How Trump won: The Role of Social Media Sentiment in Political Elections

Chong Oh  
*University of Utah, chong.oh@utah.edu*

Savan Kumar  
*University of Utah, Savankumar.m@gmail.com*

Follow this and additional works at: [http://aisel.aisnet.org/pacis2017](http://aisel.aisnet.org/pacis2017)

**Recommended Citation**
[http://aisel.aisnet.org/pacis2017/48](http://aisel.aisnet.org/pacis2017/48)
How Trump won: The Role of Social Media Sentiment in Political Elections

Completed Research Paper

Chong Oh
Department of Operations and Information Systems,
David Eccles School of Business,
University of Utah,
Salt Lake City, Utah.
chong.oh@utah.edu

Savan Kumar
Department of Operations and Information Systems,
David Eccles School of Business,
University of Utah,
Salt Lake City, Utah.
Savankumar.m@gmail.com

Abstract
The outcome of the recent US Presidential Election of 2016 shocked and baffled many. Some claimed that social media may play a larger role in influencing the outcome that expected. This study examined Twitter messages containing political discussions with references to both Trump and Clinton to uncover insights about the role of social media sentiment in political elections. We adhere to the social media analytics (SMA) framework of Fan and Gordon (2014) and the sentiment analysis taxonomy of Abbasi, Chen, and Salem (2008) as a structure to extract positive and negative sentiment from the collected tweets during the pre-election period between Nov 3 and Nov 7. The first finding reveals that Trump has an overwhelmingly larger volumes of total, positive, and negative tweets over Clinton implying a higher volume of public discourse around Trump. Secondly, the propagation of negativism towards Clinton is much more than Trump although both candidates have increasingly more negative tweets days leading up to the Election Day of Nov 8. Finally, word clouds for both candidates reveal that the Twitter public are engrossed with more negative topics against Clinton than Trump. This study clarifies the role of social media sentiment, specifically in how Trump is able to use Twitter as a conduit to reach his intended audience over and above traditional media. In addition, the influence of negative tweets seem to have a toll on Clinton creating distrust and weakening her political position especially among the working and middle-class communities that made the difference leading to Trump’s eventual victory.

Keywords: Social media analytics, sentiment analysis, text analytics, Twitter.

Introduction
The outcome of the recent US Presidential Election of 2016 shocked and baffled many, including politicians, general public, pundits, and pollsters alike. The forecasted winner, Democratic Party nominee, Hillary Clinton, was upend by the Republican Party candidate, Donald Trump. Ironically, out of the twenty-one polls reported on Nov 7, 2016, a day before Election Day, nineteen (91.5%) projected an easy Clinton win with one poll predicted as much as plus seven percentage spread (NBC News) (Real Clear Politics, 2017; Business Insider, 2016). Once the dust settled, Trump garnered an electoral vote of 304 against Clinton’s 227, despite Clinton’s popularity vote of 48% versus Trump’s 45.9%. The media in general concluded that Trump’s victory lies in the resurgence of “working class whites who felt stung by globalization and uneasy with a diversifying country where their political power seemed diminishing” (Washington Post, 2016). The significance of this demographic group was largely ignored because in the past it was unprecedented for this group to vote in such large numbers.

The role of social media in this Election cannot be ignored as many claimed that it has an influence on Trump’s unexpected victory (Ad Week, 2016). Even Trump himself declared that his social media presence played a key role in his win and that he will continue to use it as the president (CBS, 2016). He added that “his social media presence had greater leverage on the election than hundreds of millions of dollars spent by the Democrats to influence its outcome”. This is true despite the fact that Clinton’s campaign outspend Trump’s by nearly 90%; “by Oct 2016, Clinton raised $513 million and spent $450 million while Trump campaign raised $255 million and spent $239 million” (Schoen,
2016). “Analysts monitoring social media activity of both campaigns have seen the outcome of this election months before and kept talking about the massive silent voter base that was forming around the Republican nominee. Social media analysts continually sounded the alarm that all of the polls were not reflecting the actual situation on the ground in the pre-election landscape” (AdWeek, 2016).

What is the key role of social media in this election? Social media have grown in leaps and bounds. It has evolved from just being a platform for social networking to full-fledged channel of multimedia interactions which allows individuals to create, share, interact, and contribute information or ideas with virtually no limitation on the reach and growth of information. In 2016, 62% of US adults reported to get their news from social media (Pew Research Center, 2016). Twitter, for example, is a social media platform generates 303 million tweets per day on average and over 313M active monthly user from across the world. This proves to be a great source of digital information for all context, especially so for politics. Twitter blog commented that “people in the U.S. sent 1 billion Tweets about the election since the primary debates began in August of last year.” This provides a great platform for people to express their opinions about a particular topic, event, brand, organization, or even an individual and such may influence others in their networks. Tapping this information from social media may provide great insights on the public’s opinions and their correlations with economic and social outcomes.

We approach this study with the following research questions: 1) What is the role of social media (specifically Twitter) sentiment in the recent presidential election? And specifically how do we define this role through 2) volume of tweets? 3) social media sentiment? and 4) topics derived from tweets for both presidential candidates. In order to address these questions, we propose the following study in extracting social media sentiment in providing indicators of positivity and negativity in relation to political candidates’ election outcomes. Based on the Sentiment Polarity classification taxonomy of Abbasi, Chen, and Salem (2008) and Social Media Analytics (SMA) methodological framework of Fan and Gordon (2014), we built a sentiment extraction system that download tweets from Twitter API. We then proceed to perform data analysis and review the findings to uncover nuance insights pertaining to possible reasons for Trump’s victory and Clinton’s defeat.

Related Work

The topic of understanding political discourse from social media is nascent and just started to gain attention among researchers. But most work are from the communication and social science literatures. For example, Tumasjan, Sprenger, Sandner, and Welpe (2011) examined over 100,000 tweets with reference to either political party or politician from the 2009 German federal election and found that Twitter is used extensively for political deliberation and that tweet sentiment corresponded closely with voters’ political preferences. Similarly, Bekafiugo and McBride (2013) found that those with strong political partisan and those exhibiting high levels of traditional political participation tweet about politics most often which leads to the conclusion that the same political activists are online as well as offline. Likewise, Bode and Dalrymple (2014) conducted a survey of political Twitter users, in order to understand their use of the medium and their political behaviors within it. Their research results indicate that political Twitter users are more interested in and engaged in politics in general and less trusting of the mainstream media. By the same token, Dang-Xuan, Stieglitz, and Neuberger (2013) analyzed tweets from top 30 most retweeted users concerning the state parliament election in Berlin (Germany) in 2011 and found that different groups of influencers differ in expression of emotionality, appraisals, and topic variety. In addition they determined that those more emotional and with higher appraisals of political parties or politicians tend to receive more retweets implying that sentiment may drive information diffusion in political tweets in addition to individual sphere of influence.

Upon reviewing the literature, we note a scarcity of such research among Information Systems scholars despite its high visibility and practical importance. Thus upon heeding the call of Agarwal and Lucas (2005) to extend IS research to other disciplines as an inter-disciplinary research we propose this study as a contribution of cross-disciplinary work of information systems, computer science, and social science.

1 http://about.twitter.com
Research Context

Twitter is a real-time information network that connects individuals and entities to the latest stories, ideas, opinions, and news about what is interesting via microblogs (popularly known as ‘tweets’) that do not exceed 140 characters. Twitter is used essentially as a tool for obtaining information from an individual’s social networks and for sharing information with others. Scholars found that being actively involved in sending and receiving tweets may lead to a feeling of co-presence (Oh 2013). Within the Twitter environment, hashtags, represented by the # symbol, are used to categorize topics among tweets so that they are easily found in a Twitter search. This simple idea of micro-blogging service attracted users from all over the world making it one of the most successful social media platform in recent years (Dewan & Ramaprasad, 2014). The simple design of Twitter provides a platform through which users can communicate information quickly via computers and mobile devices. As a result, individuals as well as public and private organizations have been creating Twitter accounts to engage with others who may be potential consumers of their products and services. The popularity of Twitter is impressive with nearly 320 million monthly worldwide active users on Twitter and one billion monthly unique visits as of December 20152. The New York Times articles quoted “On the day of the 2016 U.S. presidential election, Twitter proved to be the largest source of breaking news, with 40 million tweets sent by 10 p.m. that day” and “Twitter’s reach on Election Day was particularly striking in the number of posts embedded outside of the service and into news sites like The New York Times, as well as entertainment-focused sites like TMZ and Perez Hilton. Even other social networks, like Facebook, reaped the benefit of news breaking on Twitter” (Isaac & Ember, 2016).

Why did we choose Twitter as the data source from other social media platforms? We find it to be the best fit for our need in monitoring high volume trending news from general public over a period of time. Scholars concluded that there are two factors that explain the variations between different social media platforms: half-life and depth of information (Weinberg & Pehliven, 2011). While half-life refers to ‘longetivity of information in terms of availability’, depth refers to ‘richness of the content, and the number and diversity of perspectives’. Thus trending content on Twitter comes in high volume but may quickly move off the screen (short half-life) while rich and engaging interactions on Facebook build relationships (high depth of information). In fact, two-thirds of US adults use Twitter for their news consumption (Pew Research, 2015).

Research Design

The focused research methodology in this study is sentiment analysis which is about the classification of direction-based text that contain opinions, emotions, appraisals, and attitudes commonly termed as sentiment (Abbasi et al., 2008; Pang & Lee, 2008). With the advent of user-generated content via Web 2.0, sentiment is rampant present in social media. We adhere to the taxonomy of sentiment polarity classification (See Table 1) introduced by Abbasi, Chen, and Salem (2008) that outlined the following tasks, features, techniques, and domains. Adhering to this taxonomy provides us with a ‘best practice’ set of guidelines for research design of sentiment analysis. We proceed to discuss the various parts of this taxonomy adapted to our study. The three tasks in our study focuses on (1) the extraction of positive and negative sentiment classes from tweets of political discourse where each tweet represents a document (2) as most tweets consist of only a single sentence due to Twitter’s 140 character limitation. In this study, the (3) source is the public’s sentiment represented in tweets and target of the sentiment focuses on the two candidates: Trump and Clinton.

There are generally four feature categories examined in the literature: syntactic, semantic, link-based, and stylistic features. This study focuses on semantic features which is the most commonly used for sentiment analysis (Abbasi et al., 2008). Semantic features are those with polarity tags, appraisal groups, and semantic orientation. They generally incorporate manual, semi, and/or fully automatic annotation techniques to add polarity (positive or negative sentiment) tags or scores to words or phrases. In the next section we discuss the approach of manual coding of a set of tweets known as the gold standard which is then used for evaluating a set of popular classifiers. The classifier with the highest accuracy is then selected to automatically code the remaining tweets for sentiment extraction and analysis.

http://about.twitter.com

---

2 http://about.twitter.com
Our study explores the sentiment analysis technique using machine learning which relies on using algorithms such as Naïve Bayes, Support Vector Machine, and Logistics classifiers code and test sentiment labels. The domain of this study is political discourse in the form of tweets surrounding the two presidential candidates: Trump and Clinton. We proceed to discuss the system design of this study in the next section.

**System Design**

We propose the following system design (Figure 1) which is based on the Social Media Analytics (SMA) methodological framework of Fan and Gordon (2014) (hereafter known as FG framework). The FG framework outlines a clear approach to social media analytics which consists of three stages: 1) capture, 2) understand, and 3) present. We find the FG to be a good fit for our study. One contribution of this study is to demonstrate an implementation of the FG framework where we outline our implementation parallel to the three stages of FG, in particular, the steps of sentiment feature extraction and sentiment analysis of tweet data.

### 1.0 Download Tweets

This study focus on capturing the social media reactions of the public to news, events, and statements related to the two 2016 US presidential candidates: Hillary Clinton and Donald Trump. The objective is to gain insights on the public sentiment extracted from social media surrounding those candidates which findings are relevant to researchers and practitioners alike. As such the first step in the FG framework – ‘capture’ is to find key identifiers, commonly known as hashtags, for Trump (#trump), Clinton (#clinton), and the election (#elections16) that are widely used in Twitter in mentioning those candidates along with news and opinions about them. Past research have shown evidence of the relevance of keywords or hashtags in conveying the value of the specific subjects such as brand commercials (Oh, Sasser, & Almahmoud, 2015), political candidates (Dang-Xuan, Stieglitz, Wladarsch, & Neuberger, 2013; Larsson & Moe, 2011), and social-political events (Choudhary, Hendrix, Lee, Palsetia, & Liao, 2012).

This stage also involves the setup of a platform to download tweets from Twitter Streaming API using Python scripts and MongoDB. Python (python.org) is an easy to learn robust programming language with plenty of packages for collecting, cleaning, and processing data to required format processing. In the current example of extracting tweets from Twitter we have used python package Tweepy (tweepy.org) which support wide range of functions with minimal lines of code. MongoDB (mongodb.com) is a document-based database that uses documents instead of tuples in tables to store data. These documents look like just like JSON objects using key-value pairs.
Since tweet object from Twitter is in JSON format, storing a tweet in MongoDB database is as easy as putting the entire content of the tweet’s JSON string in an insert statement. This reduces the performance overhead due to parsing of the data and saving to a relational data base, processing these tweets once they are stored are simple as well. MongoDB structure provides high scalability as compared to the popularly used Relational DB systems. We collect tweets sent between Nov 3 to Nov 7. Tweets from Nov 8 (Election Day) are not included in our dataset to avoid any noise from after the election result is announced. Examples of both positive and negative tweets for both Trump and Clinton are shown in Table 2.

<table>
<thead>
<tr>
<th>Type of Tweet</th>
<th>Tweet Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump-Positive</td>
<td>@realDonaldTrump HOLD on courageous VETS, #Trump HELP is on the way! #VetsForTrump #DraininTheSwamp</td>
</tr>
<tr>
<td>Trump-Negative</td>
<td>#Trump is an obvious nightmare but keep eyes on the evil #PuppeteerPence. He’s calling more shots than you may think <a href="https://t.co/jt3KW4MS9o">https://t.co/jt3KW4MS9o</a></td>
</tr>
<tr>
<td>Clinton-Positive</td>
<td>She is my president! #ImWithHer #ThePeoplesPresident</td>
</tr>
<tr>
<td>Clinton-Negative</td>
<td>Law Expert: Donna #Brazile, #Clinton May Have Committed a Federal #Crime. #lockherup #hillary #debates <a href="https://t.co/7TWRFMFTaC">https://t.co/7TWRFMFTaC</a></td>
</tr>
</tbody>
</table>

Table 2. Example tweets for both Trump and Clinton

2.0 Pre-processing

The ‘capture’ stage in the FG framework also includes the pre-processing of downloaded tweets, the extraction of features for sentiment analysis, and the next step of data analysis. Once the data is collected, values are extracted from multiple sources of data (MongoDB collections) in order to be stored to a centralized DB. During this process certain features or attributes of each tweet such as timestamp, location, and tweet message are transformed to the required format. Each tweet has a timestamp (e.g. ‘Wed Feb 22 23:51:46 +0000 2017’) and location (e.g. ‘Washington DC’) that needs to be tokenized and aggregated to know its date time and location (Note that location is a feature we did not utilized in this study). Converting the timestamps to required format for filtering and adding more filters such as language, time-zone, and geo location could give us more features to consider. The date
time information facilitates aggregating of tweets into different time periods. The focal component of the tweet is the tweet message (examples given in Table 2). The tweet message has the identifier hashtag or keyword that tells us who or what the tweet is about. Also identifying the new keywords used along with the already identified ones will give us more insight on trending terms. In addition, the stored tweet needs to be cleaned of special characters such as ‘@’, ‘#’, ‘…’ and so on, in order to be passed on to the classifier for classification. Python has regular expression operations to handle these needs. Once each tweet is free of special characters it is tokenized or broken down to words. Once tokenized the words are to be checked for stop words. These are commonly used words that are generally considered useless for sentiment analysis (such as ‘the’). Most search engines ignore these words because they are so common that including them would greatly increase the size of the index without improving precision or recall. We use NLTK, a Natural Language Processing (NLP) toolkit for text processing that comes with a stop words corpus that includes a list of 128 English stop words. This is used to filter out stop words in a tweet. Once tokenized, we create a bag of words model (Jurafsky & Martin, 2014) which is a method of information retrieval that relies on frequency of occurrence of each word in a sentence. This is used as a feature for training a classifier to improve the performance and accuracy of the model. We use a simplified bag of words model where every word is a feature in a tweet and if that word is present in a tweet, it is marked true, else is marked false.

3.0 Sentiment Extraction

Sentiment extraction starts off with preparing a set of manually coded dataset (hereafter known as the gold standard) to test the accuracy of the classifier models in the evaluation stage. We randomly select 508 tweets from the original dataset to be manual coded as positive, negative, or neutral sentiment by three experts. Any differences in manual coding are streamlined by correlating labels from the three experts, finding gaps between the labels, and performing two additional recoding attempts to reach a satisfactory Holsti’s coefficient of 0.95 which shows a high reliability of inter-coder reliability for manual coding (Holsti, 1969). The distribution of the gold standard dataset is in Table 2. Neutral labels are excluded from the gold standard as they add noise rather than improving the classifier performance. The gold standard dataset is then divided into training dataset (250 tweets) and testing dataset (108 tweets) for classification evaluation.

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>250</td>
<td>103</td>
<td>147</td>
</tr>
<tr>
<td>Testing</td>
<td>108</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>Removed</td>
<td>150</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>508</td>
<td>143</td>
<td>215</td>
</tr>
</tbody>
</table>

Table 3. The sentiment distribution of tweet count for training and testing gold standard datasets.

In the next step we build the sentiment models using the libraries available on NLTK and Scikit4. More than one classifier model is built to compare each one’s performance and accuracy by training with the training gold standard dataset and testing against the testing gold standard dataset. Once the classifier with the best accuracy is determined, the remaining of the tweets are classified as positive, negative, or neutral using this classifier. The classified tweets are then used for all our analysis.

We start with Naive Bayes classifiers, which is simple and highly scalable. Naive Bayes classifier assumes that the value of a particular feature is independent of the value of any other feature, given the class variable. Despite their naive design and apparently oversimplified assumptions, Naive Bayes classifiers work well in many complex real-world classifications. Another advantage of Naive Bayes

---

3 NLTK (Natural Language Toolkit) is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries.

4 Scikit is an open source Python library that implements a range of machine learning, preprocessing, cross-validation and visualization algorithms.
classifier is that it only requires a small number of training data to estimate the parameters necessary for performing classification.

Support Vector Machine or SVM is the next classifier we test. A SVM model uses the examples as points in space, position them in two clear categories with a distinct gap that is as wide as possible. New examples are then positioned into the same space and predicted to belong to one category versus the other based on which side of the gap they fit. SVM does not do well when the number of features are more than the number of examples (tweets) which is likely the case with our dataset. Another disadvantage of SVM is that it does not provide any probability estimates which is a challenge for us to determine the neutral examples.

We then use the Logistic regression classifier which has a number of advantages over Naive Bayes and SVM. The overly strong conditional independence assumptions of Naive Bayes mean that if two features are in fact correlated naive Bayes will multiply them both in as if they were independent, overestimating the evidence. Logistic regression is much more robust to correlated features; if two features are perfectly correlated, the classifier will assign half the weight to each side (Jurafsky & Martin, 2014). Thus when there are many correlated features, logistic regression will assign a more accurate probability than Naive Bayes. The advantage of using this classifier apart from having good accuracy is the ability to view probability scores for each tweet helping us to assess the classification of a particular sentiment. We remove all tweets with marginal probability lower than 0.6 as neutral sentiments.

Out of the most frequently occurring words in the feature list or bag of words, we find up to 5000 words/term that are chosen based on the total number of terms in the corpus. As large number of features tend to add noise we fine-tune this by testing the values for various sizes of features for training and testing corpus. We find this step to considerably improve the accuracy and performance of the classifiers. This is known as the feature reduction (Abbasi et al., 2008) in the machine learning context.

### 4.0 Evaluation and Machine Sentiment Coding

Table 4 shows the results of classification models of two sets, models 1-3 (NB, LOG, and SVM) without feature reduction and models 4-6 (NB_FR, LOG_FR, and SVM_FR) with feature reduction (down to 2000 terms). All classification are performed with ten-fold cross-validation to add classification robustness. Recall and precision are two different accuracy measures while F-measure is a mean of both recall and precision (also known as the harmonic mean of recall and precision). The results show that models 1-3 produced the same accuracy but in the feature reduction models (models 4-5) NB_FR performed worse while both LOG_FR and SVM_FR did better. We note that LOG_FR (with F-measure of 0.802) outperforms the remaining classifiers in classifying the test labels. We thus select Model 5 (LOG_FR) as the classifier of choice and proceed to automatically (machine) label all tweets in the remaining dataset using this classifier. We discuss the data distribution in the next section.

<table>
<thead>
<tr>
<th>Model</th>
<th>Recall</th>
<th>Precision</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Naïve Bayes (NB)</td>
<td>1</td>
<td>0.620</td>
<td>0.765</td>
</tr>
<tr>
<td>2 Logistic (LOG)</td>
<td>1</td>
<td>0.620</td>
<td>0.765</td>
</tr>
<tr>
<td>3 SVM (SVM)</td>
<td>1</td>
<td>0.620</td>
<td>0.765</td>
</tr>
<tr>
<td>4 Naïve Bayes with feature reduction (NB_FR)</td>
<td>0.537</td>
<td>0.705</td>
<td>0.610</td>
</tr>
<tr>
<td>5 Logistic with feature reduction (LOG_FR)</td>
<td>0.850</td>
<td>0.76</td>
<td>0.802</td>
</tr>
<tr>
<td>6 SVM with feature reduction (SVM_FR)</td>
<td>1</td>
<td>0.620</td>
<td>0.765</td>
</tr>
</tbody>
</table>

**Table 4. Classification accuracy results**

### 5.0 Data Analysis and Findings

In this section we discuss the data analysis and findings of our study. This step coincides with the ‘understand’ stage and the ‘present’ stages respectively as shown in the FG framework. We analyze 1) volume of tweets, 2) sentiment index, and 3) text analytics of public sentiment for tweets for both Trump and Clinton.

#### 5.1 Volume of Tweets

We analyze the share of attention on Twitter with regards to each candidate, Trump and Clinton, by aggregating number of tweets having the respective hashtags (#trump and #clinton). This is
consistent with the approach of previous studies such as Tumasjan et al. (2011) and Larsson and Moe (2011). Overall, our study concludes a higher number of pre-election tweets for Trump over Clinton by 219% (See Table 4). Even comparing day-to-day, Trump surpassed Clinton each day from 11/3 to 11/7. After removing tweets where both hashtags #trump and #clinton are present, we obtained 434,425 count for Trump versus 135,875 count for Clinton. This initial findings are consistent with reports from Wall Street Journal (Keegan, 2017) which found that Trump (16.8 million followers) has over 46% more followers than Clinton (11.5 million followers) and engaged (34,092 tweets sent by Trump) over three times more than Clinton (9,838 tweets sent by Clinton).

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>3-Nov</th>
<th>4-Nov</th>
<th>5-Nov</th>
<th>6-Nov</th>
<th>7-Nov</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump</td>
<td>434425</td>
<td>44537</td>
<td>86285</td>
<td>55597</td>
<td>119715</td>
<td>128291</td>
</tr>
<tr>
<td>Clinton</td>
<td>135875</td>
<td>13684</td>
<td>29896</td>
<td>15201</td>
<td>34939</td>
<td>42155</td>
</tr>
<tr>
<td>Trump-Clinton</td>
<td>298550</td>
<td>30853</td>
<td>56389</td>
<td>40396</td>
<td>84776</td>
<td>86136</td>
</tr>
<tr>
<td>(Trump-Clinton)/Clinton (%)</td>
<td>219.72</td>
<td>225.46</td>
<td>188.61</td>
<td>265.74</td>
<td>242.64</td>
<td>204.33</td>
</tr>
<tr>
<td>Trump-positive</td>
<td>93223</td>
<td>9310</td>
<td>18299</td>
<td>16071</td>
<td>24582</td>
<td>24961</td>
</tr>
<tr>
<td>Trump-negative</td>
<td>341202</td>
<td>35227</td>
<td>67986</td>
<td>39526</td>
<td>95133</td>
<td>103330</td>
</tr>
<tr>
<td>Clinton-positive</td>
<td>3260</td>
<td>157</td>
<td>232</td>
<td>487</td>
<td>403</td>
<td>1981</td>
</tr>
<tr>
<td>Clinton-negative</td>
<td>132615</td>
<td>13527</td>
<td>29664</td>
<td>14714</td>
<td>34536</td>
<td>40174</td>
</tr>
</tbody>
</table>

Table 5. Contrast of total, positive, and negative tweet volumes for Trump and Clinton from Nov 3 to Nov 7.

Scholars have confirmed that engagement in social media is correlated with brand success (Oh, Roumani, Nwankpa, & Hu, 2017) and political success (Tumasjan et al., 2011). According to Stieglitz and Dang-Xuan (2013) higher volume (scale) of public sentiment for a particular brand is likely to correlate with higher demand for the brand’s products and services. This disparity denotes a clear chasm between social media support for Trump and Clinton that probably led to a Trump victory over Clinton.

We then proceed to contrast the count of positive versus negative tweets for both candidates. We found Trump to have overwhelmingly more tweet count for both positive and negative over Clinton. This signifies the higher level of buzz surrounding Trump as compared to Clinton. We thus conclude that this correlation between higher volumes of tweets (total, positive, and negative) led to a higher likelihood of a Trump win.

5.2 Sentiment Index

We aggregate machine coded tweets into four groups 1) Trump-positive, 2) Trump-negative, 3) Clinton-positive, and 4) Clinton-negative which distribution is outlined in Figure 2. Results show that the proportion of negative sentiment is more than positive sentiment for both candidates, although, Trump’s sentiment volume is more than Clinton for both sentiment as well. Unlike Tumasjan et al. (2011) who stated that positive emotion outweigh negative emotion by 2 to 1, we find about four times more negative tweets than positive ones in total implying that the proportion of tweet sentiment is not always consistent and context dependent. Following bullishness index formulation of Antweiler and Frank (2004) that examined sentiment of stock market reviews, we generate a sentiment index for both Trump and Clinton over the five days leading up to Election Day. It is interesting to note that sentiment index for Trump is more than Clinton indicating that although the public are negative towards both candidates, the level of negative sentiment for Trump is lower than that of Clinton’s. Nonetheless both candidates’ negative tweet volumes increase dramatically towards Election Day. We thus conclude that the higher sentiment index is likely to correlate with a Trump win. Results is shown in Table 6.

\[
\text{SENTIMENT} = \ln \frac{1 + \text{TOTAL}^{\text{POSITIVE}}}{1 + \text{TOTAL}^{\text{NEGATIVE}}}
\]
We proceed to perform text analytics based on word clouds to uncover frequently used associated topics or terms generated by the public for both Trump and Clinton. In addition to its novelty, word cloud has emerged as a straightforward and visually appealing visualization method for text analytics. They are used in various contexts as a means to provide an overview by isolating text messages down to words, topics, or terms that appear with highest frequency. These terms do provide insights into important conversations and opinions from the public concerning the sources, which in our case, are the two presidential candidates. Scholars concluded that “The results indicate that users are on average more effective in spotting a specific term in an alphabetically ordered unweighted list than in an alphabetically ordered word cloud. However, frequently used terms are found more quickly in word clouds due to their larger font sizes” (Rivadeneira et al., 2007).

Hashtags are used in Twitter to classify messages, propagate ideas, and also to promote specific topics and people (Cunha et al., 2011). Essentially it is a way to participate in a conversation on a certain topic. However when these conversations are aggregated they may be substantial in affecting the outcome. Easley and Kleinberg (2010) characterize what is known as “rich-get-richer phenomenon” or “preferential attachment process” where the popularity of the most common terms tends to increase faster than the popularity of the less common ones resulting in further diffusion of information that achieve a higher level of esteem.

We start off by analyzing word cloud of Trump (see Figure 3). The term or hashtag #MAGA is the most frequently appearing term in tweets for both Clinton and Trump. #MAGA is the abbreviation for Trump’s main campaign tag line “Make America Great Again” which was first used by Trump on the day of the 2012 Presidential Election win by the previous U.S. president Obama over Republican nominee Mitt Romney. Trump’s slogan closely resembles former U.S. president Ronald Reagan’s 1980 slogan, ”Let’s Make America Great Again.” MAGA quickly became Trump’s central campaign rallying cry which is used by Trump’s supporters that trends both candidates’ tweets due to the high volume of Trump tweets overall. This common hashtag is a key factor allowing supporters to join the relevant conversations using a central topic congregating as a whole resulting in a monumental social voice influencing public sentiment.

Trump’s criticism and frequent mention of Obama along with public’s tweets using hashtag #obama resulted in the term #obama being the second most prominent hashtag in Trump’s word cloud (e.g. see an example tweet below). Other popular hashtags such as #TrumpTrain, #Trump2016, #TrumpPence16, and #TrumpPresident are all positive Trump terms signifying the public’s support for the candidate in the propagation of such tweets.
Another interesting finding is the presence of the term #brexit which is related to a combination of the terms ‘British’ and ‘exit’ in reference to the shocking and unexpected decision by the UK voters to withdraw from the EU. In hindsight, those in power, the elite folks and urban bureaucrats have grossly underestimated the strength of public sentiment. And unbelievably, similar to the outcome of Brexit, the presidential election did surprise and shock many as well.

On Trump’s negative side, the terms #NotMyPresident and #TrumpProtest are two prominent hashtags used to convey the public’s dislike towards Trump. Interestingly, the term “Not my president” was a favorite catchphrase of the Tea Party, a fundamental support group to oust the former President Obama. We also note the existence of neutral hashtags such as #ElectionNight and #USElection2016 indicating public’s participation in the conversations surrounding the event.

In examining the word cloud of Clinton, we note a higher concentration of negative terms as contrast with positive terms signifying Twitter users’ overwhelming uproar against Clinton. The more noticeable terms are #MAGA, #ImWithHer, #WikiLeak, and CrookedHillary. #MAGA seems to be mentioned the most alongside #Clinton which may be due to the lower number of Clinton’s tweets as compared to Trump’s. #ImWithHer was the most positive hashtag used to show the public’s support for Clinton.

On the negative side, we note a massive propagation of negative tweets against Clinton signifying the influential effect of negativism on voter preferences. For example, the terms #WikiLeaks and spiritcooking are related to WikiLeaks’ release of email messages involving Clinton’s campaign Chairman John Podesta. Podesta is a long-term associate of the Clintons and was President Bill Clinton’s Chief of Staff from 1998 until 2001. In March 2016, Podesta’s gmail account was hacked resulting in at least 20,000 pages of emails stolen. These emails were later obtained and release by Wikileaks. Some of the emails provide inside information on the Clinton’s campaign. This incident coincided with the closing of poll gap between Trump and Clinton although the release of emails was never acknowledged to have an effect of Clinton’s trustworthiness. Nevertheless many do believe this has a role in Clinton’s eventual loss.

The term #FBI and #comey relate to the FBI’s 2015 investigation of Clinton’s use of private email server for official communication including discussion on arrests in the Benghazi attacks. Some claimed that she violated government protocols and procedures and such act is deemed a crime. These emails were later classified as ‘SECRET’ by the FBI. In May 2016, the FBI announced that Clinton was careless in her handling of these emails but no charges were filed against her. Then in late Oct 2016, FBI Director James Comey revealed startling news when he notified Congress that FBI had discovered additional emails in a separate investigation that could be connected to the previous case. He then
declared not finding any evidence of criminality against Clinton on Nov 6, just two days before the election. This uproar was taken as an opportunity by Trump and his supporters to draw mistrust from the public over Clinton’s reputation. Some argued that the timing and the effect of this negative news played a role in Clinton’s defeat. In the national exit poll, some 45% of voters surveyed said they were disturbed by this incident (ABC News, 2016).

Other terms found associated with Clinton like #draintheswamp, #LNYHBT, and #TCOT are terms used to declare the conservative stance against the existing liberal establishment. Conservatives on numerous occasions have called to clean up government corruption with the tag ‘drain the swamp’. The hashtag #draintheswamp was later piggybacked by Trump as a campaign pledge to clean the existing government. LNYHBT is an acronym for “Let Not Your Heart Be Troubled” used by fans of conservative broadcaster Sean Hannity. The term #TCOT is an acronym for top conservatives on Twitter in reference to top conservative individuals voicing their stand against the liberal influence. Many include this hashtag to side with conservative camp.

Our analysis show overwhelming negative tweets flooding Twitter days leading up to the election. These tweets are proportionately worse off for Clinton than Trump which may have resulted in deteriorating public support for Clinton. Sadly, tweets on positive endeavors such economic plans or job creation are overcrowded with negativity.

**Conclusion**

In this IS-focused study, we implement a framework to extract sentiment from social media data of political discourse related to both candidates of the recent U.S. presidential election. Our framework covered nuance insights into the reasons how Trump won the election despite 91.5% of the polls predicting otherwise. We concluded that there is a higher volume of tweets for Trump signifying larger share of public attention, a higher proportion of negative sentiment for Clinton than Trump, and more negative trending topics against Clinton than Trump. This similar finding is also supported in Facebook, another popular social media platform, where Trump has more volume of engagements (67 million for Trump, 59 million for Clinton) and interactions (1.14 billion for Trump, 934 million for Clinton) than Hillary: engagements interactions (Washington Examiner, 2016).

The overall consensus of voting polls from major news networks showed a totally different outcome - predicting a Clinton landslide win over Trump. These polls represent public sentiment over the period leading up to election day. Table 7 below outlined ten polls sources (Real Clear Politics, 2016). Ironically, even Fox News, a conservative channel, predicted similarly. This misalignment is a puzzling question for all parties involved.

<table>
<thead>
<tr>
<th>Poll</th>
<th>Date</th>
<th>Clinton</th>
<th>Trump</th>
<th>Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bloomberg</td>
<td>11/4-11/7</td>
<td>46</td>
<td>48</td>
<td>Clinton +3</td>
</tr>
<tr>
<td>IBD/TIPP</td>
<td>11/4-11/7</td>
<td>43</td>
<td>42</td>
<td>Clinton +1</td>
</tr>
<tr>
<td>Economist</td>
<td>11/4-11/7</td>
<td>49</td>
<td>45</td>
<td>Clinton +4</td>
</tr>
<tr>
<td>LA Times</td>
<td>11/1-11/7</td>
<td>44</td>
<td>47</td>
<td>Trump +3</td>
</tr>
<tr>
<td>ABC</td>
<td>11/3-11/6</td>
<td>49</td>
<td>46</td>
<td>Clinton +3</td>
</tr>
<tr>
<td>FOX News</td>
<td>11/3-11/6</td>
<td>48</td>
<td>44</td>
<td>Clinton +4</td>
</tr>
<tr>
<td>Monmouth</td>
<td>11/3-11/6</td>
<td>50</td>
<td>44</td>
<td>Clinton +6</td>
</tr>
<tr>
<td>NBC</td>
<td>11/3-11/5</td>
<td>48</td>
<td>43</td>
<td>Clinton +5</td>
</tr>
<tr>
<td>CBS</td>
<td>11/2-11/6</td>
<td>47</td>
<td>43</td>
<td>Clinton +4</td>
</tr>
<tr>
<td>Reuters</td>
<td>11/2-11/6</td>
<td>44</td>
<td>39</td>
<td>Clinton +5</td>
</tr>
</tbody>
</table>

**Table 6. Polling results from ten news networks from Nov 1 to Nov 7.**

National exit polls as well as recent research have shown that it is the resurgence of the middle class (Pew Research, 2016) and working class white (Washington Post, 2016) communities that tip the votes towards Trump. And analysts attribute this to Trump’s ability to not just connect but also arouse these people to take the necessary action in making a political change. Trump’s campaign engaged in a manner that is provocative and at times disturbing but his message aligns well with his supporters.
The connection between exit polls results and the result of our study lies in understanding the demographics of Trump’s Twitter 20 million followers of which the major groups are 1) faith, family and football (20%), 2) conservatives (13%), and 3) suburban moms (7%) (Forbes, 2017). These three groups that total 40% of all Trump followers are generally conservative, Christians, working class, and white, which shows a consistent overlap between Trump voters and Twitter followers.

Thus we show how social media in general, and twitter sentiment specifically, play an acute role in disseminating of political messages to the general public. Social media does not necessarily replace the role of traditional media but superimposed or add to it by broadcasting information of all types, without any filter on authenticity and morality of any kind. As long as it catches enough attention and re-broadcasted by enough people, it starts to trend. Gone are the days when the only conduit to reach the public is through traditional media (e.g. newspapers, television, and radio) which is strongly controlled and influenced by those in power.

The role of social media as a medium of information dissemination and sharing must not be taken lightly. Thus we conclude that applying sentiment analysis to obtain insights from political conversations in social media surrounding political candidates is an important endeavor contributing to design of better monitoring tools for social media use. Such tools provide more in-depth insights of political discourse for prediction and forecasting.

References


