

Sentiment Analysis: A Case Study for Apparel Brands - FABINDIA v/s BIBA

Syed Aqsa Ahmed,

School of Computer Science & Engineering, VIT University, Bhopal, India
Corresponding author e-mail: [syed.aqsa2021\[at\]vitbhopal.ac.in](mailto:syed.aqsa2021[at]vitbhopal.ac.in)

Abstract: Social media is now providing a platform for users to express their opinions, where users discuss various events, services, and brands. Company officials are in continuous need of the feedback to improve the quality and quantity about their services. But, due to the bulk amount of data, it is difficult to detect the consumer's opinions about their services. This paper illustrates the problems about the data for the sentiment analysis. In addition to that, it uses text mining and NLP on the data through the machine learning techniques. My study is to find the public opinion about the top two apparel Indian brands and compare the positive and negative attitude and sentiments of common users about each brand. We found that positive reviews of Fabindia are more than the Biba brand and there is slight difference in negative reviews.

Keywords- Sentiment Analysis, NLP, Apparel brands, Brand Reputation, Machine learning, Text Mining

1. Introduction

In the current scenario, the world could be a world because of the web availability. 55% of the world population is using the Internet. Now a day's social media has been a part of people's day to day life and a good platform to share useful information efficiently and effectively. The statistics shows that 71% of the web has been used through social media by the customers. Studies show that over half customers like to read other's comments on the products before purchasing it. Hence, public opinions are the simplest source of feedback for businesses about their products and services which enables them to increase the standard factor. The social media provides the important chances to the brands i.e., marketing, selling, branding, and promotional chances. It's basically a study of people's thoughts and opinions which might be positive or negative about any product or event through the linguistic communication processing called as 'sentiment analysis'. Sentiment analysis is correlated with text mining and data processing. The importance of sentiment analysis is to assure the polarity of language by performing machine learning techniques. The purpose of my research is to use the sentiment classification methodology to the apparel datasets to find the positives and negatives of FabIndia and Biba brand. This research will illustrate the link between the consumers and also the apparel Brands. In my research, pre-processing step has been taken to perceive the high analysis results. This comparative study will assist the new researcher to analyse the social media for the sentiment detection. The main contribution of this paper:

- To conduct a comparative analysis of two Indian apparel brands.
- To find the solution, which apparel brand is most well-liked among the consumers at the bottom of sentiment analysis?

The rest of the paper is: section-2 which explains the case study in which we discuss the pre-processing of data and the implementation of the sentiment techniques at my proposed case study, section-3 gives us the experimental results of our case study. Finally, section-4 clarify conclusion.

2. Case Study

2.1 Data Collection

I used Tweepy a streaming API to retrieve the data. We retrieved 48,765 tweets by using the apparel brand's name as a keyword including "FabIndia" and "BIBA". This data is fetched in irregular dates of 2019. Table 1 is representing the total number of tweets for apparel brands. Sentiment analysis over the Twitter data can be a very challenging task because of tweet's character limitation. Such issues can impact on the lexicon search. We resolved it in the pre-processing step to form machine learning algorithms perform better.

Table 1: The total number of tweets for brand

Brand Name	No. of tweets
FabIndia	24587
Biba	24178
Total	48765

2.2. Pre-Processing:

As our dataset have 48,765 tweets but after removal of the non-English tweets, the remaining dataset contains 28,146 tweets. Then, after removing the duplicate tweets (retweets); the total remaining tweets are 23,486.

2.3 Sentiment Analysis:

Then, instead of manually tagging the review texts in the dataset, the sentiment analyzer of NLTK was used to automate the process. We set the threshold of rating 3. i.e., if a review rating is greater than or equal to 3, it is considered as positive feedback, otherwise it is considered negative feedback. In the mentioned manual, there is a flaw in intuitive tagging of not considering some neutral sentiments. Hence, we use the sentiment analyzer by NLTK.

2.4 Polarity

Using TEXTBLOB, we made a Blob list which determines the sentiment and polarity behind a particular review. Then we convert the decimal score to categorial variable

- 1) If the polarity score > 0, then sentiment is positive
- 2) If the polarity score =0, then the sentiment is neutral
- 3) If the polarity score < 0, then sentiment is negative.

Lastly, we validate the model for the extraction of the polarity percentage for three classes: positive, neutral, and negative to evaluate the results.

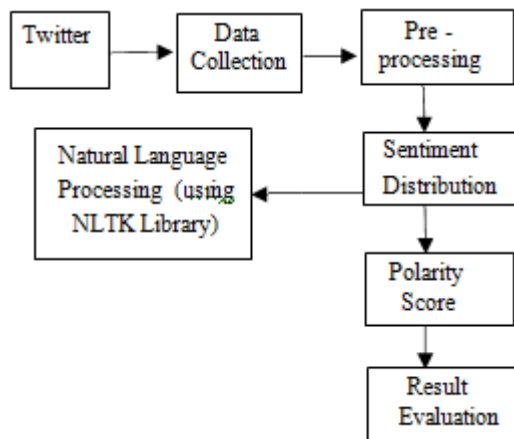


Figure 1: Process Diagram on Data

3. Results

After the implementation of sentiment classification, we got the values of sentiment distribution for FabIndia and Biba.

Figure 2 display the sentiment distribution of FabIndia and Biba. FabIndia sentiment distribution shows the positive, negative, and neutral views (41.6%), (2.5%), and (8.8%) respectively. Similarly, the distribution of Biba sentiment with positive (37.1%), negative (2.3%) and neutral (7.7%). It was found that positive reviews of FabIndia are more than the Biba. While the neutral values record the satisfaction level among the users for both brands which is less than 50% of total reviews.

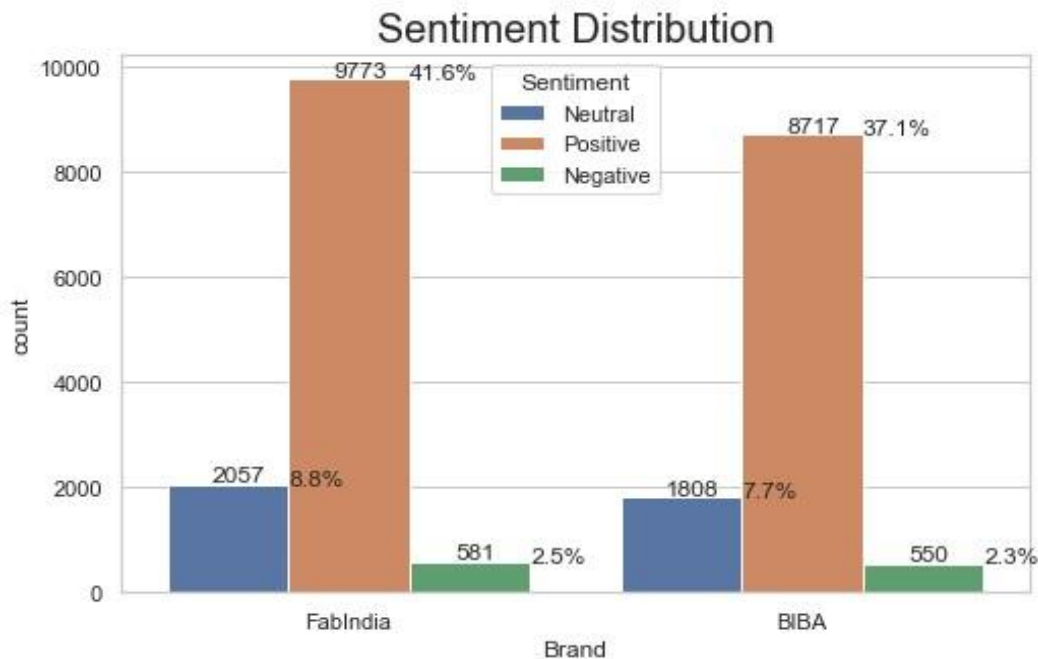


Figure 2: Sentiment Distribution of Brands

It is observed that people usually discuss the other brands as well when they want to talk about any single brand product. So, it's concluded that online consumers always attempt to compare the identical product of various brands before making a procurement decision.

Below is the word cloud for positive keywords (as shown in figure 3). Similarly, the word cloud for negative keywords (as shown in figure 4).



Figure 3: Positive Word Cloud



Figure 4: Negative word cloud

Figure 5 shows total distributions of sentiments for our apparel dataset.

Positive reviews were found greater than the negative reviews in both brands. It's a representation of most popular brands. As compared to the positive number of comments, the neutral comments of FabIndia are better than the Biba. While the more negative review of FabIndia is a disappointment for its market value as compared to Biba.

Figure 6 is shown as a comparison of both brands for sentiment distributions about the online users. It is indicating the sentiment of each brand, FabIndia and Biba.

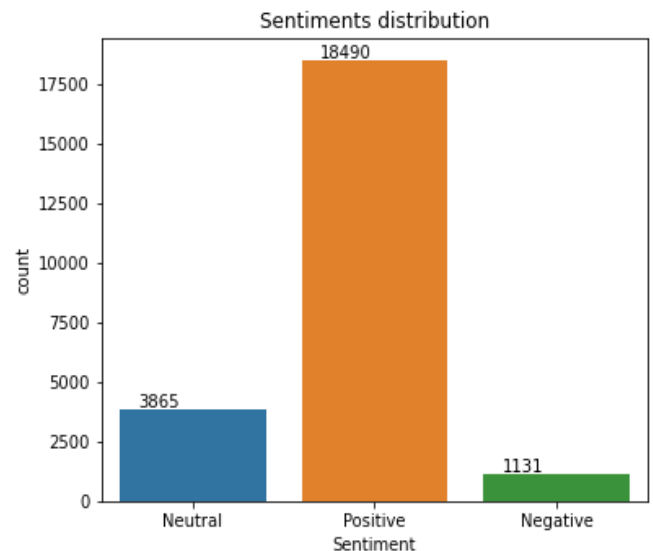


Figure 5: Sentiment Distribution of Total dataset

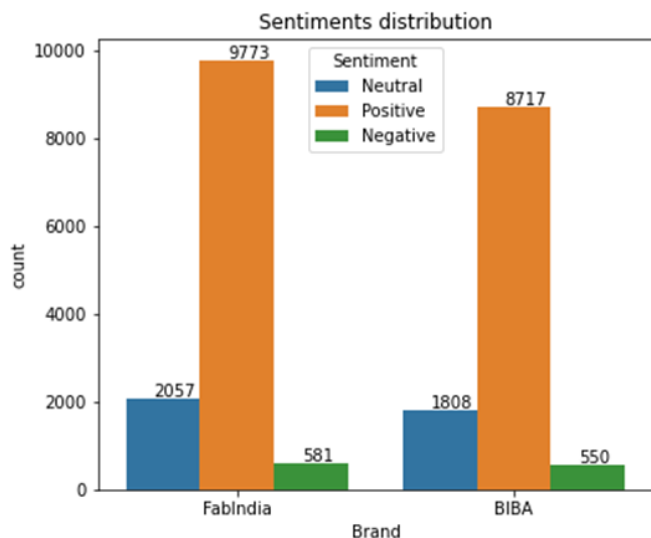


Figure 6: Sentiment distribution of each Brands.

4. Conclusion

Now a days, regular shopping is replaced by online shopping. An apparel shopping is very popular among the web users on social media. With the advancement in technologies, for customer satisfaction new methods are being introduced day by day by using the massive data, text mining and sentiment analysis. Before implementing the sentiment, we apply the pre-processing step at our data to create it much easier for our sentiment model. We conduct a comparative case study to research the sentiment for the posts about Apparel brands to seek out which brand is popular among FabIndia and Biba. We concluded from our experiments that FabIndia has more positive sentiment than Biba, while the negative polarity between them is slightly different.

5. Acknowledgments

I am very grateful and would like to appreciate our respected professor Puspinder Singh Patheja for his guideline and insightful feedbacks.

References

- [1] Abdur Rasool, "Twitter Sentiment Analysis: A Case Study for Apparel Brands" 2019
- [2] Luo and Tao, "Sentiment Analysis Based on the Domain Dictionary: A Case of Analyzing Online Apparel Reviews" 2018.
- [3] Statista.com, "number of social media users worldwide from 2010 to 2021 in Billions, Link: <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>.
- [4] DoubleClick_Inc. (2005). Understanding Buyer Search Activity as it Builds to Online Purchase. Retrieved from www.innovationmarketing.at/news/new_smodul/upload/429181378_searchpurchas_e_0502.pdf.
- [5] Hu, Bhargava, Fuhrmann, "Analyzing users' sentiment towards popular consumer industries and brands on Twitter"

- [6] Sharma and Alavi (2017), "Generating trust using Facebook-A study of 5 online apparel brands", Information Technology and Quantitative Management (ITQM2017)
- [7] Liu, B. (2012). "Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies", 5(1), 1-167. doi:10.2200/S00416ED1V01Y201204HLT016.
- [8] Kim JH, Kim M, Kandampully J. "Buying environment characteristics in the context of e-service. European Journal of Marketing". ResearchGate 2009 Sep 18;43(9/10):1188-204
- [9] Asghar MZ, Khan A, Ahmad H, Kundi FM, Ismail S (2017) "A rule-based sentiment classification framework for health reviews on mobile social media". Journal of Medical and Health Informatics, J Med Imaging Health Inf 7(6):1445-1453
- [10] Shukri, Yaghi, "Twitter Sentiment Analysis: A Case Study in the Automotive Industry", 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). 2015.
- [11] Geetika Gautam, "Sentiment Analysis of Twitter Data Using Machine Learning Approaches and Semantic Analysis". 7th International Conference on Contemporary Computing, 978-1-479951734/14/\$31.00 ©2014 IEEE
- [12] Jitendra Kumar, Sambit Bakshi, Karen L. Williams, "A model for sentiment and emotion analysis of unstructured social media text," Electron Commerce Research, (2018) 18:181-199,
- [13] Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. Processing 150(12):1-6.
- [14] R. Parikh and M. Movassate, "Sentiment Analysis of User-Generated Twitter Updates using Various Classification Techniques", CS224N Final Report, 2009.