Sentiment Analysis using Term based Method for Customers' Reviews in Amazon Product

Thilageswari a/p Sinnasamy, Nilam Nur Amir Sjaif
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia

Abstract—**Customers' review in Amazon platform plays an important role for making online purchase decision making, however the reviews are snowballing in E-commerce day by day. The active sharing of customers' experience and feedback helps to predict the products and retailers' quality by using natural language processing. This paper will focus on experimental discussion on Amazon products reviews analysis coupled with sentiment analysis using term-based method and N-gram to achieve best findings. The investigation of sentiment analysis on amazon product gain more valuable information on related text to solve problem related services, products information and quality. The analysis begins with data pre-processing of Amazon products reviews then feature extraction with POS tagging and term-based concept. e-Commerce customer's reviews normally classify different experience into positive, negative and neutral to judge human behavior and emotion towards the purchase products. The major findings discussed in this journal will be using four different classifier and N-grams methods by computing accuracy, precision, recall and F1-Score. TF-IDF method with N-gram shows unigram with Support Vector Machine learning with highest accuracy results for Amazon product customers' reviews. The score reveals that Support Vector Machine for unigram achieved 82.27% for accuracy, 82% precision, 80% Re-call and 72% F1-Score.**

Keywords—Sentiment analysis; e-commerce; term based; n-gram

I. INTRODUCTION

The huge growth of opinion sharing platform in E-commerce such as Amazon, E-bay, Shopee, Zalora and Lazada could facilitates the customers to understand better about products that been available in online platform [1] [2]. Sharing customer's experience and feedback help to predict the products and retailer quality by using natural language processing. Furthermore, customer's experience rating (1 to 5 stars) and reviews also play important role to express customers satisfaction namely 4 and 5 score represent positive attitude, score 3 represent neutral and score 1 and 2 represent negative attitude [3] [4]. Product star ratings in E-commerce used to evaluate product quality information and online customer behavior from various dimensions whether it could be positive or negative.

Amazon is one popular online platform for buying and selling online products purchase by customers [2] [5]. Customers' review and rating in Amazon platform plays important role to make a purchase decision however the reviews have grown remarkably in e-Commerce. Naturally, it

consumes plenty of time to digest and explore huge amount of reviews to make a right decision [2] [3] [6] [7]. Other than that, customer may also get indecisive to make right purchase decision especially the 3-stars in Amazon represent neutral. In addition, there are also inadequate summary of dominant reviews from E-commerce platforms for satisfied customers [8]. The semantic words in reviews also create noise to the result and low frequency problem.

The aim of this paper is to explore on quality of text representation and categorization with better classifiers as to measure and judge the important raw information. Therefore, sentiment analysis help predicts the polarity sentiments in reviews which could be positive, neutral or negative. Sentiment analysis explores expression and emotion of customers which feeds extract important information from customers' reviews [6]. Sentiment classification strategy evaluate and described the text in reviews with aspect level by identifying interest words from reviews using text mining method [6] [9], whereby sentiment analysis approach machine learning and dictionary based method [10]. Machine learning method classifies the text and dictionary method identifies polarity of words. In this case, text mining would drive the information to organize text as positive, negative and neutral. Text mining retrieves the relevant information to segregate between structured and unstructured data [11] [12]. This paper will present experimental result and analysis using term based method using data from Amazon products. The objective of this paper is to reveal the insightful meaning and identify the polarity pattern from unstructured data using sentiment analysis. It includes method as data pre-processing, feature extraction, sentiment polarity prediction and classification.

This paper would be sorted out such a manner that after this introduction, Section II comprises of related works it covers some of previous studies on sentiment analysis. Later in Section III, it elaborates text mining methods. The following Section IV will explain in detail of the approach on data and method as solution to this paper summary. The methodology used in this experiment as describe deeply in this section. Next, Section V present summary of result on discussion obtained from experiment. Section VI discuss on contribution of this paper to customers using e-Commerce. Finally, the last section concludes the findings of paper and acknowledgment.

II. RELATED WORK

This section presents several studies with different approaches to analyze sentiments in review. Many researchers

have experience to work with customers' reviewers in using sentiment analysis. Naturally Customers' reviews are very important to identify product satisfaction, product features and services for convenient customers [3] [4] [13]. The sentiment analysis result from various method helps provide effective purchase decision and reducing search cost in E-Commerce. Sentiment analysis polarize the review into positive, neutral and negative with natural language processing for judging human behavior and attitude [6]. Authors in [14] have worked with sentiment analysis using Support Vector Machine (SVM) to analyze reviews in tweets. The survey results present several techniques and methods of sentiment analysis which provides different results but if combine both methods then it produces better result. The main factors of best result are selection of methodology for data preprocessing, feature extraction phase and ratio between training and testing. Hence, author Minu with other researchers [15] has conduct experiments on mobile products reviews using machine learning algorithms and exploring the result with different datasets. POS tagging used to recognize the reviews noun, adjective, verb and adverb. The aspects are than polarize into positive and negative using TextBlob and machine learning approaches such as K-Nearest Neighbour, Support Vector Machine, Multinomial Naïve Bayesian and Bernoulli Naïve Bayesian used for predict the result. The best accuracy of mobile reviews is given by K-Nearest Neighbour and Bernoulli Naïve Bayesian. Furthermore, the sentiment analysis also implemented on different language for analyze various machine learning algorithm with N-gram as feature extraction [16]. Arabic opinion from twitter collected and associate data pre-processing with the phrases for noise reduction. N-gram which is number of terms used to search word such unigram (1), bigram (2) and trigram (3) used as feature extraction to generate frequent appealing words. By applying machine learning approaches the result present that PA and RR using unigram, bigram and trigram present best result as 99.96%. On other hand, online customer reviews from Tokopedia been analyze for understanding service quality level using sentiment analysis [17]. Whereby the unstructured data performed starting with data pre-processing, TF-IDF implemented for identifying frequent itemset and reduce noisy from dataset. Finally, Naïve Bayes classifiers used for measure accuracy, recall and precision. The result shows Naïve Bayes present very good result for classifying sentiment and the overall methods used are more effective to perform better. More details on text mining methods been discussed in Section III.

III. TEXT MINING METHODS

Some prominent researchers have used text mining methods in E-commerce such as the Amazon reviews from large textual database in the form of structured and unstructured data [12] [18]. Text mining includes process such as data collection, data pre-processing, feature extraction, classification and measurement. Even Machine learning tools been utilized to explore text mining method to organize and categorize data in further details. The process data used for predictive analysis, business application, business intelligence, decision support system and data warehouse where there is a need to refer customers' requirement [19]. There are four

method of text mining are discussed in this section such as Term based, Phrase based, Concept based and Pattern based.

A. Term Based

The dataset filtered and stemmed to obtain frequency of term used to represent as document is presented with term-by-frequency matrix [18]. While Term base method contributes terms into documents as to discover weights for each describe terms. Term frequency and inverse document frequency (TF-IDF) model generally been used to calculate frequent word in a document with inverse proportion of word whereby the model converts textual data to Vector Space Model (VSM) [20].

TF : Term Frequency refer to number of term in a documents [18] [21]. TF calculated as [22] :

$$TF(t,d) = \frac{n_t}{N_{(T,d)}} \qquad (1)$$

Nt represent number of frequency of term t in a document d, *N(T,d)* shows sum of total terms *T* in document.

IDF : Inverse Document Frequency refer to calculated (log(N/DF)) number of documents containing term as shown below [18] [21].

$$IDF = \log \frac{D}{n_d} \qquad (2)$$

D represent sum of documents in the corpus, *nd* number of document. Hence, TF-IDF calculated as:

$$TF\text{-}IDF = TF * IDF \qquad (3)$$

Yusheng Zhou with other researchers, has implemented TF-IDF algorithm to investigate helpfulness of online reviews and which later explores correlation between review title and content on review helpfulness [21]. Other than that, in social media such as Twitter and Facebook also utilize TF-IDF method to extract and categorize sentiment in reviews from unstructured data [20] [23]. Those studies indeed help decision making by identify polarity (positive, neutral, negative) of reviews using sentiment analysis method. The researcher also capitalizes TF-IDF extraction method with Amazon dataset to analysis customers' reviews on products in making decisions and improving performance of retailers [24]. Other than English language reviews, Bahasa reviews from Tokopedia website been collected to analysis and identify hidden pattern which supports companies using TF-IDF method using sentiment analysis process [17].

B. Phrase Based

Phrases based filtration rarely been implemented by researcher due low frequency, noisy appearance with synonym words and second class statistical properties [12] [25]. The advantage of phrase-based method would be less doubtful and more presentable and represents accurate result on phrase basis.

C. Concept Based

This method concerns more on relevant and extract valuable meaning of sentences using natural language processing [18] [12] [18] [25]. Concept based model consist of three components. The first component is to analyze synonym

structure of sentences by extracting verbs and arguments from text and then, to present Conceptual Ontological Graph (COG) model as one to one relationship among constituents [12] [26]. The third is to extract information using vector space model whereby it helps screening importance term in every sentences. Concept based mining measure closeness between the documents to evaluate usefulness of concept sentence level (Conceptual term frequency, ctf), document level (term frequency, tf) and corpus level (Document frequency, df).

D. Pattern Based

In text document reviews, pattern based method uses pattern deploying and pattern evolving for discover hidden pattern and trend [27]. Analysis on pattern based is discovered with method such as association rule, frequent item set mining, sequential and closed pattern mining [12]. This helps to reduces low frequency and misinterpretation by leading performance and support of related patterns. Pattern based techniques include algorithms such as Generalized Sequential Patterns (GSP), Prefix-Projected Sequential Pattern Mining (PrefixSpan), Suffix Arrays, Sequence Joining and nGram Linking [27]. Based on investigation, evaluation of those algorithms to Sequence Joining give preferably best result compared to others algorithm. The researcher works with pattern based method using association rules to determined hidden trend and sentiment in online text from social network [28].

IV. DATA AND METHODS (METHODOLOGY)

This research propose sentiment analysis techniques and term based method to identity polarity of reviews from E-commerce site whereby the customer reviews has two components, namely, ratings and reviews [5] [7] [9]. These two columns have been incorporated in this study for assessment. Fig. 1 shows methodology of sentiment analysis to be develop using Amazons' electronic customer reviews whereby the analysis process begin with data pre-processing (refer section B), than feature extraction (refer section C), sentiment classification (refer section D) and finally Evaluation score (refer section E). The method develops using python Jupiter notebook, Anaconda.Navigator 1.10.0 whereby it is free open-source distribution can be supported in windows, macOS and Linux. SciKit-Learn and Natural Language Tool Kit (NLTK) libraries used for develop this model.

A. Data Collection

The compiled data for this paper are extracted from Amazon electronic category; it is downloaded from Kaggle in English language [29]. The file was in Comma Separated Values (CSV) whereby it's convenient to consume in python. The sample data provide customers' rating of score 1 to 5 stars and reviews which was written in English language and covers electronic products which been purchased online using Amazon.com. In total, 34,633 customers' reviews were collected from the website. For each review, rating and product details are provided. The variable and description of dataset are shown in Tables I and II. Table II provides number reviews by classes from 1 to 5. The dataset has been chosen to identify polarity of reviews.
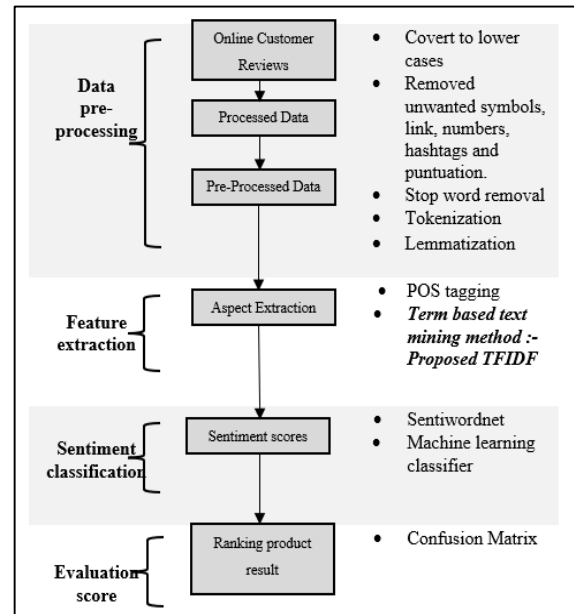

Fig. 1. Sentiment Analysis Process Flow.

TABLE I. DESCRIPTION OF DATASET

No	Variables	Description of variables	Variables choose for this study (Yes/No)	Description
1	Id	Products' ID	No	N/A
2	Name	Products' Name	No	N/A
3	Asins	Amazon standard identification numbers	No	N/A
4	Brand	Products' brand	No	N/A
5	Categories	Products' categories	No	N/A
6	Manufacturer	Amazon or Amazon Digital Services	No	N/A
7	Review Date	Date of review data	No	N/A
8	Review ID	Unique ID for review	No	N/A
9	Review Rating	Customers' rating	Yes	Dependent Variable :- The rating is given by customer to the purchase product in the form of score ranging from 1 to 5.
10	Reviewer Name	Name of the reviewer	No	N/A
11	Review Title	Title of the review	No	N/A
12	Review Text	Text of the review	Yes	Independent Variable:- The reviews is unstructured text where expressed sentiments on purchase product positive or negative.

TABLE II. CONTENT OF DATASET

Number of Reviews	34660
Number of Columns	21
Number of reviews with rating 5	23774 (68.66%)
Number of reviews with rating 4	8541 (24.67%)
Number of reviews with rating 3	1499 (4.33%)
Number of reviews with rating 2	410 (1.16%)
Number of reviews with rating 1	402 (1.18%)

B. Data Preprocessing

Pre-processing phase is very important method in sentiment analysis for present quality result and enhance accuracy of the classifier to customers, it applies to Amazon dataset as shown in Table I. It also converts unstructured data to structured data which is suitable format for feature extraction. The first step of data preprocessing is converting selected alphabets into lower cases and omit all unwanted symbols, links, numbers, hashtags and punctuation. Followed by, step as below:

- Convert to lowercase: The model will consider upper case as different words, hence converting to lowercase will remove noise in dataset.

- Removal of unwanted symbols, links, numbers, hashtags and punctuation: By removing those details can reduce the feature space and which does not help in performance of result.

- Stop word removal: Stop words will is not required for analysis, hence, it been removed from reviews for simplify of the text and improve performance of result. Example of stop words are 'a', 'is', 'are' and 'that'.

- Tokenization: The process of breaking the sentences into phrases and words.

- Lemmatization: Method for switching the words into root meaning. For example, 'used' to 'use'.

- POS tagging: Part-of-speech tagging where it will identify each words of reviews from noun, adjective, verb and adverb.

After preprocessing, feature extraction is takes place for sentiment analyses.

C. Aspect Extraction

Part-of-Tagging (POS) is preferred method to identify each words of review noun, adjective, verb and adverb [15] [30]. Table III and Table IV show sample of POS tagging results after preprocessing.

Based on word segmentation, the keywords from text been identified for further process. Then, term-based text mining method was performed to generate most frequent itemset with TF-IDF model. In between, N-gram play important role to represents number of texts as unigram (1), bigram (2) and trigram (3) [16] [20]. When applying N-gram it can represent numbers of word needed in frequent itemset. Fig. 2, 3 and 4 show comparative evaluation and analysis on most frequent words based on N-gram features. The summary result present that 'great' is most frequent unigram words used in Amazon

product dataset after removing all noisy or unwanted words whereas based on Fig. 3 'easy use' is most frequent bigram words and Fig. 4 'amazon fire tv' show most frequent trigram words in selected dataset.

TABLE III. SAMPLE RESULT -1 FROM DATA PRE-PROCESSING

Original Review	great for beginner or experienced person. Bought as a gift and she loves it
After convert to lower cases remove unwanted symbols	great for beginner or experienced person bought as a gift and she loves it
After removal of stop words	great beginner experienced person bought gift loves
After Lemmatization	['great', 'beginner', 'experienced', 'person', 'bought', 'gift', 'loves']

TABLE IV. SAMPLE RESULT-2 FROM DATA PRE-PROCESSING

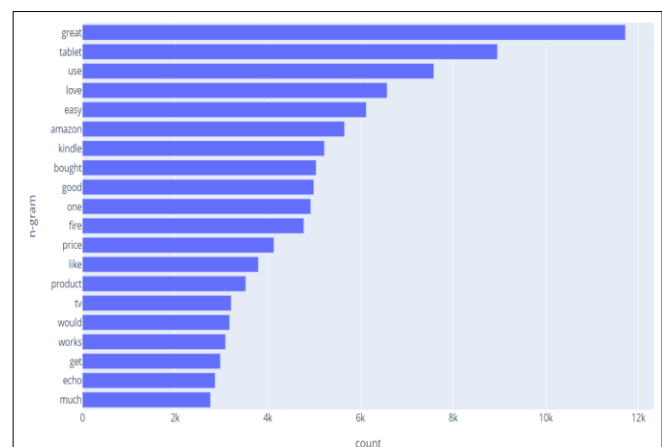|	
Original Review	great for beginner or experienced person. Bought as a gift and she loves it
Star Rating	5
After POS tagging	[('great', 'JJ'), ('beginner', 'NN'), ('experienced', 'VBD'), ('person', 'NN'), ('bought', 'VBD'), ('gift', 'NN'), ('loves', 'NNS')]

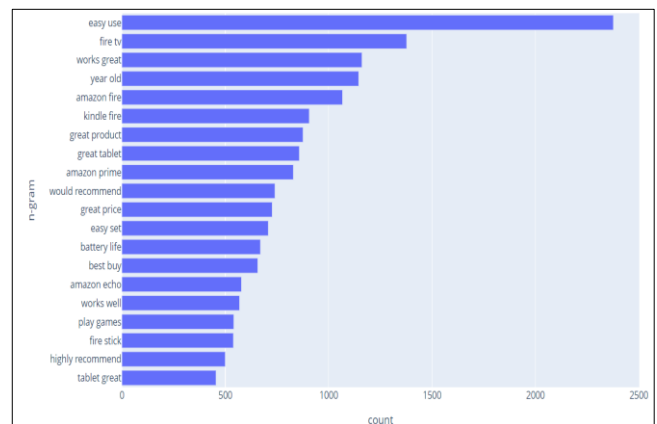Fig. 2. Most Frequent Unigram Words.
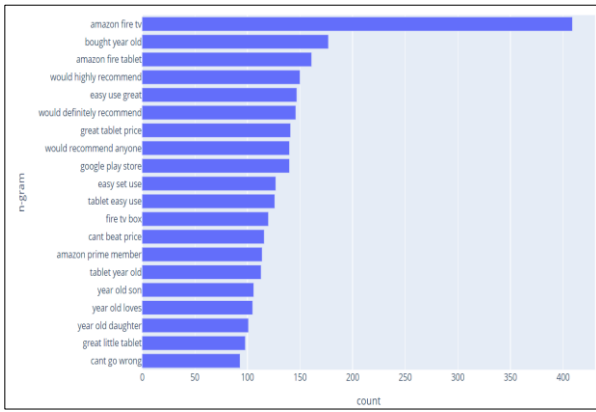

Fig. 3. Most Frequent Bigram Words.

Fig. 4.   Most Frequent Trigram Words.

D. Sentiment Classification

Lexicon based method SentiWordNet is used for expressed sentiment in words by scoring each words [31] [32] [33] [34]. SentiWordNet is like dictionary assign to each synset of WordNet with English language with positive and negative scores. Over 100,000 English words are in SentiWordNet for sentiment approaches with positive and negative scores. Based on POS tagging, for each synset are assign to scores from 0 to 1. The final positive and negative score of each sentence calculated with below equation:

$$\text{pos\_score} = \sum_{i=i}^{n} pos\_score\_senti(i) \qquad (4)$$

$$\text{neg\_score} = \sum_{i=i}^{n} neg\_score\_senti(i) \qquad (5)$$

$$\text{senti\_score} = \begin{cases} \frac{pos\_score}{tota\_score} \, if \, pos > neg \\ \frac{neg\_score}{total\_score} \, if \, neg > pos \end{cases} \qquad (6)$$

Total senti score below and equal to 0 is consider negative reviews and above 0 is consider as positive reviews. Based on Table IV total score of review is 0.875 which consider as positive review. In line with this finding, diverse machine learning classification methods been used to evaluate sentiment on reviews. First of all, training and testing samples are divided as trained sample 70% and test sample 30%. While four classifier Naïve Bayes, Support Vector Machine, Decision Tree and K-Nearest Neighbour are applied for find accuracy, precision, F1-score and recall. Machine learning (ML) approach is focus on future decision making to manage textual data with intelligence whereby different ML algorithms provide different result for comparison [9] [16] [35].

$$\text{Precision} = \frac{(True\ Positive)}{True\ Positive\ + False\ Positive} \qquad (7)$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive\ + False\ Negative} \qquad (8)$$

$$F_{Score} = \frac{2*Precision*Recall}{Precision+Recall} \qquad (9)$$

E. Evaluation Score

After comparing four machine learning approaches in Amazon reviews the result is evaluated with confusion matrix hence it present true positive (TP), false positive (FP), true negative (TN) and false negative (FN).

TP : The output where it correctly predicts the positive class.

FN : The output where incorrectly predicts the negative class and mislabeled as negative.

FP : The output incorrectly predicts the positive class whereby mislabeled as positive.

TN : The output where correctly predicts the negative class.

V. RESULT AND DISCUSSION

As discussed in Section III, the paper develop model on term-based sentiment analysis process with Amazon products. The textual data analyzed using python programming language and the result obtained from experiments shown in Table V. After data pre-processing, feature extraction demonstrates frequent itemset for polarize the reviews into sentiments [6] [9]. Data pre-processing first that required conducted data analysis for reduction of noise and for perform best result [5] [20]. In order detected overall sentiment of dataset 4 different type classification methods like Naïve Bayes, Support Vector Machine, Decision Tree, K-Nearest Neighbour approach with trained and tested using Amazon products dataset. Based on 4 machine learning classifier and by N-gram applied Decision tree show highest accuracy and Support Vector Machine show highest precision, recall and F1-Score. The result indicates Support Vector Machine model performs well with proposed method with Amazon product dataset. In between, unigram show us better performance whereas trigram shows us poor performance. The N-gram weight was most traditional features applied compared the best accuracy as the result present unigram perform well compared to other N-gram [16]. Hence, TF-IDF features with N-gram show different results as different machine leaning models. The approach is efficient and more robust for process and evaluates unstructured Amazon product reviews [17]. The similarity performance is only seen in Naïve Bayes model for precision, recall and F1-score. Table V shows average summary result for positive, negative and neutral classes for different machine leaning models with TF-IDF and N-gram.

TABLE V.   SUMMARY OF PERFORMANCE WITH TF-IDF AND N-GRAM

N-Grams	Accuracy (%)			Precision (%)			Recall (%)			F1-Score (%)		
	1	2	3	1	2	3	1	2	3	1	2	3
Naïve Bayes	78.869	79.01	79.23	63	63	63	79	79	79	70	70	70
Support Vector Machine	82.27	79.22	79.04	82	79	67	80	79	79	72	70	70
Decision Tree	99.98	99.91	99.59	78	73	74	77	64	39	78	67	45
K-Nearest Neighbour	63.33	33.83	58.75	75	66	70	50	19	29	56	19	33

VI. CONTRIBUTION

Online ratings and reviews evaluate customer view and influence sales performance. Another important contribution is had been information show attitude and behavior of customers toward purchase products to make decision which led to customers' satisfaction. Sentiment detection from online reviews also influences prospective customers on online decision purchasing and a better understanding on products. The classification of reviews positive, negative and neutral evaluate subjective information and describe nature of opinion in better way. Other than that the manipulation of unstructured data being processed with intelligent technology is for present quality text information for judging human behavior. Hence, based on analyze result retailers can improve their products' price, quality and services.

VII. CONCLUSION

There are two main dimensions for identifying sentiments in Amazon online products with reviews and star ratings. The proposed text preprocessing and machine learning model with TF-IDF and N-gram method have been used for investigate performance of Amazon products dataset as the result as discuses in section 5. The most important insight from this study would be classification of reviews in positive, negative and neutral for present overall performance result. For future work, different text mining method might improve accuracy result for Amazon product dataset. Furthermore, different sentiment analysis techniques such as hybrid deep learning models can be proposed together with N-gram features to perform better result and comparison. In overall, we believe all features applied in this study would improve the performance of E-commerce sites.

REFERENCES

[1] C. C. B. T. Alexander Ligthart, "Systematic reviews in sentiment analysis: a tertiary study," Springer, p. 57, 2021.

[2] D. H. G. Tayybaha Quyyam, "Sentiment Analysis of Amazon Customer Product Reviews: A Review," IJSRED, vol. 4, no. 1, p. 32, 2021.

[3] A. S. M. AlQahtani, "Product Sentiment Anlaysis for Amazon Reviews," IJCSIT, vol. 13, p. 16, 2021.

[4] S. S. A. B. U. Rahul Ramachandrana, "Exploring the relationship between emotionality and product star ratings in online reviews," ScienceDirect, p. 10, 2022.

[5] M. J. Budhwar and P. Singh, "Sentiment Analysis based Method for Amazon Product Reviews," IJERT, 2021.

[6] T. Sinnasamy and N. N. Amir Sjarif, "A Survey on Sentiment Analysis Approaches in e-Commerce," IJACSA, vol. 12, p. 6, 2021.

[7] U. Ahmed Chauhan, M. T. Afzal, A. Shahid, M. Abdar, M. E. Basiri and X. Zhou, "A comprehensive analysis of adverb types for mining user sentiments on amazon product reviews," World Wide Web, p. 20, 2020.

[8] Y. A. Amrani, M. Lazaar and K. E. El Kadiri, "Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis," ScienceDirect, p. 10, 2018.

[9] S. Wassan, X. Chen, T. Shen, M. Waqar and N. Jhanjhi, "Amazon Product Sentiment Analysis using Machine Learning Techniques," REVISTA ARGENTINA, p. 9, 2021.

[10] D. A. B. Tanvi Hardeniya, "Dictionary Based Approach to Sentiment Analysis - A Review," IJAEMS, p. 6, 2016.

[11] P. Carracedo, R. Puertas and L. Marti, "Research lines on the impact of the COVID-19 pandemic on business. A text mining analysis," ScienceDirect, p. 8, 2020.

[12] A. Raut and P. V. Shinde, "Effective Methods and Techniques in Text Mining," IJRITCC, vol. 5, no. 3, p. 4, 2017.

[13] H. S.Choi and S. Leon, "An Empirical Investigation of Online Review Helpfulness: A Big Data Perpective," ScienceDirect, p. 12, 2020.

[14] M. Ahmad, S. Aftab and N. Hameed, "Sentiment Analysis using SVM : A Systematic Literature Review," IJACSA, vol. 9, p. 7, 2018.

[15] M. P Abraham and U. K. Reddy, " Feature Based Sentiment Analysis of Mobile Product Reviews using Machine Learning Techniques," IJATCSE, vol. 9, p. 8, 2020.

[16] D. Gamal, M. Alfonse, E.-S. M. E1-Horbaty and A.-B. M.Salem, "Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-Gram Features," ScienceDirect, p. 9, 2019.

[17] P. Kencana Sari, A. Alamsyah and S. Wibowo, "Measuring e-Commerce service quality from online customer review using sentiment analysis," IOP, p. 7, 2018.

[18] S. D. K. and D. S. , "The Comparison of Term Based Methods Using Text Mining," IJCSMC, vol. 5, no. 9, p. 5, 2016.

[19] A. Masood Khan and K. Rahat Afreen, "An approach to text analytics and text mining in multilingual natural language processing," ScienceDirect, p. 3, 2020.

[20] R. Ahuja, A. Chug, S. Kohli, S. Gupta and P. Ahuja, "The Impact of Features Extraction on the Sentiment Analysis," ScienceDirect, p. 7, 2019.

[21] Y. zhou, S. Yang, Y. Li, Y. chen and J. Yao, "Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining," ScienceDirect, p. 11, 2020.

[22] M. Mujahid, E. Lee, F. Rustam, P. B. Washington, S. Ullah, A. Ahmad Reshi and I. Ashraf, "Sentiment Analysis and Topic Modeling on Tweets about Online Education during COVID-19," MDPI, p. 25, 2021.

[23] S. T. Y. L. H. S. Z. G. X. Y. J. B. C. C. and X. H. , "Interpreting the Public Sentiment Variations on Twitter," IEEE, vol. 26, p. 13, 2014.

[24] V. Raghuraman, S. Pattanayak, A. VK and A. C Patil, "Sentiment Analysis on Amazon Product Reviews," IEEE, p. 6, 2020.

[25] R. and S. , "A Study Of Text Mining Methods, Applications, and Techniques," IJESRT, p. 6, 2017.

[26] S. Hassan, F. Karray and M. Kamel, "Concept Mining using Conceptual Ontological Graph (COG)," ResearchGate, p. 11, 2014.

[27] P. Ozdzynski and D. Zakrzewska, "Using Frequent Pattern Mining Algorithms in Text Analysis," Information systems in Management, p. 10, 2017.

[28] J. A. Diaz-Garcia, M. Ruiz and M. J.Martin Bautista, "Non-Query-Based Pattern Mining and Sentiment Analysis for Massive Microblogging Online Texts," IEEE, p. 17, 2020.

[29] Datafiniti, "Kaggle," kaggle, 2020. [Online]. Available: https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products. [Accessed 2022].

[30] S. Vanaja and M. Belwal, "Aspect-Level Sentiment Analysis on E-Commerce Data," IEEE, p. 5, 2018.

[31] S. K. Kochhar and U. Ojha, "Index for objective measurement of a research paper based on sentiment analysis," ScienceDirect, p. 7, 2020.

[32] S. A. Firmanto and R. Sarno, "Prediction of Movie Sentiment based on Reviews and Score on Rotten Tomatoes using SentiWordnet," IEEE, p. 28, 2018.

[33] N. M. Elfajr and R. Sarno, "Sentiment Analysis using Weighted Emoticons and SentiWordNet for Indonesian Language," IEEE, p. 26, 2018.

[34] P. Mehta and S. Chandra, "Parameter Tuning in Updating the Sentiment Polarity of Objective Words in SentiWordNet," IEEE, p. 43, 2015.

[35] N. Nandal, R. Tanwar and J. Pruthi, "Machine learning based aspect level sentiment analysis for Amazon products," ResearchGate, p. 8, 2020.