Social Data Analytics for Forecasting Electoral Outcomes

Khalid Ait Hadi, Rafik Lasri, Abdellatif El Abderrahmani

***Abstract*: Social networks are ubiquitous, so much so that they are used to conduct polls on various societal issues. Thus, multiple domains have been subject of studies aimed at making predictions based on signals captured on online social networks. Recently, many works have focused on exploring the potential of social media platforms to capture social trends and forecast voting outcomes in electoral consultations. This paper comes in this wake and aims at presenting a strategy of predicting election results using big data collected from the Twitter platform.**

***Index Terms*: Big Data, Social Networks, Elections, Prediction, Bidirectional Gated Recurrent Unit, Convolutional Neural Network.**

I. INTRODUCTION

With an attractive pluralism and an ever-increasing evolution, social media have become veritable broths of information, vital tools of speech and the unavoidable expression channels for a large pleiad of the general public, who do not hesitate anymore to launch the debate, to express their opinions, their support or discontent.

In recent years, a plethora of researchers have focused on the discovery and analysis of the political potential of big data extracted from social networking platforms. In this register, two main branches of research have emerged: (1) social movements and phenomena analysis [6, 7], and (2) forecast and study of electoral outcomes [4, 13].

The first direction explores the correlations and linkages between social media and real popular mobilizations and decorticates the role of social media in the collective behaviours' conception, the thought communities' creation and the social movements' emanation. The second branch of research focuses on the electoral process and attempts to decipher the real underpinnings and macroscopic factors of electoral operations while seeking to determine popularity ratings and predict electoral votes' results. This last area of research is experiencing, actually, an unusual infatuation due to the primordial interest of the political opinion poll, and the limit of conventional approaches, previously used for gathering public opinion, that are laborious and overlong. This, as opposed to the enormous possibilities offered by social big data that have become the new substitute for habitual surveys and can be used as an interactive and real-time instrument to keep an eye on voters' sentiments and opinions in a certain electoral process and to discern users' political propensities, which require time and effort to gather and compile all responses in long surveys.

It's in this wake that fits the present work which aims, in fine, to introduce a machine learning model dedicated to predicting the results of an electoral vote based on users' opinions on social networks.

The remainder of this work is organized as follows. Section II is devoted to discussing the background and related work. The structure of the proposed methodology is detailed in Section III. Section IV is dedicated to the preparation of data intended for analysis. Finally, experiments are performed and discussed in Section V.

Revised Manuscript Received on June 05, 2019.

Khalid Ait Hadi: Laboratory of Sciences and Advanced Technologies, Department of Computer Sciences, Polydisciplinary Faculty of Larache, Abdelmalek Essaadi University, Morocco.

Rafik Lasri: Laboratory of Sciences and Advanced Technologies, Department of Computer Sciences, Polydisciplinary Faculty of Larache, Abdelmalek Essaadi University, Morocco.

Abdellatif El Abderrahmani: Laboratory of Sciences and Advanced Technologies, Department of Computer Sciences, Polydisciplinary Faculty of Larache, Abdelmalek Essaadi University, Morocco.

II. BACKGROUND AND RELATED WORK

Over the years, social media have become very attractive and engaging by dint of its interactivity with political content, principally in electoral context. Recently, several studies have looked at the power of social media platforms, intending to capture societal trends and predict voters' intentions and inclinations in a given region. The ability to predict voting intentions constitutes a "commodity" highly sought after by decision makers and the motivation of researchers in the election prediction based on social big data field promises to be continually growing. Various tools, models and algorithms, more or less accurate, have been produced for the election campaigns observation and the electoral outcomes forecast. In this area of research, we can evoke the work done in [8], where Chi-Square Automatic Interaction Detection (CHAID) decision tree, Naive Bayes and Support Vector Machine (SVM) algorithms were applied to data extracted from Twitter. Furthermore, Natural Language Processing (NLP) techniques have been combined with Rapid miner Discover and Text tool in [9] to discuss election outcomes by confronting sentiments analysis of Twitter data to conventional opinion polls. In [2], hybrid approach based on K-Nearest Neighbors (KNN) algorithm has been applied to microblogging data from Twitter so that to predict users' sentiments. On another side, regression analysis techniques with root mean squared error have been exploited in [11] to treat the strength of Twitter in forecasting election results.

Deep learning models have been also investigated for election prediction task and it has been proven in [5] that Long Short Term Memory (LSTM) network leads to overcome some limitations of traditional approaches successfully.

Through this literature review, it appears that a part of these techniques often fail to handle complex features implied in election forecast, while LSTM model suffers from an organic imperfection which consisting of its unsuccess to capture text's local features and its inability to properly take into account sequence correlation information when processing unstructured matter.

Here, the combination of Bidirectional Gated Recurrent Unit (BiGRU) and Convolutional Neural Network (CNN) seeks to overcome these obstacles by presenting a design with improved accuracy rate and better convergence than the LSTM model.

III. ARCHITECTURE OF THE PROPOSED MODEL

The use of deep learning models is widespread in the machine learning field and among others for the electoral prediction problem. They are known by their ability to surmount the limits of certain usual approaches, like decision tree and random forest algorithms, and to grasp more efficiently the complex particularities encountered during electoral forecast processes.

Election prediction task requires a skillfully model to identify and evaluate features, and these are the same goals that aims to achieve Natural Language Processing (NLP) where deep learning techniques are broadly used.

In this paper and to perform the task of electoral results forecasting, we introduce an enhanced Deep Neural Network (DNN) to capture consequential interconnection information and local features by incorporating Bidirectional Gated Recurrent Unit (BiGRU) with a Convolutional Neural Network (CNN), which beefs up adequate text representation for sentiment classification task and furthermore ameliorates the text learning capacity.

The proposed combination also aims to sidestep the fragmentation of the concatenated framework of text, by handing over the output of feature layer to the BiGRU and CNN ones, and consequently procure concluding representation of the introduced phrase.

In the ensuing, we will explain the two parts of our algorithm.

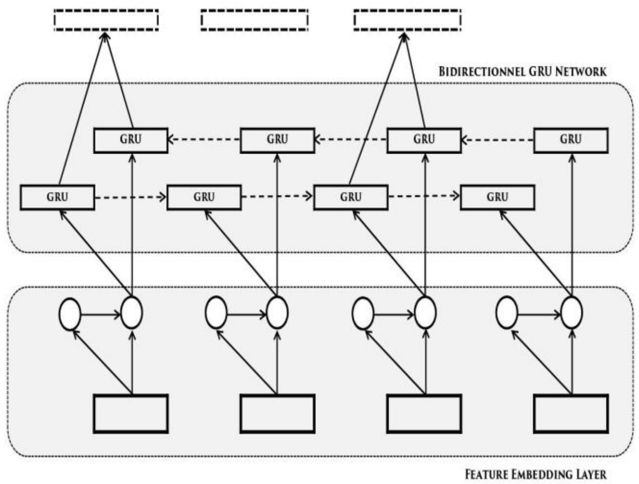

Fig. 1. Architecture of the BiGRU model

As shown in *Fig. 1*, BiGRU handles data in both directions with forward and backward sequences. The results of both directions are then combined in the output. Compared to the unidirectional case, this scheme has advantageous design in the sense that it takes into consideration both the future (direct) and the past (reverse) contexts.

Introduced in [1], GRU is considered as an improved version of the Long Short Term Memory (LSTM) network [3] model, which allows overcoming problems faced by standards recurrent neural networks (RNN), such as vanishing gradient and long term dependency problems.

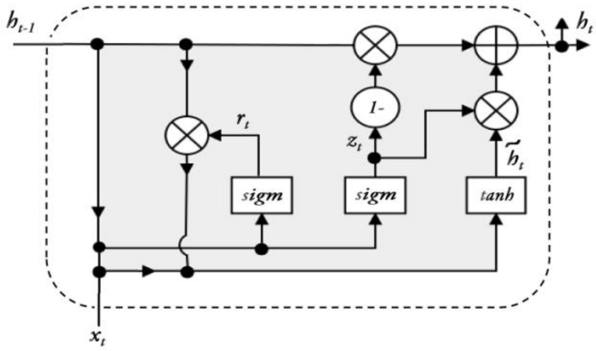

Fig. 2. Architecture of the Gated Recurrent Unit (GRU)

GRU is designed so that every recurrent unit grasps the correlations between the different time scales, and has gating units that regulate the current of information inside the memory. It uses update gate z_t and reset gate r_t to hold information upon plenty of time periods so as to impact a future time one.

As shown in *Fig. 2*, for an input x_t in time-step t, hidden state h_t is calculated by iterating the following equations:

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$z_t = \sigma(W_z * [h_{t-1}, x_t])$$
$$r_t = \sigma(W_r * [h_{t-1}, x_t])$$
$$\tilde{h}_t = \tanh(W_{\tilde{h}} * [r_t \otimes h_{t-1}, x_t])$$
$$h_t = (1 - z_t) \otimes h_{t-1} + z_t \otimes \tilde{h}_t$$

where $*$ denotes matrix multiplication, $\otimes$ is elementwise multiplication, σ is the sigmoid activation, and $W_{(\bullet)}$ denotes appropriate sized matrix.

Hence, in the BiGRU layer, data is processed in both directions with forward $\overrightarrow{h_t}$ and backward $\overleftarrow{h_t}$ hidden layers. The output of the BiGRU is the stepwise sum of the forward and backward states.

Although the sentence representation attained by the BiGRU network preserves the sequence information of the phrase, it is, however, confronted with a lack of flexibility to capture local features and forecast the sentiment polarity. To attenuate this impediment and to avert underrating local features and sequence information when processing unstructured text, we combine the BiGRU layer with Convolutional Neural Network (CNN) [10] one. The advantage of CNN in this way is its capability of recovering and retrieving the local features in multi-dimensional circumstances. Precisely, the CNN layer, which includes convolutional and max-pooling layers, can process along the word dimension to convolve with several kernels of various widths.

IV. DATA PREPARATION

This section is concerned with the introduction of preliminary modules for acquisition and preprocessing the collected data. The goal is to build a reliable dataset for analysis purpose.

A. Data gathering

The data have been collected using Twitter's streaming API. Twython python library has been used for linking the Twitter API and to obtain data intended for analysis. We focused on 2016 US presidential election, and we only employed two candidates' names Hillary and Trump as keywords to collect tweets during October 07, 2016 to November 07, 2016.

B. Data preprocessing

Data preprocessing is a primordial phase in text mining, which provides excellent quality of text classification and significantly diminishes computational intricacy [12]. In the present step we perform tweets preprocessing and cleaning. The following processes are conducted: removing hashtags, screen name, URLs, punctuations, symbols, numbers and special characters, stemming and tokenization.

C. Training and testing phases

In this step, we proceed to the division of the dataset into training and testing parts, with ratio 80:20.

The training module seeks to preliminary extract features, which will permit the algorithm to spot the required features, useful for identifying sentiment polarity of the tweet. To achieve this task, we use convolutional layers with residual connections, layer normalization, and maxout non-linearity, which give much better efficiency than the standards models, like Bidirectional Long Short Term Memory (BiLSTM). For this purpose, we employ the spaCy solution for python.

The testing dataset is used to assess the conduct of the model saved by the training phase. It occupies the remaining 20 percent of the original dataset. Through this phase, our model's results are measured in comparison with the awaited output.

V. EXPERIMENTAL RESULTS AND DISCUSSION

To validate the efficiency of our approach, we establish a comparison, basing on the accuracy rate, between the proposed algorithm and other tools: LSTM and Naïve Bayes.

	Proposed method	LSTM	Naïve Bayes
Accuracy Rate	89.35%	82.54%	69.82%

Table 1. Performance of the proposed approach

It is seen from Table 1 that combination between BiGRU layer and Convolutional Neural Network performs more accurate results compared with alternate techniques.

Now, we take on the task to predict the 2016 US presidential election results by performing sentiment analysis of the data. We consider positive, negative and neutral classes to extract the sentiment for both candidates Donald Trump (C_1) and Hillary Clinton (C_2).

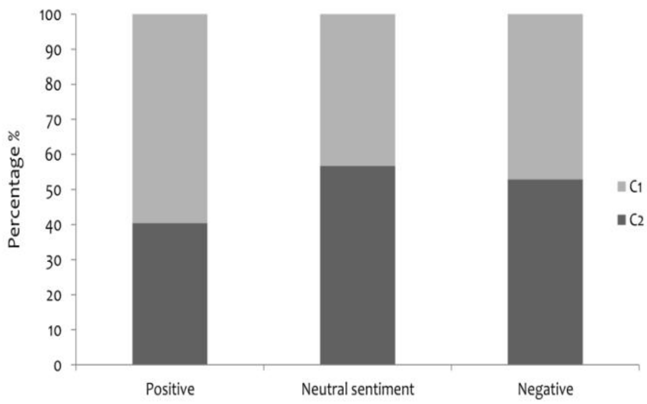

Fig. 3. Percentage on number positive, neutral and negative tweets using BiGRU-CNN approach

In the following, we consider the rate of positive sentiments in relation to the opinionated data for each candidate, calculated as follows:

$$Rate(C_i) = \frac{P_{C_i}}{D_{C_i}}$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

where P_{C_i} and D_{C_i} denote respectively the number of positive and opinionated (positive, neutral or negative) tweets concerning the candidate C_i.

The computation of these values gives rise to the following results:

	Donald Trump	Hillary Clinton
Rate **(C_i)**	39.73%	26.73%

Table 2. Positive candidates' appreciation rates

The table above provides an appropriate idea about the most appreciated candidate who is, in this case, Donald Trump, the winner of the 2016 US presidential election.

Consequently, this paper presents an enhanced approach that has shown efficient and promising results, and which can be explored to investigate various cases of electoral outcomes forecasting based on data extracted from social networks.

REFERENCES

1. K. Cho et al., Learning phrase representations using RNN Encoder–Decoder for statistical machine translation, Proc. of the 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1724–1734, 2014.
2. S. Goyal, Sentiment analysis of Twitter data using text mining and hybrid classification approach, International J. of Engineering Development and Research, 2017.
3. S. Hochreiter and J. Schmidhuber, Long short-term memory, Neural Comput., 9, pp. 1735–1780, 1997.
4. K. Jaidka et al., Predicting elections from social media: a three-country, three-method comparative study, Asian Journal of Communication, 2018.
5. S. JayaKumar et al., Election result prediction using Deep Learning Techniques, International J. of Engineering and Advanced Technology (IJEAT), 8 (4), pp. 74–77, 2019.
6. J. T. Jost et al., How social media facilitates political protest: information, motivation, and social networks, Advances in Political Psychology, 39 (1), 2018.
7. S. Lee, The role of social media in protest participation: the case of candlelight vigils in South Korea, International J. of Communication, 12, pp. 1523–1540, 2018.
8. T. Mahmood et al., Mining Twitter big data to predict 2013 Pakistan election winner, 16th International Multi Topic Conference (INMIC), pp. 49–54, 2013.
9. D. J. S. Oliveira et al., Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls, Journal of Information Technology & Politics, 2016.
10. T. N. Sainath et al., Improvements to Deep Convolutional Neural Networks for LVCSR, IEEE international conference on acoustics, speech and signal processing, pp. 315–320, 2013.
11. M. Safiullah et al, Social media as an upcoming tool for political marketing effectiveness, Asia Pacific Management Review, 22, pp. 10–15, 2017.
12. S. Vijayarani, J. Ilamathi and M. Nithya, Preprocessing techniques for text mining - An Overview, Int. J. of Com. Sci. & Comm. Net., 5(1), pp. 7–16, 2015.
13. L. Wang and J. Q. Gan, Prediction of the 2017 French Election based on Twitter data analysis using term weighting, 10th Computer Science and Electronic Engineering Conference (CEEC), pp. 231–235, 2018.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication