**Twitter and the Kaepernick Ad Campaign’s Impact on Nike**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cartney Thompson**  School of Information Studies  Syracuse University  cdthomps@syr.edu | |  |  | |
|  | |  | | |

**Abstract**

***The purpose of this project was to distinguish common topics on Twitter related to Nike before and after the Colin Kaepernick Nike campaign and their effect on Nike brand sentiment. Using data from Twitter about Nike, a week’s worth of tweets were collected, scored for sentiment about Nike, and modeled into topics using Latent Dirichlet Allocation (LDA). Lastly, the topics and sentiment were compared pre vs. post start of Nike Kaepernick campaign. Nike’s brand sentiment on Twitter changed after the Kaepernick campaign, resulting in an increase in the number of negative to positive tweets.***

**1 Introduction**

Sentiment is defined as “a view of or attitude toward a situation or event; an opinion.”[[1]](#footnote-1) The Nike campaign featuring Colin Kaepernick certainly includes a lot of varying opinions about Nike as a brand as a result of including Kaepernick in their campaign. Colin Kaepernick, a former quarterback for the San Francisco 49ers of the National Football League, garnered national attention for knelling during the national anthem before games as a protest against police brutality within the African American community. His demonstration lasted from August 2016 up until his last game in the NFL, January 2017. His action of kneeling during the national anthem, remains a topic in political discourse to this day.

Twitter has become a resource for brands to reach a large audience quickly. With more than 326 million monthly active users as of October 2018[[2]](#footnote-2), brands flock to Twitter to take advantage of this reach. For the Kaepernick campaign, Nike released one Tweet to kick-off the campaign (see Figure 1).



*Figure 1: Image of Nike Kaepernick campaign on Twitter*

Even before the launch of the Nike Kaepernick campaign, the sheer mention of the name Colin Kaepernick led to divisive discussions. According to an article from the Atlanta Journal Constitution, Kaepernick was one of the most controversial figures in 2016 and 2017.[[3]](#footnote-3)

In this paper, I examine Nike’s brand sentiment on Twitter prior to the launch of the Nike Kaepernick campaign and after the launch of the Kaepernick campaign and compare Nike’s brand sentiment after the campaign. I also explore key topics that were part of the discussion about Nike during these two time periods.

**Method**

**2.1 Data Acquisition**

Data was acquired utilizing twitterscraper, a python script to scrape tweets from Twitter based on specified criteria. [[4]](#footnote-4) A benefit of using twitterscraper script is by default, retweets are excluded. For our analysis, we are less interested about the number of retweets a tweet receives but the content of the original tweet. The criteria in retrieving tweets included the following:

* Randomly sampled Tweets that included “Nike” or JustDoIt in the tweet. This allows to only focus on tweets that were about Nike specifically. While I will acknowledge some tweets about specific brands do not include the brand name in the tweet, for the purpose of this analysis we will exclude those tweets.
* Collection of tweets before launch of Kaepernick campaign: August 21st, 2018 – August 27th, 2018. This was a random week before the launch of the Kaepernick that was long enough before the campaign launch, and close enough to get an understanding of brand sentiment about Nike before start of Kaepernick campaign, In total this data set returned 22,964 tweets.
* Collection of tweets after launch of Kaepernick campaign: September 4th, 2018 – September 10th, 2018. This data set returned 268,716 tweets.

The data set included several fields. For the analysis, we had to remove several variables from the data set that were not useful. These fields included html and url links, number of replies, retweets, and likes, user publicly available twitter names. After removal of these fields, we were left with two data sets: a collection of tweets before start of Kaepernick campaign and another collection of tweets post start of Kaepernick campaign.

Additional processing was required on both the before and after tweets. Any https, urls, and picture links contained within the actual tweets were removed. Tweets that contained links and were identified as being duplicative in its text to another tweet, were identified as a spam tweet and removed.

**2.2 Determining Sentiment**

Sentiment was manually determined for each tweet. In order to do this, a sample of our roughly 1,000 tweets each from before and after data sets were used for training and testing a model to determine sentiment. Sentiment, for this analysis, was positive, negative, or neutral about the Nike brand in a tweet. Table 1 illustrates how sentiment about Nike was manually determined.

|  |  |
| --- | --- |
| **Tweet** | **Nike Sentiment** |
| Whoa, slow down ma'am. Can we knock this Serena thing out first? We don't have the capacity to focus on more than one thing at a time. | Neutral |
| Seriously, in past days I was considering to buy a pair of Adidas that I really like, but no no. It’s now my commitment to add a pair of swoosh’s to my stock everytime Ican.#JustDoTheRightThing#JustDolt | Positive |
| Right there with Krista.  Nike clearly stands out as unAmerican now. | Negative |

*Table 1: Example neutral, positive, and negative tweets*

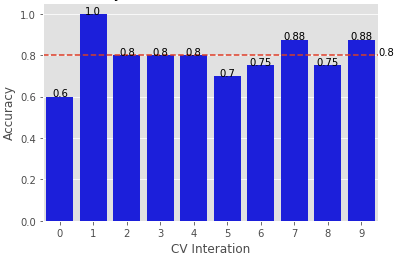
In the first tweet, Nike as a brand is not positioned as positive or negative, therefore the tweet is determined as neutral. In the second tweet, the tweeter writes their approval of Nike as a brand, even suggesting a preference for Nike over its competitor Adidas. In the third example, the tweeter states that Nike is un-American now, a definite negative portrayal.

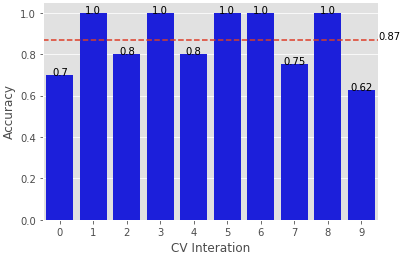
**2.2.1 Vectorization**

The vectorization method chosen for the data set was simple count of words. This vectorization method was used both for the model to determine sentiment and for developing the topics. When determining sentiment, the count of words method performed significantly better (in accuracy) compared to any sort of weighting. Also, for sentiment only, a boolean method of counting (if the word was in a document then a 1, 0 if not) was used for vectorization since tweets contain few words and a raw count would not be best for building a model. This seemed to help with model accuracy. An n-gram range of 1-3 was chosen since tweets are so short (average tweet in this data set contained around six words). Lastly, stop words were removed and all words were made lower case.

**2.2.3 Classifier Selection**

The classifier chosen for sentiment analysis was linear SVM. When compared to Naïve Bayes, accuracy was much higher utilizing SVM to determine sentiment (see Figure 2). Validation of algorithm was determined using 10-fold cross validation (average accuracy of 80% for Naïve Bayes and 87% for SVM). Since accuracy was higher for SVM, the SVM algorithm was used to classify sentiment for the remaining tweets. Tuning the classifier involved selecting the optimal values for C. Changing the value for C resulted in a small drop in accuracy the higher the value. The SVM algorithm performed best when the value of C was tuned to a value of 0.2. Any C value above 0.5 oddly resulted in pretty big drop off in model accuracy.



**** *Figure 2: Multinomial Naïve Bayes (top) versus Support Vector Machines (bottom) sentiment classification*

**2.3 Key Topics**

To identify key topics before and after launch of Nike Kaepernick campaign, I utilized Latent Dirichlet Allocation or LDA for topic modeling. Topic modeling on the tweets allows us to understand what are key themes/topics highlighted by the tweets.

In order to determine the optimal number of topics for both before and after tweets, grid search method tis used to perform this task. Grid search for topic modeling is used to iterate through different settings to determine which setting maximizes log likelihood estimation. The settings we will iterate through is the number of topics. For simplicity and to avoid having to many topics, the number of topics will iterate between 3 and 15.

**3 Results**

**3.1 Word Analysis**

When analyzing top words prior to the launch of the Kaepernick campaign (excluding Nike and Justdoit as they appear in almost every tweet), the most common words are displayed in a word cloud in Figure 3.



*Figure 3: Word cloud of top words associated with Nike on Twitter prior to Kaepernick campaign.*

When analyzing words after the launch of the Kaepernick campaign, the top terms are different with the exception of the word “one” as illustrated in figure 4. Interestingly, the term “one” before the Kaepernick campaign was often associated with best, greatest and athletes, signaling that users were often tweeting about Serena Williams.



*Figure 4: Word cloud of top words associated with Nike on Twitter after launch of Kaepernick campaign.*

In contrast, the word “one” after the Kaepernick campaign was a top word as well, although in a different sense. One was often associated with the following words: “day”, “another”, and “best”.

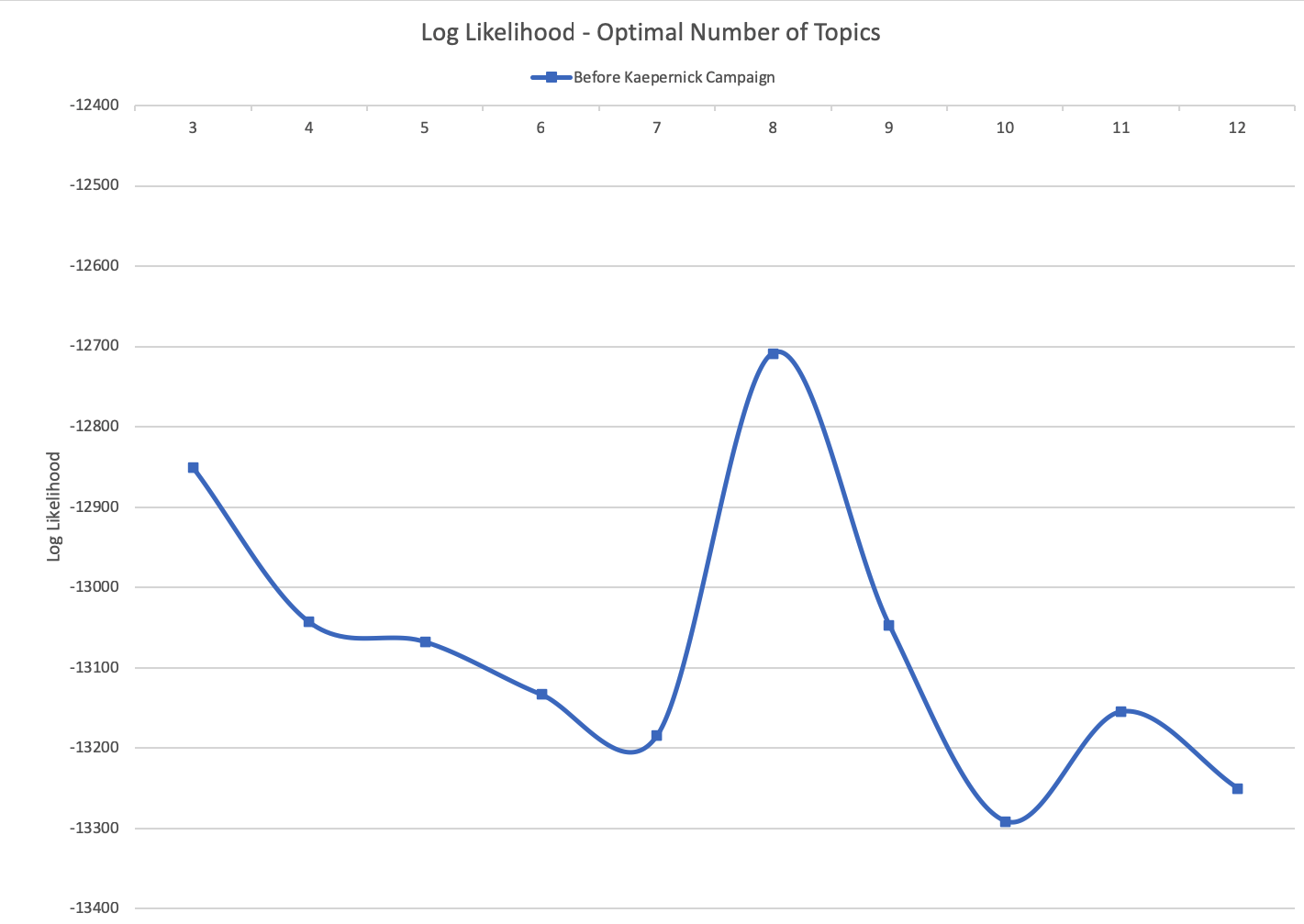
**3.2 Sentiment**

Examining the sentiment as shown in Figure 5, the percentage of tweets prior to launch of campaign had a much higher percentage of positive to negative tweets. This number was around 4.08 positive tweets to negative tweet. This ratio may have been influenced by the positive Serena Williams Nike campaign that had been running during this time frame. However, the number of postive to negative tweets is almost one to one (1.11) for the first week of the Colin Kaepernick Nike campaign.

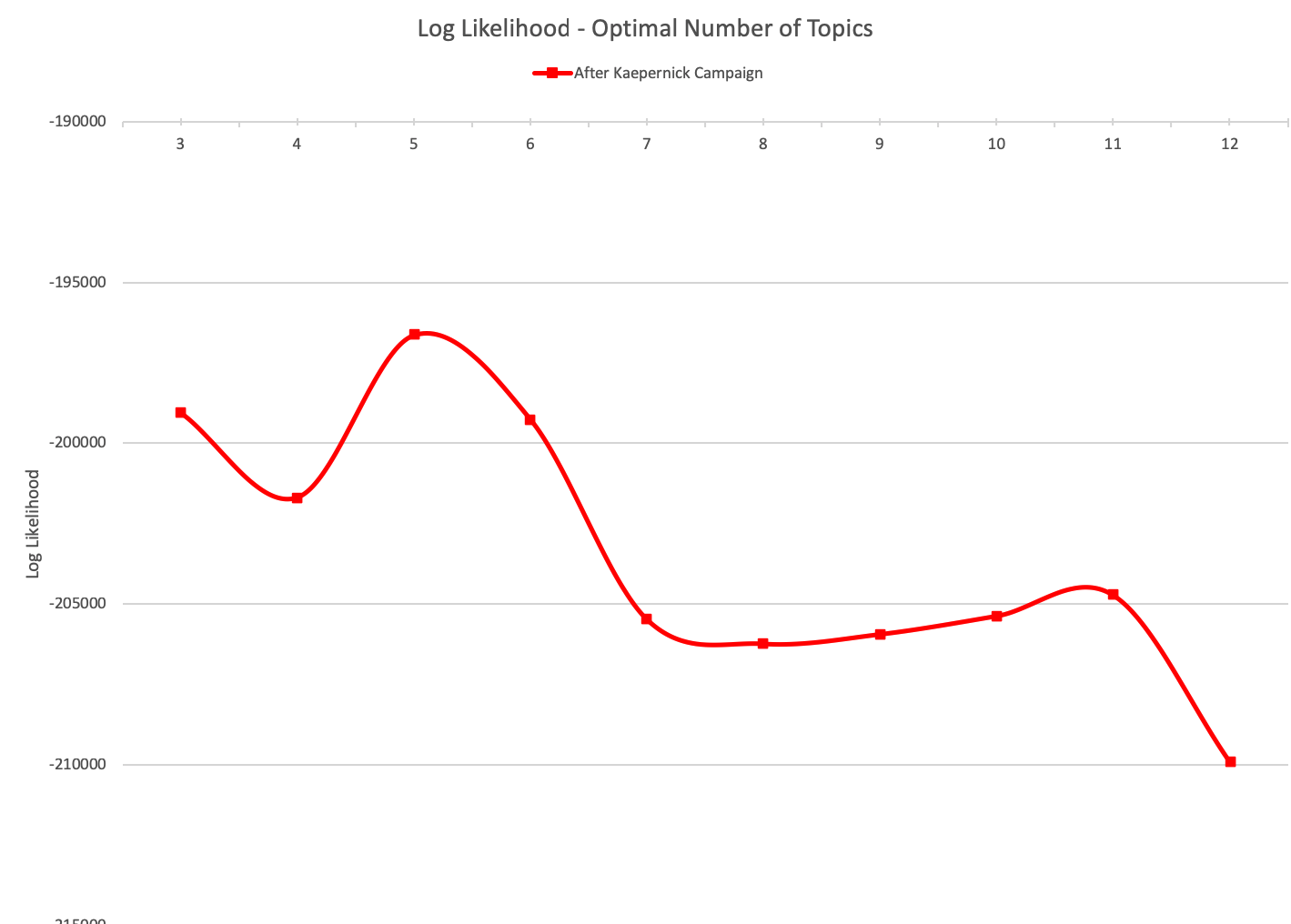
*Figure 5: Percentage of tweets by sentiment pre vs. post Kaepernick campaign launch.*

**3.2 Topic Models**

In order to pick the number of topics, grid search is performed. Grid search iterates through all LDA parameter combinations and selects the optimal number of topics based upon the highest log likelihood. Figures 6 and 7 show the optimal number of topics for before and after Kaepernick campaign respectively. For tweets before campaign launch, the optimal number of topics is 8. For tweets after Kaepernick campaign launch, optimal number of tweets is 5.



*Figure 6: Optimal number of topics for Nike tweets before Kaepernick campaign. Optimal number is 8.*



*Figure 7: Optimal number of topics for Nike tweets after Kaepernick campaign launch. Optimal number is 5.*

Running our LDA model first on the tweets before the Kaepernick campaign launch, there are a total of 8 different topics. Table 2 lists those topics, the key words based on counts, and the number of tweets associated with those topics.

|  |  |  |  |
| --- | --- | --- | --- |
| Topic Number | Topic | Key words | Number of Tweets |
| 0 | Air Jordan Shoes | air, jordan, know, white | 11,206 |
| 1 | A New Era for Nike | good, new, time | 1,390 |
| 2 | New Air Max Shoes Running Shoes | air max, need, running, wear | 1,172 |
| 3 | Shoe Availability | black, check, size | 1,000 |
| 4 | What People will do for Air Max | air max, make, people | 726 |
| 5 | Serena Williams Ad Campaign | great, got, serena, tennis | 1,393 |
| 6 | Great Day for Women | great, day, like, women | 1,277 |

*Tables 2: List of topics before Kaepernick campaign*

Table 3 below lists topics after the start of the Kaepernick Nike campaign. There are clear differences in the topics of the tweets pre-campaign compared to post campaign. Before, tweets were mostly about the Serena Williams ad campaign or the new Air Jordan shoes recently released. After, mostly about Kaepernick, race, and protest/boycott Nike.

|  |  |  |  |
| --- | --- | --- | --- |
| Topic Number | Topic (Inferred) | Key words | Number of Tweets |
| 0 | Black People Buy Nike, Don't Support Trump | black, people, buy, right, don't, trump | 152,789 |
| 1 | Kaepernick Fans/Supporters | like, really, kaepernick | 22,781 |
| 2 | Nike Campaign is Throught Provoking | campaign, good, make, shoes, think, time, way | 29,729 |
| 3 | Nike Needs and Loves Money | believe, great, know, need, nfl, way, love, money | 15,487 |
| 4 | Protest Kaepernick and Boycott Nike | kepernick, new, protest, stand, support, boycott | 27,005 |

*Table 3: List of topics after launch of Kaepernick campaign*

**4 Conclusion**

Sentiment about Nike grew negative after the introduction of the Kaepernick campaign. While the ratio of positive tweets to total tweets was similar pre vs. post campaign launch, the number of negative tweets after Kaepernick campaign launched grew significantly. As a result of the growth of negative tweets post campaign, the topics associated with Nike changed from more product and positive focused, to injecting politics and ethnicity into the discourse.**5 References**

Sholom M Weiss, Nitin Indurkhya, and Tong Zhang. *Fundamentals of Predictive Text Mining*. New York, Springer

David Biel, Andrew Ng, and Michael Jordan. *Latent Dirichlet Allocation*

<http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>

1. <https://www.merriam-webster.com/dictionary/sentiment> [↑](#footnote-ref-1)
2. <https://www.fastcompany.com/90256723/twitters-q3-earnings-by-the-numbers> [↑](#footnote-ref-2)
3. <https://www.ajc.com/entertainment/2016-most-controversial-figures/NhLJfLYURpT7O9psqvbzwI/>

   <https://www.ajc.com/news/national/year-review-controversial-figures-from-2017/0uSy8WGJjONvbsixmxHIYI/> [↑](#footnote-ref-3)
4. <https://github.com/taspinar/twitterscraper> [↑](#footnote-ref-4)