

ANALYTICS SPECIALIZATIONS & APPLICATIONS, 2021/2022

Coursework II: Brand Analysis Using Twitter

University of Nottingham Business School

Student Number: 20325050

Course Code: BUS14370

Date: May 03, 2022.

1. Executive Summary

The world has experienced phenomenal growth in digital media platforms due to increasing penetration of social media in people's lives. Information on platforms such as Twitter is disseminated all over the world in seconds. Users of any social media platform have power to share their thoughts and information with their followers from any corner of the world. Twitter has been in limelight recently as Billionaire Elon Mask is trying to purchase the social media site for its potential in business and marketing. Given the importance of the microblogging site, this study has chosen blogs related to the world 'Nike' to see what people think about the sports and fashion brand. The study is done on almost 5000 tweets from people all around the globe. The purpose is to find out sentiments about the brand. The same analysis is done for Nike's close competitor, Adidas, to see what people think about these two brands in the same time period, i.e., from 27th April to 29th April 2022. Not surprisingly, they both compete neck to neck with same percentage of positive sentiments on twitter. However, Nike has slightly more negative sentiments. Most common words in Nike's positive tweets contain its three most famous shoes: Dunk, Air and Kobe. Adidas does not have its shoe articles in its most positive tweets. However, in the negative sentiments, late delivery and size are common words for both brands. In neutral tweets, Nike's collaboration with famous Tennis star Serena Williams - to lunch people of color is fashion industry - is dominating. Afterwards, the study has found potential micro-influencers based on their followers count, total tweets, and location. Nicole Behman from USA is selected to be a micro-influencer biggest market of Nike. Nevertheless, this study is conducted for just three days data. The company should analyze tweets for at least a year to check public perception in a comprehensive way.

2. Methodology

Tweepy, a python library for extracts data, is used to collect blogs for word Nike. All the tweets are in English as Tweepy api cursor allows us to extract data in the chosen language. Authentication is handled through Tweepy. Then, National Language Processing (NLP) tools such as TextBlob and NLTK are used to pre-process the text data. These are pre-processing stages before sentiment analysis.

- I. Cleaning: Tweets have a lot of unwanted characters. Therefore, extensive cleaning gets rid of punctuation and other undesired symbols such as hashtags, hyperlinks, and a like.
- II. Tokenization: sentences are broken into words using this tokenization method.
- III. Stop words removal: these words – I, we, they, us, about, with , for, etc. - do not hold much value individually. So, stop words are removed.
- IV. Stemming: it involves reducing the words to their stem(basic) form to make it simpler to do text analytics. For example, playing is reduced to play.
- V. Lemmatization: It is an advanced form of stemming and reduce the word to its dictionary form.

Furthermore, Textblob's polarity score is applied to the processed data to separate good, bad, and neutral tweets. Akash explains polarity in his articles: "The polarity and subjectivity of a statement are returned by TextBlob. The range of polarity is [-1,1], with -1 indicating a negative sentiment and 1 indicating a positive sentiment." (Akash, 2021) Whereas, zero show a neutral sentiment. The data of separated sentiments gives us insights about the topics in the form of wordcloud. All of this is done after removing duplicate rows. More repeated words in each of the three segment gives the company insights about the good or bad perception in the public.

Machine learning is also applied on the text. Text is converted into number using counter vectorizer method. Then three classification models are used to predict the classification of tweets.

3. Data Collection Section

Twitter allows to extract extensive data. For this study, 5435 unique tweets for the “Nike” were downloaded on 28th and 29th April 2022. Each of these tweets has an author_id with name and screen name. Other information for each tweet includes verification status, description, location, followers count of the account, listed count, favorite count, statuses count, language, date of creation, the text of the tweet, re-tweet count, get hashtags, source, and favorite count. There are 4162 unique user accounts tweeting about Nike in the given period. Most of the tweets were generated from the USA in general, and California in particular. Surprisingly, Kenya’s capital, Nairobi, came third location with the greatest number of tweets. Kent_nike is a twitter account which posted the most – 60 times- about the word Nike. However, this is an unverified account with no location and just 1000 followers. Thus, this account with most tweets is insignificant. After processing data, Polarity column is added into our data frame to do sentiment analysis. This column helps in sifting positive, negative, and neutral sentiments. Table 1 shows data comparison. Noticeably, accounts with more than 100,000 followers are talking more about Nike than Adidas.

Brand or Name Searched	Data range	No. of Tweets	No. of unique accounts/users twitting	Twitter Accounts with > 100,000 followers commenting on the brand	Location with most tweets
Nike	28 – 29 April 2022	5435	4162	84	United States
Adidas	27 & 29 April, 2022	5181	4219	54	United States

Table 1: A comparison of data used for this study

4. Analysis and Description of Results

4.1. Brand Exploration

I. Distribution of positive and negative tweets.

After cleaning and pre-processing data, sentimental analysis for both the shows the following results.

Brand Name searched	Sentiment Found			Text Words Per Sentiment(avg)		
	Positive	Negative	Neutral	Positive	Negative	Neutral
Nike	893	336	1595	10.44	10.1	8.40
Adidas	832	295	1505	9.85	10.48	8.8

Table 2

Table 2 shows that the two competitors are contending on twitter as well. Since the data has almost equal number of tweets, their sentiments are also neck to neck. Both have almost similar number of positive tweets. However, as shown in figure 1 and 2, Nike has slightly more negative sentiment than Adidas. Nevertheless, they both have same positivity on the social media platform. Twitter accounts are slightly more neutral about Adidas than Nike.

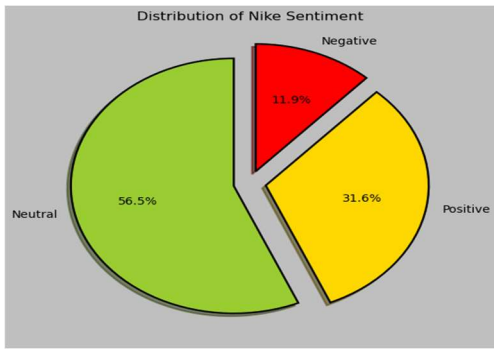


Figure 1

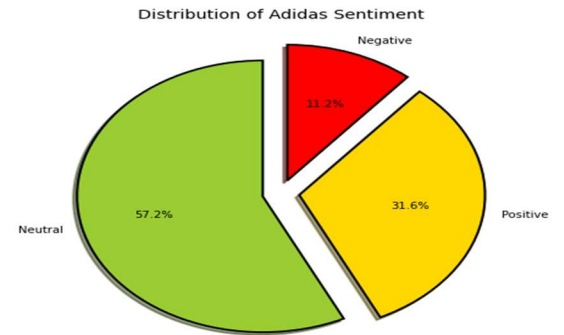


Figure 2

II. Topic Detection

a. Word cloud of positive sentiment



Figure 3: Nike's words in positive tweets



Figure 4: Adidas words in positive tweets

b. Word cloud of Negative Sentiment

[illegible]

Figure 5: Nike's Negative Sentiment

c. Word cloud of Neutral Tweets

Similarly, Adidas Yeezy is the topic of Adidas neutral comments. It is significant for the fact that the German company has partnered with Kanye West, an American designer and rapper, to launch a high-end sneaker lineup. (Hanbury, 2018) Presence of these public figures on twitter helps the brands in reaching out to a whole range

[illegible]

Figure 6: Nike's Neutral sentiment

III. Machine Learning Models:

After analyzing corpus approach, it is important to apply machine learning models on our processed data. Text is converted into number for machine learning using CountVectorizer method. As a supervised learning task to predict the sentiment of the text, three classification models are used: Logistic Regression, Support Vector Regressor and Random Forest. Logistic regression is used after hyperparameter tuning. As shows in below figure 7, Random Forest is best with the accuracy score of 90 percent.

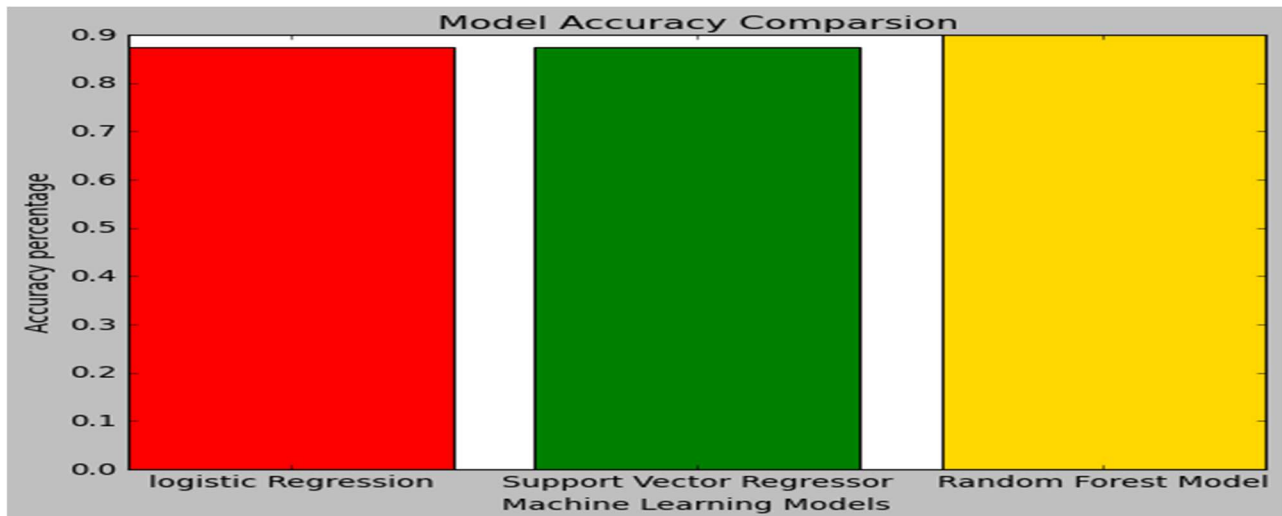


Figure 7: ML models

4.2. Micro-influencer

Park and Lee describe micro-influence as someone with followers between 10,000 and 100,000 in their article in the Journal of Marketing. They think “micro-influencers are more persuasive than mega-influencers because endorsements by micro-influencers (versus mega-influencers) bestow higher perceptions of authenticity on the endorsed brand, which “rubs off” from the perceptions regarding influencer authenticity.” (Park, et al., 2021) Not only authentic, but micro-influencers are also cost efficient as well. Kim says that their “Retweeting brand posts to their followers makes it possible to exchange information about brands more quickly and easily.” (Kim, et al., 2014)

Even though Nike is a big brand with global reach, it cannot neglect the efficacy of micro-influencers. Given its outreach, the sports company can find a plethora of micro-influencers in different locations. Even though there were 36 accounts who positively tweeted about Nike and had followers in the range of 10,000 to 100,000, top 5 are mentioned below.

Author name	Screen name	Follower’s count	Location
MoreSneakers.com	more_sneakers	84713	Paris, France / Worldwide
SneakerAlert	SneakerAlertHD	77096	United States
nic b	NicoleBehnam	71863	United States
The Drop Date	Thedropdate	62523	United Kingdom
Kicks Deals Canada	KicksDealsCA	48037	CANADA

These micro-influencers have presence in four countries. As Nike is a big company, it can afford to choose more influencers in big metropolitan cities. Choosing one is difficult, but Nicole Behnam with more than 72000 followers seems more appropriate because of her fans and her recognition as an individual. Besides, as United States is the biggest economy with highest consumer power and a population of 330 million people, It's the most appealing location for hiring Nicole Behnam as a micro-influencer. Also, she has a verified account with an aim to advise people. Interestingly, she has tweeted 8157 times with her account; this shows her active participation on twitter. Lastly, her last tweet about Nike in our data base says: "@HighwaterKyle I love Nike. Just an example."



Figure 8

5. Conclusion, Insights and Recommendation:

In short, the snapshot study of Nike using microblading site is revealing. It examines the current content of the brand in the social media. Inferred information is meaningful for improving both conventional and social media strategy of the brand. Nike's famous articles such as Air max were trending in the specific period. Those tweeting from different locations give the company immediate feedback on their experience. Overall, the sentiment is neutral. The company shall use the valuable lesson from current topics to improve its brand awareness and engagement. Moreover, it should take care of delivery system as free and late shipping are also being discussed actively on twitter.

Furthermore, it is observed from this study that the analysis of three days tweets is not enough to shape the long-term strategy of the company. Microblogs of at least one year should be studied to seek more meaningful trends all over the year or seasons. Tweets for occasions such as Champions League and Football World cup will have data representing people's emotions with football and the brand Nike. Same can go for Christmas's shopping spree. Thus, the company should analyze tweets for a long period to get seasonally inclusive sentiments.

References

- Akash, 2021. *Making Natural Language Processing easy with TextBlob*. [Online]
Available at: <https://www.analyticsvidhya.com/blog/2021/10/making-natural-language-processing-easy-with-textblob/>
[Accessed 30 04 2022].
- Alcorn, C., 2021. *CNN Business*. [Online]
Available at: <https://edition.cnn.com/2021/08/17/business/nike-serena-williams-swdc/index.html>
[Accessed 01 05 2022].
- Elsevier, 2014. Kim, Eunice; Sung, Yongjun; Kang, Hamsu. *Computers in Human Behavior*, p. 19.
- Hanbury, M., 2018. *Business Insider*. [Online]
Available at: <https://www.businessinsider.com/how-kanye-west-made-yeezy-brand-a-success-2018-4?r=US&IR=T>
[Accessed 01 05 2022].
- Kim, E., Sung, Y. & Kang, H., 2014. Brand followers' retweeting behavior on Twitter: How brand. *Computers in Human Behaviour*, p. 25.
- Park, J., Lee, J. M. & Seo, Y., 2021. David and Goliath: When and Why Micro-Influencers Are More Persuasive. *Journal Of Advertising*, Volume 50, p. 584.