



Sentiment Analysis of Earphone Based on Customer Reviews on Amazon

¹Dr. P. C. Gita, ²Pavan. J. K.

¹Associate Professor,
Kristu Jayanti College, Bangalore
Email - pcgita@gmail.com
²II MBA, School of Management
Kristu Jayanti College, Bangalore

Abstract: A key analysis tool for determining the emotions of clients is sentiment analysis, often known as opinion mining. Sentiment analysis aids us in knowing the customers' attitudes toward products and helps to categorize them as positive, negative, and neutral. A smart product manager may effectively utilize this information to improvise the features of their products and also benchmark their products with that of the Industry's Best. This exercise will enable them to develop items that are well-received by the general public, allowing the business to outperform its rivals. The study aims to build a prediction model for Sentiment analysis through Sentiment distribution in percentage, Word Cloud, Accuracy of classifiers and ranking of the products based on positive sentiments. The study concluded that Real me Buds Wireless in-Ear was the product with highest positive sentiments among the customers and Logistic Regression classifier was found to be the most consistent classifier with highest accuracy. Sound and Sound Quality emerged as the two main features the customers appreciated in ear phones.

Keywords: Amazon, Earphone, Customer Review.

1. INTRODUCTION:

Internet technology has transformed the way organizations do their business and the way consumers access their products. Due to this digital transformation, there emerged a plethora of e-commerce sites which made purchase of the product possible with a click of the mouse. Organizations and their agencies were able to understand the behavior of consumers by observation and purchase behavior. With the advent of online purchase, the direct connect between the customer and the seller is lost. But it is very crucial for the marketers to understand the attitudes, opinions and sentiments of the customers towards their products. The outcome of web 2.0 of the internet era transformed the internet from simple information to opinion source. (Schmalegger & Carson, 2008) An online customer has the choice to voice his opinion, thoughts, emotions and feelings about his purchase experience in online websites in the form of reviews. (Litvin et al 2008). Reviews have been gaining high significance in today's scenario as purchase are getting modified to online purchases. These reviews are considered significant from the marketer's perspective as well as consumer's perspective. (Carson 2008) Marketer is interested to know the opinion and attitude of consumers towards their products and competitor's product whereas consumers are interested in deciding the product of purchase based on these reviews.

Sentiment analysis determines the opinions and attitudes of the customers towards a particular product or brand. Sentiment analysis (opinion mining) analyses the sentiments, opinions, attitudes and emotions of consumers towards goods, services, issues, events, organizations and employers. (Bing Liu 2012). Sentiment analysis at a business tries to examine the feelings of consumers as they relate to marketing campaigns or goods that are being offered for sale online. It gains importance when it comes to customer loyalty, consumer happiness, the effectiveness of advertising and promotion, and product acceptability. Understanding people's opinions and sentiments from text in online media is a cumbersome task as the data is unstructured, difficult to retrieve, needs lot of cleaning and analysis of emotions is highly complicated.

Importance of Analysis of Online reviews:

Understanding customer and product experience feedback from various channels with the use of customer review analysis will help the managers find insightful business information. Employing a variety of customer review sources to find out about customers and products allows organisations to get more precise, detailed, and accurate findings. Sentiment analysis to drive customer review analysis provides strategic, practical business insights thanks to machine learning's seamless automation of emotion mining from many sources. Types of Review Analysis are explained below:

**Emotional Analysis:**

A Python package called Text2emotion was created to detect the relevant emotions that are embedded in text data. Text2Emotion can discern the emotions from the provided text data.

Text2emotion is a tool that may be used to analyze any text data, identify the emotions present within, and produce an output in the form of a dictionary. Basic emotion categories like "excited," "angry," "happy," "sad," "surprised," "fear," and so on fit it nicely.

Picture Analysis:

The processing of an image into essential components in order to extract important and intelligible information is referred to as image analysis. When the data is in.jpeg or.png format, such analysis can aid in understanding the data as a whole.

Emoticons Analysis:

In order for our prediction model to grasp which emotion each emoji represents; we must process the emoji in all of their many forms and assign them a meaning. The programme is trained to recognize which emoji is associated with which emotion, allowing the sentiment of a customer to be investigated

Earphones is one of the most commonly available products after the mobile penetration in India. It is a common sight to see people on the roads talking, listening to music, watching a movie, watching YouTube channel online using their earphones. As it has become an essential product which is purchased along with a mobile phone and most commonly purchased product online, reviews of earphones are chosen for sentiment analysis in this study. For this study reviews related to seven top brands of earphones are chosen for analysis from Amazon India website.

The study aims to perform analysis of reviews of customers related to earphones in Amazon India website. The study includes the following steps – to create a lexicon of text components from the entire reviews, next we conduct sentiment analysis which is used to classify customer reviews into positive, negative and neutral. We rank the classifiers based on the accuracy to predict the accurate classifier for sentiment analysis. This introduction is followed by review of literature, process of sentiment analysis , analysis of text data and finally concludes the research findings and elaborates on the scope for future research.

Related works:

This section deals with the nuances of Sentiment analysis that is relevant for this study. Sentiment analysis studies the natural language and identifies opinions, attitudes, emotions and evaluations. (Wiebe.J, 1994).Sentiment analysis determines the polarity of the words or phrases such as negative, positive or neutral. This kind of automatic retrieval of polarity is pioneered by Hatzivassiloglou and McKeown (1997) by using machine learning methods to construct lexicon of sentiment terms. Wiebe.J (1994) used adjectives and its phrases whereas Riloff, E. et al (2003) used nouns to study lexicons. Qiu et al. (2009) proposed an approach for retrieving a large number of sentiment words and assigning polarity. In the early days, sentiment analysis targetted document level analysis and later studies started working on words and phrases for deeper analysis of sentiment. Pang et al (2002) used common machine learning methods for analysis - maximum entropy classification, support vector machine and naïve Bayes and SVM emerged to perform better and accurate than the remaining methods. Sentiment analysis have been performed using different classifiers and algorithms. In a study by Janhavi et al 2018, datamining algorithm ROCK was used for clustering and CART algorithm was used for classification. The reviews were classified by subjectivity / objectivity and negative / positive opinion of buyers using Decision tree algorithm and Naïve Bayes (Gurneet Kaur et al 2018). In another sentiment analysis study by Nguyen et al (2018), the methods used were Random Forest, Classification and regression tree, least absolute shrinkage and selector operator (LASSO) logistic regression based on bag of words text mining approach. Sentiment analysis of reviews from amazon were conducted by Dadhichl et al (2021) using Hybrid Rule based approach.

In another study by brownfield et al 2020 deep learning models such as multilayer perceptron (MLP) and recurrent neural network (RNN) are compared with traditional machine learning methods such as Naïve Bayes analysis and Support Vector machines (SVM). Haque et al 2018 in their study on sentiment analysis used supervised machine learning methods such as random forest, Decision tree, linear regression, Naïve Bayesian, Stochastic gradient Descent (SGD) and Support vector machine classifier (SVMC). Budhwar et Singh (2021) used a hybrid methodology that combines Naïve Bayes Technique, LSTM and KNN classifiers for better results. Raju (2020) identified CatBoosting with TF-IDF and Logistic Regression with Count Vectorizing as best performing models. Xiao et al (2021) in their study performed Sentiment analysis using Texblob a natural language processing tool in python. Katic & Milicevic (2018) combined traditional classifiers such as bag of grams, bag of words and their TF-IDF variants with Deep learning models



such as Simple Long short term memory (LSTM) recurrent neural network, Convulated Neural networks (ConvNets) Support Vector machines (SVM) and Logistic Regression in their study on Sentimental Analysis

Based on the above reviews the classifiers are chosen for this study to determine the predictive model with highest accuracy.

2. METHODOLOGY:

Data Source: The data is collected from the Amazon India website which is then subject to preprocessing steps.

Sample Size: Overall 3500 reviews were collected from Amazon India websites for the seven brands of earphones

Reviews of Brands under earphones which are considered for this Study are:

1. One plus Bullets Z2 Bluetooth Wireless
2. Boulton Audio Z Charge Bluetooth Wireless
3. Infinity (JBL) Glide 120, In Ear Wireless Earphones
4. Real me Buds Wireless in-Ear Bluetooth Earphones
5. P Tron Tangent Lite Bluetooth 5.0 Wireless in Ear Earphones
6. P Tron Tangent beat in-Ear Bluetooth 5.0 Wireless Headphones
7. Real me Buds Wireless 2 Neo Bluetooth

Scraping the web

Web scraping, often known as web data extraction, is a method for obtaining data from internet pages. The content in internet can be directly accessed by web scraping software or code using either the Hypertext Transfer Protocol (HTTP) or a web browser (WWW). The information is then displayed in its entirety, whether it contains words or not. The data for this study, which are product reviews from Amazon, were obtained using Python programming.

A tool called Beautiful Soup is used to scrape data from websites. It provides Pythonic paradigms for iterating through, searching through, and modifying the parse tree on top of an HTML or XML parser.

Natural Language Processing (NLP)

To make computers understand the natural language spoken by humans is the foremost goal of natural language processing because computers cannot currently understand natural language. Python package that can be used for NLP is called NLTK, or the Natural Language Toolkit. Unstructured, human-readable text makes up a sizable amount of the data that can be analysed. One must pre-process the data before they may analyze it programmatically.

Data preprocessing

When dealing with textual data, it is common to find that the data is highly cultural and unstructured. And in order to do the analysis on it, we must first clean it and make it understandable to the algorithms.

The product reviews scraped from Amazon included extraneous data such as special symbols, emoticons, punctuation, and typos, among other things. This needs to be removed from the dataset, and the following actions were taken to do this:

Punctuation Elimination:

Punctuations are important when it comes to natural language, but for a Machine it has no meaning. So, in order to analyze the text efficiently we remove the unwanted punctuations like commas (,), full stops exclamation marks (!), (.), question marks (?), semi-colons (;), colons (:), speech marks (“,”) and apostrophes (‘).

Eliminating HTML tags:

HTML tags may be included in text data that is scraped from websites. These HTML tags add nothing to the text data. It is crucial that we get rid of it before performing an analysis on it.

Removal of Special Characters:

Special characters such as - (hyphen) or / (slash) have no relevance in sentiment analysis, so they must be removed to make the data considerably lighter to for in. Depending on the use case, characters are eliminated.

Removing Alphanumeric words:

Only text data is evaluated in a Sentiment analysis. In this situation, we preserve only alphabetical data and eliminate all digits and alphanumeric data to make the data light and efficient for analysis.

Tokenization:

Tokenization splits raw text into little pieces, often known as tokens. Such tokens are crucial since they aid in comprehending the context or in building the NLP model. It aids in interpreting the meaning of the text by analyzing the word sequence.

**Removal of Stop words:**

Since computers cannot currently understand natural language, the main objective of the field of natural language processing is to make authentic human discourse comprehensible by computer programmes.

Lower casing:

Lower casing is the simple act of turning a word to lower case. When written in lower case, words like "Hook" and "hook" have the same meaning, but the vector space model depicts them as two different words (resulting in more dimensions). This maintains the consistency of our data, which makes it simpler for us to work with it and do more accurate analyses.

Lemmatization:

Typically, one word in a text may appear in several different contexts, such as crop and cropped (past participle) or slice and slices (plural). Text normalization assists in reducing word variants to a single, dependable root form. This allows the data to become more consistent.

Python:

Python is a commonly used general purpose object oriented, interactive and interpreted high level programming language. It was created by Guido van Rossum during the period 1985- 1990. Python supports multiple programming paradigms, such as Object Oriented, Procedural and Functional programming language. Python source code is available under the GNU General Public License (GPL). Python is in great demand today in organisations also require programmers well versed in python.

Sentiment Classification: Using Python as the programming language, sentiment classification of the collected reviews was performed. The reviews were classified based on subjectivity and polarity into three types - positive, negative and neutral sentiments.

Word Cloud: Word cloud is an electronic image which shows the words occurring in the reviews. The most occurring words appear in larger size when compared to the rest. The words appear in different sizes proportional to the frequency of their occurrence. In the study, using word cloud the frequency of the words occurring in the reviews were studied. Word cloud highlights the most sought-after feature of a particular brand

Classifiers:

A machine learning Classifier orders and classifies data into classes. A classifier is the algorithm or the rules to classify data which are used by machines. A classification model is the outcome of classifier's machine learning. Classifier trains the model and the model finally classifies the data. Classifiers may be supervised, unsupervised or semi supervised. Sentiment analysis is a supervised machine learning example where classifiers are trained to analyze text and output the text into the class: Positive, Neutral, or Negative.

In our study we have used the following classifiers for the classification of the sentiments

Naïve Bayes: The Naive Bayes classifier approach combines the probabilities of different words and classes to determine the probabilities of various classes assigned to diverse texts. Naive Bayes is one of the simplest and fastest classification methods available in Python for a sizable amount of data. Numerous applications, including spam filtering, text categorization, sentiment analysis, and recommendation systems, have called for the usage of naive Bayes. The Bayes probability theorem for unknown class prediction is used by Nave Bayes to categories the data and display the results.

Random Forest Classifier: An approach for supervised machine learning called Random Forest Classifier is based on ensemble learning. In ensemble learning, the same algorithm is employed repeatedly to help build a more potent prediction model, or numerous joins are built from a single join of different types of accessible algorithms. The Random Forest Classifier gets its name from the fact that the Random Forest Algorithm mixes several different algorithms of the same kind, or many decision trees. This creates a forest of trees. The collection of different tree-structured classifiers is another way to define the Random Forest Classifier. It's a more advanced and intricate variant of Bagging that includes randomization. RF splits each node using the most effective and powerful predictor among a subset of several predictors randomly picked at that node, rather than the optimal split among all variables.



Support vector machine Classifier: SVM Classifier, also known as Support Vector Machine, is an algorithm that identifies the best decision boundary between vectors that belong to a particular group/category and vectors that do not. This is applicable to any type of vector that encodes any type of data.

Logistic Regression Classifier: Logistic regression is named after the principle of predictive modelling that it employs as a regression technique. Logistic Regression is most commonly used to predict the dependent variable from a given collection of independent factors. In models with a twofold scenario, the results are often defined as 0's

Decision Tree Classifier: A supervised learning method that is used to solve classification and regression problems. It is a tree-structured classifier, where internal nodes denotes a dataset's features, branches denotes the decision-making process, and each leaf node denote the classification result.

KNN Test Classifiers: k-nearest neighbor's algorithm, KNN or k-NN, is a supervised learning classifier that employs nearness or proximity to produce predictions or classifications about the grouping of a single data point.

After classification of the sentiments, confusion matrix is used to identify the number of right classifications and misclassification to understand the accuracy of each classifier

Confusion Matrix

Confusion matrix is used to analyze the categorization outcomes from the top- performing model. A table that summarizes the results of a classification algorithm and identifies the types of misclassifications that have occurred during the analysis is called a confusion matrix. In other words, it aids in exposing the classifier's muddled thinking about diverse groups of various classes. A confusion matrix displays the True Label and the anticipated Label in its rows and columns, respectively. It is exceedingly improbable that a perfect classifier would have large numbers on the major diagonal and zeroes everywhere else.



Fig: Process of Sentiment Analysis



3. ANALYSIS:

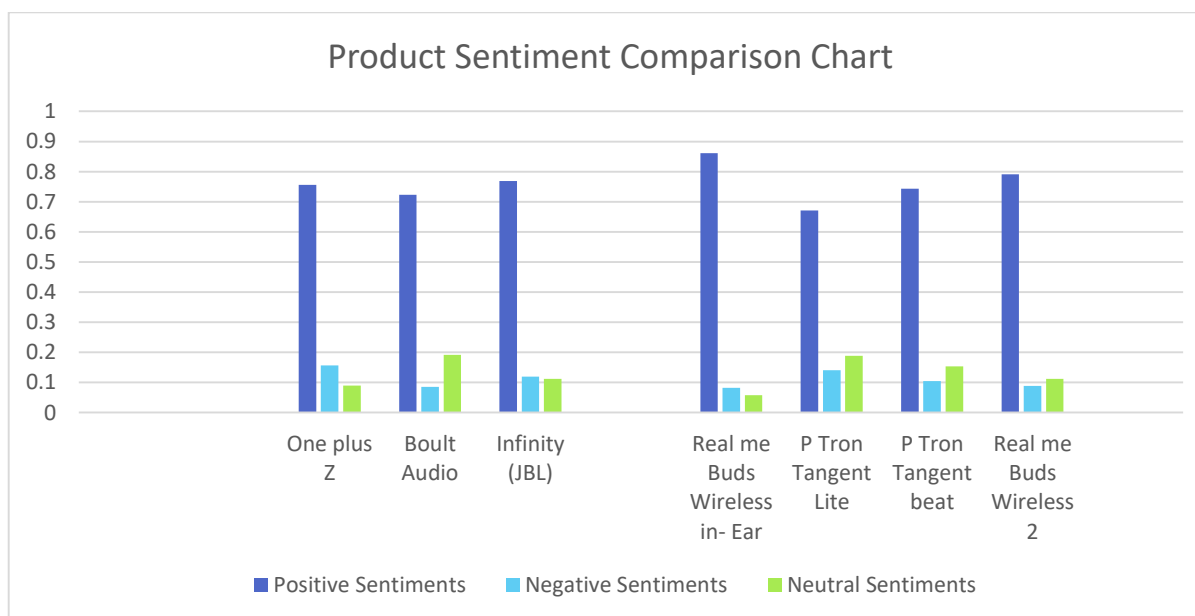
Product Recommendation

Table:1 Product Sentiment Comparison

Hypothesis H1: There is no Significant difference in terms of distribution of nature of sentiments among the product reviews

Products	Positive Sentiments	Negative Sentiments	Neutral Sentiments	Rank
One plus Z	75.6%	15.6%	8.9%	4
Boult Audio	72.3%	8.5%	19.1%	6
Infinity (JBL)	76.9%	11.9%	11.2%	3
Real me Buds Wireless in- Ear	86.1%	8.2%	5.8%	1
P Tron Tangent Lite	67.1%	14.1%	18.8%	7
P Tron Tangent beat	74.3%	10.4%	15.3%	5
Real me Buds Wireless 2	79.1%	8.8%	11.2%	2

Table1: Comparison of sentiment classification of the reviews



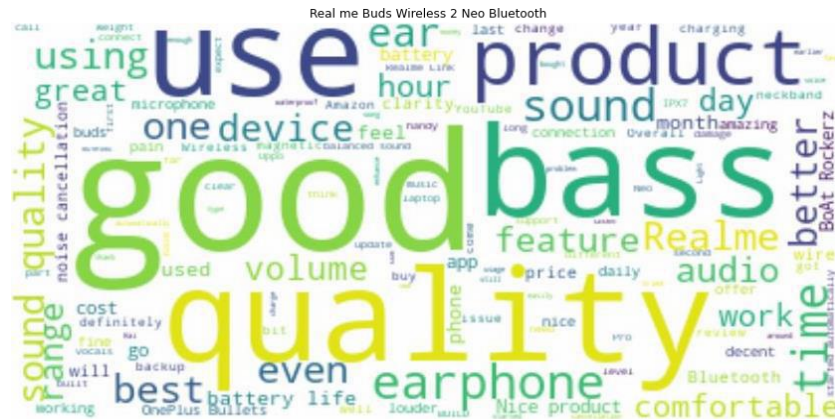
(Fig 1: Sentiment classification of brands reviews)

Based on the positive sentiments the rank was given to various products where 1 being the best and 7 being the worst of the lot. It was clear that Real me Buds Wireless in-Ear was the most liked product from the list of products by the consumers based on their sentiment review with 86.1% Positive sentiments towards the product. Hence, it will be better to go

Comparison of Key Features of Earphones using Word Cloud

Hypothesis

H2: There is no significant difference between the commonly referred Key features of various Product Reviews.



(Fig 1: Word Realme Buds Wireless 2 Neo Bluetooth showing repeated words frequency)

Products	Major Key Features commonly referred by the consumers.
One plus Z	Neckband and Sound Quality.
Boult Audio	Good Sound and Sound Quality.
Infinity (JBL)	Sound and Bass.
Real me Buds Wireless in-Ear	Good Quality, Good Bass and Sound.
P Tron TangentLite	Quality, Sound Quality and Bluetooth.
P Tron Tangentbeat	Product Quality, Sound Quality and Good
Real me Buds Wireless 2	Good Bass, Quality and Sound.

Table 2: Key features identified from threviews using word Cloud

As we can see from the above table, Sound and Sound Quality are the most commonly referred by the consumers which tells us about the most common requirement by today’s consumers from their Earphones. Each product has a different set of key features which is commonly talked about by the customers which proves that we reject H2 as there is a significant difference between the commonly referred Key Features of various Product Reviews.

Overall Rank for Classifiers (Least is better)

Hypothesis: H3: There is no significant difference between the performace of various Classifiers.

Table 1: Rank of Classifiers

Rank	One plus Z	Boult Audio	Infinity (JBL)	Real me Buds Wireless in-Ear	P Tron Tangent Lite	P Tron Tangent beat	Real me Buds Wireless 2	Overall Rank
NB	5	1	5	6	2	3	4	2.6
RFC	4	4	4	4	4	2	2	2.4
SVM	3	6	1	1	1	1	3	1.6
LR	1	2	2	2	5	1	1	1.4
DTC	5	5	3	3	6	2	3	2.7
KNN	2	3	6	5	3	1	2	2.2

Logistic Regression was found to be the most accurate classifier among all the classifiers with the rank score of 1.4 points followed by SVM with a score of 1.6, followed by KNN Test with 2.2 points, followed by Random Forest Classifier with 2.4, Naïve Bayes followed by with 2.6. and with least points of 2.7 From Decision Tree Classifier Therefore, H3 is rejected as there clearly is a significant difference between the performance of various classifiers among our dataset of Product Reviews.



4. Findings and Conclusion:

The bar plot and product sentiment comparison indicates high positive sentiments for all chosen brands of earphones. Among the seven brands, Realme Buds Wireless in ear had emerged at the first rank with maximum positive sentiments and minimum negative sentiments. The least of all brands was P Tron Tangent lite with least positive sentiments.

The reviews of each product under student highlighted different features under word cloud. But the most common features that was often repeated among all the brands reviews are sound and sound quality.

Finally, the classifiers are ranked to find out the best prediction model and Logistic Regression was found to be the most accurate classifier among all the classifiers with the rank score of 1.4 points

In this study sentiment analysis was performed using sentiment distribution to understand the polarity of the reviews towards earphone brands. Later to identify the key features that have occurred in the reviews of the brands word cloud was performed and sound and sound quality emerged as the most commonly occurring words in the product reviews. The study also aimed to identify the most accurate predictive model by using various algorithms such as naïve bayes (NB), Random Forest Classifier (RFC), Support vector machines classifier (SVM), Logistic Regression (LR), Decision tree classifier (DTC) and Kth Nearest neighbour classifier (KNN). The algorithm that emerged to be more accurate in this study was Logistic Regression.

With the rapid enhancement of online sales in the recent times, the need for sentiment analysis of online reviews is gaining more attention. Identifying the polarity of the online reviews is the first step to understand online consumer behaviour and perception towards online products. The key feature identification from reviews using word cloud helps marketers understand the most preferred features by customers, also help marketers compare the attractive features among competitors. This also provides insights to marketers to enhance their products or come out with new variants of the product. Sentiment analysis can be performed using any algorithm. Knowing the most accurate algorithm enables the analysts to use the right tool in future analysis.

However, the sentiment analysis performed in this study classified sentiments based on polarity and not emotions which is highly required for sentiment requirement. Hence future studies can focus on finer classification of emotions in sentiment analysis

REFERENCES:

1. Brownfield, S., & Zhou, J. (2020, October). Sentiment Analysis of Amazon Product Reviews. In *Proceedings of the Computational Methods in Systems and Software* (pp. 739-750). Springer, Cham.
2. Dadhich, A., & Thankachan, B. (2021). Sentiment Analysis of Amazon ProductReviews Using Hybrid Rule-based Approach.
3. Haque, T. U., Saber, N. N., & Shah, F. M. (2018, May). Sentiment analysis on large scale Amazon product reviews. In *2018 IEEE international conference on innovative research and development (ICIRD)* (pp. 1-6). IEEE.
4. Hatzivassiloglou, V and McKeown, K R, (1997) Predicting the semantic orientation of adjectives, *Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL*, 174–181.
5. Katić, T., & Milićević, N. (2018, September). Comparing sentiment analysis and document representation methods of Amazon reviews. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 000283-000286). IEEE.
6. Litvin, S.W., Goldsmith, R.E. & Pan, B. (2008). Electronic Word-Of-Mouth in Hospitality and Tourism Management. *Tourism Management*, 29(3): 458-68.
7. Nguyen, H., Veluchamy, A., Diop, M., & Iqbal, R. (2018). Comparative study of sentiment analysis with product reviews using machine learning and lexicon-based approaches. *SMU Data Science Review*, 1(4), 7.
8. Pang, B, Lee, L, and Vaithyanathan, S, (2002) Thumbs up? Sentiment classification using machine-learning techniques, *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* (2002) 79–86.
9. Qiu, G, Liu, B, Bu, J and Chen, C, (2009) Expanding domain sentiment lexicon through double propagation, *Proceedings of the 21st International Joint Conference on Artificial Intelligence* (Morgan Kaufmann, San Francisco, 2009) 1199–1204.
10. Riloff, E, Wiebe, J, and Wilson, T, (2003) Learning subjective nouns using extraction pattern bootstrapping, *Proceeding of the 7th Conference on Natural Language Learning* (2003) 25–32.
11. Schmallegger, D. & Carson, D. (2008). Blogs in Tourism: Changing Approaches to Information Exchange. *Journal of Vacation Marketing* 14(2): 99-110.
12. Subroto, A., & Christianis, M. (2021). Rating prediction of peer-to-peer accommodation through attributes and topics from customer review. *Journal of Big Data*, 8(1), 1-29.
13. Wiebe, J. (1994) Tracking point of view in narrative, *Computational Linguistics* 20(2) 233–287.
14. Xiao, Y., Qi, C., & Leng, H. (2021, March). Sentiment analysis of Amazon product reviews based on NLP. In *2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* (pp. 1218-1221). IEEE.