, \mathrm{pos_t} = |\mathrm{neg_t}| \,\} \qquad (2)$$

$$\mathrm{Net_x} = \frac{\mathrm{s_x}}{\sum_{j=1}^{n} \mathrm{s_j}} \% \qquad (3)$$

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\mathrm{e_i}| \qquad (4)$$

The Facebook data has not been much utilized for predicting electoral results. However, a research predicted the congressional polls by using a proposed model, Senate Vote = ƒ (partisan voting index + incumbency + participation advantage) and data collected from Facebook. Where, senate vote is the percentage of forecasted votes won by either two of the major parties, Partisan Vote Index (PVI) is the past election results. The metric from Facebook was used to calculate the incumbency and participation advantage. The incumbency was added in the proposed Facebook model to overcome the limitation of PVI. The participations factor was used to enable the model for forecasting trends. It represents the real-time success rate of each political campaign. The components attached with the participation variable include likes, active post engagements, and time slices. The track of fans and the post engagements has been continuously kept by the Facebook pages. The metrics used by the research are expected to be influential for conducting political campaigns to outreach potential supporters. The results of the study reported that Facebook data accurately predicted the outcomes for US Senate elections 2012. It is recommended to further investigate the usefulness of Facebook data for predicting electoral outcomes to verify the effectiveness and accuracy of the proposed model [23].

A predictive model based on Naïve Bayesian approach using Twitter data has been proposed to predict the electoral result for the US presidential elections 2016. The Twitter tweets have been collected over the period of Three months, from December till February. After doing simple pre-processing of data the model for sentiment prediction has been reported to achieve 95.8% accuracy. The model accuracy was tested using 10 fold cross validation. F1 Score shows the accuracy of model for predicting positive sentiments whereas ~F1 donates the accuracy of model for classification of negative sentiments. The accuracy for negative sentiments is represented by equation (5). The candidate having the greater volume of positive sentiments was considered to be the selected nomine for a particular party in elections. The error (E), equation (7) was calculated by doing the comparison of positive sentiments with the results of opinion polls by considering the number of nominees (e_i). In the equation (6) $1 \leq i \leq n$ where n is number of total nominees, Po represents poll and Pre represents predictions. The reported accuracy of model was 98.5% but the actual polls results were reported to be predicted with 54.8% accuracy. The reported results, the prediction made by the model might not be accurate when compared with the results of actual polling data as the election was reported as ongoing [24].

$$\sim F_1 = 2 \times \frac{\frac{TN}{TN+FN} \times \frac{TN}{TN+FP}}{\frac{TN}{TN+FN} + \frac{TN}{TN+FP}} \quad (5)$$

$$e_i = |Po_i - Pre_i| \quad (6)$$

$$E = \frac{\sum_i^n e_i}{n} \quad (7)$$

The accuracy of classification of Twitter tweets using sentiment analysis has been questioned by many researchers. To enhance the accuracy of classification a two-staged Machine Learning model was proposed for the creation of training dataset by keeping into consideration the contextual features from the collected data from Twitter. Out of 60 thousand, 30 thousand processed tweets were used as the training dataset as neutral tweets were discarded to enhance the quality of results. The Machine Learning algorithms including NB and SVM were used to classify the tweets according to polarity. The accuracy achieved by using SVM was more than NB algorithm. The candidate predicted to win the election was based on, $\text{Ratio} = |P| \ / \ |T|$ where P represents number of positive tweets and T represents the total number of tweets. The proposed model was reported to be more accurate when used in combination with statistical tools and offline approaches i.e. votes casted [25].

A study investigated the effectiveness of information sharing by political candidates over Twitter for polling and predicting electoral outcomes. A dataset of nearly 3 million Twitter tweets has been collect in 2016 over the period of three months from September to November to perform analysis at both state and national level. The election prediction was performed only at the state level by using approaches of winner-take-all and by the polling count based on sentiment analysis and deep learning. The volume of tweets and the number of positive tweets per candidate has been used as the metric for prediction of electoral results for twenty-one states. Deep neural network was used to automatic labeling of data, as the volume of data was high and the manual labeling was not possible. Python API along and Tensorflow were used to implement and excute CNN. The study reported that using polling data along with sentiments adds value to the predictions at the national level. However, the volume of tweets used for the prediction of electoral results of individual candidates is not an efficient parameter [26].

To validate the reliability of Twitter data in planning and predicting electoral results Linguistic Inquiry and Word Count (LIWC) approaches has been utilized by researchers for federal elections in Germany. The results extracted by the analysis of approximately 100,000 tweets reported that Twitter is very frequently used for political discussions. The tweets collected were related to the six parties form the German Parliament and prominent politicians form these political parties. The sentiment form the tweets were extracted using automatic text analysis software, LIWC2007 which works on the basis of an internal psychometric dictionary. To process the tweets using LIWC English dictionary all tweets were automatically translated from German to English. To clearly interpret the results in addition to radar charts distance measures d for various combinations of political candidates and parties were calculated from the perspective of 12 dimensions. The equation (8) below represents the distance measure d for all dimensions. The value of the i-th dimension for any political entity p is referred as di,p . The low value of d represents the political parties has similar profiles. The relationship of political parties has also been considered after measuring the distance among them. Only two parties CDU, CSU were taken to demonstrate these relationships. The strength of relationship was measured by taking in consideration the tweets in which one party mentions or refers the other party. The equation (9) below was used to calculate the relative frequency (f), share (CDU, CSU) was the joint share reported. The results although concluded that the sample of Twitter tweets and the dataset used were not the true representative of German elections, but however the activity before elections can be used to predict election outcomes [9].

$$d = \sum_{d=1}^{n} \frac{\left| d_{i,p} - \left(\sum_{p=1}^{n_p} d_{i,p}\right) / n_p \right|}{n_d} \Big/ 12 \quad (8)$$

$$f = \frac{\text{share(CDU,CSU)}}{\text{P(CDU|CSU)+P(CSU|CDU)/2}} \quad (9)$$

The effectiveness of social media in predicting electoral results has been investigated for general elections 2013 in Pakistan. Naïve Bayes (NB) and K-Nearest Neighbors (KNN) were used as classification algorithms. The accuracy of data was data was tested by taking 40% of data as a testing data, whereas rest of data was used for the purpose of training classifiers. The accuracy calculated of two only two classes positive and negative, was 70%. But the accuracy eventually drops to 50% when the neutral tweets were included. The sentiment analysis for classing tweets as positive, negative and neutral has been performed and the results reported were appeared contradictory when compared with actual results declared by the Election commission of Pakistan. The PTI has 79.29 positive tweets but get the actual polled votes of 20.32%, whereas PMLN has 57.50% positive tweets but get the actual polled votes of 39.35%. Hence, Twitter data was reported as a not reliable and non-accurate source for predicting the electoral results in the case for Pakistan general elections 2013. The need of other approaches along with existing approaches was suggested to utilize for accurate prediction of electoral results [27].

A system has been proposed for the real-time analysis of sentiments extracted using Twitter data for electoral results of 2012 elections in the US. It is further explored how the sentiments vary and influenced by the occurrence of various events. To test the proposed sentiment model the data for training and testing was generated by crowdsourcing for performing sentiment annotation on political data. The analysis and prediction of results took multiple days to report the results using naïve Bayes model and unigram features. The accuracy of the proposed model for accurately classifying the data into four categories, positive, negative, neutral and unsure was 59%. This was much better than the accuracy of previous classifiers which was appeared to be 56%. The proposed approach also accomplished in accurately classification of data for each category. The advantage of the proposed system was the instant and timely predictions of results by real-time monitoring of Twitter that will be useful in making the future plans [28].

A bilinear model has been proposed to predict the voting intentions using Twitter data. The two datasets of social media content were generated for UK and Austria. A structured multi-task model has been also utilized to overcome the problem of formulating groups. There are multiple ways in which information can be incorporated into regression model. One way by simply adding the features, but this will lead to failure due to increasing size of features set. The grouping of tweets was appeared as a one possible solution. The naïve techniques failed where users use same words to communicate there indentions and opinions. The bilinear model is represented by equation (10) where u and β are taken as parameters whereas the matrix m x p for each word count is represented by X. The accuracy of proposed model has been evaluated by using a real-life case study to emulate voting intention prediction. The data from 10 polls have been used to test the model whereas data from 5 polls have been used to test the proposed model. The results demonstrated that the proposed approach performed much better than the traditional approaching being utilized for predicting voting intentions [29].

$$f(X) = u^{T}X\omega + \beta \tag{10}$$

The Summary of the studies discussed in related work section is presented in Table I.

TABLE I. SUMMARY OF LITRATURE REVIEWRD

Reference	Description	Data Sources		Techniques		Outcomes	
		Twitter	**Facebook**	**Sentiment Analysis**	**others**	**Can predict**	**Cannot Predict**
Tumasjan et al., [9]	Validated the reliability of Twitter data in planning and predicting electoral results Linguistic Inquiry and LIWC.	✓			✓		✓
Chung & Mustafaraj, [19]	Examined the effectiveness of the previously proposed methods for predicting electoral results using new dataset.	✓		✓	✓		✓
Wang et al., [28]	Performed real-time analysis of sentiments extracted using Twitter data for electoral results of 2012 elections in the US.	✓		✓		✓	
Lampos et al., [29]	Proposed a bilinear to predict the voting intentions using Twitter data.	✓			✓	✓	
Razzaq et al., [27]	Examined the effectiveness of social media in predicting electoral results for general elections 2013 in Pakistan.	✓		✓			✓
MacWilliams, [23]	Predicted the congressional polls by using a proposed model, Senate Vote = *f* (partisan voting index + incumbency + participation advantage).		✓		✓	✓	
Cameron et al., [21]	Studied the relationship between the size of the social network of candidates and the chances of winning the elections.	✓	✓		✓	✓	
Tunggawan & Soelistio, [24]	Predicted the electoral result for the US presidential elections 2016.	✓			✓		✓
Ramteke et al., [25]	Enhanced the accuracy of classification by keeping into consideration the contextual features.	✓		✓		✓	
Anuta et al., [18]	Investigated the accuracy of polls in comparison with sentiment analysis.	✓		✓			✓
Jaidka et al., [22]	Evaluated the predictive power of social media data for inferring electoral results for three countries, Pakistan, India, and Malaysia.	✓		✓		✓	
Hinch, [20]	Performed numerical comparisons for the slogans and visualized using WordCloud for predicting election results of US elections 2016.	✓			✓		✓
Heredia et al., [26]	Investigated the effectiveness of information sharing by political candidates over social media for polling and predicting electoral outcomes.	✓		✓		✓	

III. Discussion

Many attempts have been made by researchers to present a reliable solution for the problem of predicting electoral results. The solutions proposed for predicting election results are highly dependent on the techniques and data sources used. This classification of the techniques and data sources used by previous studies is represented in Figure 2 and Figure 3 respectively. The techniques are grouped into categories of volumetric analysis, sentiment analysis and social network analysis. The techniques associated with sentiment analysis are sub classified into supervised and unsupervised approaches. This survey only focused and discussed the studies that used data collected from social media. Twitter is appeared to be the most frequently used social media for political studies and predicting electoral results, whereas Facebook data is used by very few studies until now. Historical election results and opinion polls are among the other sources of data that are also being utilized in addition to social media data. The sentiment is the most used approach used for data analysis and electoral prediction by using Twitter tweets.

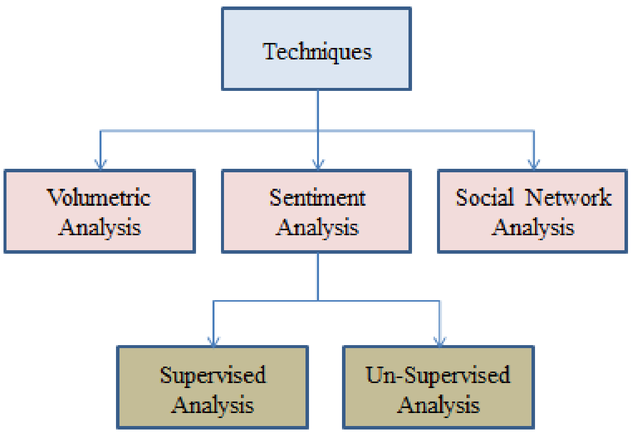

Figure 2: Classification of Techniques

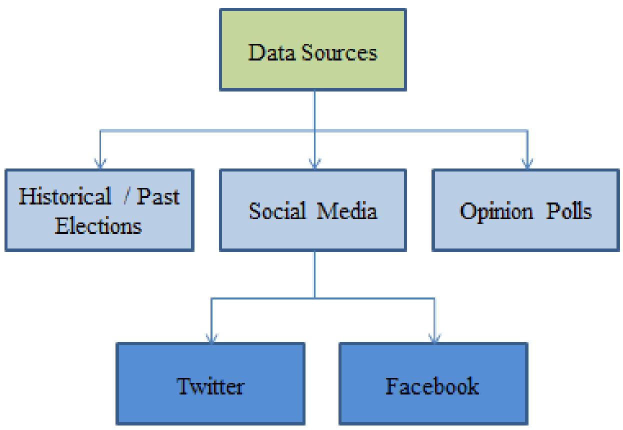

Figure 3: Classification of Data Sources

There are many limitations and open challenges associated with each technique and data source being utilized by previous studies. The few limitations and open challenges which needed to be addressed by future studies are as follow:

- *Language Barrier:* The data collected form most of the Twitter and Facebook is in languages other than English as well. The technique of sentiment analysis is only capable to process English text to label the polarity. Due to this the data is required to translate into English language or discard the text in any other language. This language was faced by a study tried to predict the electoral results in three countries i.e. Pakistan, India and Malaysia. But the data collected from Twitter accounts from Malaysia was 77% in Malay and only 23% was in English [22]. Hence in situations like this where majority of data got discarded prediction accuracy will suffer and questionable.
- *Misclassification:* The Twitter tweets processed using approach of sentiment analysis faced the problem of misclassification, when a positive tweet is labeled as a negative tweet and vice versa.
- *Data Imbalance:* The sentiment analysis approach greatly used by previous studies can also lead to the problem of data imbalance, in which majority of tweets are allocated to one class when compared to the other class.
- *Data Reliability:* The data being used for making predictions matters a lot. Garbage in garbage out (GIGO), the predictions made on a reliable data will be reliable and the predictions made by using unreliable data will be definitely unreliable. The data collected from Twitter is reported to be unreliable for making electoral predictions by various studies. Similarly, many limitations are attached to the historical elections results and opinion polls that needed to be addressed.

IV. Conclusions

The topic of predicting elections is gaining the attention of researchers as the utilization of social media data is making an important place due to its real-time nature and easy availability. Many attempts have been made by researchers to explore and validate the reliability of social media data in predicting election results. Mostly, twitter data was used by researchers in making predictions of electoral outcomes by checking the polarity of words using sentiment analysis. It is observed that very few studies utilized data collected from Facebook in making electoral predictions. Majority of studies found the social media data effective in making electoral predictions. Whereas few of the studies also discussed and reported that social media data cannot be relied upon. Similarly, sentiment analysis is the most common approach being adopted by researchers. But most of the studies also argued about the limitations attached to the approach of sentiment analysis and suggested to consider the linguistic and contextual feature as well. In the future work, the data sources other the social media will be discussed along with the comparison of data sources and approaches used.

References

[1] Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). 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.

[2] Lovaglio, P. G., Cesarini, M., Mercorio, F., & Mezzanzanica, M. (2018). Skills in demand for ICT and statistical occupations: Evidence from web- based job vacancies. Statistical Analysis and Data Mining: The ASA Data Science Journal, 11(2), 78-91.

[3] Bilal, M., Malik, N., Khalid, M., & Lali, M. I. U. (2017). Exploring Industrial Demand Trend's in Pakistan Software Industry Using Online Job Portal Data. University of Sindh Journal of Information and Communication Technology, 1(1), 17-24.

[4] Nawaz, M. S., Bilal, M., Lali, M. I., Ul Mustafa, R., Aslam, W., & Jajja, S. (2017). Effectiveness of social media data in healthcare communication. Journal of Medical Imaging and Health Informatics, 7(6), 1365-1371.

[5] Tiwana, M. S., Javeed, F., Lali, I. U., Dar, H., & Bilal, M. (2018). Comparative Analysis Of Context Based Classification Of Twitter. Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA 2018), IEEE.

[6] Boutet, A., Kim, H., & Yoneki, E. (2012). What's in Your Tweets? I Know Who You Supported in the UK 2010 General Election. ICWSM, 12, 411-414.

[7] Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. Social Science Computer Review, 31(6), 649-679.

[8] Livne, A., Simmons, M. P., Adar, E., & Adamic, L. A. (2011). The Party Is Over Here: Structure and Content in the 2010 Election. ICWSM, 11, 17-21.

[9] Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm, 10(1), 178-185.

[10] Fisher, S. D., Ford, R., Jennings, W., Pickup, M., & Wlezien, C. (2011). From polls to votes to seats: Forecasting the 2010 British general election. Electoral Studies, 30(2), 250-257.

[11] Gibson, R., & Lewis-Beck, M. S. (2011). Methodologies of election forecasting: Calling the 2010 UK "hung parliament". Electoral Studies, 30(2), 247-249.

[12] SKORIC, M., Liu, J., & Lampe, C. (2015, May). Gauging public opinion in the age of social media. In International Communication Association Annual Conference.

[13] Bermingham, A., & Smeaton, A. (2011). On using Twitter to monitor political sentiment and predict election results. In Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011) (pp. 2-10).

[14] Ward, I. A. N. (2008). Kevin07. Labor's pitch to Generation YouTube. Social Alternatives, 27(2), 11.

[15] Langlois, G., Elmer, G., McKelvey, F., & Devereaux, Z. (2009). Networked publics: The double articulation of code and politics on Facebook. Canadian Journal of Communication, 34(3).

[16] Wills, D., & Reeves, S. (2009). Facebook as a political weapon: Information in social networks. British Politics, 4(2), 265-281.

[17] Túñez, M., & Sixto, J. (2011). Social networks, politics and commitment 2.0: Spanish MPs on Facebook. Revista Latina de Comunicación Social, 66, 210-246.

[18] Anuta, D., Churchin, J., & Luo, J. (2017). Election bias: Comparing polls and twitter in the 2016 us election. arXiv preprint arXiv:1701.06232.

[19] Chung, J. E., & Mustafaraj, E. (2011, April). Can collective sentiment expressed on twitter predict political elections?. In AAAI (Vol. 11, pp. 1770-1771).

[20] Hinch, J. (2018). # MakeAmericasPollsGreatAgain: Evaluating Twitter as a Tool to Predict Election Outcomes. The Geographical Bulletin, 59(1), 45-54.

[21] Cameron, M. P., Barrett, P., & Stewardson, B. (2016). Can social media predict election results? Evidence from New Zealand. Journal of Political Marketing, 15(4), 416-432.

[22] Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2018). Predicting elections from social media: a three-country, three-method comparative study. Asian Journal of Communication, 1-21.

[23] MacWilliams, M. C. (2015). Forecasting congressional elections using facebook data. PS: Political Science & Politics, 48(4), 579-583.

[24] Tunggawan, E., & Soelistio, Y. E. (2016, October). And the winner is…: Bayesian Twitter-based prediction on 2016 US presidential election. In Computer, Control, Informatics and its Applications (IC3INA), 2016 International Conference on (pp. 33-37). IEEE.

[25] Ramteke, J., Shah, S., Godhia, D., & Shaikh, A. (2016, August). Election result prediction using Twitter sentiment analysis. In Inventive Computation Technologies (ICICT), International Conference on (Vol. 1, pp. 1-5). IEEE.

[26] Heredia, B., Prusa, J. D., & Khoshgoftaar, T. M. (2018). Social media for polling and predicting United States election outcome. Social Network Analysis and Mining, 8(1), 48.

[27] Razzaq, M. A., Qamar, A. M., & Bilal, H. S. M. (2014, August). Prediction and analysis of Pakistan election 2013 based on sentiment analysis. In Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 700-703). IEEE Press.

[28] Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations (pp. 115-120). Association for Computational Linguistics.

[29] Lampos, V., Preoţiuc-Pietro, D., & Cohn, T. (2013). A user-centric model of voting intention from Social Media. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (Vol. 1, pp. 993-1003).