REPORT : PREDICTING VOTING INTENTION

Goal:

Our focus is to generate report on the papers which focuses on predicting the political views and results of elections based on the social media trends specifically Twitter in promising direction.. The research papers which are chosen for this report are "*Predicting the 2011 Dutch Senate Election Results with Twitter*" by *Erik Tjong Kim Sang* and "*On Using Twitter to Monitor Political Sentiment and Predict Election Results*" by *Adam Bermingham and Alan Smeaton*. The first research paper focuses on whether by improving the quality of document collection and by performing sentiment analysis on entity counts of the tweets, the prediction is considerably improved, the second paper focuses on understanding the political sentiments of the nation during Irish elections using sampling and various modelling approaches.

Introduction :

Till the advent of social media in a huge scale, lot of standard and traditional methodologies such as polls using human intervention or opinion polls by voting on TVs and radios are used which had their own disadvantages such as time consumption and error prone. The questions these research papers dealt with are whether opinion polls could be conducted based on the information collected by Twitter and how might we analyse this data to produce results that approximate what can be achieved through traditional market research.

In the first paper, more strongly the approach is to test whether by counting the Twitter messages mentioning the political party names, can we accurately predict the election outcome. Also, to investigate the factors that influence the predictions based on the Dutch tweets. In the second paper, the system uses various techniques to produce live real time interface during election. Using the volume and sentiment data from this system, the number of sampling approaches and methods of modelling political sentiments are reviewed. The evaluation of error is done with respect to the polls as well as with respect to the election results.

Methodologies:

In the first paper, a prediction is done with manual sentiment analysis while the second paper is the continuation of building an automatic sentiment analysis in order to predict the results. Here, the data collecting involves searching for words from a list of hundred high-frequent Dutch words and hashtags using filter stream provided Twitter, which in result returned some false positive tweets that contains words from other languages similar to Dutch. In order to improve the results, they used a language guesser to filter out the non-Dutch tweets using n-gram models of texts and excluded tweets that do not contain names of political parties. In the data the search is done based on two variants i, e the abbreviated version and the full name, allowing for minor punctuation and capitalization variation. Most of the party names are identified with 100% accuracy but there are few ambiguous names as well as tweets in terms of political inclination which were removed. Manual annotation is done on the collected data to find the political inclination instead of automated sentiment analysis and though the predicted number of seats are close to the manual/traditional predictions, the output set is very small and there's room for improvement. The challenges that they faced is the difference between the structure of the tweets and the polls where people can have one vote each where as on Twitter people can have multiple tweets about political parties, out of which few tweets can be ambiguous in terms of political sentiments. The other issue is the demographic of the population that operates Twitter whose age groups cannot be taken into account for proper opinion generation. These issues are further solved by considering only the first tweet out of multiple tweets by a user, using mathematical evaluations on the demographic to change the weight of an opinion for a better fit. Although these techniques worked, since the sentiment analysis is manual, it might be considered as a traditional approach with cost and time consumption as disadvantages.

On the other hand, the second research paper deals with approaches and challenges by including automatic sentimental analysis based on supervised learning. It also collected tweets by searching for party names, their abbreviations, and election hashtag. In order to improve the results, it filtered out tweets about polls results, minority parties and independent candidates. The standard measure of error in predictive forecasting is Mean Absolute Error (MAE), defined as the average of the errors in each forecast where as n is the number of forecasts and *ei* is the difference in actual result for the *ith* forecast.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$

MAE is used to compare Twitter-based predictions with polls as well as with the results of the election. Volume based measure is used to define as the proportional share of party which is mentioned in a set of tweets for a given time period:

$$SoV(x) = \frac{|Rel(x)|}{\sum_{i=1}^{n}|Rel(i)|}$$

Here, *SoV(x)* is the share of the volume for a given party, *x* in a system of *n* parties, |Rel(i)| is the number of tweets relevant to party i. The score for the parties are the proportions summing to 1 and are easily compared to poll results. The sets of documents tested against are separated as *Time-based, Sample size-based, Cumulative, Manual.* The tweets in each set of annotation session were taken from different time periods in order to develop as diverse set as possible.The annotation categories consists of three sentiment classes such as positive, negative and mixed and one non sentimental class as neutral. There are three other classes namely unannotatable, non-relevant, unclear. If due to some economic crisis, a specific negative political sentiment dominates for a particular time, supervised learning will have a tough time in forming unbiased opinion for a long term. To reduce such errors, the machine learning algorithm optimises for F-measure which balances precision and recall across the classes. The feature vector consists of unigrams, the tokenizer used optimses user generated content such as emoticons which are used to predict an emotion of the tweet. The topics, URLs and usernames which are biased are filtered in this approach. If there’s a single emotion towards most of the parties either positive or negative, it might be difficult to gravitate towards one. Hence two approaches are followed here such as for inter-party the volume based measure is modified whereas for intra party, the log ratio sentiment is applied.

Here, regression is used to the inter-party and intra- party measures, trained on poll data.

$$y(x) = \beta_v SoV(x) + \beta_p SoV_p(x) + \beta_n SoV_n(x) + \beta_s Sent(x) + \varepsilon$$

Challenges and Outcomes:

The outcome predicts that when the model is computed and compared with respect to the last 1000 tweets and the cumulative tweets over time, the cumulative tweets give more accurate results because of the large amount of data. The challenges that were not tackled are the opinions of the political parties which are not on twitter are not considered, the activeness of a party with respect to another might influence the polls but the results might vary. These are few of the reasons for the errors in the output.

Conclusions:

The first paper showed that the results obtained by just counting the tweets that mention political parties is not sufficient to obtain good predictions. Compared to the first research paper, the second paper's results are way more authentic because of strong mathematical approaches in calculating errors. However, it is unclear whether confining ourselves to sentiment and volume data will allow us to approach levels of acceptable accuracy for reliable measurement. The representativeness and the potential for adversarial activity should be addressed in a credible way for practical implementations.

References:

- Adam Bermingham and Alan Smeaton. 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), pages 2–10, Chiang Mai, ailand, November 2011. Asian Federation of Natural Language Processing. URL http://www.aclweb.org/anthology/W11-3702.
- Erik Tjong Kim Sang and Johan Bos. Predicting the 2011 Dutch senate election results with Twitter. In Proceedings of the Workshop on Semantic Analysis in Social Media, pages 53–60, Avignon, France, April 2012. Association for Computational Linguistics. URL http://www.aclweb.org/anthology/W12-0607.

Report by
Abdullah Almuqhim (9954873662)
Arpan Mittal(9755980875)
Siri Chandana Pisupati (9971282366)