ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

NLTK AND SUPERVISED LEARNING APPROACH TO ASPECT BASED SENTIMENT ANALYSIS OF GLOBAL BRANDS IN THE CONTEXT OF "ATMANIRBHAR BHARAT ABHIYAN"

Beena Kapadia[1] and Dr Amita Jain[2]

[1]Research Scholar, School of Computer Application and Technology, Career Point University, Kota, Rajasthan, India

ORCID ID: [0000-0002-5064-3480]

[2]Research Associate, School of Computer Application and Technology, Career Point University, Kota, Rajasthan, India

ORCID ID: [0000-0002-1118-3389]

ABSTRACT

To be self-reliant India, we first require improvement in the quality of some of our products. This paper focuses on the features or qualities of the global brand, which make it superior. If Indians can make the quality product with affordable price, then Indians may certainly buy Indian products rather than buying global brands. Also, the Indian product can become a brand and that brand can be exported. Many Indians prefers buying Indian brands for whatever product they buy. But still sometimes it becomes irresistible to buy global products. To find those factors or features we have used Aspect-based sentiment analysis in this paper. Rather than only finding the sentiment score using either rule based or based on any machine learning algorithm directly, this paper first attempts to detect the quality features what buyers find in the global product using aspect-based sentiment analysis. Along with those features or aspects, we find the sentimental value of each aspect. GitHub offers a huge repository of community published data. we have used the data collected from the website https://github.com/AFAgarap/ecommerce-reviews-analysis/blob/master/Womens%20Clothing %20E-Commerce%20Reviews.csv We have also used Natural Language Toolkit (NLTK) built-in library of Python to find frequency of words. We have also used supervised machine leaning algorithms like support vector machine classifier, multinomial naïve based classifier, Random Forest classifier and K nearest neighbor classifier. We have also worked on Hybrid Ensemble Learning Model, which gave us the accuracy .99 for all models, as it is not directly applied to the review text, but it was applied to the aspects. great, perfect, super, best, vibrant are the top five features of global brands considered by the respondents.

Keywords: Brands, Natural Language Toolkit, aspect-based sentiment analysis, features, supervised machine leaning, support vector machine classifier, multinomial naïve based classifier, Random Forest classifier and K nearest neighbor classifier, Hybrid Ensemble Learning Model.

1. INTRODUCTION

Brands can be helpful in strengthening a person's identity. Hence, consumers use brands as instruments for reflecting their personality and individual goals. The self-concept as a vital component of emotional aspects, boosts attitudinal loyalty towards a brand. (K. P., Zeynep, G. C., 2007) Vocal for Local is not just about buying the Indian brands only being an Indian as far as possible but also to bring the brand up to that quality, which we can export to other countries. To promote the idea to be Vocal for local, we should first understand about what is that factors or features, which make non-Indian brands so popular, even being it extremely costly sometimes and even the money we spend will be shared with non-Indians.

We have studied for the features of non-Indian products using aspect-based sentiment analysis, which make them superior over Indian products. We have collected the data from the website https://github.com/AFAgarap/ecommerce-reviews-analysis/blob/master/Womens%20Clothing %20E-Commerce%20Reviews.csv We performed aspect-based analysis on amazon data to find the attributes and their description. We then calculated the sentiment from each description using VADER sentiment analysis. As it is too difficult to include every description into consideration for counting each description's frequency, we have just taken top 30 description, whose sentiment is > 0.5.

2. RESEARCH OBJECTIVE

To find out features of non-Indian products which make them superior over Indian products.

3. LITERATURE REVIEW

In this paper, we have used NLTK for frequency counter, aspect-based sentiment analysis, word cloud and various Machine leaning algorithms like Support Vector Machine (SVM) algorithm, Multinomial Naïve Bayes algorithm and RandomForestClassifier.

A. NLTK for frequency counter:

NLTK is the platform for building Python programs working with natural language. It provides interfaces to over 50 corpora and lexical resources as well as a suite of text processing libraries for classification, parsing, semantic reasoning. The FreqDist() function of NLTK (nltk.org, 2023) to count the frequency of features. A frequency distribution records the number of times each outcome of an experiment has occurred.

B. Sentiment Analysis and its Types:

Sentiment Analysis is one of the most popular applications of Natural Language Processing. It is the process of identifying and categorizing opinions expressed in any text to determine whether the writer's attitude towards a particular topic, product, hotel, movie etc. is positive, negative, or neutral. The data for performing sentiment analysis can be taken from various resources like face book, twitter, article and hence it can be in different format like xml file, JSON file, csv file or simple text file. (K. Beena, 2021).

The analysis of the sentiment, or in general of the opinion, can be performed at any of the four different levels of detail, from the most generic to the most specific, such as document-level, sentence-level, aspect-level, and concept-level. (Hemmatian F.,2019) The document-level analysis (Pontiki M.,2016) is used to understand the polarity of a whole document product (Behdenna S.,2018), news article (Shirsat VS,,2017), a post, a tweet (Gurini D.,2013) reviews are few examples of document level sentiment analysis. Here, the information is quite general, as it finds the polarity of many sentences as a unique positive or negative. Lexicon-based approaches for document-level analysis typically identify the polarity of each term and then aggregate their polarity scores to classify the sentiment of the whole document.

Considering the level of importance of each word in this kind of approach could improve the accuracy of the classification. sentence-level SA (Appel O.,2016).

The aim is to understand if a sentence has a positive or negative polarity. Usually, a sentence level approach can be used also to analyze the sentiment of a complete document by analyzing its sentences and using an aggregation operator to combine the polarity of each sentence to obtain the global polarity of the document (Fang X, 2015). An analysis at the documents or sentence-level provides useful insights regarding the opinion of users about specific entities, but they do not allow to understand which elements, features, or aspects of such entities have influenced the (positive or negative) opinion. To perform such a kind of analysis, aspect-level sentiment analysis should be performed (Hu M, Liu B, 2004).

This level of analysis clearly provides a more accurate result if compared with document and sentence-level analysis. The idea of concept level sentiment analysis is to have machines able to deeply understand the natural language and in particular, to understand emotions, opinions, sentiments. Such kind of approaches are based on semantic analysis techniques able to identify and analyze the concepts in a text.

C. TextBlob Sentiment Analysis:

It is a simple python library -API used especially for sentiment analysis. Textblob sentiment analyzer returns the two properties for a particular given sentence: one of them is the Polarity, which is a float value between [-1,1], where -1 indicates negative sentiment and +1 indicates positive sentiments. Though VADER gives better outcome compared to TextBlob for KNN.(F. Fazrin,2022)

D. Spacy Sentiment Analysis:

spacytextblob is a pipeline component that enables sentiment analysis using the TextBlob (A. K. Singh, 2021)

E. VADER Sentiment Analysis

VADER is abbreviation for “Valence Aware Dictionary and sEntiment Reasoner” and is available under the MIT License. The VADER tool was released in 2014. It uses a lexicon driven approach and additional heuristics for rating the input. It offers consistent ratings and requires no training data, as VADER is not a machine learning approach. It achieved remarkable scores for multiple domains such as tweets, movie or product reviews. (Hutto, C.J.,2014).

VADER is a rule-based sentiment analysis tool to express the sentiments on a given text. ([Yang, S. E. (n.d.),2019] It is used to label the dataset into positive or negative sentiment score, based on whether its value exceeds 0.5 or not.

The compound score can be calculated as the sum of all lexicon ratings which are normalized ones. (NIPS-2019) VADER works better than TextBlob for the text, taken from either social media or any web sources. (V. Bonta, 2019)

F. Word Cloud:

The Goole Cloud Natural Language API (n.d.,2023) Retrieved from Natural Language AI: https://cloud.google.com/natural-language uses machine learning techniques to support different tasks of natural language analysis, among which entity recognition, content classification, and sentiment analysis. In particular, the library supports both sentiment analysis and aspect-based sentiment analysis, at a document or sentence level. The polarity of the sentiment is given in the range [–1; +1].

G. Support Vector Machine Algorithm:

Support Vectors are simply the coordinates of individual observation. The SVM classifier is a frontier which best segregates the two classes (hyper-plane/ line). In SVM, it is required to select the hyper-plane which segregates the two classes better. If we can’t find linear hyperplane between the two classes, then it is required to introduce additional feature. (S. Vanaja, , 2018)

H. Multinomial Naïve Bayes Algorithm:

Multinomial Naïve Bayes classifier works on the concept of term frequency (TF). TF means the number of times the word occurs in a document. Multinomial Naïve Bayes is used to find two facts that whether the word exists in a document or not as well as that words frequency in that document. (Tyagi, A., 2019).

I. Random Forest Algorithm:

Random Forest algorithm is a supervised machine learning algorithm used for classification. It works based on the concept of ensemble learning, in which number of decision trees get various subsets of the data set. Here, all decision trees predict the output for the new data and the final class of that new data is considered as most of the outcome predicted from all decision trees. (Sarraf, T., 2020)

J. K Nearest Neighbor Algorithm:

The K Nearest Neighbor algorithm takes all the given data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a category which is most suitable. That is, this algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much like the new data.

This algorithm first selects the number K of the neighbors (to get the exact k-value we need to test the model for each expected k-value.). Then, it calculates the minkowski distance of K number of neighbors. After that, it takes the k Nearest Neighbors as per the calculated minkowski distance. Among the identified k neighbors, it counts the number of the data points in each category. It then assigns the new data points to that category for which the number of the neighbor is maximum. (Pramesti, D., 2020)

K. Hybrid Ensemble Learning Model:

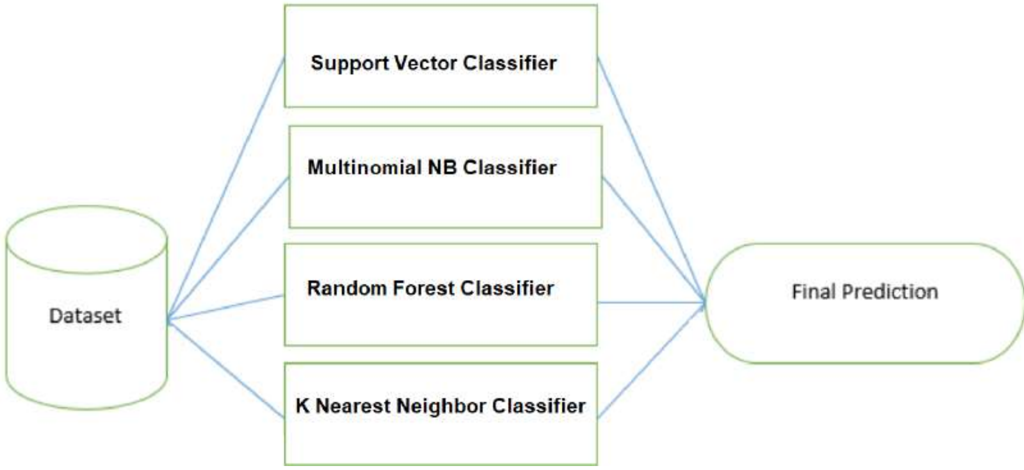

Figure 1: Hybrid Ensemble classification Model.

For developing Hybrid Ensemble classification Model, we defined each of the four machine learning models – SVM, MNB, RFC and KNN - 5 times that results in a combination of a total of 20 weak learners as shown in figure 1. Finally, we used Max Voting Classifier method where the class which has been predicted mostly by the weak learners will be the final class prediction of the ensemble model. (Dr. Vaibhav Kumar, 2021).

4. RELATED WORK

A typical model of NLTK and Supervised Learning Approach to Aspect Based Sentiment Analysis of Global Brands:

Figure 2 shows A typical model of NLTK and Supervised Learning Approach to Aspect Based Sentiment Analysis of Global Brands.

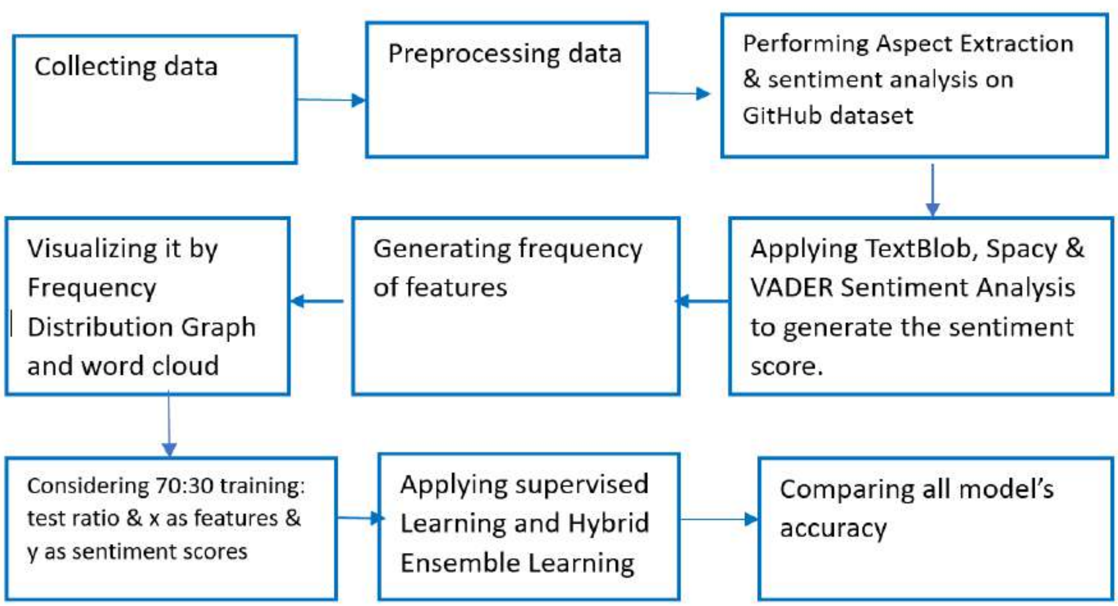

Figure 2: Typical Aspect Based Sentiment Analysis and NLTK and supervised learning model.

A. Collection of data:
We have used GitHub dataset 'Womens Clothing E-Commerce Reviews.csv' file for our study, which has 23486 rows. The URL for this file is https://github.com/AFAgarap/ecommerce-reviews-analysis/blob/master/Womens%20Clothing%20E-Commerce%20Reviews.csv

B. Preprocessing data:
We have applied all data preprocessing techniques like,

C. Performing Aspect Extraction & sentiment analysis on GitHub dataset:
We first fetched the data and pre-processed the data. Next, we converted each input sentence into tokens. If a word has some specific meaning, that is if it is a semantic argument-nsubj, and if it is a noun then it is considered as the aspect name. If any next token in the sentence is empty or some adverb and adjective or simply adjective, then it is considered as aspect extraction. E.g. for the sentence – The Product is very attractive. Aspect: Product and Description: very attractive. Next. We found its sentiment value using Vader Sentiment Analysis. And consider only those features, whose sentiment value is greater than +0.5 for the popular feature.

We performed aspect-based analysis on GitHub data using to find the attributes and their description from 'review Text' column of 'Womens Clothing E-Commerce Reviews.csv' file. it looks like below:

Aspect and its Description Extraction

```
{'aspect': 'top', 'description': 'great'}, {'aspect': 'size', 'description': 'great'},
```

We have taken just the description into consideration. As shown in the example 'perfect' can be one of the features that temps user to buy that product. We have found sentiment score of each description as shown above. As it is too difficult to take every description into consideration, we have just shown the first 10 descriptions along with its sentiment score, whose sentiment is > 0.5.

Below is shown the features of non-Indian Brands, which, along with its sentiment score.

```
ASPECT EXTRACTION AND ITS VADER SCORE

['great', 'perfect', 'perfect', 'great', 'perfect', 'super', 'great', 'perfect', 'perfect', 'vibrant', 'great',
[0.6249, 0.5719, 0.5719, 0.6249, 0.5719, 0.5994, 0.6249, 0.5719, 0.5719, 0.5267, 0.6249, 0.5267, 0.5719, 0.5106
[0.5, 0.5, 0.6, 0.8, 1.0, 0.5, 0.6, 0.5, 0.5, 1.0, 0.6, 0.6, 0.7, 0.5, 0.8, 1.0, 0.7, 0.8, 0.5, 0.5, 1.0, 1.0,
[0.5, 0.5, 0.6, 0.8, 1.0, 0.5, 0.6, 0.5, 0.5, 1.0, 0.6, 0.6, 0.7, 0.5, 0.8, 1.0, 0.7, 0.8, 0.5, 0.5, 1.0, 1.0,
```

D. Applying TextBlob, Spacy and VADER Sentiment Analysis:
There are three separate APIs available in Python as Textblob, SpacyTextBlob and VADER sentiment analysis to find the words polarity which gives values from +1.0 to -1.0. The five categories of sentiment considered in this paper. Being 5 as most positive, whose values considered as 0.5 to 1 polarity, 4 is positive, whose values considered as greater than zero and less than 0.5, 3 to be the neutral, considered as zero polarity, 2 as the negative sentiment with polarity between zero (exclusive) polarity and -.5 (exclusive) polarity and 1 considered as most negative sentiment with polarity between -0.5 to -1.0, both inclusive. On the basis of VADER sentiment analysis, the frequency distribution of features has been calculated and wordCloud is designed. . To draw the frequency distribution graph of features against the count, the first 15 features we have considered.

E: Applying Supervised Learning Algorithms and Hybrid Model:
After finding the frequency distribution and wordCloud, the X is considered as the aspect- feature and y is considered the sentiment analysis score (5 to 1) as per the textblob API and the ML algorithms applied to the split of training and test data with the ratio 70:30. There are four ML algorithms we studied and also the hybrid model we applied. In the similar fashion, we studied the Spacytextblob and the VADAR sentiment analysis algorithm.

TextBlob and supervised Learning Algorithms:

As shown in figure 3, for the SVM- TextBlob, total 458, 5506, 13443, 7441, and 7558 are classified correctly with sentiment 5, 4, 3, 2 and 1 respectively, whereas 2, 4, 989, 9, 7, 11, 10, 25, 29 and 4 data misclassified. As per MNC - TextBlob, total 457, 5507, 13447, 7396, and 7552 are classified correctly with sentiment 5, 4, 3, 2 and 1 respectively, whereas 2, 1, 7, 99, 3, 1, 1, 2, 25, 46, 20, 34 and 5 data misclassified. As per RFC - TextBlob, total 457, 5510, 13443, 7435, and 7564 are classified correctly with sentiment 5, 4, 3, 2 and 1 respectively, whereas 2, 1, 2, 95, 8, 2, 4, 3, 14, 10, 28, 24, and 3 data misclassified. As per KNN - TextBlob, total 457, 5452, 13435, 7344, and 7500 are classified correctly with sentiment 5, 4, 3, 2 and 1 respectively, whereas 3, 7, 157, 1, 7, 2, 6, 19, 94, 30, 3, 81 and 7 data misclassified.

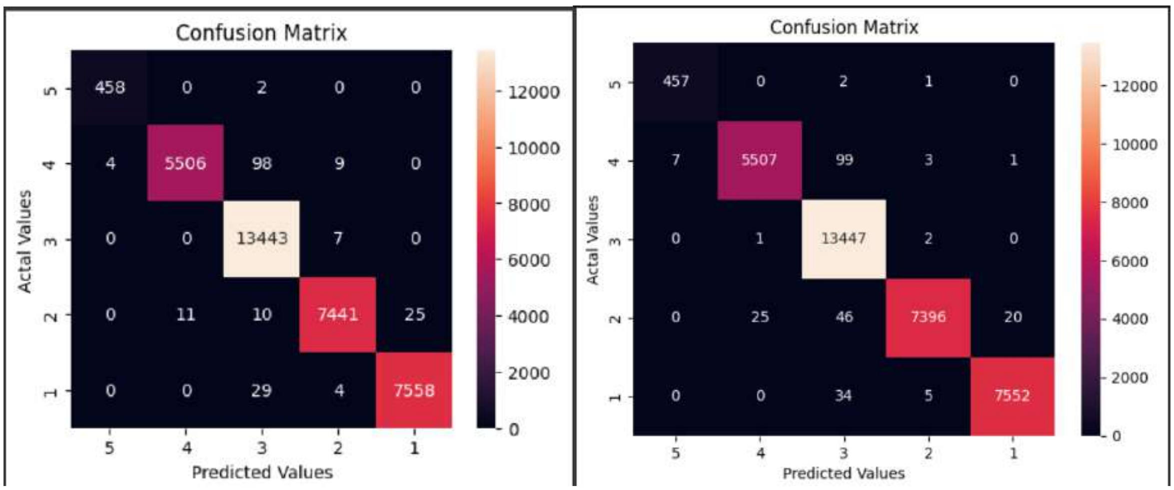

Figure 3: SVM and MNB – TextBlob

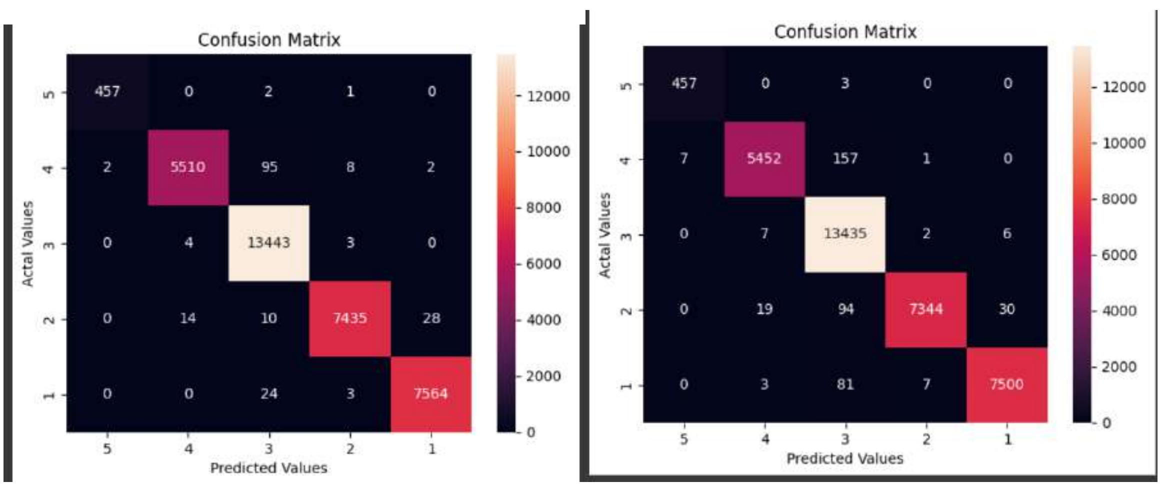

Figure 4: RFC and KNN – TextBlob

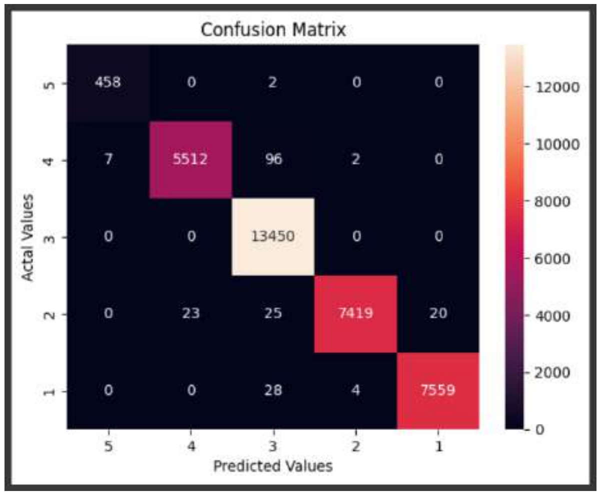

Figure 5: Hybrid Ensemble Learning Model – TextBlob

As per Hybrid Ensemble Learning Model - TextBlob, total 458, 5512, 13450, 7419, and 7559 are classified correctly with sentiment 5, 4, 3, 2 and 1 respectively, whereas 2, 7, 96, 2, 23, 25, 20, 28 and 4 data misclassified, which is lowest of all.

Spacy and supervised Learning Algorithms:

As per SVM -Spacy, 34406 data classified correctly whereas 199 data misclassified. As per MNB -Spacy, 34359 data classified correctly whereas 246 data misclassified. As per RFC -Spacy, 34409 data classified correctly whereas 196 data misclassified. As per KNN -Spacy, 33918 data classified correctly whereas 687 data misclassified. As per Hybrid Ensemble Learning Model -Spacy, 34400 data classified correctly whereas only 205 data misclassified.

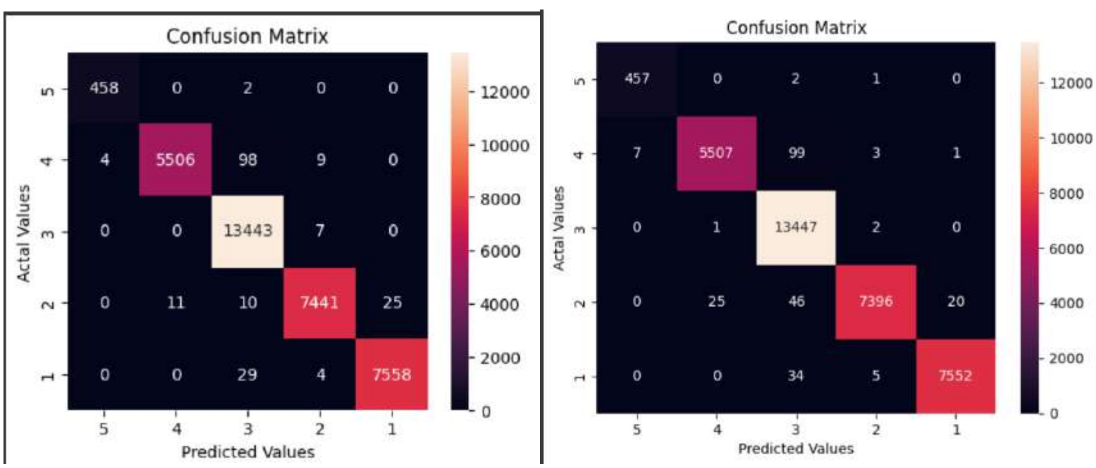

Figure 6: SVM and MNB – SpacyTextBlob

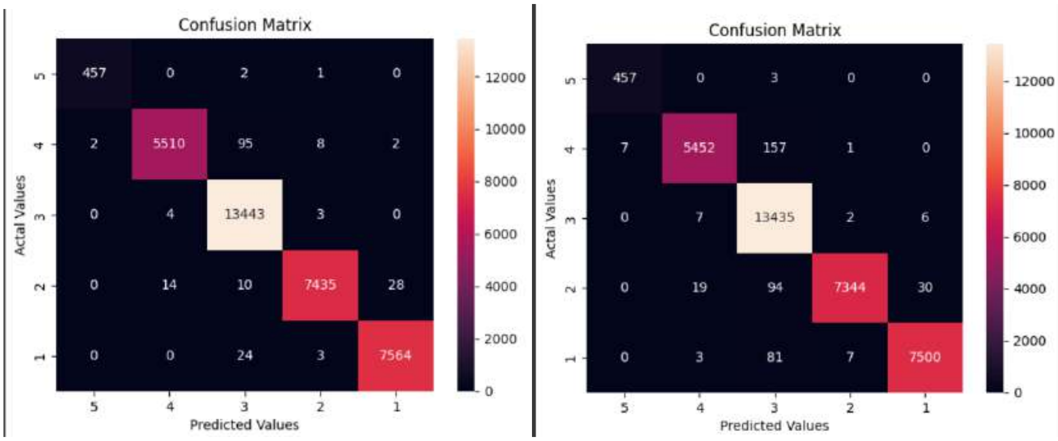

Figure 7: RFC and KNN – SpacyTextBlob

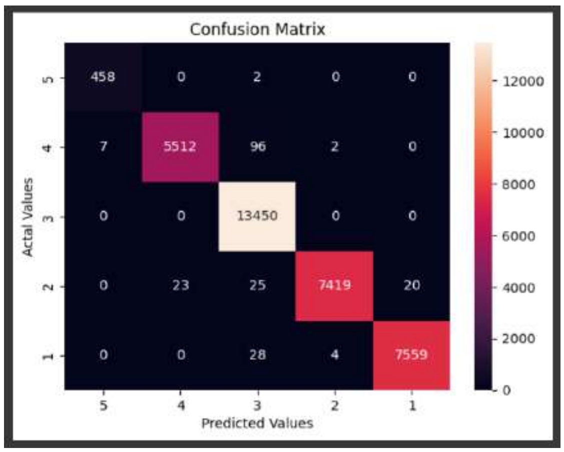

Figure 8: Hybrid Ensemble Learning Model – SpacyTextBlob

VADER and supervised Learning Algorithms:

As per SVM -VADER, 34514 data classified correctly whereas 91 data misclassified. As per MNB -VADER, 34483 data classified correctly whereas 122 data misclassified. As per RFC -VADER, 34543 data classified correctly whereas 62 data misclassified. As per KNN -VADER, 34359 data classified correctly whereas 246 data misclassified. As per Hybrid Ensemble Learning Model -VADER, 34532 data classified correctly whereas only 73 data misclassified.

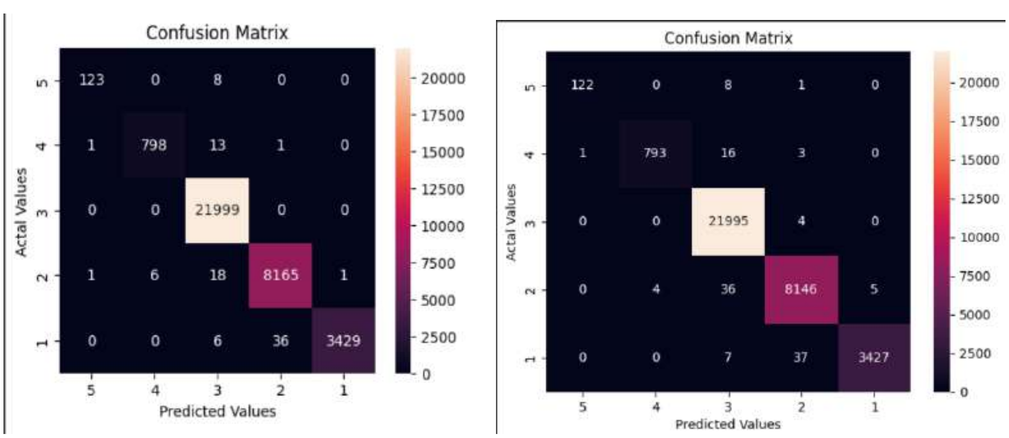

Figure 9: SVM – VADER and MNB – VADER

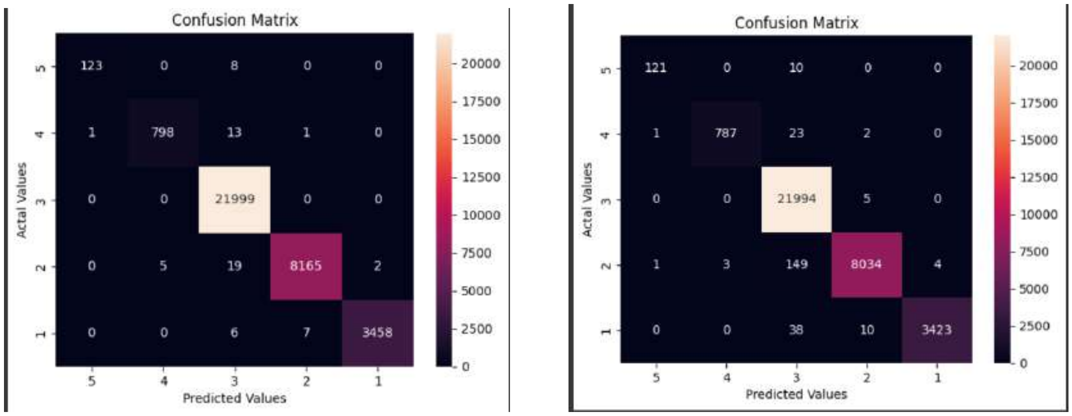

Figure 10: RFC and KNN – VADER

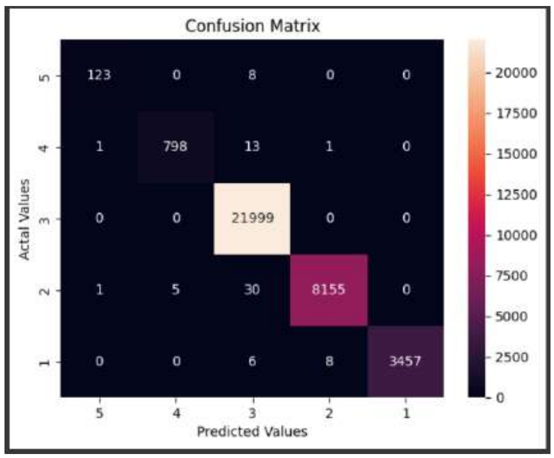

Figure 11: Hybrid Ensemble Learning Model – VADER

F: NLTK Frequency Counter

In NLTK, the nlkt package has one function FreqDist(), which one parameter as list corpus, which accepts list and finds frequency of each unique word. The output from frequency counter what we received is as shown in figure 12.

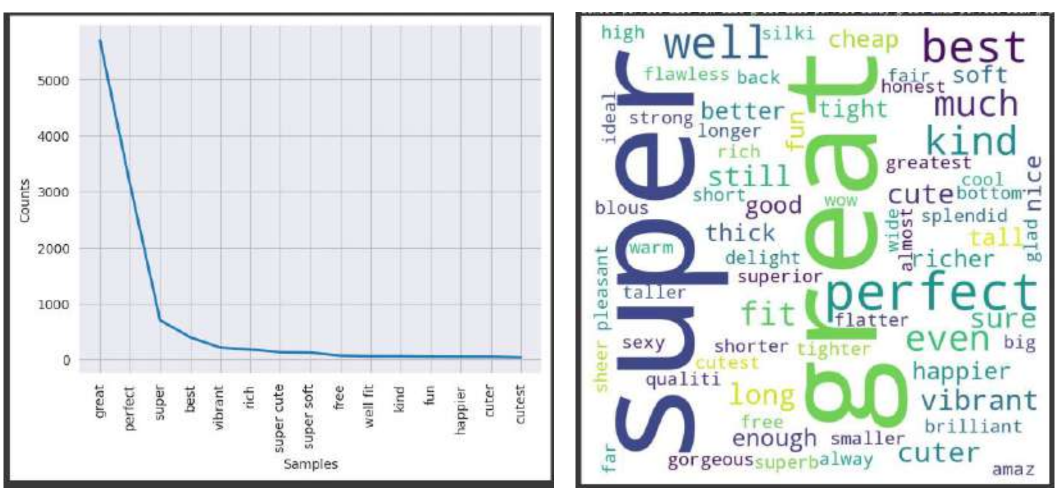

Figure 12: Line chart and WordCloud

The figure shows that the most positive features of what the respondents found as per the dataset of GitHub in global brand cloths are great, perfect, super, best, vibrant, rich, super cute, free, well fit, kind, fun, happier, cuter and cutest. The frequency of each of the features are as: {'great': 5696, 'perfect': 3129, 'super': 703, 'best': 397, 'vibrant': 215, 'rich': 179, 'super cute': 130, 'super soft': 122, 'free': 68, 'kind': 57, 'well fit': 57, 'fun': 50, 'cuter': 48, 'happier': 48, 'cutest': 36}

5. CONCLUSION AND FUTURE WORK

The line chart has been plotted with plot() function of Python. The WordCloud() function imported from wordcloud package. The line chart shows the most popular features people liked in global brands in decreasing order. WordCloud is a visualization technique for text data. In the WordCloud each word is picturized with its importance in the context or its frequency.

Algorithm/accuracy	TextBlob	SpacyTextBlob	VADER
SVC	0.9940	0.9941	0.9970
Multinomial Naïve Bayes	0.9927	0.9927	0.9961
Random Forest	0.9945	0.9945	0.9980
KNN	0.9870	0.9870	0.9928
Hybrid Ensemble Model	0.9937	0.9937	0.9976

Table 1: Comparative data of TextBlob, SpacyTextBlob and VADAR sentiment analysis

As per table 1, the conclusion can be drawn that TextBlob and SpacyTextBlob yields almost the same results. The VADAR sentiment analysis has shown more accuracy than the other two APIs. So further work can be done only using VADER sentiment analysis. There is not much improvement in using Hybrid Ensemble learning model because we have classified the aspects and not the direct reviews. Because we classified the aspects and on the aspects we applied the ML algorithms, It has given wonderful accuracy in the base work itself. Very little fine tuning could be done with the hybrid ensemble learning model. But even though the scope of improvement was not there, the hybrid ensemble learning model has shown minor improvement in all categories of sentiment analysis.

REFERENCES

[1] Swaminanthan, V., Winterich, K. P., Zeynep, G. C. (2007), ‘My’ Brand or ‘Our’ Brand: The Effects of Brand Relationship Dimensions and Self- Construal on Brand Evaluations, Journal of Consumer Research, 34 (2), pp. 248-59.

[2] *NLTK Documentation*. (n.d.). Retrieved from https://www.nltk.org/_modules/nltk/probability.html

[3] Kapadia, Beena & Jain, Amita. (2021). Analysis of Papers Based on Sentiment Analysis Applications on E-Commerce Data. 10.1007/978-3-030-73603-3_30.

[4] Hemmatian F, Sohrabi MK (2019) A survey on classification techniques for opinion mining and sentiment analysis. Artif Intell Rev 52(3):1495–1545. https://doi.org/10.1007/s10462-017-9599-6

[5] Pontiki M, Galanis D, Papageorgiou H, Androutsopoulos I, Manandhar S, AL-Smadi M, Al-Ayyoub M, Zhao Y, Qin B, De Clercq O, Hoste V, Apidianaki M, Tannier X, Loukachevitch N, Kotelnikov E, Bel N, Jiménez-Zafra SM, Eryiğit G (2016) SemEval-2016 task 5: aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), Association for Computational Linguistics, San Diego, CA, pp 19–30. https://doi.org/10.18653/v1/S16-1002

[6] Behdenna S, Barigou F, Belalem G (2018) Document level sentiment analysis: a survey. EAI Endorsed Trans Context Aware Syst Appl. https://doi.org/10.4108/eai.14-3-2018.154339

[7] Shirsat VS, Jagdale RS, Deshmukh SN (2017) Document level sentiment analysis from news articles. In: 2017 international conference on computing, communication, control and automation (ICCUBEA), pp 1–4. https://doi.org/10.1109/ICCUBEA.2017.8463638

[8] Gurini D, Gasparetti F, Micarelli A, Sansonetti G (2013) A sentiment-based approach to twitter user recommendation. In: CEUR Workshop proceedings, CEUR-WS, vol 1066, conference of 5th ACM RecSys workshop on recommender systems and the social web, RSWeb 2013- Co-located with the 7th ACM conference on recommender systems, RecSys 2013; Conference Date: 13 October 2013; Conference Code:111232

[9] Appel O, Chiclana F, Carter J, Fujita H (2016) A hybrid approach to the sentiment analysis problem at the sentence level. Knowl Based Syst 108:110–124. https://doi.org/10.1016/j.knosys.2016.05.040

[10] Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):5. https://doi.org/10.1186/s40537-015-0015-2

[11] Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD international conference on knowledge discovery and data mining, Association for Computing Machinery, New York, NY, USA, KDD '04, pp 168–177. https://doi.org/10.1145/1014052.1014073

[12] (n.d.). Retrieved from Natural Language AI: https://cloud.google.com/natural-language

[13] Hutto, C.J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of social media Text. ICWSM.

[14] Yang, S. E. (n.d.). Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment. In M. 1.-1. IMECS 2019 (Ed.). ISBN: 978-988-14048-5-5

[15] https://github.com/AI4Finance-Foundation/*Financial-News-for-Stock-Prediction-using-DP-LSTM-NIPS-2019 national information processing system* Li Yang stock prediction using financial news in 33rd conference, on Neural Processing Information System, Vancouver, 2019.

[16] V. Bonta, N. K. N. J. for *Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis, Asian Journal of Computer Science and Technology* vol. 8 pp. 1 -6, 2019.

[17] https://github.com/AFAgarap/ecommerce-reviews-analysis/blob/master/Womens%20Clothing%20E-Commerce%20Reviews.csv as the dataset as on 18th Feb 2024 1:00pm.

[18] Satuluri Vanaja, M. B. (2018). Aspect-Level Sentiment Analysis on E-Commerce Data. Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018) (p. 5). Bengaluru: Amrita School of Engineering.

[19] Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. 2019 International Conference on Automation, Computational and Technology Management (ICACTM).

[20] Singh, S. N., & Sarraf, T. (2020). Sentiment Analysis of a Product based on User Reviews using Random Forests Algorithm. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence).

[21] Irawaty, I., Andreswari, R., & Pramesti, D. (2020). Development of Youtube Sentiment Analysis Application using K-Nearest Neighbors (Nokia Case Study). 2020 3rd International Conference on Information and Communications Technology (ICOIACT).

[22] Dr. Vaibhav Kumar (2021). A Hands-on Guide To Hybrid Ensemble Learning Models, With Python Code, AI Mysteries.

[23] F. Fazrin, O. N. Pratiwi and R. Andreswari, "Comparison of K-Nearest Neighbor and Logistic Regression Algorithms on Sentiment Analysis of Covid-19 Vaccination on Twitter with Vader And Textblob Labeling," 2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA), Denpasar, Bali, Indonesia, 2022, pp. 39-44, doi: 10.1109/ICSINTESA56431.2022.10041609.

[24] A. K. Singh and A. Verma, "An Efficient method for Aspect Based Sentiment Analysis Using SpaCy and Vader," 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2021, pp. 130-135, doi: 10.1109/CSNT51715.2021.9509650.