Sentiment Analysis of Influential Messages for Political Election Forecasting

Oumayma *Oueslati*[1], Moez Ben *Hajhmida*[1],
Habib *Ounelli*[1], and Erik *Cambria*[2]

[1]University of Tunis El Manar, Tunisia
[2]Nanyang Technological University, Singapore

Abstract. In this paper, we explore the use of sentiment analysis of influential messages on social media to improve political election forecasting. While social media users are not necessarily representative of the overall electors, bias correction of users messages is critical for producing a reliable forecast. The observation motivates our work is that people on social media consult the messages of each other before taking a decision, this means that social media users influence each other. We first built a classifier to detect politically influential messages based on different aspects (messages content, time, sentiment, and emotion). Then, we predicted electoral candidates votes using sentiment degree of influential messages. We applied our proposed model to the 2016 United States presidential election. We conducted experiments at different intervals of times. Results show that our approach achieves better performance than both off-line polling and classical approaches.

1 Introduction

Nowadays, writing and messaging on social media is a part of our daily routine. Facebook, for example, enjoys more than one billion daily active users. The exponential growth of social media has engendered the growth of user-generated content (UGC) available on the web. The availability of UGC raised the possibility to monitor electoral campaigns by tracking and exploring citizens preferences [1]. Jin et al. [2] stated that analyzing social media during an electoral campaign may be more useful and accurate than the traditional off-line polls and surveys. This approach represents not only a more economical process to predict the election outcome but also a faster way to analyze such a massive amount of data.

Thus, many studies proved that analyzing social media based on several indicators led to a reliable forecast of the final result. Some works [3,4] have relied on simple techniques such as the volume of data related to candidates. More recent works tried to provide a better alternative to the traditional off-line poll using sentiment analysis of UGC [5,6]. Whatever the used technique, addressing the data bias is a crucial phase which impacts the quality of the outcome. While social media contents are not necessarily all relevant for the prediction, an appropriate technique to bias UGC is needed.

In this paper, we propose a sentiment analysis based approach to predict the political elections by relying only on influential messages shared on social media. Social influence has been observed not only in political participation but many other domains such as health behaviors and idea generation [7]. To the best of our knowledge, our work is the first investigating politically influential messages to forecast election outcome. According to Cialdini and Trost [8], social influence occurs when an individual's views, emotions, or actions are impacted by the views, emotions or actions of another individual. By analogy, the political influence was achieved thanks to the direct interaction with voters through social media platforms. The politicians tweet on Twitter and post on Facebook to receive voters feedbacks and understand their expectations. Hence, we built a classifier to select the influential messages based on content, time, sentiment, and emotion features.

To compute sentiment features, we adopted a concept-level sentiment analysis Framework largely recommended in literature [9] called SenticNet [10]. To extract emotion features, we built an emotional lexicon based on the Facebook reactions. For each electoral candidate, the number of votes was predicted using sentiment polarity and degree of influential messages figuring in the candidate official Facebook page. We applied the proposed approach to the 2016 United States presidential election. To evaluate the prediction quality, we mainly considered two kinds of ground truths for comparison: the election outcome itself and polls released by traditional polling institutions. We also compared our approach with classical approaches merely based on data volume. Experiments were conducted at different intervals of time. Results showed that using Influential messages led to a more accurate prediction. In term of structure, the rest of the paper is organized as follows: Section 2 explores the current literature; Section 3 addresses the research methods; Section 4 presents the results discussion and implications; lastly, Section 5 gives a synopsis of the main concluding remarks.

2 Related Works

2.1 Sentiment Analysis

Social media platforms have changed the way that people use the information to make a decision. They tend to consult the reviews of each other before making their choices and decisions. Sentiment analysis in social media is a challenging problem that has attracted a large body of research. In [11], authors investigated the impact of sentiment analysis tools to extract useful information from unstructured data ranging from evaluating consumer products, services, healthcare, and financial services to analyzing social events and political elections.

Cambria et al. [10] have introduced SenticNet which is a concept-level sentiment analysis framework, consisting of 100,000 concept entries. SenticNet acts as a semantical link between concept-level emotion and natural word-level language data. Five affiliated semantic nodes are listed following each concept. These nodes are connected by semantic relations, four sentics, and a sentiment polarity value. The four sentics present a detailed emotional description of the concept they belong to, namely sensitivity, aptitude, attention, and pleasantness. The sentiment polarity value is an integrated evaluation of the concept sentiment based on the four parameters. The sentiment polarity provided by SenticNet is a float number in the range between -1 to 1.

Many applications have been developed by employing SenticNet. These applications can be exploited in many fields such as the analysis of a considerable amount of social data, human and computer interactions. In [12], Bravo-Marquez et al. used SenticNet to build a sentiment analysis system for Twitter. In [13], authors used SenticNet to build an e-health system called iFeel which analyze patients opinions about the provided healthcare. Another study by Qazi et al. [9] recommended SenticNet to extract sentiment features. Encouraged by these works, we also used SenticNet framework to extract sentiment features from the extracted message.

2.2 Election Forcasting Approaches

Forecasting elections in social media have become the latest buzzword. Politicians have adopted social media, predominantly Facebook and Twitter, as a campaigning tool. On the other hand, the general public has widely adopted social media to conduct political discussions [14]. Hence, Bond et al. [15] affirm that social media content may influence citizens political behavior. Sang and Bos [16] stated that many studies have proven that analyzing social media using several techniques and based on different indicators led to a reliable forecast of electoral campaigns and result.

Tumasjan et al. in [4] were the first using Twitter to predict the outcome of German Federal election. They used a simple technique based on counting the number of tweets that a party get. Although their success in predicting the winner of the 2009 German Federal Elections, their simple technique get many critics. Jungherr et al. [17] highlighted the lack of methodological justification. Furthermore, Gayo-Avello [5,18] stressed making use of sentiment analysis to produce more accurate results. In [5], Gayo-Avello reported a better error rate when using sentiment analysis (17.1% using volume, 7.6% using sentiment). Consequently, many works have taken his advice such as [19,20,6,21].

Addressing the data bias is an essential phase in predicting an electoral outcome [22,23]. Social media users are not necessarily representative of the overall population. However, many works such as [4,17] did not proceed by biasing data. Some others works such as [22,5] attempted to reduce the bias according to user age and geolocation. They attempted to improve the overall view of the electorate. However, the authors reported that the success was minimal and the improvement was somewhat marginal. A very recent work by Arroba et al. [23] explore the geographic weighting factors to improve political prediction results. They stated that geographic weighting along with sentiment polarity and relevance led to a better outcome.

3 Proposed Method

In this section, we introduce the approach used to build our model, shown in Figure 1. Our methodology consists of a series of steps that range from the extraction of Facebook user messages (FUMs) to the election prediction process. Our work is influenced by the advice of [5]. Instead of merely relying on the volume (the number of messages the candidate receive), we have used sentiment analysis in our methodology along with the attempt to bias data by selecting only influential messages. We applied this methodology to the last presidential election of the U.S. The presidential election took place on November 8, 2016 with two favorite candidates: The Republican Donald Trump, and the Democratic Hillary Clinton. Republican Donald Trump lost the popular vote to Democrat Hillary Clinton by more than 2.8 million votes.

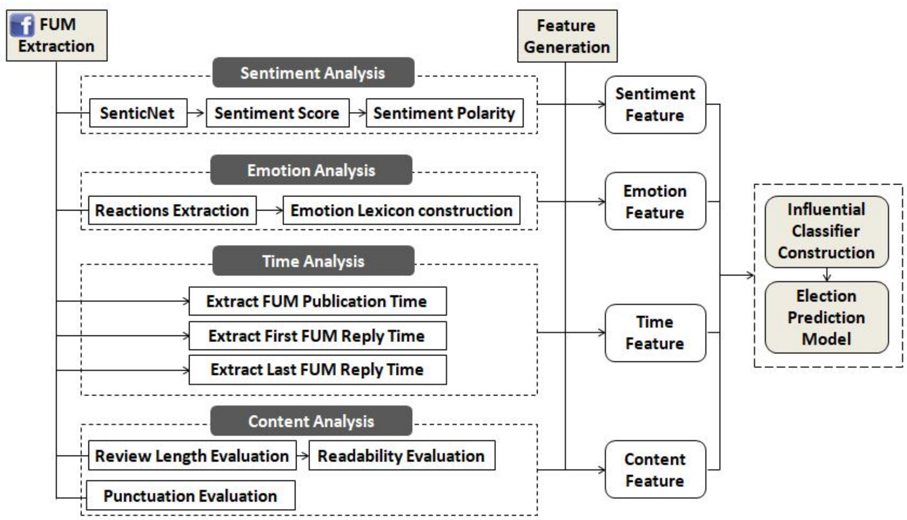

Fig. 1. Workflow of the proposed model.

3.1 Data Collection

Twitter is the most used to predict election outcome thanks to the ease that Twitter platform gives to extract data. To choose our data source, we compared Facebook and Twitter in term of data quality and platform popularity. Many previous studies [24,25,18] found that Twitter data was unreliable to predict electoral outcomes. It is mainly due to selecting tweets unrelated to the candidates. Selecting tweets based on a manually constructed list of keywords certainly led to a loss of relevant information. Though tweets may not comprise any keywords from the pre-defined list, it does not mean that they are necessarily irrelevant.

In contrast, Facebook provides official candidates pages which allow having a large sample of relevant data independently from keywords. It also provides more information about the text message and does not limit the user to a specific number of characters. Twitter limits their users to a 240-characters which forces users to express their opinions briefly and sometimes partially. Next, favorable statistics on the U.S. Facebook users encouraged us to rely on it to have an accurate electoral prediction. The total Facebook audience in the United States amounted to 214 million users, where more than 208 million users are older than 17 years[1]. We extracted data from candidates' official Facebook pages. Namely, we extracted FUMs along with: Users responses on the FUM, Users reactions (Likes, Love, Haha, Wow, Sad, and Angry), and Timestamps (FUM publication time, First FUM reply time, and Last FUM reply time). The collection was directly done from public verified Facebook pages with a self-made application, using the Facebook Graph API[2]. Data collection was conducted within one year before the presidential election in November 2016 so that we can experiment our model over several periods of times (one year before the election day, six months before, one week before, etc.).

In the first pre-processing step, we deleted URLs, empty messages, non-English messages, and duplicated row data. Hence, if a message is duplicated but has different metrics, we kept it. For example, the following message: "WE NEED TRUMP NOW!!!" appeared three times in our raw data but each time with different numbers of likes and replies, so we have considered it. After the data cleaning step, we kept 10k messages from Hillary Clinton official Facebook page and 12k messages from Donald Trump official Facebook page.

3.2 Feature Generation

This subsection describes our features to characterize influential messages. Based on the definition of the social influence stated by Cialdini and Trost [8]: *"Social influence occurs when an individuals views, emotions, or actions are impacted by the views, emotions or actions of another individual"*, we designed four kinds of features (sentiment, emotion, time, and content). In total, we designed 20 features to characterize whether the message is influential or not.

Sentiment Feature: We conducted the sentiment analysis task using SenticNet. We attributed to each FUM a sentiment score between 1 and -1. Practically, SenticNet is inspired by the Hourglass of emotions model [26]. In order to calculate the sentiment score, each term is represented on the ground of the intensity of four basic emotional dimensions namely sensitivity, aptitude, attention, and pleasantness. In our work, we computed the sentiment features using the Sentic API[3]. The sentiment features are as follow: (1) **OSS** which is the Overall Sentiment Score of the FUM, (2) **SSPMax** which is the Sentiment Score of the most Positive term in the FUM, (3) **SSNMax** which is the Sentiment Score of the

[1] www.statista.com/statistics/398136/us-facebook-user-age-groups/
[2] www.developers.facebook.com/tools/explorer/
[3] http://sentic.net/api

most Negative term in the FUM, (4) **SSNMax** which is the Sentiment Score of the most Negative term in the FUM, and (5) **SSNMin** which is the Sentiment Score of the least Negative term in the FUM.

Emotion Feature: To extract the emotion features, we first built an emotion lexicon based on Facebook reactions. Kumar and Valdamni [27] stated that in social networks, if someone reacted to a public post or review (message), it means that the person has positive or negative emotions towards the entity in question. So emotions may be denoted explicitly through reviews and messages or implicitly through reactions. In our work, we explore Facebook reactions to construct an emotion lexicon. This lexicon would allow emotion extraction from any FUM, even if the FUM did not receive any reaction yet.

There are six reactions that Facebook users use to express their emotions toward a message (Like, Love, Haha (laughing), Wow (surprised), Sad, and Angry). From the collected data we selected all the messages which had received any reaction. After cleaning the review and deleting stop words, based on the reactions, we selected terms reflecting emotions. So we got a list of emotional terms, and based on reactions count, we attribute a score for each term. For example, the term 'waste' appears in two messages with (5, 10) Like, (0, 1) Love, (12, 0) Haha, (3, 18) Wow, (0, 40) Sad, and (30, 7) Angry. So the term 'waste' has in total 15 Like, 1 Love, 12 Haha, 18 Wow, 40 Sad, and 37 Angry. We normalized the score through the sum of all reaction (123 in this example).

Lastly, the emotion features were extracted based on the constructed emotion lexicon. The first emotional feature **EMT** evaluates the presence of EMotional Terms in the FUM (the number of emotional terms divided by the number of all terms). The six others features are **LKR**, **LVR**, **LGR**, **SPR**, **AGR**, and **SDR**. They represent the Like ratio, the Love ratio, the Laugh ratio, the Surprise ratio, the Anger ratio, and the Sadness ratio respectively.

Time Feature: The time aspect is important to analyze. Indeed, we generated two time-features to evaluate users engagement towards FUMs:

$$LCF = \frac{LastPostedReplyTime - MessagePulicationTime}{ElectionPredictionTime - MessagePulicationTime}$$

$$RCT = 1 - \frac{FirstPostedReplyTime - MessagePulicationTime}{ElectionPredictionTime - MessagePulicationTime}$$

The feature Life cycle (**LCF**) measures how much the message persists and remains popular by knowing how long the content can drive user attention and engagement. The Life cycle value is comprised between zero and one. The feature Reaction Time (**RCT**) evaluates the time that a FUM makes to start receiving responses. This feature allows knowing if a message has rapidly engaged the users and drove their attention.

Content Feature: The generated content features attempt to evaluate the quality of the message content. A message which is not clear and readable cannot be influential. Hence, content features include (1) **NBC** which is the Number of Characters in the FUM, (2) **NBW** which is the Number of Words in the FUM, (3) **NBS** which is the Number of Sentences in the FUM, (4) **NBWS** which is the Number of word Per Sentence, (5) **NBSE** which is the Number of

Spelling Errors in the message, and (6) **ARI** which is the Automated Readability Index. ARI is a measure calculated as following: $ARI = 4.71 * (\frac{CharCount}{WordCount}) + 0.5 * (\frac{WordCount}{SentCount}) - 21$. This score indicates the US educational level required to comprehend a given text. The higher the score, the less readable the text [28].

3.3 Influential Classifier Construction

In our work, we propose to reduce the data bias based on messages influence rather than who wrote the message. As social influence has been observed in political participation [15], we built a classifier to select only politically influential FUM which make others actions and emotions impacted by the actions and emotions of the FUM writer.

To build our classifier we need a labeled dataset. While it is too expensive to label influential message manually, we selected messages which got many responses from other users. If the message and the responses have approximately the same sentiment polarity (positive or negative), the message is marked as influential. On the other hand, if the message and its responses have different sentiment polarity, the message is marked as no-influential. We did a manual revision for the messages having the same sentiment polarity as their responses but a margin that exceeds 0.5 regarding score. Through this technique of semi-automatic labeling, we got a labeled dataset of 1561 messages: 709 labeled influential and 852 labeled non-influential.

3.4 Election Outcome Prediction Model

We used the methods by [5] with some changes. While Gayo Avello et al. counted every positive message and every negative message, we included only the influential one. Then, the predicted vote share for a candidate *C1* was computed as follows:

$$\frac{infPosSent(C1)+infNegSent(C2)}{infPosSent(C1)+infNegSent(C1)+infPosSent(C2)+infNegSent(C2)}$$

C1 is the candidate for whom support is being computed while *C2* is the opposing candidate. Therefore, *infPosSent(C)* and *infNegSent(C)* are respectively, the number of positive influential and the number of negative influential messages multiplied by their sentiment score.

4 Results and Findings

First of all, we compared the performance of several supervised classification algorithms to select the best one. Subsequently, relying on the best algorithm, we derived our prediction model and evaluated its performance. In our experimentation, we used machine learning algorithms from scikit-learn package, tenfold cross-validation to improve generalization and avoid overfitting.

4.1 Learning Quality

To obtain a model that reasonably fits our objective, we performed the learning phase through several supervised classification algorithms. Then, we selected the best algorithm regarding accuracy (ACC), F-measure (F1) and AUC. To better understand classifiers performance, we examine how classifiers label test data. Therefore, we focus on True Positive (TP) and True Negative (TN) rates generated by each classifier. Classifiers performance are reported in Table 1.

Classifier	ACC	F1	AUC	%TP	%TN
NN	73.48	72.76	74.34	83.78	64.91
RBF SVM	55.54	69.35	51.85	11.57	92.14
DT	82.51	84.10	82.29	79.83	84.74
RF	**89.11**	**90.02**	**89.02**	**88.01**	**90.02**
ANN	52.72	64.86	49.98	20.03	79.93
NB	62.91	70.38	61.11	41.47	80.75
LR	75.53	75.79	76.07	81.95	70.19

Table 1. Performance comparison of various classification algorithms.

In term of ACC, F1, and AUC, the Random Forest (RF) achieved the best performance followed by Decision Tree (DT), Logistic Regression (LR), Nearest Neighbors (NN), and RBF SVM; while Naive Bayes (NB) and Artificial Neural Network (ANN) perform poorly. Regarding the TP and TN rates, Random Forest (RF) also achieved the best rates. Moreover, Random Forest was the best classifier realizing the right balance between the two classes (88.01% as TP and 90.02% as TN).

In Figure 2, we plot in the same graph the ROC curve of each classifier. Upon visual inspection, we observe that the curve of Random Forest classifier is closer than other curves to the upper-left corner of the ROC space. This proves that Random Forest classifier has the best trade-off between sensitivity (TP rate) and specificity (1 FP rate). Random Forest shows the best performance to predict the Influential class with minimal false Positives correctly. After that, we used the Random Forest classifier for further classifications. For sentiment features, the overall sentiment (OSS) of the FUM is the most important followed by the sentiment score of the most negative term, and the sentiment score of the least positive term (SSNMax and SSPMin).

4.2 Features Quality

In this subsection, the relevance of the generated features through its prediction strength. We draw the features importance plot in Random Forest classification, as shown in Figure 3. We notice that features related to FUM sentiment are the most important, followed by the features related to the content, and the features related to the time.

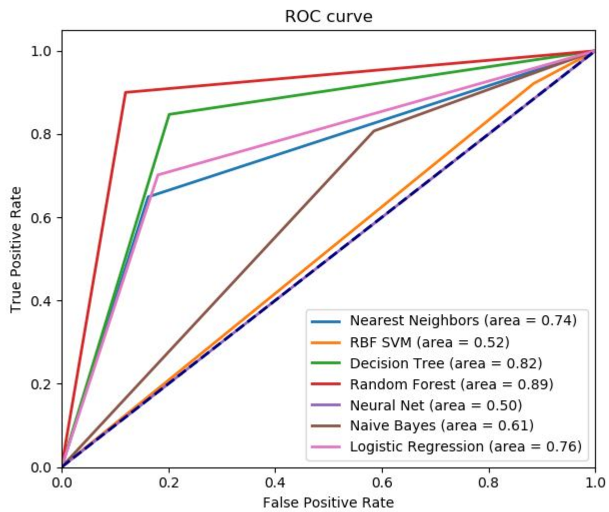

Fig. 2. ROC curve of the different classifiers.

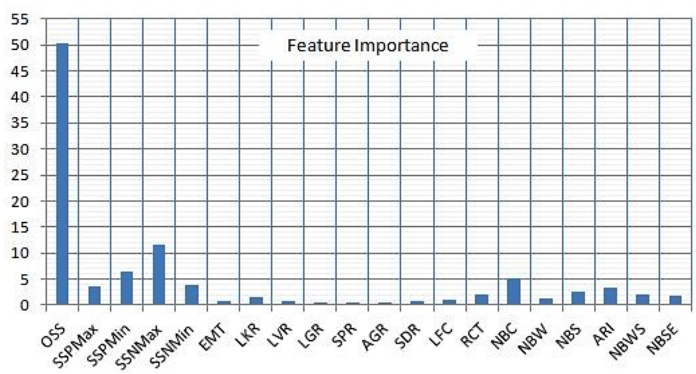

Fig. 3. Feature importance in Random Forest classification.

We find that strongly negative FUM tend to be more attractive and influential than strongly positive FUM. This finding is in line with observations for the features related to emotions. The rate of likes (LKR) is the most important followed by the rate of emotional terms presence (EMT) and sadness rate (SDR). We state that the Like button is ambiguous. Before October 2015, the other reactions did not exist. Only the Like reaction was available which made it overused to explain positive and negative emotions. Even after introducing the rest of emotional reactions, the Like is still overused. Therefore, LKR reflects users engagement toward the FUM more than the emotion that users give off. In contrast, we note that the sadness rate (SDR) is more decisive than the love rate (LVR). This observation also affirms that strongly negative FUMs that imply negative emotion tend to be more influential than FUMs implying positive emotion like Love.

For content features, features related to the FUM length (NBC and NBS), and readability (ARI) are the most important. We find that brief FUM cannot be influential as a long FUM. However, the FUM must be readable and comprehensible by a wide range of peoples to be influential. We find that the ARI measure performs well in the context of social media because it was designed based on the length indicators. However, the spelling error rate (NBSE) is not critical in social media because people tend to use colloquial and invented words and to make some frequent mistakes. Lastly, for the time features, we find that the feature RCT is more important than the feature LC. The FUM making less time to engage users tend to have a longer life cycle. The first replies on the FUM reflect if the FUM would be influential or not.

4.3 Predicting Election Outcome Quality

In order to quantify the difference between the prediction and the ground truth, we relied on the *Mean Absolute Error* (MAE) like the vast majority of previous works. The *MAE* is defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \mid (P_i - R_i) \mid$$

wehre n is the number of candidates, P_i is the predicted vote percentage of the candidate $\boldsymbol{i}$ and R_i is the true election result percentage of the candidate $\boldsymbol{i}$.

We applied our approach to different time intervals. We also tried other previous approaches to evaluate better the contribution of influential message selection and sentiment analysis: Message Count (**MC**), Message Sentiment (**MS**), and our Influential Message Sentiment (**IMS**). Results are presented in Table 2.

	1 Year	6 Months	3 Months	1 Month	2 Weeks	1 Week
IMS	**01.29**	01.52	03.02	**00.88**	**01.10**	01.42
MS	03.00	03.83	06.33	02.71	02.00	02.50
MC	06.92	08.72	15.53	07.16	06.07	06.80

Table 2. MAE at different time intervals.

The best *MAE* done by the most well-known polling institutes[4] is **2.3** by *Reuters/Ipos*. However, the worst *MAE* is **4.5** by *LA Times/U.S.C Tracking*. Our approach was capable of achieving a **0.88** by choosing influential messages posted one month before the election day, **1.10** by selecting influential messages published two weeks before, and **1.52** by selecting influential messages published six months before.

[4] www.realclearpolitics.com

Also, compared to the MC approach and MS approach, our approach was more accurate by achieving an error rate inferior to one. However, the best error rate achieved by MC was 6.07 and the best error rate achieved by MS was 2. Relying only on data volume led to the highest error. When considering the sentiment add to the volume, the MAE slightly decreases. And especially after removing non-influential messages the MAE is considerably improved. Furthermore, scoring each vote by the strength of the expressed sentiment helps the prediction model to ignore weak messages.

To better visualize the difference between approaches, we illustrate the MAE in Figure 4. We noted that that independently of the time interval, relying only on data volume always led to the highest error. We also noted that predicting the election outcome one year before the day of election achieve a good performance compared by others time interval. Exploring the candidates' online presence strategy before the election is relevant to conclude how well the candidate worked on his/her public image. Based on the success of this last, we can accurately predict the election result. There is mainly two kinds of online presence strategy; the long-term (1 year) and the short-term (less than one month). However, the more the historical record of time is reduced the more the forecast performance got worst.

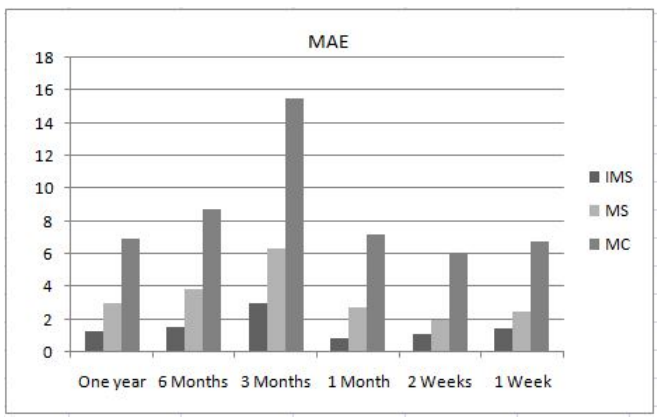

Fig. 4. MAE overview on different time intervals by different approaches.

The error rate is reduced while forecasting is one year and one month before the election day. In contrast, the error is enormous six months and one week before the election day. That is to say exploring partially the historical record is like analyzing an online political strategy by half. Moreover, few days before the election day, the noise is present more than any period.

5 Conclusion

In this paper, we proposed a novel model for election forecasting using sentiment analysis of influential messages. We collected data from Facebook graph API. Then, we constructed a classifier to select only the influential messages based on messages content, time, sentiment, and emotion. Random Forest algorithm has shown the best classification performance. We applied our model to the 2016 United States presidential election. We demonstrated that it is reliable to predict election results based on sentiment analysis of influential messages. Also, we demonstrated that data bias is appropriately addressed with influential messages selection. We found that our approach was capable of achieving better MAE than both off-line poll and classical approaches. In the future, we plan to continue our work on performing sentiment analysis of influential messages using other modalities such as influence degree definition.

References

1. Woodly, D.: New competencies in democratic communication? blogs, agenda setting and political participation. Public Choice **134** (2008) 109–123
2. Jin, X., Gallagher, A., Cao, L., Luo, J., Han, J.: The wisdom of social multimedia: using flickr for prediction and forecast. In: Proceedings of the 18th ACM international conference on Multimedia, ACM (2010) 1235–1244
3. Williams, C., Gulati, G.: What is a social network worth? facebook and vote share in the 2008 presidential primaries. In: American Political Science Association. (2008)
4. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M.: Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm **10** (2010) 178–185
5. Gayo Avello, D., Metaxas, P.T., Mustafaraj, E.: Limits of electoral predictions using twitter. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, Association for the Advancement of Artificial Intelligence (2011)
6. Burnap, P., Gibson, R., Sloan, L., Southern, R., Williams, M.: 140 characters to victory?: Using twitter to predict the uk 2015 general election. Electoral Studies **41** (2016) 230–233
7. Romero, D.M., Reinecke, K., Robert Jr, L.P.: The influence of early respondents: information cascade effects in online event scheduling. In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, ACM (2017) 101–110
8. Cialdini, R.B., Trost, M.R.: Social influence: Social norms, conformity and compliance. (1998)
9. Qazi, A., Raj, R.G., Tahir, M., Cambria, E., Syed, K.B.S.: Enhancing business intelligence by means of suggestive reviews. The Scientific World Journal **2014** (2014)
10. Cambria, E., Poria, S., Hazarika, D., Kwok, K.: SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings. In: AAAI. (2018) 1795–1802

11. Cambria, E., Song, Y., Wang, H., Howard, N.: Semantic multidimensional scaling for open-domain sentiment analysis. IEEE Intelligent Systems **29** (2014) 44–51
12. Bravo-Marquez, F., Mendoza, M., Poblete, B.: Meta-level sentiment models for big social data analysis. Knowledge-Based Systems **69** (2014) 86–99
13. Araújo, M., Gonçalves, P., Cha, M., Benevenuto, F.: ifeel: a system that compares and combines sentiment analysis methods. In: Proceedings of the 23rd International Conference on World Wide Web, ACM (2014) 75–78
14. Strandberg, K.: A social media revolution or just a case of history repeating itself? the use of social media in the 2011 finnish parliamentary elections. New Media & Society **15** (2013) 1329–1347
15. Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D., Marlow, C., Settle, J.E., Fowler, J.H.: A 61-million-person experiment in social influence and political mobilization. Nature **489** (2012) 295–298
16. Sang, E.T.K., Bos, J.: Predicting the 2011 dutch senate election results with twitter. In: Proceedings of the workshop on semantic analysis in social media, Association for Computational Linguistics (2012) 53–60
17. Jungherr, A.: Tweets and votes, a special relationship: the 2009 federal election in germany. In: Proceedings of the 2nd workshop on Politics, elections and data, ACM (2013) 5–14
18. Gayo-Avello, D.: " i wanted to predict elections with twitter and all i got was this lousy paper"–a balanced survey on election prediction using twitter data. arXiv preprint arXiv:1204.6441 (2012)
19. Franch, F.: (wisdom of the crowds) 2: 2010 uk election prediction with social media. Journal of Information Technology & Politics **10** (2013) 57–71
20. Ceron, A., Curini, L., Iacus, S.M., Porro, G.: 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** (2014) 340–358
21. Caldarelli, G., Chessa, A., Pammolli, F., Pompa, G., Puliga, M., Riccaboni, M., Riotta, G.: A multi-level geographical study of italian political elections from twitter data. PloS one **9** (2014) e95809
22. Choy, M., Cheong, M.L., Laik, M.N., Shung, K.P.: A sentiment analysis of singapore presidential election 2011 using twitter data with census correction. arXiv preprint arXiv:1108.5520 (2011)
23. Arroba Rimassa, J., Llopis, F., Muñoz, R., Gutiérrez, Y., et al.: Using the twitter social network as a predictor in the political decision. In: 19th CICLing conference. (2018)
24. Bermingham, A., Smeaton, A.: 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). (2011) 2–10
25. Metaxas, P.T., Mustafaraj, E., Gayo-Avello, D.: How (not) to predict elections. In: Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on, IEEE (2011) 165–171
26. Cambria, E., Livingstone, A., Hussain, A.: The hourglass of emotions. In Esposito, A., Vinciarelli, A., Hoffmann, R., Muller, V., eds.: Cognitive Behavioral Systems. Volume 7403 of Lecture Notes in Computer Science. Springer, Berlin Heidelberg (2012) 144–157
27. Kumar, R., Vadlamani, R.: A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems **89** (2015) 14–46
28. Smith, E.A., Kincaid, J.P.: Derivation and validation of the automated readability index for use with technical materials. Human factors **12** (1970) 457–564