

Chapter 31

Topic Modeling as a Tool to Gauge Political Sentiments from Twitter Feeds

Debabrata Sarddar

University of Kalyani, India

Raktim Kumar Dey

 <https://orcid.org/0000-0002-7942-5233>

Simplex Infrastructures Limited, India

Rajesh Bose

 <https://orcid.org/0000-0002-0967-455X>

Simplex Infrastructures Limited, India

Sandip Roy

 <https://orcid.org/0000-0002-5447-803X>

Brainware University, India

ABSTRACT

As ubiquitous as it is, the Internet has spawned a slew of products that have forever changed the way one thinks of society and politics. This article proposes a model to predict chances of a political party winning based on data collected from Twitter microblogging website, because it is the most popular microblogging platform in the world. Using unsupervised topic modeling and the NRC Emotion Lexicon, the authors demonstrate how it is possible to predict results by analyzing eight types of emotions expressed by users on Twitter. To prove the results based on empirical analysis, the authors examine the Twitter messages posted during 14th Gujarat Legislative Assembly election, 2017. Implementing two unsupervised clustering methods of K-means and Latent Dirichlet Allocation, this research shows how the proposed model is able to examine and summarize observations based on underlying semantic structures of messages posted on Twitter. These two well-known unsupervised clustering methods provide a firm base for the proposed model to enable streamlining of decision-making processes objectively.

DOI: 10.4018/978-1-6684-7472-3.ch031

INTRODUCTION

The emergence of micro blogging sites has ushered in an immense wave of outpouring of public sentiment, views, and opinions. Peoples across the globe can, with a few keystrokes, make their views known to others of diverse social, religious, cultural and geographic backgrounds. No longer is the word of mouth or ideas confined to a localized section of society. The power and reach of the almost all-encompassing worldwide web have now made it possible for millions to view a microblog in seconds and at almost a press of a button irrespective of geographical distances involved.

Analysis of tweets related to politics can reveal trends. Evaluation of trends is imperative for ascertaining breaking news, generate friend recommendations, conduct data mining on opinions espoused by users, and record sentiments (Alvarez-Melis, 2016). Natural Language Processing (NLP) allows to classify opinions expressed through posts on Twitter and Facebook. NLP is comprised of opinion mining and text mining (Zhang & Liu, 2017). The impact of mining data from microblogging sites is significant. Insofar as analyzing the views on new musical releases, new product launches, motion picture releases, and similar other entertainment features, mining of tweets has been shown to be instrumental in helping formulate strategies accordingly (Mohey & Mohamed, 2016).

Research into vernacular languages of the likes of Hebrew and Hindi have also demonstrated that it is possible to analyze microblogging feeds in terms of positive, negative and neutral opinions (Hacohen-Kerner & Badash, 2016), (Arora, 2013). People across the globe have embraced the fact that social media tools accord the widest reach for their views and opinions on any subject. Nowhere is it more abundantly clear than the subject of politics and gauging acceptance of a political leader or the government in power in a country. Interviews of politicians conducted by journalists and reporters are posted on YouTube and broadcast on television. Interviews of public are also recorded and shown. However, public opinions and views expressed either on talks shows or captured using opinion polls are difficult to transform into searchable records. This challenge is easier to overcome in cases where opinions and views are shared on microblogging and social media sites such as Twitter and Facebook (Haddia et al., 2013). Twitter recorded 500 million microblogs or “tweets” per day (Mamgain et al., 2016). This was at a rate where a typical Twitter user has a daily limit of 2400 tweets capped at 140 characters per tweet. Given this extremely large volume of data that offers a fertile ground to harvest opinions and views of millions of users across communities. Twitter is regarded as one of important sources for researchers in the field of sentiment and opinion analysis (Gage, 2018). As a type of process to mine data, Sentiment Analysis is used in conjunction with Natural Language Processing (NLP). This type of analysis allows for classification into three forms of opinions, viz., “positive”, “negative”, or “neutral” (Bose et al., 2018), (Medhat et al., 2014). This allows researchers to attach numerical values for assessment following observations of reactions of the public to events, people, communities and politics. Extensive research and analysis have been conducted into political sentiments based on Twitter data. In all cases involving analyses of political sentiments, the primary objective has been to predict election results by gauging opinions expressed by individuals either in groups or solo. In a debate conducted live on TV, Barack Obama and John McCain – the then two candidates then vying for the presidency of the United States of America, asked the audience to post comments on Twitter. This TV show called “Hack the Debate” ushered in a new era following the victory of Barack Obama. For the first time ever, a microblogging site was observed to reflect the sentiments of voters and to transform into a potent channel for communicating the sentiments of the voters of arguably the world’s most powerful country (Shamma et al., 2009a). That experience was key in recognizing that analyzing political sentiments would, at last, grow into a

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/topic-modeling-as-a-tool-to-gauge-political-sentiments-from-twitter-feeds/312701

Related Content

Smart Grid

Nikhil Swaroop Kaluvala and Abbe Forman (2013). *International Journal of E-Politics* (pp. 39-47).

www.irma-international.org/article/smart-grid/78378

Self-Organization in Social Software for Learning

Jon Dron (2010). *Social Computing: Concepts, Methodologies, Tools, and Applications* (pp. 312-320).

www.irma-international.org/chapter/self-organization-social-software-learning/39729

Life-Value of Young Smartphone Users

Chin-Tsai Lin, Sih-Wun Wang, Chuan Lee and Yi-Hsueh Chen (2016). *International Journal of Virtual Communities and Social Networking* (pp. 46-56).

www.irma-international.org/article/life-value-of-young-smartphone-users/168628

Stakeholder Participation in Investigating the Impact of E-Commerce Upon the Value Chain

Richard Taylor (2008). *Social Simulation: Technologies, Advances and New Discoveries* (pp. 323-341).

www.irma-international.org/chapter/stakeholder-participation-investigating-impact-commerce/29273

An Abstract Framework for Modeling Argumentation in Virtual Communities

Tudor Groza, Siegfried Handschuh, John G. Breslin and Stefan Decker (2009). *International Journal of Virtual Communities and Social Networking* (pp. 37-49).

www.irma-international.org/article/abstract-framework-modeling-argumentation-virtual/34094