Analyzing Sentiment in Terms of Online Feedback on Top of Users' Experiences

Mohammed Alonazi
Department of Information Systems, College of Computer Engineering and Sciences,
Prince Sattam bin Abdulaziz University, Al-Kharj, 16273, Saudi Arabia

***Abstract*—Since most businesses today are conducted online, it is crucial that each customer provide feedback on the various items offered. Evaluating online product sentiment and making suggestions using state-of-the-art machine learning and deep learning algorithms requires a comprehensive pipeline. Thus, this paper addresses the need for a comprehensive pipeline to analyze online product sentiment and recommend products using advanced machine learning and deep learning algorithms. The methodology of the research is divided into two parts: the Sentiment Analysis Approach and the Product Recommendation Approach. The study applies several state-of-the-art algorithms, including Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, Bidirectional Long-Short-Term-Memory (BI-LSTM), Convolutional Neural Network (CNN), Long-Short-Term-Memory (LSTM), and Stacked LSTM, with proper hyperparameter optimization techniques. The study also uses the collaborative filtering approach with the k-Nearest Neighbours (KNN) model to recommend products. Among these models, Random Forest achieved the highest accuracy of 95%, while the LSTM model scored 79%. The proposed model is evaluated using Receiver Operating Characteristic (ROC) - Area under the ROC Curve (AUC). Additionally, the study conducted exploratory data analysis, including Bundle or Bought-Together analysis, point of interest-based analysis, and sentiment analysis on reviews (1996-2018). Overall, the study achieves its objectives and proposes an adaptable solution for real-life scenarios.**

Keywords*—Sentiment analysis; product review; machine learning; recommendation system; collaborative filtering; exploratory data analysis

I. Introduction

In this era of modern computational technology, technological advancement can be seen everywhere; even the business sector is taking the initiative to enhance its revenue in computing and technology [1]. The term sentiment analysis has extensively been utilized to track out social media, allowing businesses to extract hidden information from the recorded data or identify critical information before coming into the limelight [2]. Thinking about giant tech companies such as Facebook, Google, Apple, and Microsoft, they have a huge amount of datasets. Every day, lots of data is being recorded to their central database. Besides, manually analyzing these data is time-consuming [3]. With a massive amount of dataset, it is quite difficult to manually extract meaningful insights that can help them make a business oriental decision. Turning into a product-based company, it can be stated that the product-based company tends to develop its product and launch it into the market [4]. If the thing is grocery or jewelry items so in this case, users will be giving their opinion based on the items whether the product is caught their attention or not. Suppose the clients explore a large e-commerce platform like Amazon.com. In this case, it is noticeable that the end-users threw their comments or reviews related to a specific item. Other individuals take themselves towards the advertisement phenomenon [5]. In turn, this brings us to a big question, whether it or not possible to handle such massive amounts of ratings manually and extract the business insights. So, the automated system can be the possible solutions to overcome these issues [6]. Thus, the impact of information on user sentiment and physical environments is not limited to modern technology.

The computational approach can be taken into consideration. Nowadays, machine learning algorithms have widely been utilized in biomedical imaging, forecasting things or even critical disease prognosis [7]. For the case of product analysis, it has shown their promising performance beforehand. Researchers are now using computing power to take their analysis to a satisfactory level from where meaningful insights can be extracted easily by analyzing a large number of datasets [8]. The priority of this research is to analyze the product sentiment using machine learning algorithms and propose a recommendation system for the stakeholder to make a better decision while doing online business. In this research, conventional Machine Learning (ML) algorithms were adopted to analyze the online product, and our study will significantly contribute to the research community. This is the motivation of this proposed study.

On the other hand, there are three contributions have been addressed in this study, and these are following:

- Three types of data analysis have been completed and through which business owners can make a variety of decisions towards their product.
- Various machine learning algorithms were applied to check their credibility to analyze the Sentiment, and satisfactory accuracy was turned out to be successful. This research also ensures the robust machine learning pipeline that achieved a good accuracy, and the concentrated model can be deployed to a webserver to achieve sustainable goals.
- The different assessment pointers bend assessed the proposed show, and at long last, a proposal framework has been submitted by coordination overall sifting strategy. By taking after the proposal framework, the

This study is supported via funding from Prince Satam bin Abdulaziz University project number (PSAU/2023/R/1444).

partner will be able to supply important data to their enlisted client, which can improve the request for specific things.

The manuscript is classified into six interconnected sections. Section II presents the exciting works with research gap analysis. Section III depicts the overall methodology of the proposed system with proper discussion. Section IV shows the results associated with our proposed solutions, including data-driven analysis and approaches. Section V represents the observations and discussion on the results. Finally, the conclusion of the research with future work will be discussed in Section VI.

II. Literature Review

This section illustrates the background study of the previous works related to the proposed model. In this section, the research gaps have been extracted with proper discussion. Many great contributors have traced fruitful online product contributions or sentiment analysis contributions.

The authors of the paper in [9] worked on an efficient way to optimize the accuracy of the sentiment analysis in Egyptian Arabic. The proposed work was identically based on the conventional semantic orientation and machine learning techniques, and the authors had achieved the highest accuracy of 92.98% while working with Support Vector Machine (SVM). The purpose of the article in [10] was to examine the attitudes of buyers regarding electrical devices by analyzing various sale tweets. The experimental results of the proposed research will be valuable to a variety of business organizations in making business decisions that will ultimately increase the sales of the products they offer. The author of the paper also claimed that they had achieved the highest accuracy of 86%, 91%, and 91% with the Logistic Regression (LR) in the phone, laptop, and television, respectively.

The paper in [11] aimed to extract the text features into the semantics of words. The authors adopted a Word Sense Disambiguation (WSD) technique to extract the features from the reviewing sentences. A supervised learning approach has been adopted to analyze the product reviews and utilized 10-fold cross-validation to validate the results. The authors had significantly optimized the performance by 10.6%, while precision was 10.9% higher and recall was 9.2% higher than baseline approaches. The author of the paper in [12] had presented a significant comparison among several conventional deep learning-based models for word embedding in product sentiment analysis. Thus, they adopted data augmentation techniques to enrich the dataset and classify it into identical classes. The research also claimed they found the highest accuracy of 96% while working with CNN-RNN based BI-LSTM algorithms.

In paper [13] proposed an Adaptive Neuro-Fuzzy Inferences System (IANFIS) model to produce a way of analyzing the sentiment of online products. The method is identically based on natural language processing to track the user's opinion. The authors classified the dataset into three interconnected parts: contents, grades, and collaborations. Then, they applied deep learning algorithms to make a prediction on the negative and positive comments from the users. The research also performs a comparison among the existing solutions. The paper in [14] aimed to implement a model for analyzing the sentiment of the users in movie reviews. The authors extracted the feeling and feedback from existing text patterns. The models had the ability to detect several types of feeling like negative, positive, and even neutral. To accomplish this goal, they utilized different machine learning algorithms and classifiers, and mechanisms of natural language processing.

The authors in [15] have presented a machine learning-based online product sentiment analysis. In this work, they showed the labeled product reviews in several websites with the help of supervised and unsupervised (lexicon-based) based algorithms. The models were then applied to the iPhone 5s reviews collected from the existing popular online shops. The authors further extracted the combination of unigram and bigram features, which placed the best results while dealing with machine learning-based classifiers.

Paper in [16] identified three subtasks that must be addressed: the definition of the target; the separation of good and bad news content from good and bad sentiment expressed on the target; and the analysis of clearly marked opinion that is defined explicitly, without the need for interpretation or the use of world knowledge. The authors in [17] created a new strategy that combines previous approaches to provide the best coverage results and competitive agreement. They had also proposed iFeel, a free Web service that provides an open API for retrieving and comparing findings from several sentiment methods for a given text. In paper [18], researchers categorized movie reviews using features based on these taxonomies paired with traditional "bag-of-words" features, and reported 90.2 percent accuracy. Furthermore, they discover that some types of assessment appear to be more critical for sentiment classification than others.

The contributions of the papers in [19] had only focused on the development of a notable features selection on online product sentiment analysis. But the researchers didn’t focus on the correct terms of algorithms and algorithm tuning to optimize sentiment analysis accuracy level. In sharp contrast, the manuscript presented three forms of data analysis. These analyses will allow business owners to make several judgments regarding their specific products. Various machine learning methods were used to check their reliability and determine adequate accuracy. This research also assures a robust machine learning pipeline that the condensed model can be deployed to a webserver to fulfill long-term objectives for product sentiment analysis.

Based on the literature review, several studies have been conducted to improve sentiment analysis and product recommendation using machine learning techniques. The authors of one study achieved the highest accuracy of 92.98% in sentiment analysis of Egyptian Arabic using Support Vector Machine. Another study analyzed buyer attitudes towards electronic devices through sale tweets and achieved high accuracy of 86%, 91%, and 91% for phone, laptop, and television respectively using Logistic Regression. One study aimed to extract text features through Word Sense Disambiguation and achieved a significant performance improvement compared to baseline approaches. Lastly, a study compared several deep learning-based models for word embedding in product sentiment analysis and achieved an accuracy of 96% using a CNN-RNN based BI-LSTM algorithm.

Overall, the studies show the effectiveness of machine learning techniques in sentiment analysis and product recommendation. These approaches can help businesses make data-driven decisions to improve sales and customer satisfaction. However, there is still room for further research to optimize the accuracy and efficiency of these techniques.

III. Methodology

This section presents the overall design of the proposed model, including the illustration. Fig. 1 illustrates a block diagram of the proposed research through which this study was conducted. By looking at Fig. 1, it can be observed that this research was carried out through three interconnected stages. The experimental dataset was collected and preprocessed to model the data in the first stage. In the second phase, product sentiment analysis was accomplished. Finally, in the third stage, a recommendation system is proposed as the priority of this research is to provide a model that can analyze the product sentiment and transform the traditional business into a data-driven approach.

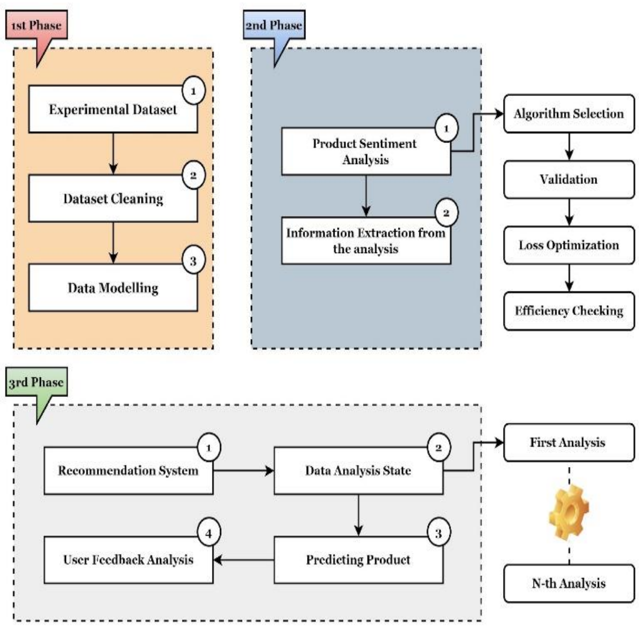


Fig. 1. Block diagram of the proposed research, including the experimental analysis to product recommendation approach.

A. Sentiment Analysis Approach (SAP)

This section explains the steps of approach, including experimental data, data preparation and concentrated research algorithms in more details.

1) Experimental dataset: From May 1996 through July 2018, Amazon's "Clothing, Shoes, and Jewelry" category received 2.5 million product ratings and information from 2.5 million customers. This collection includes reviews (scores, description, and sentiment comments), product metadata (descriptions, controls and monitors, pricing, branding, and image attributes), and links [20].

2) Data preparation: Perusing different JSON records from a single JSON record, 'ProductSample.json,' and including them in the list in such a way that each list of the list has the substance of a single JSON record [21]. Following that, iterate over the list, loading each index as JSON, extracting the data from each index, and creating a list of Tuples containing all the data from the JSON files. Again, each cycle begins with a clean JSON file converted to the right JSON format using some substitutions. Finally, a data frame is created using the list of Tuples obtained in the previous step.

3) Concentrated research algorithms (CRA): The Concentrated Research Algorithms (CRA) indicates the suggested model that has been adopted in this study to analyze the research data [22]. It is to be specified that various conventional techniques were applied in this investigation; among them, the Naïve Bayes and Decision tree algorithms were found to be satisfactory [23]. So, the mathematical interpretation and model optimization procedure were highlighted in this section. These models have a benchmark performance that appeared in the previous research. In addition, model performance is depending on the data distribution, furthermore, the Naïve based and Decision tree are capable enough to handle the product sentiment analysis data.

- Naïve Bayes Algorithm: Nave Base Classifier is a classification-type machine learning algorithm [24]. This algorithm is based on the Base Theorem. Simply put, the base theorem is a method of determining the probability of one event (X) occurring and another event (Y) occurring. If clouds are seen in the sky, there is a possibility of rain. The base theorem can be mathematically written as,

$$\frac{P(y)P(y)}{P(X)} \tag{1}$$

The above equation is a simple base equation used to determine probability in the case of a conditional event only. In practice, most datasets are multivariate, in which case the equation becomes a bit more complicated. Then we can write the equation like this:

$$P(y \mid x_1, \dots, x_n) = \frac{P(x_1|y)P(x_2|y)\dots P(x_n|y)P(y)}{P(x_1)P(x_2)\dots P(x_n)} \tag{2}$$

A few of the highlights of Naïve Base Classifier: It is exceptionally simple to execute and works moderately quick, works well indeed on small datasets, gives a small less exactness than other calculations, and all traits in Naïve Base are considered commonly autonomous but within the genuine world Isn't.

- Decision Tree: Both classification and regression problems can be solved with the classification and regression tree or CART algorithm [25]. In short, many people call it the Decision Tree. The decision tree looks a lot like the branches of a tree, which is why the word 'tree' is associated with its name. The decision trio starts from the 'root node' just as the tree starts from the root. From the root node, the branches of this tree spread through different decision conditions; such nodes are called decision nodes, these nodes are called leaf nodes after making a final decision. Other Parameters of the Decision Tree: Splitting - The process of moving a dataset across a series of variables, starting from the root node, is called splitting [26]. Entropy - Entropy is the amount of chaos. When the tree is split, the amount of data of the same type/class in each node is purity. All the data in a pure node are of the same class. The lower the purity, the higher the entropy. Again, the lower the entropy, the higher the purity. Information Gain - The measure of righteousness is information gain [27]. The higher the information gain, the purer nodes the tree can create. Guinea Index - Guinea is the probability of all node members being in the same class. This value ranges from 0 to 1. Guinea value 0 means all the members of that node belong to the same category, and Guinea value 1 means the members of that node are randomly distributed or of different classes, i.e., entropy is much higher. If the value of Guinea is 0.5, then the members of the two classes are equal (if the number of classes is 2) [28].

$$Entropy = p(A)log(p(A) - p(B)log(p(B \quad (3)$$

Information Gain= Entropy Before Split - Entropy After Split

$$Gini\ Index\ = 1 - \sum(P(x = k))^\wedge 2 \quad (4)$$

B. *Product Recommendation Approach (PRA)*

This section highlighted the product recommendation procedure. It can be said that the PRA is vital towards business transformation because user behavior and pattern cannot easily be identified if the stakeholder did not design any recommendation system. By looking at Fig. 2, it is noticeable that a flow chart has been proposed in terms of the recommendation system. The collaborative filtering approach is selected that will filter out the user ID based on age, gender, location and rating score etc. After having all of that information, the system will make a comparison set for the specific users. However, the detailed sequence and consequences are shown in Fig. 2. The diagram in Fig. 2 suggests that the proposed pipeline aims to improve e-commerce and enhance customer experience by using a collaborative filtering method for item recommendation. It implies that the pipeline could help businesses identify the most relevant products for their customers, thus improving their overall shopping experience. The use of collaborative filtering suggests that the pipeline may leverage the behavior of similar users to provide personalized recommendations, ultimately leading to increased customer satisfaction and potentially higher sales. Overall, the diagram title hints at a promising approach to improving the online shopping experience and driving business growth.

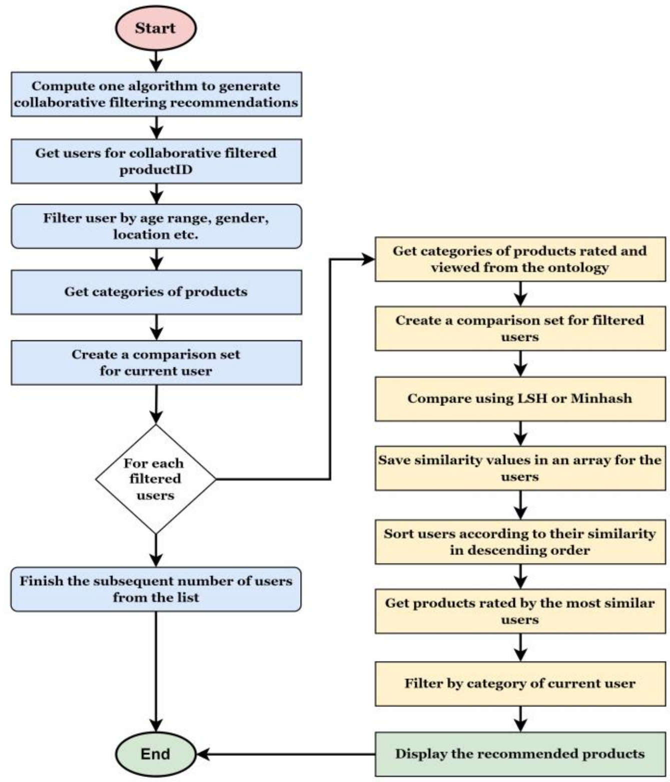
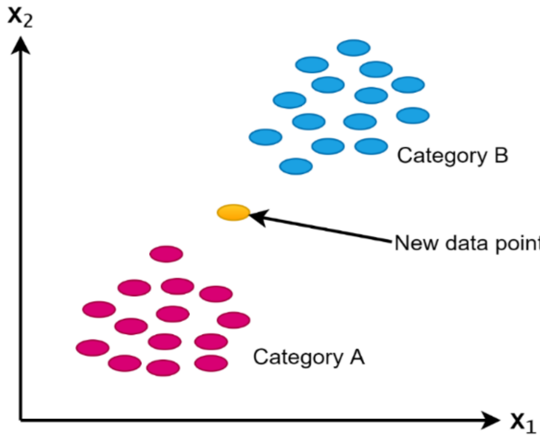


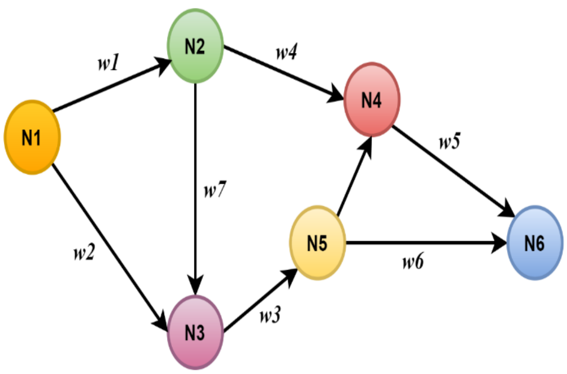
Fig. 2. Efficient pipeline for the item proposal approach towards commerce change and client design acknowledgment through the collaborative sifting method.

The research system utilizes several methods like K-Nearest Algorithm (KNN) [29], the Jaccard's coefficient, the Dijkstra algorithm, and the cosine similarity. The aim is to suggest based on users' behavior patterns. In recommendation systems, the most common types are the Collaborative Filtering Method (CFM), the Content-Based Filtering Approach (CFA), and Hybrid Recommendation System (HRS) [30]. This filtering approach generally focuses on collecting and analyzing user experience information, behaviors, or interests and predicting what they would like based on similarity with other users. The collaborative sifting approach's imperative advantage is that it does not depend on machine analyzable substances and can accurately prescribe complex things without requiring an "understanding" of the thing itself.

A typical recommendation engine processes data over the following four steps: selection, storage, analysis, and filtering. We have applied the K-Nearest Algorithm (KNN), the Jaccard's coefficient, the Dijkstra algorithm, and the cosine similarity to forecast the shortest path from the user’s current location to the user’s desired destination and as well as to suggest places to the users based on the rating. In Figure 3 a) shows the corresponding KNN's cluster filtering working procedure and b) shows the Dijkstra algorithm ‘s workflow.



(a)



(b)

Fig. 3. a) Perform K –nearest algorithm's cluster filtering, b) Dijkstra algorithm.

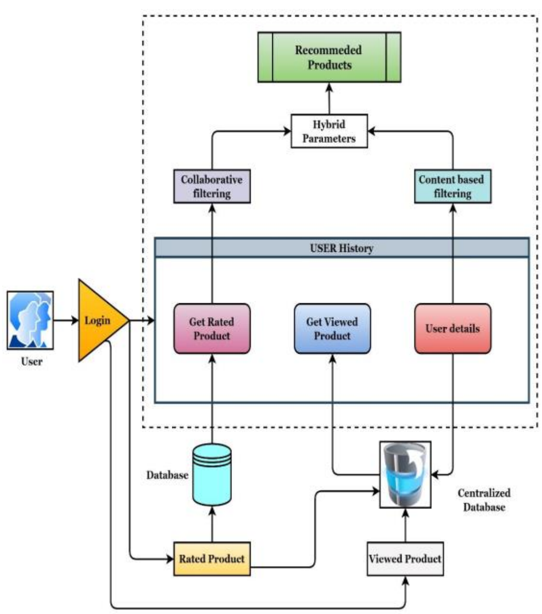


Fig. 4. Architecture diagram of the product recommendation system using various filtering methods.

In Fig. 4, it can be stated that the suggested recommendation engine is the result of a number of interconnected methodologies being used. In order to utilize the system, the user must first complete their registration before being allowed to proceed. After logging into a system, a log file will be created automatically to keep track of the user's patterns of behavior. At this point, we have implemented a cooperating filtering method as well as a content-based filtering approach. Users' comments and product ratings will be taken into consideration by the recommendation system, which will take action based on both forms of data. After obtaining all of the necessary parameters, the system will proceed to the next stage of proposing a specific product to a user who has expressed interest in it. The architecture also assures that collaborative and content-based filtering are the middle layer of this architecture that are the responsible for the product sentiment analysis. However, the hybrid parameters will then send to the model for identifying the sentiment.

IV. Result Analysis

The result analysis section is categorized into several parts: Classification Metrics Interpretation (CME), Measuring the Efficiency, Interpretation of Sentiment Analysis, Observation & Discussion. The precision of expectations from the classification calculations is evaluated by applying a classification report. The report illustrates the exactness, review, and f1-score of the key classification measurements per lesson. These measurements are computed by utilizing genuine and untrue positives and genuine and wrong negatives. The measurements comprise of four components: genuine positive, untrue positive, genuine negative, wrong negative, and wrong negative. The taking after Condition (1), (2), (3), and (4) was considered for finding the exactness, review, and f1-score. In Table I, the classification report of the machine learning calculation, is depicted. In this table, we can clearly observe that the machine learning algorithms like Random Forest (RF) and Logistic Regression provide better results, such as 95% and 94%, compared to the other algorithms like LSTM and CNN_LSTM. This is because for this dataset, we have found that low-cost classifiers work far better at deep computation because of the small size of the dataset. Thus, we have achieved the highest accuracy from conventional classifiers.

Precision: It is the relationship between the true positive estimate of the model and the overall positive estimate (both accurate and wrong) .It is articulated as:

$$Precision\ (P) = \frac{TP}{TP+FP} \tag{5}$$

Recall / Sensitivity: The probability of being capable of predicting is a positive ratio. It is given in mathematical form as:

$$Recall\ (R) = \frac{TP}{TP+FN} \tag{6}$$

F1-score: As a general rule, the harmonic mean for Accuracy and Review provides a much better; a significantly better; higher; a stronger and more enhanced gauge than the Precision Metric of the incorrectly categorized occurrences. It is given, mathematically, as:

$$F1 - \text{score} = 2 \cdot \frac{Precision.Recall}{Precision + Recall} \tag{7}$$

Accuracy: It is the sum of all the cases in which the predictions were right. It is given as:

$$Accuracy\ (A) = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

TABLE I. CLASSIFICATION REPORT OF THE CONCENTRATED ALGORITHMS

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naïve Bayes (NB)	93	92	93	93
Logistic Regression (LR)	94	93	94	94
SVM	93	93	93	93
Decision Tree (DT)	91	89	89	90
Random Forest (RF)	95	94	95	95
BI-LSTM	76	70	71	70
CNN_LSTM	77	77	75	76
Stacked LSTM	76	76	76	76
LSTM	79	78	78	77

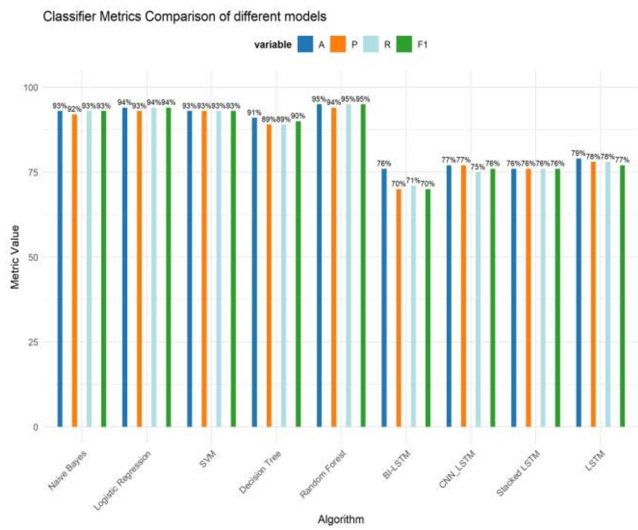
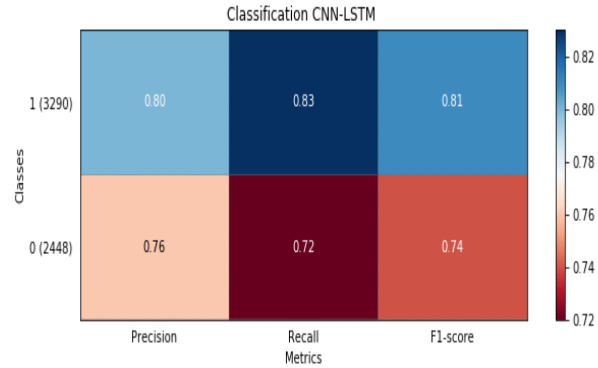


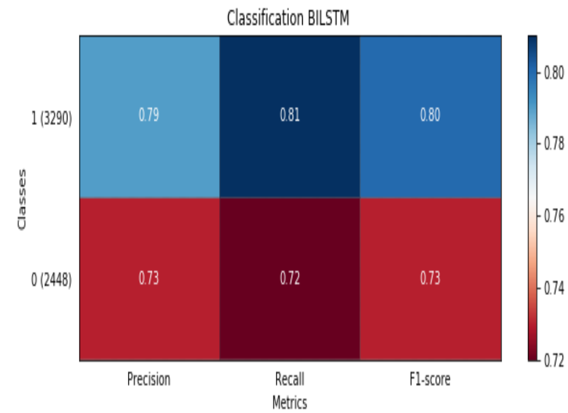
Fig. 5. Performance analysis from the different models.

On the other hand, Fig. 5 highlights the confusion matrix on top of the Random Forest and LSTM model. The confusion matrix consists of four values: TP, TN, FP, and FN. We can identify the total sensitivity and specificity ratio by following the Confusion matrix. This is another model evaluation indicator, and in the field of data science, this matrix has been utilized extensively to measure a specific model. The confusion matrix consists of four values, True positive, True negative, False positive, and False negative are the four values in the confusion matrix. Fig. 6 (a) and (5) illustrates the significant proportion of true positive and false negative values.

ROC-AUC curve is used to determine how good a model is. This evaluation indicator distinguishes the positive and negative data points from the dataset. If the ROC curve goes to 1.0, the model can accurately differentiate the positive and negative data points. Fig. 7 (a) and (b) are almost close to 1.0 or the area under the curve. It can be stated that this model is applicable to use in real life.

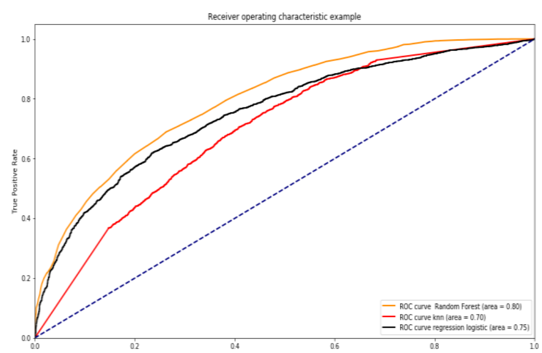


(a)

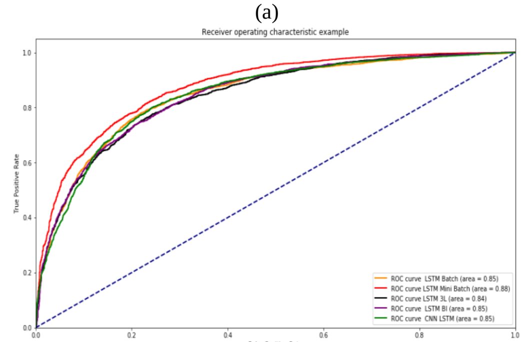


(b)

Fig. 6. (a) Confusion matrix on CNN+LSTM (b) Visualizing the confusion matrix on top of the BI-LSTM model towards product sentiment analysis.



(a)



(b)

Fig. 7. (a) Evaluating curve on top of the Random Forest model (b) Measuring the model and visualizing the ROC-AUC curve on top of the LSTM model.

A. Interpretation of Sentiment Analysis

This analysis is divided into three stages: Sentiment analysis on reviews (1996-2018), exploratory analysis on product reviews (1996-2018). This is of particular note since it is the one with the greatest number of prior sales on Amazon, as determined by an examination of "Bundles" or "Bought-Together."

B. Analysis 1: Sentiment Analysis on Reviews (1996-2018)

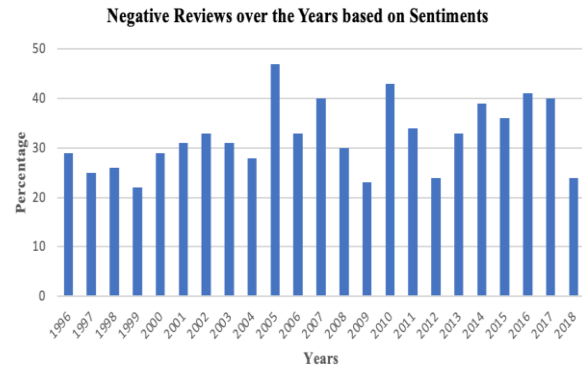
Fig. 8 indicates the exploratory product data analysis in terms of user feedback. In Fig. 8 (a) and Fig. 8 (b), the sentiment analysis on user reviews have been demonstrated for the year of between 1996 to 2018. It is noticeable that, the year 2000 was the least number and 2001 reach out at the highest number of negative product reviews based on the sentiment. At the same time, the year 2000 was the top year where significant number of positive reviews based on the users sentiment was recorded. After that, the percentage was decreased and remained same for the next consecutive years. Besides, Fig. 7 (c) and (d) shows the word cloud visualization based on the positive and negative reviews where it the keywords are tagged and, in this way, it would be convenient to find out the negative and positive keywords from the dataset.

For Positive Word Cloud, some well-known terms such as adored, idealized, decent, wonderful, best, outstanding, and so on were used to describe the goods. A large number of people who were polled were pleased with the prices of things sold on Amazon. Bra, coat, bag, and outfit are some of the most commonly discussed goods. Disappointment, terrible fit, horrible deformity, return and etc. are some of the well-known adjectives used to describe the things. Some of the most talked-about goods were shoes, binoculars, bras, batteries, and so on. Predominant item in terms of how others feel about it. There are 953 positive reviews for the Speak Unisex Chuck Taylor Classic Colors Sneaker. There are 672 positive reviews for the Talk Unisex Chuck Taylor All-Star Howdy Best Dark Monochrome Sneaker. There are 65 negative reviews for the Yaktrax Walker Footing Cleats for Snow and Ice. There are 44 negative reviews for the Speak Unisex Chuck Taylor Classic Colors Sneaker. Welcome to the Best Dark Monochrome Sneaker, with a total of 247 honest reviews. [31], which state achieved that 72.7 % was positive, 5 % was negative, and 22.3 % was neutral. Overall, Sentiment for reviews on Amazon is on the positive side as it has very few negative sentiments.

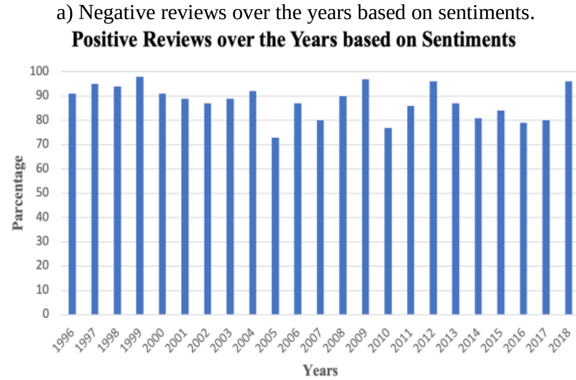
The drift for Rate of Audit over a long time, positive surveys rate has been lovely reliable between 70-80 all through a long time. Negative surveys have been diminishing recently since final three a long time; possibly they worked on the administrations and issues.

C. Analysis 2: Point of Interest-based Analysis

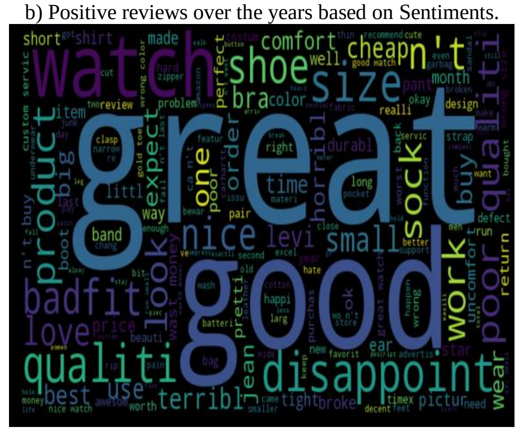
The lexical density of a language is a concept in computational linguistics that measures the structure and complexity of human conversation in [31]. Functional and content words are used to calculate a written or spoken composition's lexical density.



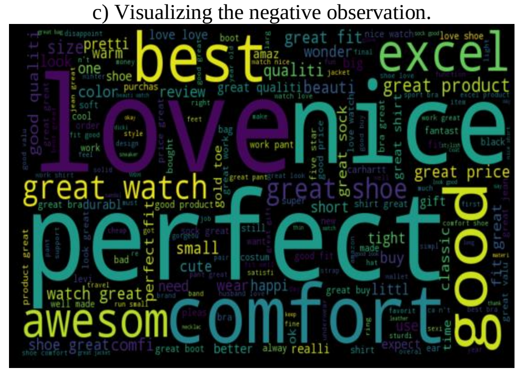
a) Negative reviews over the years based on sentiments.



b) Positive reviews over the years based on Sentiments.



c) Visualizing the negative observation.



d) Visualizing the postive observation.

Fig. 8. a) Negative reviews over the years based on the Sentiment (b) Positive reviews based on the Sentiment (c) Visualizing the negative observation through the word cloud visualization approach (d) Illustrating the positive observation with the help of word cloud visualization.

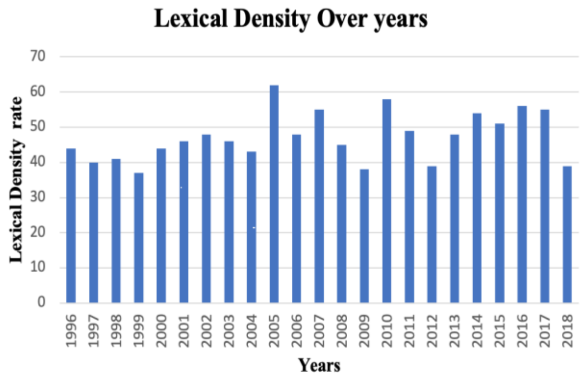


Fig. 9. Lexical density over years.

By looking at the Fig. 9, it can be observed that lexical density over years have been displayed. In 2018, the significant amount just over the 40 but in 1996, there is a downward trend at nearly 36 and remained steady for the next subsequent years.

D. *Analysis 3: Bundle' or 'Bought- Together' based Analysis*

In the Table II, Bundle or Bought together based analysis has been interpreted in terms of up vote, helpful rating, total votes, and percentage. Based on the reviewer ID, it is clearly demonstrated the helpful rating, and these has been illustrated owing to the fact that for the case of analysis, product sentiment analysis and review records are considerably required so that essential information can be extracted.

Taking a close look at the Fig. 10, helpfulness and average length have been displayed. The findings show that the effectiveness of review length is influenced by product category; longer evaluations are more helpful for think products. Furthermore, review helpfulness is linked to the degree of consistency between individual review ratings and total product ratings. In contrast the Fig. 11 shows the data analysis with correlation between the ASIN.

E. *Analysis 3: Exploratory Data Analysis*

1a) The common survey rating for the foremost commonly checked items is between 4.5 and 4.8, with little variation. 1b) Whereas there's a little converse affiliation between the recurrence level of ASINs and regular audit appraisals for the primary four ASINs, this relationship isn't critical since the normal survey for the prior four ASINs is evaluated between 4.5 and 4.8, which is respected greatly by and large audits. 2a) As illustrated within the bar chart (beat), ASINs with lower frequencies have much more change in their routine survey evaluations on the point-plot chart (foot), as demonstrated by the length of the vertical lines. As a result of the tall fluctuation, we accept that our research's normal audit evaluations for ASINs with lower frequencies are not imperative. 2b) On the other hand, we assume that the lower frequencies for ASINs are inferable to lesser quality items. 2c) Moreover, the final four ASINs have no change due to their significantly lower frequencies. Whereas the survey appraisals are a culminate 5.0, we ought not to consider these audit appraisals' significance due to the lower recurrence as shown in 2a). Based on our information examination between ASINs and audits. Rating, we have taken note that numerous ASINs with the common event had huge fluctuations; in this way, we decided that these moo event ASINs are not imperative in our think about due to the little test measure. Additionally, we found nearly no interface between ASINs and surveys. Rating in our relationship considers which is reliable with our findings.

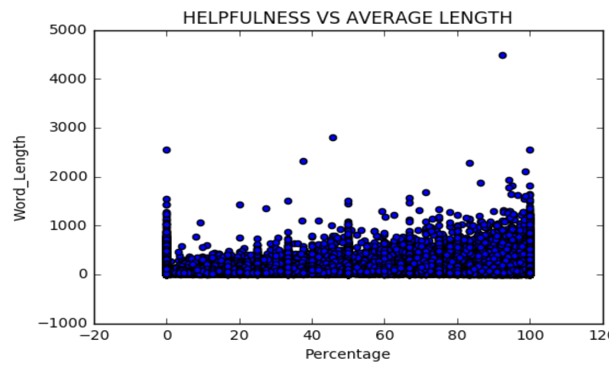


Fig. 10. Helpfulness and average length.

TABLE II. EXPLORARYTORY DATA ANALYSIS WITH THE DATASET AND THEIR FEATURES

Reviewer ID	Rating helpful	Upvote	Total Votes	Percentage	Rating
A2XVJBSRI3SWDI	5.0	0.0	N/A	0.0	N/A
A2G0LNLN79Q6HR	4.0	1.0	0.0	N/A	N/A
A2R3K1KX09QBYP	2.0	1.0	100.0	N/A	N/A
A19PBP93OF896	1.0	1.0	100.0	N/A	N/A
A19PBP93OF896	0.0	0.0	0.0	N/A	N/A
A0000188NWOSI5X2PMS	0.0	N/A	0.0	1.0	N/A
A000063614T1OE0BUSKUT	N/A	N/A	N/A	0.0	5.0
A00031045Q68JAQ1UYT	N/A	N/A	N/A	4.0	N/A
A00028781NF0U7YEN9U19	N/A	N/A	N/A	0.0	5.0
A00031045Q68JAQ1UYT	N/A	N/A	N/A	100.0	1.0

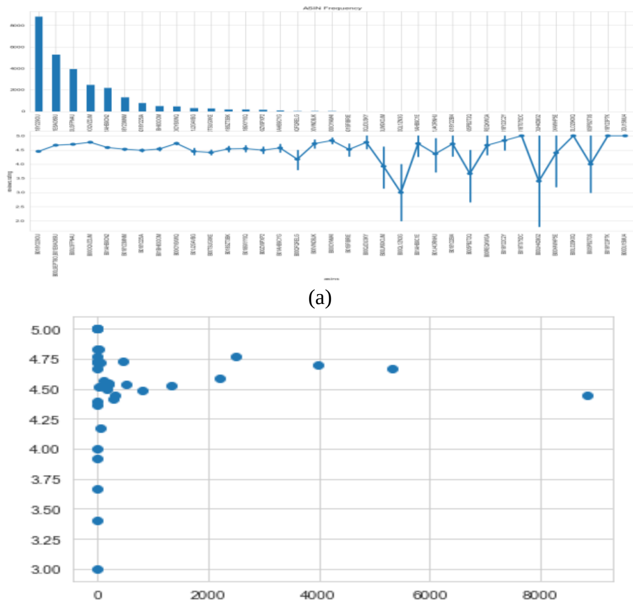


Fig. 11. (a) Exploratory data analysis on top of the ASIN (b) Correlation analysis between ASIN.

V. Discussions

The foremost regularly looked into products have routine audits within the 4.5 to 4.8 extend, with small fluctuation. Even though there's a slight converse relationship between the ASINs recurrence level and normal audit evaluations for the primary 4 ASINs, this relationship is immaterial. The normal survey for the prior 4 ASINs is evaluated between 4.5 to 4.8, which large surveys consider great. For ASINs with lower frequencies, we see that they're comparing normal audit evaluations on the point-plot chart (foot) have an essentially bigger change, as appeared by the length of the vertical lines. As a result, we recommend that the normal audit appraisals for ASINs with lower frequencies are not critical for our examination due to high variance.

On the other hand, due to their lower frequencies for ASINs with lower frequencies, we recommend that this result from more second-rate quality items 2c). Moreover, the final 4 ASINs have no fluctuation due to their altogether lower frequencies. Even though the audit evaluations are a culminate 5.0, we ought not to consider the centrality of these review evaluations due to lower recurrence. I am able to see that certain items have much more reviews than others based on the ASIN investigation, proposing a greater deal for those items. Ready to see that the ASINs have "right-tailed" dissemination, demonstrating that specific things have bigger deals, which can be connected to the higher recurrence of ASINs within the audits. Moreover, we took the log of the ASINs to normalize the information so that we seem to get a more nitty-gritty to see each ASIN and see that the dissemination is still "right-tailed." In our study, product sentiment analysis has been carried out towards in the year between 1996 to 2018. Three types of data analysis have been completed and through which business owners can make a variety of decisions towards their particular product. The different assessment pointers bend assessed the proposed show, and at long last, a proposal framework has been submitted by coordination overall sifting strategy. By taking after the proposal framework, the partner will be able to supply important data to their enlisted client, which can improve the request for specific things.

VI. Conclusion and Future Work

Nowadays, product sentiment is very important because, when a business is run online, it is important for every user to recommend their various products through pattern recognition. In order to use cutting-edge machine learning and deep learning algorithms to evaluate online product sentiment and make recommendations, a thorough pipeline is needed. This research proposes a pipeline for analyzing the online product. Also, a recommendation system has been presented through which a similar product can be filtered out for users. The study approach comprises two distinct components, namely the sentiment analysis approach and the product recommendation approach. The study uses appropriate hyperparameter optimization techniques to apply a number of cutting-edge algorithms, such as Naïve Bayes, Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, Bidirectional Long-Short-Term Memory (BI-LSTM), Convolutional Neural Network (CNN), Long-Short-Term Memory (LSTM), and Stacked LSTM. The k-Nearest Neighbors (KNN) model is combined with the collaborative filtering approach in the study to make product recommendations. Of these models, the Random Forest model had the highest accuracy (95%), followed by the LSTM model (79%). The Area under the ROC Curve (AUC), also known as the Receiver Operating Characteristic (ROC) curve, is used to evaluate the proposed model. In addition, the study carried out exploratory data analysis on reviews (1996–2018) using point-of-interest-based analysis, sentiment analysis, and bundle or bought-together analysis. Overall, the study meets its goals and suggests a flexible fix for practical situations.

Different machine learning and deep learning algorithms were applied to analyze the sentiment in this research. The Random Forest was found to be satisfactory through the investigation and can recommend any product effectively. In addition, three types of analysis have been carried out in this study. This research has several limitations, such as experimenting with just one dataset, but experimenting on multiple datasets was required, which we will complete in the future. In addition, a software system will be developed where a recommendation system will be integrated. Also, various loss optimization formulas will be applied to ensure model efficiency. Evaluation indicator approaches will later justify the type of model followed for recommendation in this phase.

References

[1] G. S. Blair, K. Beven, R. Lamb, R. Bassett, K. Cauwenberghs et al., “Models of everywhere revisited: a technological perspective,” Environmental Modelling & Software, vol. 122, pp. 104521, 2019.

[2] N. Mantel and W. Haenszel, “Statistical aspects of the analysis of data from retrospective studies of disease,” National Cancer Institute, vol. 22, no. 4, pp. 719–748, 1959.

[3] B. Jiang, “Investigating the market strategy of smart food’s latest product based on business analysis models," in 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021), Guangzhou, China, vol. 43, pp. 2218-2223. 2021.

[4] A. Schwarz, S. Isaksson, U. Källman and M. Rusner, "Enabling patient safety awareness using the Green Cross method: A qualitative description of users' experience," Journal of clinical nursing, vol. 30, no. 5, pp.830-839, 2021.

[5] D. Gavilan and M. Gema, "Exploring user's experience of push notifications: a grounded theory approach," Qualitative Market Research: An International Journal, vol. 3, pp. 35-57, 2022.

[6] S. Abuelenin, S. Elmougy and E. Naguib, "Twitter sentiment analysis for arabic tweets," in International Conference on Advanced Intelligent Systems and Informatics, London UK, pp. 467-476 .2017.

[7] S. Kumar, V. Koolwal and K. K. Mohbey, "Sentiment analysis of electronic product tweets using big data framework," Jordanian Journal of Computers and Information Technology, vol. 5, pp. 43-59, 2019.

[8] B. Setya Rintyarna, R. Sarno and C. Fatichah, "Semantic features for optimizing supervised approach of sentiment analysis on product reviews," Computers, vol. 8, p. 55, 2019.

[9] V. Balakrishnan, Z. Shi, C. L. Law, R. Lim and Y. Fan, "A deep learning approach in predicting products' sentiment ratings: a comparative analysis," The Journal of Supercomputing, vol. 4, pp. 1-21, 2021.

[10] P. Sasikala and L. Mary Immaculate Sheela, "Sentiment analysis of online product reviews using dlmnn and future prediction of online product using IANFIS," Journal of Big Data, vol. 7, pp. 1-20, 2020.

[11] Y. Zhu, J. Jiang, W. Han, Y. Ding and Q. Tian, "Interpretation of users' feedback via swarmed particles for content-based image retrieval," Information Sciences, vol. 375, pp. 246-257, 2017.

[12] W. Muhammad, M. Mushtaq, K. N. Junejo and M. Y. Khan, "Sentiment analysis of product reviews in the absence of labelled data using supervised learning approaches," Malaysian Journal of Computer Science, vol. 33, pp. 118-132, 2020.

[13] L. Monsurrò and L. Dezi, "Elderly experience of smart objects: how technology and family support can make senior users overcome their limits," In 2021 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Marrakech, Morocco, pp. 1-5, 2021.

[14] K. Vembandasamy, R. Sasipriya and E. Deepa, "Heart diseases detection using naive bayes algorithm," International Journal of Innovative Science, Engineering & Technology, vol. 2, pp. 441-444, 2015.

[15] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," Journal of Applied Science and Technology Trends, vol. 2, pp. 20-28, 2021.

[16] Z. Zhang, "Introduction to machine learning: k-nearest neighbours," Annals of translational medicine, vol. 4, pp. 345- 356, 2016.

[17] M. Almaliki, C. Ncube and R. Ali, "The Design of adaptive acquisition of users feedback: an empirical study," 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), Marrakech, Morocco, pp. 1-12, 2014.

[18] N. Sherief, W. Abdelmoez, K. Phalp and R. Ali, "Modelling users' feedback in crowd-based requirements engineering: an empirical study," In IFIP Working Conference on The Practice of Enterprise Modelling, pp. 174-190. Springer, Cham, 2015.

[19] S. Francesco, Lorenzo, B. Cristian, L. Danza, M. Ghellere et al., "Integrated method for personal thermal comfort assessment and optimization through users' feedback, IoT and machine learning: A case study," Sensors, vol. 18, no. 5, pp. 1602, 2018.

[20] S. Manikandan, A. Delphincarolinarani, C. Rajeswari, T. Suma and D. Sivabalaselvamani, "Recognition of font and tamil letter in images using deep learning," Applied Computer Science, vol. 17, no. 2, pp. 90–99, 2021.

[21] P. N. Tan, M. Steinbach and V. Kumar, "Introduction to data mining," in The Pearson. Upper Saddle River, NJ, USA: Pearson,vol. 54, pp. 37–69, 2006.

[22] A. Saurav, "Amazon Review Data (2018)," Amazon review, 06 March, 2022, Available: https://nijianmo.github.io/amazon/index.html.

[23] K. McCartan, A. Danielle, F. Harris and S. David, "Seen and not heard: The service user's experience through the justice system of individuals convicted of sexual offenses," International journal of offender therapy and comparative criminology, vol. 65, no. 12, pp. 1299-1315, 2021.

[24] D. Gavilan and G Martinez-Navarro, "Exploring user's experience of push notifications: a grounded theory approach. Qualitative Market Research," An International Journal, vol. 4, 2022.

[25] P. Sajjadi, L. Hoffmann, P Cimiano, and S. Kopp, "A personality-based emotional model for embodied conversational agents: Effects on perceived social presence and game experience of users," Entertainment Computing, vol. 32, pp.100313, 2019.

[26] B. Sovacool, J. Osborn, M. Martiskainen and M. Lipson, "Testing smarter control and feedback with users: Time, temperature and space in household heating preferences and practices in a Living Laboratory," Global Environmental Change, vol. 65, pp.102185, 2021.

[27] W. Wanyuan, Y. Jiang and W. Wu, "Multiagent-based resource allocation for energy minimization in cloud computing systems," IEEE Transactions on Systems Man and Cybernetics: Systems, vol. 47, no. 2, pp. 205–220, 2016.

[28] S. Kumar and M. Singh, "Big data analytics for healthcare industry: Impact, applications, and tools," Big Data Mining and Analytics, vol. 2, no. 1, pp. 1–20, 2019.

[29] K. Sahu, F. A. Alzahrani, R. K. Srivastava and R. Kumar, "Evaluating the impact of prediction techniques: Software reliability perspective," computers," Materials & Continua, vol. 67, no. 2, pp. 1471–1488, 2021.

[30] J. Athinarayanan, V. S. Periasamy, M. Alhazmi, K. A. Alatiah and A. A. Alshatwi, "Synthesis of biogenic silica nanoparticles from rice husks for biomedical applications," Ceramics International, vol. 41, no. 1, pp. 275–281, 2015.

[31] S. O. Proksch, W. Lowe, J. Wäckerle and S. N. Soroka, "Multilingual sentiment analysis: A new approach to measuring conflict in legislative speeches," Legislative Studies Quarterly, vol. 44, no. 1, pp. 97–131, 2019.