The Thirty-First International Florida
Artificial Intelligence Research Society Conference (FLAIRS-31)

Location-Based Twitter Sentiment Analysis for Predicting the U.S. 2016 Presidential Election

Brian Heredia, Joseph D. Prusa, Taghi M. Khoshgoftaar
Florida Atlantic University
bheredia@fau.edu, jprusa@fau.edu, taghi@cse.fau.edu

Abstract

We seek to determine the effectiveness of using location-based social media to predict the outcome of the 2016 presidential election. To this aim, we create a dataset consisting of approximately 3 million tweets ranging from September 22nd to November 8th related to either Donald Trump or Hillary Clinton. Twenty-one states are chosen, with eleven categorized as swing states, five as Clinton favored and five as Trump favored. We incorporate two metrics in polling voter opinion for election outcomes: tweet volume and positive sentiment. Our data is labeled via a convolutional neural network trained on the sentiment140 dataset. To determine whether Twitter is an indicator of election outcome, we compare our results to the election outcome per state and across the nation. We use two approaches for determining state victories: winner-take-all and shared elector count. Our results show tweet sentiment mirrors the close races in the swing states; however, the differences in distribution of positive sentiment and volume between Clinton and Trump are not significant using our approach. Thus, we conclude neither sentiment nor volume is an accurate predictor of election results using our collection of data and labeling process.

Introduction

United States politics primarily revolves around two political parties: the Republicans and the Democrats, which both hold their own primaries to determine their candidate for the general election. On election day, every eligible U.S. citizen over the age of 18 is allotted a vote. In the U.S., the votes cast by the public do not actually elect the president, that decision is made through the electoral college. The electoral college was established in the U.S. constitution to allow influence both from Congress and the popular vote. Every state is allocated a number of electors equal to the sum of its number of U.S. Senators and its representatives. The number of senators is always two, while the number of representatives can vary depending on the size of the state's population determined in the official U.S. Census. In total, there are 538 electors and the candidate is declared the victor when they have a majority of these electors. Most states operate under a winner-take-all approach, where the candidate with the majority of votes is awarded all the electors for that state. States are generally considered to be one of two types: swing states or favored states. Swing states are battlegrounds where candidates are evenly matched in votes based on polling, while favored states are those where one candidate has a large advantage. Typically, campaigns focus on swing states, as allocating time and resources into favored states can be viewed as a waste due to a candidate's already large advantage or disadvantage.

For the 2016 general election, we focus on the two frontrunners: the Republican candidate Donald Trump and the Democratic candidate Hillary Clinton. A voter's decision depends on many factors, ranging from policy choices to personal beliefs. Due to this, the most effective method of predicting election outcomes is through public opinion polling. Polling refers to the process of asking a set list of questions to a subset of the population in order to gauge public opinion. As the 2016 election was one of the most popular elections, with 136 million Americans participating, many different polls were conducted by different organizations to determine public opinion on candidates (RCP 2016). Polling is conducted at the national level and also at the state level, to determine the candidate's standing for that state. While national polling did correctly estimate popular vote, this alone does not indicate the victor, as the electoral college determines the president-elect.

Twitter is a social media platform that allows users to publish short micro-blog posts, called tweets, and view other users' tweets. For this reason, Twitter has become a platform for politicians to release information and communicate with their constituents. In the 2008 general election, presidential campaigns leveraged Twitter to earn popularity and votes. This trend was amplified in the 2016 election, where both candidates took to Twitter to disseminate information related to their campaigns and to deter voters from voting for the other candidates. In response, many voters took to Twitter to express their views on the candidates and defend their choice or debase their opponent. These tweets can be used for opinion mining, as they convey the sentiment of the author. Text mining has been shown to be a strong predictor for many domains, such as box office results (Meador and Gluck 2009), spam review detection (Heredia et al. 2016) (Heredia et al. 2017), and even predicting elections in smaller countries, such as the United Kingdom (U.K.) in 2010, Ireland in 2011, and Germany in 2009 (Boutet, Kim, and Yoneki

2012), (Bermingham and Smeaton 2011), (Tumasjan et al. 2010).

Our primary investigation is to determine whether Twitter can be an effective predictor for the 2016 presidential election. To this aim, we collected approximately 3 million tweets between September 22nd and November 8th. We train a deep convolutional neural network on the sentiment140 tweet dataset and use the resulting model to label sentiment in our tweets. We use location data to place tweets in their respective states and two approaches are explored for determining election outcome: volume of tweets referring to each candidate and positive sentiment of tweets towards each candidate. Volume of tweets has been shown to correlate well with election outcome in related studies (Boutet, Kim, and Yoneki 2012), (Tumasjan et al. 2010), while our recent study found sentiment performed better than volume at polling the U.S. 2016 election (Heredia, Prusa, and Khoshgoftaar 2017).

The remainder of this paper is organized as follows. The related work section presents previous studies regarding social media mining in the political domain. Methodology presents our experimental framework, including dataset, neural network architecture, and approach. A discussion of our results and discussions are presented in the Results section, and the final section presents our conclusions and possible avenues for future work.

Related Work

In politics, understanding public opinion plays a large role in anticipating election outcomes. Research into the effects of social media on political discourse has become necessary due to the rising trend of social media politics.

Twitter has been used as the main platform for communication between those holding political office and the general public. A study by Boutet et al. proposed an algorithm for determining the political party of a Twitter user during the 2010 U.K. election cycle (Boutet, Kim, and Yoneki 2012). A total of 419 trending topics were chosen for tweet collection. The authors manually labeled self-identified party members as a ground truth dataset. A Support Vector Machine and two Bayesian classification methods were proposed for the identification of user party. Their results showed the Bayesian-Volume classifier significantly outperformed the remaining classifiers.

A study by Bermingham and Smeaton explored the predictive effects of Twitter on the 2011 Irish election (Bermingham and Smeaton 2011). During this study, the authors developed "Twitter Tracker," a software allowing them to tap into content on Twitter pertaining to the election. A total of 32,578 tweets relevant to the five main parties were collected by searching for the party names, abbreviations, and the election hashtag. The authors compared their Twitter-based predictions against nine polls collected during the election. The sets of tweets used were separated into time-based, sample size-based, cumulative, and manual. Nine annotators were trained before the date of the election and used for labeling sentiment in tweets. The tweets used to train these annotators were taken from different points in time as to develop a diverse training corpus. Authors found the addition of sentiment analysis increased the accuracy of their predictions.

Similar to the above study, Tumasjan et al. explored the power of Twitter and tweet sentiment in the 2009 German national parliament election (Tumasjan et al. 2010). The authors collected 104,003 tweets in the weeks leading up to the election. The Linguistic Inquiry and Word Count 2007 software was used to extract sentiment from tweets. Authors found that political discussions are usually driven by a small number of heavy users (80+ tweets). This study also found that the volume of tweets mirrors the result of the elections.

A study by Anuta et al. focused on exploring the accuracy of polls in the 2016 U.S. election and comparing them to the accuracy of Twitter as a polling source (Anuta, Churchin, and Luo 2017). Five polls from eight different sources were gathered between January and the election day in November. The study was split into two categories, popular and state votes. Polls were adjusted to represent a two-party system and sentiment labels were added to tweets using the VADER python package. For the state votes, nine states were chosen and only two polling sources were used. Of the states, three were Clinton favored, three were Trump favored, and the final three were swing states. Tweet collection was limited to users who appeared in a previously created dataset of tweet IDs, and may not be representative of the total or state populations. The authors presented time series data of the bias for each poll and tweets and show that as the election date draws nearer, the bias lowers. They found twitter yielded a popular vote bias of 3.4% in favor of Trump (higher than any poll) and, across all nine states, an average bias of 2.4% towards Trump.

Our study will differ from these related works in multiple ways. For one, we use supervised learning to create a sentiment classifier which labels our tweets; specifically, a deep neural network (Prusa and Khoshgoftaar 2017). Our study provides a more comprehensive look into electoral results by examining 21 states; 11 swing states, five Clinton leaning, and five Trump leaning. Our election tweets are labeled with two emotions: positive and negative. We use our election dataset to predict the electoral vote using volume of tweets, shown to be a powerful predictor in almost all related works, and also using sentiment conveyed in the tweets. We compare our results to actual state results for the 2016 election.

Methodology

Datasets

In our experiments, we use tweets collected during the election cycle and the sentiment140 corpus (Go, Bhayani, and Huang 2009), which is a collection of tweets labeled based on the emoticons present. Eight emoticons were used to label tweets as either positive or negative and once tweets were labeled, emoticons were removed from the tweets. The final dataset consists of 1.6 million tweets, with 800K positive and 800K negative instances. In our experiment, this dataset is used to train a convolutional neural network (ConvNet) for the purpose of determining sentiment in our tweet election data.

	HyperParameters				Num. of Trainable Parameters
Conv Layers	Number of Filters	Filter	Pool	Activation	Max Instance Length (374)
conv1	128	3×3	–	relu	1,280
conv2	128	3×3	–	relu	147,584
conv3	128	3×3	2×2	relu	147,584
conv4	128	3×3	–	relu	147,584
conv5	128	3×3	–	relu	147,584
conv6	128	3×3	2×2	relu	147,584
FC Layers	Number of Neurons			Activation	
fc1	512			relu	12,190,208
fc2	512			relu	262,656
fc3	2			softmax	1,026
Total Trainable Parameters					13,193,090

Table 1: Neural network layers, hyperparameters and trainable parameters

The election dataset consists of approximately 3 million instances collected from September 22nd to November 8th (election day). No restrictions on location or user were implemented; however, query terms were limited to topics related to either the Republican or Democratic candidate. Text was extracted from these tweets and labeled for sentiment using the previously mentioned ConvNet model. Tweets that contained mentions of both candidates were removed from the dataset before labeling. Tweets were then sectioned into their respective state based on the location data found in the user profile, tweets from users with no location data were removed from the dataset. Twenty-one states were selected, with 11 being swing states (CO, FL, IA, MI, MN, NC, NH, NV, OH, PA, VA), five favored for Clinton (CA, NY, MD, MA, HI), and five favored for Trump (AL, OK, TN, WV, WY). These favored states were chosen because they had the largest discrepancies between the candidates in terms of vote percentage. The remaining 11 states were identified as swing states by major polls. Only 21 states were included in our study due to the limited availability of location data in the collected tweets. The final dataset consists of 716,282 tweets from 21 states.

We did not collect data regarding the candidates from minor parties running for presidency, as we are only interested in the two front-runners. This should cause no difference when comparing to election results as there were no states won by independent candidates during the 2016 election. Two approaches are used: winner-take-all and shared elector count. For the shared elector count, data was normalized based on the amount of electors in the state and, either volume or positive sentiment originating from said state.

Neural Network Architecture and Text Embedding

Due to the size of our election dataset, labeling tweets by hand is not feasible; thus, we opt for a state-of-the-art deep neural network approach. We elect to use a ConvNet architecture which has been demonstrated to be effective for text analysis, specifically at the character-level. (Prusa and Khoshgoftaar 2017), (Zhang and LeCun 2015). ConvNets consist of one or more convolutional layers, max pooling, and dense fully connected layers in some sequence. Table 1 presents the breakdown of our neural network by layer, with the number of trainable parameters for each layer. Our architecture consists of six convolutional layers with max pooling every three layers. The output is then passed through two fully connected layers and into the classification layer. The ConvNet was implemented using TensorFlow 1.1.0 with the Python API (Heredia, Prusa, and Khoshgoftaar 2017) (Prusa and Khoshgoftaar 2017).

Neural networks require numeric inputs, meaning the text needs to be pre-processed before being passed as input to the neural network. To complete this requirement, we use character embedding, which converts text into an image-like matrix representation. We opt for log(m) embedding using a 256-character alphabet, as that encompasses the full UTF-8 character list (Prusa and Khoshgoftaar 2017). Log(m) embedding takes a character c in an instance of length m and creates a vector representation of c by converting to its corresponding numeric value and then calculating the binary representation of that number. This creates a vector of size 8, which is then combined with the remaining character vectors in the instance to form a matrix of size $8 \times m$.

As convolutional layers require uniform input, the data was padded to the size of the largest instance, in our case 374 characters. The model was trained over the course of 20 epochs using a 90/10 train/test split in the data, with the data being randomly shuffled at the start of every epoch. This resulted in an accuracy of 84% for the final model. We note that due to our labeling method, some tweets may be mislabeled, however the majority of tweets are shown to be labeled correctly.

Results

Previous studies have found tweet volume to be a good indicator of election results and political leanings (Tumasjan et al. 2010), (Boutet, Kim, and Yoneki 2012); however, these studies are conducted in countries that do not employ the electoral college process. We utilize the ratio of Trump to Clinton tweets as a metric for volume and a ratio of positive Trump to Clinton tweets as a metric for sentiment. To determine the effectiveness of our approaches, we compare our volume and sentiment ratios to the election results. This is conducted across the full dataset and for each of our 21 individual states. We explore the differences and similarities between swing states and favored states. The results for the 2016 presidential election declared Trump the victor with 306 electors (56.88%) across 30 states, leaving Clinton with the remaining 232 electors (43.12%) across the other

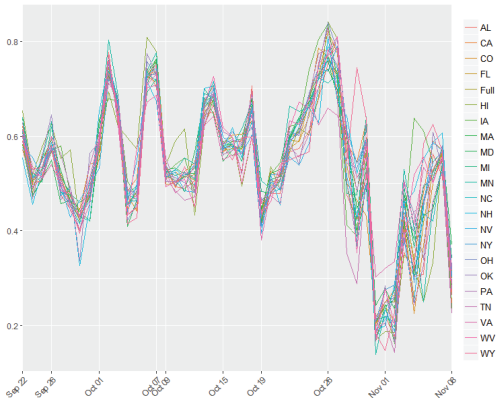

Figure 1: Visualization of the daily volume ratio of Trump tweets to Clinton tweets (Trump:Clinton) for all 21 states and the full dataset.

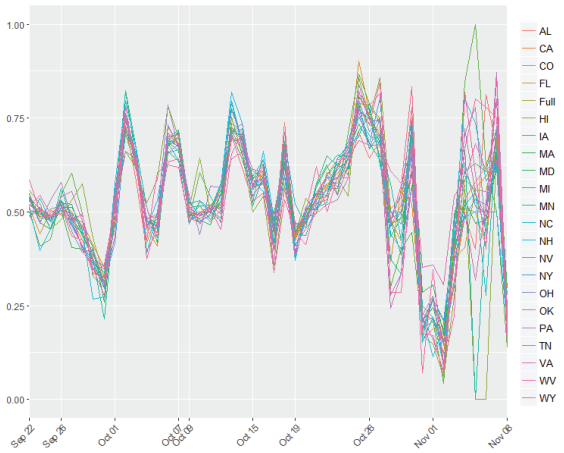

Figure 2: Visualization of the daily positive sentiment ratio of Trump tweets to Clinton tweets (Trump:Clinton) for all 21 states and the full dataset.

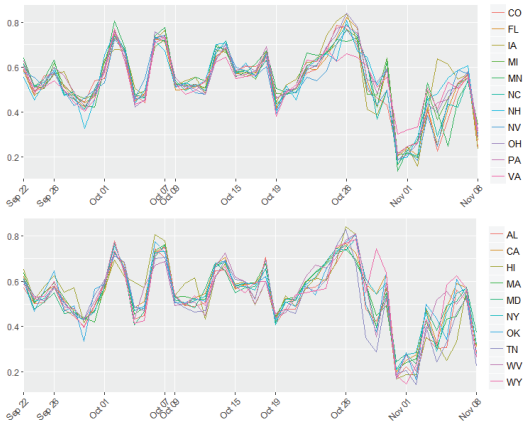

Figure 3: Visualization of the daily volume ratio of Trump tweets to Clinton tweets (Trump:Clinton) of tweets for swing (top) and favored (bottom) states.

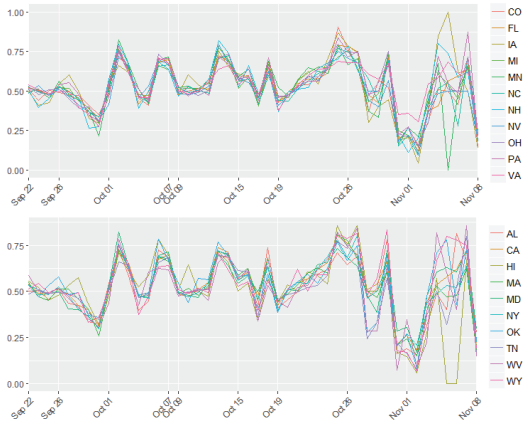

Figure 4: Visualization of the daily positive sentiment ratio of Trump tweets to Clinton tweets (Trump:Clinton) of tweets for swing (top) and favored (bottom) states.

20 states and the District of Columbia (D.C.). When looking at the states chosen in our study, Trump had 139 (47.93%) electors and Clinton had 151 (52.07%) electors.

Figures 1 and 3 depict the daily volume ratio of Trump to Clinton tweets, while Figures 2 and 4 depict the daily positive sentiment ratio of Trump to Clinton tweets. The y-axis represents the Trump:Clinton ratio, meaning a larger value indicates more tweets, or more positive tweets, about Trump. Figures 1 and 2 depict this across all 21 states and the full dataset, while Figures 3 and 4 displays the ratios across both swing (top) and favored (bottom) states separately.

The number of tweets regarding the candidates seems to fluctuate greatly across the election period. When comparing all graphs, we observe the ratio in tweets mentioning either candidate follows a trend across the majority of the election time frame regardless of state, volume or sentiment. It is interesting to note that after the first debate (Sep 26th) the number of tweets related to Clinton increases and the majority of news outlets agree that Clinton won the first debate. On October 7th, the video of Trump musing about sexually assaulting women was released to the public and we see a sharp decrease in the following days. For the second debate (Oct 9th), polls conducted by Politico, CNN, NBC, Gallup, and Fox news all found Clinton to be the victor. However, Twitter shows a steady ratio in the days after this debate, with a small increase for Clinton on Oct 13th. According to the polls, Clinton was again victorious in the third debate (Oct 19th), although this time by a smaller margin. Twitter does not depict this trend, as in the days after the 19th there is only a large steady increase for Trump. Generally, tweets follow the same pattern regardless of state, with few deviations, until the final days leading towards the election. In these days, sentiment seems to exhibit greater variation than volume. To examine volume and sentiment as predictors of election result, we take two approaches: a shared elector count and a winner-take-all approach.

Table 3 displays the volume, positive sentiment, and percentage of positive sentiment per candidate before normal-

State	Class	Electoral votes	Trump volume	Clinton volume	Winner-Take-All	Trump positive	Clinton positive	Winner-Take-All
AL	Trump	9	4.58	4.42	Trump	4.37	4.62	Clinton
CA	Clinton	55	29.69	25.31	Trump	28.01	26.99	Trump
CO	Swing	9	4.74	4.26	Trump	4.48	4.52	Clinton
FL	Swing	29	15.15	13.85	Trump	14.54	14.46	Trump
HI	Clinton	4	2.13	1.87	Trump	2.02	1.98	Trump
IA	Swing	6	3.24	2.76	Trump	3.08	2.92	Trump
MA	Clinton	11	5.89	5.11	Trump	5.51	5.49	Trump
MD	Clinton	10	5.27	4.73	Trump	5.01	4.99	Trump
MI	Swing	16	8.54	7.46	Trump	8.08	7.92	Trump
MN	Swing	10	5.43	4.57	Trump	5.14	4.86	Trump
NC	Swing	15	7.84	7.16	Trump	7.30	7.70	Clinton
NH	Swing	4	2.12	1.88	Trump	2.01	1.99	Trump
NV	Swing	6	3.12	2.88	Trump	2.98	3.02	Clinton
NY	Clinton	29	15.47	13.53	Trump	14.55	14.45	Trump
OH	Swing	18	9.48	8.52	Trump	9.01	8.99	Trump
OK	Trump	7	3.73	3.27	Trump	3.65	3.35	Trump
PA	Swing	20	10.68	9.32	Trump	10.10	9.90	Trump
TN	Trump	11	5.58	5.42	Trump	5.30	5.70	Clinton
VA	Swing	13	6.77	6.23	Trump	6.58	6.42	Trump
WV	Trump	5	2.63	2.37	Trump	2.45	2.55	Clinton
WY	Trump	3	1.57	1.43	Trump	1.51	1.49	Trump
Full		290	153.51	136.49	Trump	145.65	144.35	Trump

Table 2: Normalized electors by volume and sentiment for both shared elector vote and winner-take-all

State	Class	Trump Vol	Clinton Vol	Trump Pos	Clinton Pos	Trump pos/vol	Clinton pos/vol
AL	Trump	**7,180**	6,918	2,958	**3,122**	.4120	**.4513**
CA	Clinton	**71,967**	61,369	**29,329**	28,265	.4075	**.4606**
CO	Swing	**10,220**	9,196	4,073	**4,112**	.3985	**.4472**
FL	Swing	**50,544**	46,215	**21,066**	20,959	.4168	**.4535**
HI	Clinton	**2,174**	1,904	**863**	849	.3970	**.4459**
IA	Swing	**4,366**	3,709	**1,749**	1,658	.4006	**.4470**
MA	Clinton	**16,050**	13,924	**6,408**	6,386	.3993	**.4586**
MD	Clinton	**12,364**	11,104	**5,097**	5,071	.4122	**.4567**
MI	Swing	**13,193**	11,539	**5,281**	5,178	.4003	**.4487**
MN	Swing	**7,037**	5,931	**2,801**	2,648	.3980	**.4465**
NC	Swing	**16,799**	15,361	6,574	**6,937**	.3913	**.4516**
NH	Swing	**4,914**	4,363	**1,988**	1,982	.4046	**.4543**
NV	Swing	**7,844**	7,227	3,142	**3,187**	.4006	**.4410**
NY	Clinton	**59,297**	51,826	**24,242**	24,080	.4088	**.4646**
OH	Swing	**17,449**	15,681	**7,161**	7,145	.4104	**.4556**
OK	Trump	**5,049**	4,431	**2,171**	1,997	.4300	**.4507**
PA	Swing	**18,602**	16,222	**7,521**	7,380	.4043	**.4549**
TN	Trump	**9,907**	9,627	4,069	**4,388**	.4107	**.4558**
VA	Swing	**37,064**	34,102	**15,770**	15,399	.4255	**.4516**
WV	Trump	**2,999**	2,709	1,186	**1,238**	.3955	**.4570**
WY	Trump	**4,131**	3,774	**1,697**	1,668	.4108	**.4420**

Table 3: Non-normalized winner-take-all results

ization. Items in boldface indicate the victor in the corresponding state. This table shows that Trump has more tweets pertaining to him across every state, thus the winner-take-all approach grants Trump a unanimous victory. Since volume of tweets is imbalanced, we explore positive sentiment of a candidate over tweet volume for that candidate. When examining the percentage of positive tweets over all tweets for a candidate, we observe a flip in state turnout, where Clinton is victorious in all states using the winner-take-all approach. This indicates there were more positive tweets about Clinton than Trump across all states when you account for volume of tweets for each candidate. As for positive sentiment alone, Clinton has more positive tweets in AL, CO, NC, NV, TN, and WV, granting her 55 (18.97%) electors and leaving Trump with the remaining 235 (81.03%) electors. These results do not mirror election results as two of the three approaches result in a candidate winning every state. The final approach awards Clinton only five states; of which, three are Trump favored and two are swing states. These results are mirrored when the data is normalized, shown in Table 2. This indicates volume, positive sentiment, and the percentage of positive sentiment for each candidate are not a good model for predicting election results when using our data and a winner-take-all approach.

The shared elector count approach requires normalizing the data by state elector count and allowing for both candidates to take a percentage based on their total tweets within each state. Table 2 depicts the normalized values for the electors from each state based on volume and positive sentiment. These results indicate when using volume, out of the 290 electors allotted across the states, Trump receives 153.51 (52.93%) electors and Clinton receives 136.49 (47.07%) electors and while using sentiment Trump receives 145.65 (50.22%) electors and Clinton receives 144.35 (49.78%) electors. If we extrapolate this to total electors, Trump gets 284.76 and Clinton gets 253.23 electors when using volume, and 270.18 and 267.82 when using sentiment, respectively. It should be noted, sentiment predicts 7 of the 11 swing states correctly, regardless of approach.

A Wilcoxon-Mann-Whitney test (Mann and Whitney 1947) was performed to compare volume and sentiment, across swing states and favored states. We elected to perform this test as our data does not follow a normal distribution. Table 4 displays the results of the test. We find that neither factor was statistically significant, indicating neither volume nor sentiment can be considered an accurate predictor of election outcome per state with our data. Even so, the trend in sentiment values mirrors the close race that was the 2016 election in swing states. During the election, swing states were won by very small margins, for example MI was won by only 0.23%, and all swing states exhibit this small

margin between both candidates using sentiment.

Factor	Approach	W	p-value
Swing	volume	71	.5190
Favored	volume	57	.6305
Swing	sentiment	64	.8469
Favored	sentiment	51	.9705

Table 4: Wilcoxon-Mann-Whitney test results

Conclusion

In this study, we explore the effectiveness of using location-based tweets for determining the results of the 2016 presidential election. Three million tweets were collected from September 22nd to November 8th related to either Donald Trump or Hillary Clinton. Twenty-one states were chosen for further analysis, 11 swing states and 10 favored states. Tweet volume and positive sentiment were compared to the general election results to determine the effectiveness of using Twitter to predict state-level results. Our data was labeled using a deep convolutional neural network trained on the sentiment140 dataset.

When using the winner-take-all approach, volume and percentage of positive sentiment per candidate both grant unanimous victories for Trump and Clinton, respectively. When using positive sentiment, Clinton only obtains 18.97% of the electors, indicating none of these approaches to be a good model for predicting election results. When applying the shared elector count method, volume and sentiment both produce values similar to election results. Our approach using sentiment as a predictor of election outcome produces values closer to the election results for the 11 swing states chosen, but neither volume nor sentiment depicts the overwhelming advantage present in favored states. Regardless of approach, sentiment in tweets predict seven of the 11 swing states correctly. Tweets do not seem to be a viable predictor for performance in states where there is a clear advantage for a candidate with our approach. This may be due to demographics, as Twitter users may not be a representational sample of the population. In addition, the 2016 election was found to have Twitter bots present, which artificially increased the number of tweets for both candidates (Howard, Kollanyi, and Woolley 2016). Furthermore, the differences in distribution between tweets originating in swing states and favored states, with either volume or sentiment, is not significant when using our data and labeling process. Thus, neither volume nor sentiment is considered an accurate predictor of election outcome per state with our collection of tweets and approaches.

An avenue for future work could be bot detection, as the 2016 election is believed to have been influenced by social bots (Howard, Kollanyi, and Woolley 2016). Another option would be to alternate our approach by incorporating more states, changing our labeling method to get a fuller picture of the election, or limiting tweets to one per account.

Acknowledgement

We acknowledge partial support by the NSF (CNS-1427536). Findings in this paper are the authors' and do not reflect the views of the NSF.

References

Anuta, D.; Churchin, J.; and Luo, J. 2017. Election Bias: Comparing Polls and Twitter in the 2016 US Election. *arXiv preprint arXiv:1701.06232*.

Bermingham, A., and Smeaton, A. F. 2011. On using Twitter to monitor political sentiment and predict election results. In *Sentiment Analysis where AI meets Psychology (SAAIP) Workshop at the International Joint Conference for Natural Language Processing (IJCNLP)*.

Boutet, A.; Kim, H.; and Yoneki, E. 2012. What's in your tweets? I know who you supported in the UK 2010 general election. In *The International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Go, A.; Bhayani, R.; and Huang, L. 2009. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford* 1–12.

Heredia, B.; Khoshgoftaar, T. M.; Prusa, J. D.; and Crawford, M. 2016. Cross-domain sentiment analysis: An empirical investigation. In *2016 IEEE International Conference on Information Reuse and Integration (IRI)*, 160–165.

Heredia, B.; Khoshgoftaar, T. M.; Prusa, J. D.; and Crawford, M. 2017. Improving detection of untrustworthy online reviews using ensemble learners combined with feature selection. *Social Network Analysis and Mining* 7(1):37.

Heredia, B.; Prusa, J. D.; and Khoshgoftaar, T. M. 2017. Exploring the effectiveness of twitter at polling the united states 2016 presidential election. In *Collaboration and Internet Computing (CIC), 2017 IEEE 3rd International Conference*, 283–290.

Howard, P. N.; Kollanyi, B.; and Woolley, S. 2016. Bots and automation over twitter during the us election. *Computational Propaganda Project: Working Paper Series*.

Mann, H. B., and Whitney, D. R. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* 50–60.

Meador, C., and Gluck, J. 2009. Analyzing the relationship between tweets, box-office performance and stocks. *Methods*.

Prusa, J. D., and Khoshgoftaar, T. M. 2017. Improving deep neural network design with new text data representations. *Journal of Big Data* 4(1):7.

RCP. 2016. Real clear politics - election 2016 presidential polls. Accessed: 08-2017.

Tumasjan, A.; Sprenger, T. O.; Sandner, P. G.; and Welpe, I. M. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM* 10(1):178–185.

Zhang, X., and LeCun, Y. 2015. Text understanding from scratch. Technical report, Cornell University.