TWITTER TEXT SENTIMENT ANALYSIS OF AMAZON UNLOCKED MOBILE REVIEWS USING SUPERVISED LEARNING TECHNIQUES

R. Bharathi
Research Scholar, Department of Computer Science & Engineering, Annamalai University, Annamalai Nagar, Chidambaram, TamilNadu, India.
bharathirm17765@gmail.com

Dr. R. Bhavani
Professor, Department of Computer Science & Engineering, Annamalai University, Annamalai Nagar, Chidambaram, TamilNadu, India.
bhavaniaucse@gmail.com

Dr. R. Priya
Professor, Department of Computer Science & Engineering, Annamalai University, Annamalai Nagar, Chidambaram, TamilNadu, India.
pryknd@yahoo.com

Abstract

Due to the drastic increase in usage of Internet and e-commerce, online shopping becomes very common in day to day lives. One of the popular e-commerce websites is Amazon, which allows users to post Reviews and Ratings regarding the purchased products. It finds useful for other customers and companies to improve the product quality. Since the manual extraction of sentiments from User Product Reviews is a tedious process, accurate and Automated Sentiment Analysis (SA) tools are essential. This research work aims to examine the performance of different Machine Learning (ML) models in analyzing the sentiments of Amazon Product Reviews. Initially, the Product Reviews are pre-processed in different ways to transform the data into a useful format. Besides, Term Frequency – Inverse Document Frequency (TF-IDF) vectorizer is used to derive Feature Embeddings. Finally, three Machine Learning (ML) models namely Gaussian Naive Bayes (GNB), Logistic Regression (LR), and Support Vector Machines (SVM) are used for Sentiment Analysis (SA). The Performance Validation of the Machine Learning (ML) models is performed using Benchmark Dataset from Kaggle Repository. The Experimental Results reported that the Support Vector Machine (SVM) model has resulted in better Performance over the other Machine Learning (ML) models interms of Different Measures.

***Keywords*: Amazon Product Reviews, Sentiment Analysis (SA), Machine Learning (ML), Term Frequency – Inverse Document Frequency (TF-IDF) vectorizer, e-commerce**

1. Introduction

Amazon is one of the biggest online vendors in the Globe. People frequently gaze over the product and review the product beforehand buying the product on Amazon [1]. But the Reviews on Amazon are not necessarily of product but a combination of Product or Service Review (Product Company related or Amazon related) and Product Review. The consumer is misled as the complete sentiment (Rating Classification) that Amazon provides is a common one and there is no bifurcation between a Product and Service Reviews [2]. The presented approach reasonably segregate Product and Service Reviews, along with this it categorizes the Review as Feature Review when the client discusses certain Product Feature. A Featured Review is nothing but a Product Review, the method also provides sentiment of the text about the Product Feature. For instance, when the client writes in his review, "the camera for this phone is quite well.", then categorizes Camera Feature as Positive. While purchasing some products online, the Rating and Reviews of current users puts higher impacts.

The Extraction of User Reviews is hectic process and time consuming as the nature of Reviews is unstructured; thus, the technique of Sentimental Analysis (SA) is utilized [3]. The Primary Objective is to explore the Scores of Sentiment and explore the Reviews. Most people depend on User-Generated Content for making Decisions.

Sentiment Analysis (SA) tells the user whether the data about the product is satisfying or not beforehand they purchase it [4]. Firms and Marketers take advantage of this Data Analysis to understand the services or products so that it is provided as per the User Requirement. Usually, Sentiment Analysis (SA) is done at Distinct Levels changeable from Coarse Level to Fine Level. Fine Level copes with Attribute Level Sentiment Analysis and Coarse-level Sentiment Analysis copes with defining the sentiment of whole document [5]. Sentence-level Sentiment Analysis comes in between these two.

Sentiment Analysis (SA) could be implemented by a Machine Learning-based approach or Knowledge-based method. The Machine Learning-based approach could be further categorized as Supervised and Unsupervised Learning under Knowledge-based method, we could analyze the sentiment either by using Lexicon method or Language Processing Algorithm [6]. In Lexicon-based method, employs Sentiment Dictionary using opinion word and matches it with the information to define the Polarity. The Language Processing Algorithm is utilized for extracting Features such as Phrases, Word Frequency, Parts of Speech Tags, and Opinion Words [7]. But the Supervised Machine Learning (ML) algorithm learns the Polarity (Positive, Negative, or Neutral) of the Reviews from a data that is primarily categorized by a human [8]. Then, Scores are allocated to the Opinion Word according to Positive or Negative the words contained in the Dictionary [9]. The Lexicon and Machine Learning (ML) based approaches are the widely employed method for Sentimental Classification. However, it has been stated that this technique doesn't perform well in Sentimental Classification. This is due to the nature of Opinionated Text that needs more understanding of text when the existence of certain keywords might be the key for Precise Classification [10].

This study focuses on the Performance Analysis of three different Machine Learning (ML) models for examining the Sentiments involved in the Amazon Product Reviews. Primarily, the Product Reviews are pre-processed in diverse ways in converting the data into a useful format. In addition, Term Frequency – Inverse Document Frequency (TF-IDF) vectorizer is used to derive Feature Embeddings. Moreover, three Machine Learning (ML) models namely Gaussian Naive Bayes (GNB), Logistic Regression (LR), and Support Vector Machine (SVM) are used for Sentiment Analysis (SA). The Experimental Result Analysis of the Machine Learning (ML) models is carried out using Benchmark Dataset from Kaggle repository.

2. Related works

Alharbi et al. [11] estimated Distinct Deep Learning (DL) methods for precisely forecasting the feedback of Customer-based Smartphone Reviews attained from Amazon. The prediction is depending on analyzing this Review and categorizing them into three classes. Distinct Deep Learning (DL) approaches have been evaluated and implemented namely Recurrent Neural Network (RNN) with its four variants, such as Gated Recurrent Unit (GRNN), Long Short Term Memory Network (LSTM), Update Recurrent Unit (UGRNN), and Group Long RNN (GLRNN). Fu et al. [12] presented an Effective Review Attributes based Sentimental Pair Correlation method which evaluates the customer comment. Afterward pre-processing the comment information of Smartphones and constructing Attribute Dictionaries, the presented approach conducted a clustering process of Attribute and Sentimental Pair to accomplish precise evaluation of attribute for exploring data from User Comments.

Mukherjee et al. [13] presented an end-to-end Sentiment Analysis (SA) method for Handling Negation, and Negation Scope Marking and Negation Identification. The method presented a Customized Negation Marking Approach for performing experiments on Sentiment Analysis and explicit Negation Detection with Distinct Machine Learning (ML) approaches like SVM, Naive Bayes (NB), Recurrent Neural Network (RNN), and Artificial Neural Network (ANN) on Sentiment Analysis (SA) of Amazon Review. Sanchez-Franco et al. [14] developed an ambitious analysis of Non-Structured and Natural Narratives on Google Home and Amazon Echo. Also, to recognize major aspect that impacts the assessment of Intelligent Process Automation (IPA), this technique applies Machine Learning (ML) algorithm based Structural Topic Modelling, Sentiment Analysis (SA), Text Summarisation and XGBoost Regression, Cluster Analysis, etc.

Rehman et al. [15] presented a Hybrid Mechanism with Long Short Term Memory (LSTM) that uses Word2Vec technique for Word Embedding and Deeper Convolution Neural Network (CNN) method called as Hybrid CNN-LSTM method for overcoming the Sentiment Analysis (SA) problem. Firstly, employ Word to Vector (Word2Vec) method for training first word embedding. The Word2Vec translate the text strings to vector of numerical value, compute distance among words, and make groups of similar word related meaning. After embedding is implemented where the presented method integrates set of features which are filtered via Global Max-Pooling and Convolution Layers with Long-Term Dependency. In [16], Conventional Methods namely Term Frequency – Inverse Document Frequency (TF-IDF) variants, Bag of Words (BoW), and Bag of n-grams integrated with Linear Classification models and Deep Learning (DL) approaches like word-based ConvNets and the LSTM-RNN have been employed.

3. The Proposed Model

In this Research Work, the Sentiment Analysis (SA) of the Amazon Product Reviews is examined using different Machine Learning (ML) models. It involves three major processes namely Preprocessing, Feature Extraction, and

Classification, as illustrated in Fig. 1. Initially, the Product Reviews are pre-processed and unwanted data are removed. Then, the TF-IDF vectorizer is used to generate Feature Vectors. At last, Machine Learning (ML) models are used for the Classification of Sentiments.

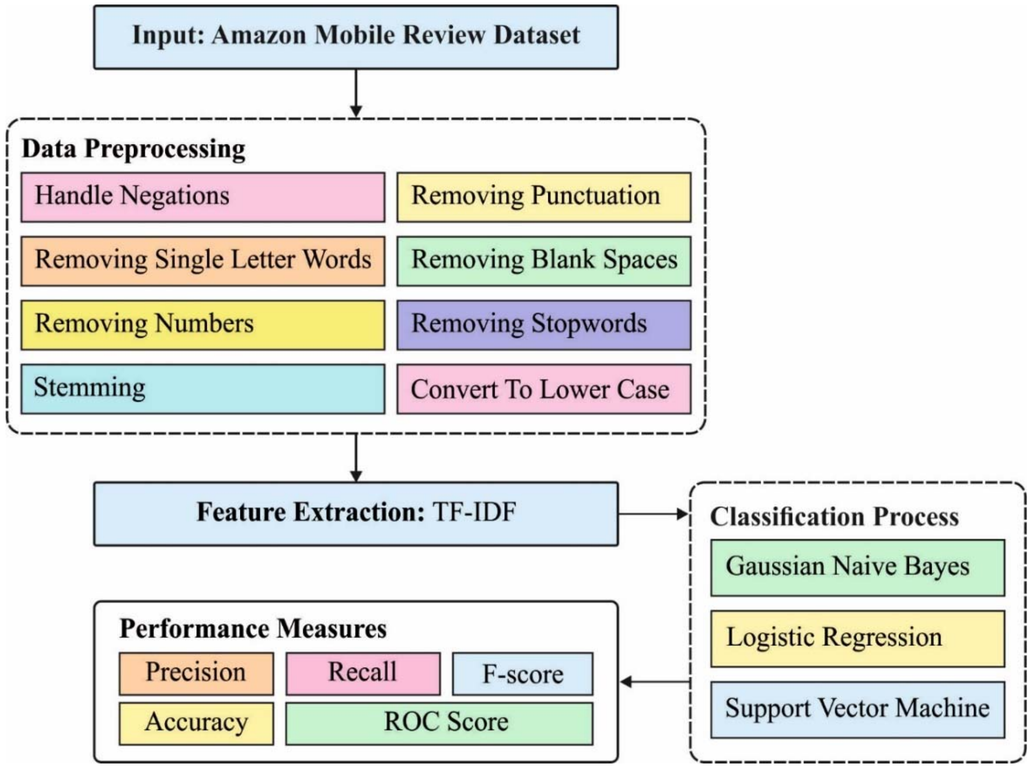


Fig. 1. Workflow of Proposed Model

3.1. *Pre-processing*

At the Initial Stage, Pre-processing is done to organize the data before Classification Process. In this Research Work, the steps involved in Pre-processing are Handling Negations, Removing Punctuation Marks, Removal of Single Character Words, Removal of Empty Spaces, Removal of Numbers, Stemming, Lower Case Conversion, and Removal of Stop Words. Once the Product Reviews are pre-processed, they are fed into the TF-IDF model for further processing.

3.1.1. Handle Negations

The Negations present in the reviews should be removed.

3.1.2. Convert To Lower Case

The reviews with upper case letters are converted to lower case.

3.1.3. Removing Punctuations

The reviews that contains punctuations have been removed.

3.1.4. Removing Single Letter Words

The reviews that contains unwanted single letter words have been removed.

3.1.5. Removing Blank Spaces

The reviews that contains unwanted blank spaces have been removed.

3.1.6. Removing Numbers

The reviews that contains unwanted numbers have been removed.

3.1.7. Removing Stopwords

The reviews that contains Stopwords have been removed.

3.1.8. Stemming

Stemming is the process of reducing a word to its word stem that affixes to suffixes and prefixes or to the roots of words known as lemma. Stemming is important in Natural Language Understanding (NLU) and Natural Language Processing (NLP) to free up the memory space.

Pre-processing Techniques	**Input**	**Output**
Handling Negations	Met all of my expectations. I can't complain at all. Great price!	Met all of my expectations. I cannot complain at all. Great price!
Convert To Lower Case	I feel so LUCKY to have found this used (phone to us & not used hard at all), phone on line from someone who upgraded and sold this one. My Son liked his old one that finally fell apart after 2.5 + years and didn't want an upgrade!!Thank you Seller, we really appreciate it & your honesty re: said used phone. I recommend this seller very highly & would but from them again!!	I feel so lucky to have found this used (phone to us & not used hard at all), phone on line from someone who upgraded and sold this one. my son liked his old one that finally fell apart after 2.5 + years and didn't want an upgrade!!thank you seller, we really appreciate it & your honesty re: said used phone. i recommend this seller very highly & would but from them again!!
Removing Punctuations	I'm really disappointed about my phone and service. The phone went out on me over a week ago. Instead of handling it when issue first surfaced. I've been told to do this and do that. Now I'm stick with no phone	I m really disappointed about my phone and service the phone went out on me over a week ago instead of handling it when issue first surfaced I ve been told to do this and do that now i am stick with no phone
Removing Single Letter Words	I'm really disappointed about my phone and service. The phone went out on me over a week ago. Instead of handling it when issue first surfaced. I've been told to do this and do that. Now I'm stick with no phone	'm really disappointed about my phone and service the phone went out on me over a week ago instead of handling it when issue first surfaced have been told to do this and do that now am stick with no phone
Removing Blank Spaces	The charging port was loose. I got that soldered in. Then needed a new battery as well. $100 later (not including cost of purchase) I have a usable phone. The phone should not have been sold in the state it was in.	the charging port was loose got that soldered in then needed a new battery as well 100 later not including cost of purchase have usable phone the phone should not have been sold in the state it was in
Removing Numbers	The charging port was loose. I got that soldered in. Then needed a new battery as well. $100 later (not including cost of purchase) I have a usable phone. The phone should not have been sold in the state it was in.	the charging port was loose got that soldered in then needed a new battery as well later not including cost of purchase have usable phone the phone should not have been sold in the state it was in.
Removing Stopwords	I feel so LUCKY to have found this used (phone to us & not used hard at all), phone on line from someone who upgraded and sold this one. My Son liked his old one that finally fell apart after 2.5+ years and didn't want an upgrade!! Thank you Seller, we really appreciate it & your honesty re: said used phone. I recommend this seller very highly & would but from them again!!	feel lucky found used phone us used hard phone line someone upgraded sold one son liked old one finally fell apart years want upgrade thank you seller really appreciate honesty re said used phone recommend seller very highly
Stemming	I feel so LUCKY to have found this used (phone to us & not used hard at all), phone on line from someone who upgraded and sold this one. My Son liked his old one that finally fell apart after 2.5+ years and didn't want an upgrade!! Thank you Seller, we really appreciate it & your honesty re: said used phone. I recommend this seller very highly & would but from them again!!	feel lucky found used phone us used hard phone line someone upgraded sold one son liked old one finally fell apart years want upgrade thank you seller really appreciate honesty re said used phone recommend seller very highly

Table 1. Example of Preprocessing Techniques

3.2. *Term Frequency – Inverse Document Frequency (TF-IDF) Vectorizer*

Term Frequency (TF) is defined as how often a term occurs in a document. In Natural Language Processing (NLP), terms correspond to words or phrases. Terms can also be represented as any token in text. Inverse Document Frequency (IDF) is the inverse of the document frequency which measures the informativeness of term t. Here term t refers to a single word in the review(text). Inverse Document Frequency (IDF) is defined as the inverse of the number of times a specific term has appeared in the entire text. Term Frequency -Inverse Document Frequency (TF-IDF) [17] is a group of Term Frequency (TF) and Inverse Document Frequency (IDF) . An IDF extend in Document Frequency (D_F) that implies the amount of documents where term takes place. It can be presented dependent upon the assumption that term which follows in any documents are considered that very essential than those which occur in further document [17]. The IDF value of certain terms is attained as demonstrated from the subsequent Eq. (1)

In Eq. (1), $D_F(t, Doc_c)$ implies the D_F value of terms t from corpus Doc_c. The symbol in Eq. (1) stands for the entire amount of documents from corpus Doc_C. To avoid infinity of any extreme cases, the equation has sometimes optimizing as illustrated under the subsequent:

$$IDF(t, doc, Doc_c) = \log \frac{|Doc_c| + 1}{D_F(t, Doc_c) + 1}. \tag{1}$$

Next, Eq. (1) extends the IDF technique with more TF values into calculation. The presented group with TF and IDF is most famous term weight technique such as TF-IDF. The typical infrastructure of TF-IDF is illustrated as:

$$TF - IDF(t, doc, Doc_c) = TF(t, doc) * IDF(t, doc, Doc_c). \tag{2}$$

In Eq. (2), $TF - IDF(t, doc, Doc_c)$ signifies the weight of terms t of document doc from corpus Doc_c, but TF (t, doc) implies the TF value of terms t from the document doc.

3.3. *Overview of Machine Learning (ML) Models*

In this section, a set of three Machine Learning (ML) models are applied for Sentiment Analysis (SA) in examining the Amazon Product Reviews.

3.3.1. Gaussian Naive Bayes (GNB) Model

Naive Bayes classification model computes the probability of a provided sample (instance) going to a particular class [18]. To provide a sample X defined as their Feature Vector $(x_1, \ldots, x_n)$ and Class Labels y, the Conditional Probability $P(y|X)$ is formulated as product of simpler probability utilizing the Naive independence consideration using Bayes concept as given below:

$$P(y|X) = \frac{P(y)P(X|y)}{P(X)} = \frac{P(y)\prod_{u=1}^{n} P\,(x_u|y)}{P(X)}. \tag{3}$$

At this point, the target y can have 2 values whereas $y = 1$ implies the hot spot residue and $y = 0$ refers to the non-hot spot residue. X for one residue (one sample) is a Feature Vector with similar size to define their feature utilizing higher frequency mode created as GNM. Eq. (4) is resolved to utilize the Quadratic Programming approaches and the static Karush-Kuhn-Tucker state. For instance, X has equivalent to vector collected of u^{th} element z_{ku} to i^{th} residue from an order if only one higher frequency mode z_k was utilized. When the 3 Higher Frequency Modes, represented as z_1, z_2, and u_3, are taken as to account, the vector X is (z_{1u}, z_{2u}, z_{3u}) for residue u from protein series. In addition, when the window size of 3 interms of the residue u has adapted, X develops $(z_{1u-1}, z_{1u}, z_{1u+1}, z_{2u-1}, z_{2u}, z_{2u+1}, z_{3u-1}, z_{3u}, z_{3u+1})$. As $P(X)$ is constant to a provided sample, the subsequent rule was adapted for classifying the sample whose class is unknown:

$$\hat{y} = \arg\max_{y} P(y) \prod_{u=1}^{n} P\,(x_u|y), \tag{4}$$

where "arg" represents the value of y to the above formula is maximizing; that is, if $P(y = 1)\prod_u P\,(x_u|y = 1)$ is superior to $P(y = 0)\prod_u P\,(x_u|y = 0), \hat{y} = 1$; else, $\hat{y} = 0.$ Besides, if the possibility of features (for instance, $P(x_u|y)$) has assumed that Gaussian, an NB classification is named GNB [18]. Due to the easier computation characteristics related to other further sophisticated techniques, GNB was extremely executed for predicting problems from Bioinformatics. During this case, GNB has mostly been utilized for training the methods by inputting the maximum frequency mode for identifying hot spot residue.

3.3.2. Logistic Regression (LR) Model

The Logistic Regression (LR) technique is a connection amongst the Categorical Response Variables and Covariate [19]. Particularly, there is Linear Group of Independent Variables with log-odds of probabilities of case from Logistic Method. The Binary LRs evaluate the possibility which is a feature of a Binary Variable has presented, provided the value of covariates. Let Y refers to the Binary Response Variable whereas $Y_j = 1$ when the character has presented and $Y_i = 0$ when the character has absent and the data $[Y_1, Y_2, \ldots, Y_n]$ are independent.

Assume π_i be present the probability of success. Moreover, Assume $x = (x_1, x_2, \dots, x_p)$ as a group of Explanatory Variables that are Discrete, Continuous, or Group of Combined Discrete and Continuous. Afterward, the logistic operation to π_i has provided as:

$$logit\ (\pi_i) = log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{i,p}; \quad (5)$$

where

$$\pi_i = \frac{exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{\mathrm{p}} x_{i,p})}{1 + exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{\mathrm{p}} x_{i,p}))} = \frac{exp(x_i'\beta)}{1 + exp(x_i'\beta)} = \Lambda(x_i'\beta) \quad (6)$$

At this point, π_i indicates the probability which an instance is from a provided type of dichotomous response variables, usually named as "success probability" and, obviously, $0 \le \pi_i \le 1$. $\Lambda(\cdot)$ represents the logistic cdf, with $\lambda(z) = e^z/(1+e^{-z}) = 1/(1+e^{-z})$ and β^s stands for the vector of parameters that are evaluated. The formulated $(\frac{\pi_i}{1-\pi_i})$ is named the odds ratio/relative risk.

3.3.3. Support Vector Machine (SVM) Model

The Support Vector Machine (SVM) is most extremely utilized as a Data Learning Tool in past years. It can be generally utilized for addressing a Binary Pattern Classifier Problem [20]. The SVM which supports binary classification and separating data points into two classes. The Binary SVM generates a group of hyperplanes from infinite dimension space that is separated into 2 types of representations like the Linear and Non-Linear SVM. Primary, it can be assumed a binary classifier issue; the trained dataset $= \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_l, y_l)\}, y_i \in \{-1,1\}, x_i \in R^d$, where x_i refers the data point and the corresponding y_i is their calculated labels. The l represents the amount of elements from the trained dataset. The Linear SVM defines the optimum separate margin with resolving the subsequent optimized tasks:

$$Minimize\ \{\frac{1}{2}|w|^2 + C\sum_{i=1}^{l} \varepsilon_i\}, \varepsilon_i \ge 0 \quad (7)$$

$$y_i(w^T x_i + b) \ge 1 - \varepsilon_i, i = 1,2, \dots, l, \quad (8)$$

where C means the penalty value, ε_i are positive slack variable, w indicates the normal vector, and b denotes the scalar quantity. The minimal problem is decreased by utilizing the Lagrangian Multiplier α_i that achieve optimal based on the Karush- Kuhn-Tucker condition. In Mathematical Optimization, the Karush-Kuhn-Tucker (KKT) conditions, also known as the Kuhn-Tucker conditions, are first derivative tests (sometimes called first-order necessary conditions) for a solution in nonlinear programming to be optimal, provided that some regularity conditions are satisfied. If $\alpha_i > 0$, then the matching data x_i is called as Support Vectors (SVs), and so, the Linear Discriminate Function has written with optimum Hyperplane Parameters w and b from the subsequent formula:

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_i\, y_i x_i^T x + b\right). \quad (9)$$

By Unconstrained Dual Form, Eq. (9) is converted as ,

$$Maximize\ \left\{\sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i\, \alpha_j y_i y_j x_i x_j\right\}, \quad (10)$$

$$C \ge \alpha_i \ge 0, i = 1, \dots, l, \textstyle\sum_{i=1}^{l} \alpha_i\, y_i = 0. \quad (11)$$

The Resultant Solution W has formulated as a Linear Combination of Trained Vector and the b is written as the Average of every SV is illustrated in

$$W = \sum_{i=1}^{l} \alpha_i\, y_i x_j,$$

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (Wx_i - y_i), \quad (12)$$

where N_{SV} refers to the amount of Support Vectors.

4. Performance Measures

4.1. *Accuracy*

Accuracy is defined as the percentage of Reviews that are classified correctly divided by the total number of Reviews.

$$ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

Here TP, TN, FP and FN are True Positives, True Negatives, False Positives, and False Negatives respectively.

4.2. *Precision*

Precision is defined as the percentage of Positive Reviews that predict truly divided by the total number of Reviews that are classified positive.

$$PPV = \frac{TP}{TP + FP} \tag{14}$$

4.3. *Recall*

Recall is defined as the measure of percentage of the Reviews that classify positively divided by the total number of Reviews which are truly positive.

$$RC = \frac{TP}{TP + FN} \tag{15}$$

4.4. *F1-Score*

F1-score is defined as the combination of both Precision and Recall.

$$F1 - score = 2 * (PR * RC)/(PR + RC) \tag{16}$$

4.5. *ROC AUC-Score*

AUC-ROC curve is a performance measurement for the classification problems at various threshold settings. ROC is a probability curve and AUC represents the degree or measure of separability. It defines how much the model is capable of distinguishing between classes.

5. Experimental Validation

The presented technique is simulated on Python 3.6.5 tool. In this Research Work, Amazon Reviews: Unlocked Mobile Phones dataset from Kaggle repository is used for Validation Process. It includes a total of 413840 reviews of Unlocked Mobile Phones vended on Amazon.com. In this Research Work, a set of 1000 Reviews are taken under each class (Positive, Negative, and Neutral). It includes distinct fields such as

- Product Titles,
- Brands,
- Price,
- Ratings,
- Reviews Text, and
- Review Votes.

Fig. 2 illustrates a set of three Confusion Matrices produced by the GNB, LR, and SVM Classifiers. Fig. 2a indicated that the GNB model has identified 690 Reviews as Negative, 500 Reviews as Neutral, and 780 Reviews as Positive. In addition, Fig. 2b denoted that the LR model has classified 712 Reviews as Negative, 576 Reviews as Neutral, and 843 Reviews as Positive. Similarly, Fig. 2c indicated that the SVM model has recognized 925 Reviews as Negative, 840 Reviews as Neutral, and 944 Reviews as Positive.

Fig. 3 shows the Precision-Recall Curve Analysis of the three DL models. The Results indicated that the GNB model has resulted in ineffective outcomes with the lower values of Precision-Recall under three class labels. Eventually, the LR model has gained certainly improved performance over the GNB model. However, the SVM model has outperformed the GNB and LR models with the maximum values of Precision-Recall.

e-ISSN : 0976-5166
p-ISSN : 2231-3850

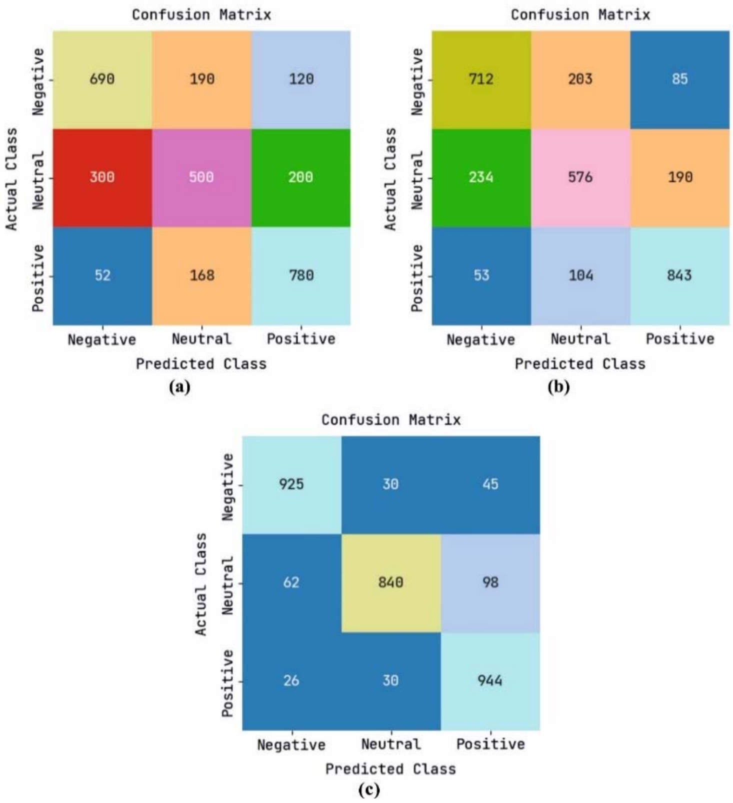


Fig. 2. Confusion Matrices of (a) GNB, (b) LR, (c) SVM

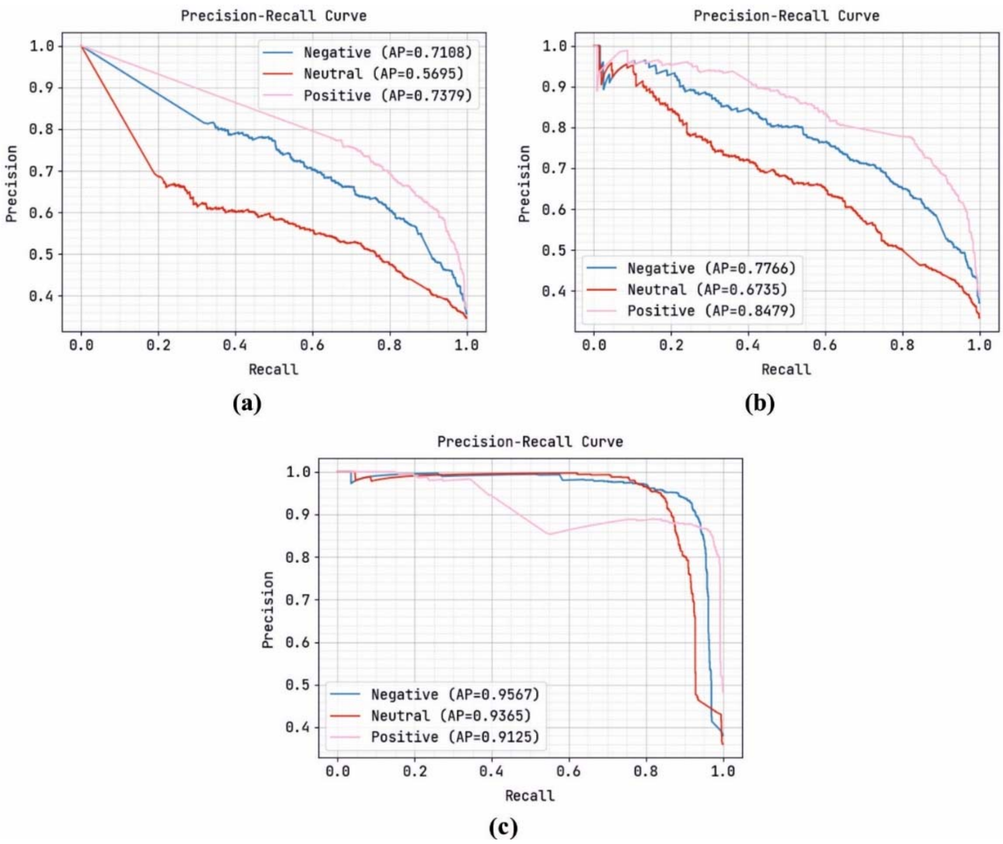
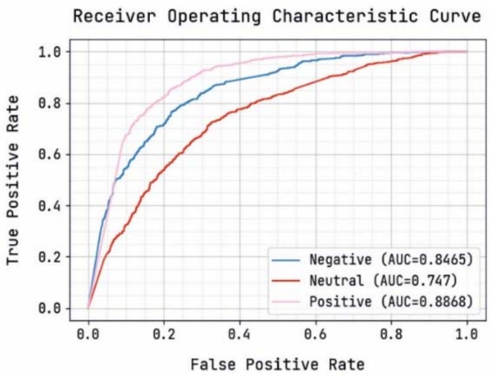
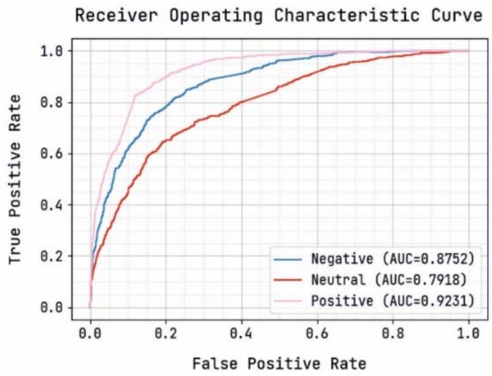


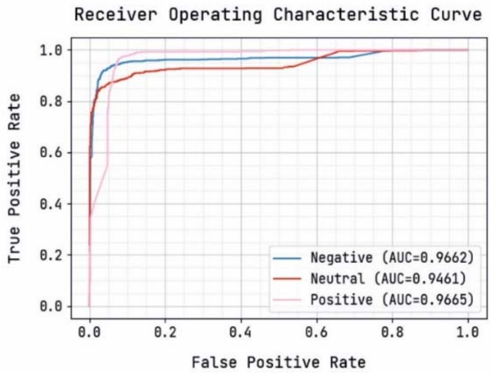
Fig. 3. Precision-Recall Curves of (a) GNB, (b) LR, (c) SVM



(a)



(b)



(c)

Fig. 4. ROC Curves of (a) GNB, (b) LR, (c) SVM

Fig. 4 demonstrates the ROC analysis of the different ML models on the test Kaggle dataset. The Results indicated that the GNB model has obtained ROC of 0.8465, 0.747, and 0.8868 on the Classification of Negative, Neutral, and Positive Sentiments. Followed by, the LR model has reached slightly increased ROC of 0.8752, 0.7918, and 0.9231 on the Classification of Negative, Neutral, and Positive Sentiments. However, the SVM model has showcased better outcomes with the ROC of 0.9662, 0.9461, and 0.9665 on the Classification of Negative, Neutral, and Positive Sentiments.

Measures	Gaussian NB	LR Algorithm	SVM Algorithm
Accuracy	66.00	71.25	90.00
Precision	65.00	71.16	91.00
Recall	66.00	71.30	90.00
F1-Score	65.00	71.50	90.00
ROC Score	83.00	86.00	96.00

Table 2. Overall Classification Performance of Different ML models

In order to demonstrate the overall Sentiment Classification Performance of the ML models, a Comparative Analysis is made in Table 1. Fig. 5 investigates the Sentiment Classification Results of the three ML models interms of Precision, Recall, and Accuracy. The Figure reported that the GNB model has shown Poor Performance over the other two methods with the Precision of 66%, Recall of 65%, and Accuracy of 66%. Followed by, the LR model has gained moderately improved Performance with the Precision of 71%, Recall of 71%, and Accuracy of 71%. But the SVM model has reached improved Performance with the Precision of 90%, Recall of 91%, and Accuracy of 90%.

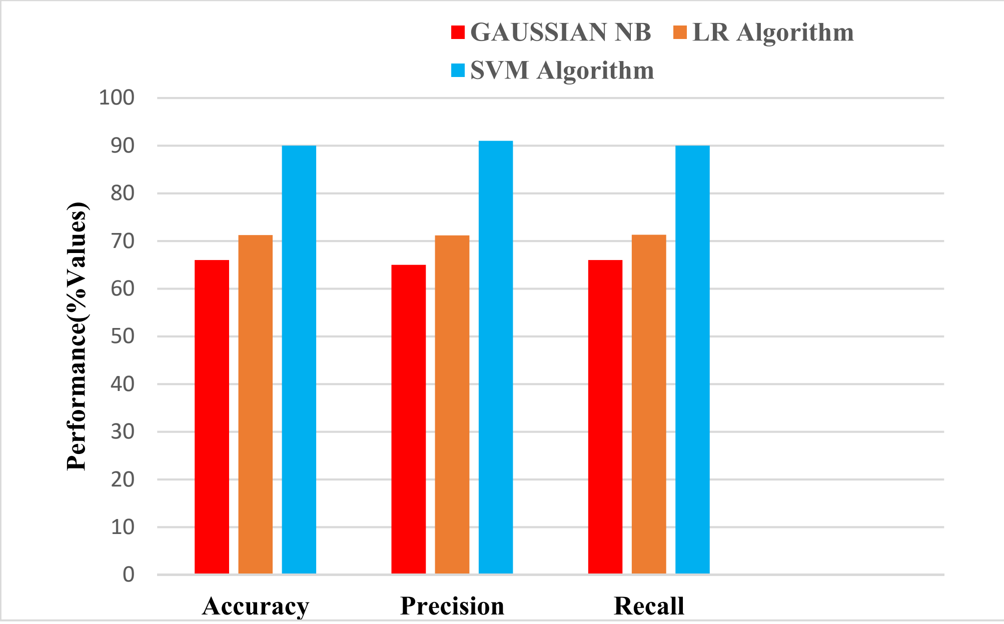


Fig. 5. Comparative Sentiment Analysis (SA) Results of Different Machine Learning (ML) Models-I

Fig. 6 examines the Sentiment Classification Results of the three ML models interms F1-score and ROC score. The Results portrayed that the GNB model has demonstrated Ineffectual Results over the other two methods with the F1-score and ROC score of 65% and 83% respectively. In line with, the LR model has offered certainly increased Performance with the F1-score and ROC scores of 71% and 86% respectively.

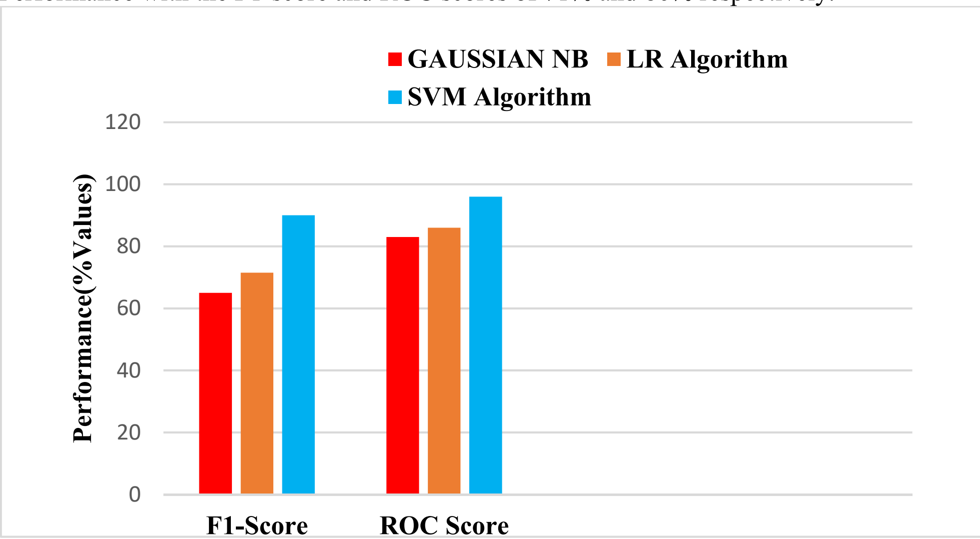


Fig. 6. Comparative Sentiment Analysis (SA) Results of Different Machine Learming (ML) Models-II

However, the SVM model has outdone the existing techniques with the F1-score and ROC score of 90% and 96% respectively. From the detailed Results and Discussion, it is noticeable that the SVM model can exhibit Effective Sentiment Analysis (SA) outcomes on Amazon Product Reviews.

6. Conclusion

In this Research Work, the Sentiment Analysis (SA) of the Amazon Product Reviews is examined using different Machine Learning (ML) models. It involves three major sub-processes such as Pre-processing, Feature Extraction and Classification. Initially, the Product Reviews are pre-processed and unwanted data are removed. Then, the Term Frequency – Inverse Document Frequency (TF-IDF) vectorizer is used to generate Feature Vectors. At last, Machine Learning (ML) models are used for the classification of sentiments. The Experimental Validation of the Machine Learning (ML) models is performed using Benchmark dataset from Kaggle repository. The Experimental Results reported that the Support Vector Machine (SVM) model has resulted in better Performance over the other Machine Learning (ML) models interms of different measures. In Future, the Data Clustering Techniques can be integrated to improve the Classification Results. Besides, the Advanced Deep Learning (DL) models can be utilized for increasing the Classification Performance.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Wassan, S., Chen, X., Shen, T., Waqar, M. and Jhanjhi, N.Z., 2021. Amazon product sentiment analysis using machine learning techniques. Revista Argentina de Clinica Psicologica, 30(1), p.695.
[2] Dadhich, A. and Thankachan, B., 2022. Sentiment Analysis of Amazon Product Reviews Using Hybrid Rule-based Approach. In Smart Systems: Innovations in Computing (pp. 173-193). Springer, Singapore.
[3] Xiao, Y., Qi, C. and 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.
[4] Rao, M.V. and Sindhu, C., 2021, March. Detection of Sarcasm on Amazon Product Reviews using Machine Learning Algorithms under Sentiment Analysis. In 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) (pp. 196-199). IEEE.
[5] Singh, S.K. and Sachan, M.K., 2021. Classification of code-mixed bilingual phonetic text using sentiment analysis. International Journal on Semantic Web and Information Systems (IJSWIS), 17(2), pp.59-78.
[6] Hajek, P., Barushka, A. and Munk, M., 2021. Neural Networks with Emotion Associations, Topic Modeling and Supervised Term Weighting for Sentiment Analysis. International journal of neural systems, 31(10), p.2150013.
[7] Li, X., Sun, X., Xu, Z. and Zhou, Y., 2021, July. Explainable Sentence-Level Sentiment Analysis for Amazon Product Reviews. In 2021 5th International Conference on Imaging, Signal Processing and Communications (ICISPC) (pp. 88-94). IEEE.
[8] Lee, W.S., Ng, H., Yap, T.T.V., Ho, C.C., Goh, V.T. and Tong, H.L., 2021. Attention Models for Sentiment Analysis Using Objectivity and Subjectivity Word Vectors. In Computational Science and Technology (pp. 51-59). Springer, Singapore.
[9] Geetha, M.P. and Renuka, D.K., 2021. Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. International Journal of Intelligent Networks, 2, pp.64-69.
[10] Chen, J., Chen, Y., He, Y., Xu, Y., Zhao, S. and Zhang, Y., 2021. A classified feature representation three-way decision model for sentiment analysis. Applied Intelligence, pp.1-13.
[11] Alharbi, N.M., Alghamdi, N.S., Alkhammash, E.H. and Al Amri, J.F., 2021. Evaluation of sentiment analysis via word embedding and RNN variants for Amazon online reviews. Mathematical Problems in Engineering, 2021.
[12] Fu, X.L., Wu, J., Chen, J. and Liu, S., 2019. Attribute-Sentiment pair correlation model based on online user reviews. Journal of Sensors, 2019.
[13] Mukherjee, P., Badr, Y., Doppalapudi, S., Srinivasan, S.M., Sangwan, R.S. and Sharma, R., 2021. Effect of negation in sentences on sentiment analysis and polarity detection. Procedia Computer Science, 185, pp.370-379.
[14] Sánchez-Franco, M.J., Arenas-Márquez, F.J. and Alonso-Dos-Santos, M., 2021. Using structural topic modelling to predict users' sentiment towards intelligent personal agents. An application for Amazon's echo and Google Home. Journal of Retailing and Consumer Services, 63, p.102658.
[15] Rehman, A.U., Malik, A.K., Raza, B. and Ali, W., 2019. A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. Multimedia Tools and Applications, 78(18), pp.26597-26613.
[16] Katic, T. and Milicevic, 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.
[17] Jiang, Z., Gao, B., He, Y., Han, Y., Doyle, P. and Zhu, Q., 2021. Text classification using novel term weighting scheme-based improved TF-IDF for Internet media reports. Mathematical Problems in Engineering, 2021.
[18] Zhang, H., Jiang, T. and Shan, G., 2016. Identification of hot spots in protein structures using Gaussian network model and Gaussian Naive Bayes. BioMed research international, 2016.
[19] Joshi, R.D. and Dhakal, C.K., 2021. Predicting type 2 diabetes using logistic regression and machine learning approaches. International Journal of Environmental Research and Public Health, 18(14), p.7346.
[20] Chao, C.F. and Horng, M.H., 2015. The construction of support vector machine classifier using the firefly algorithm. Computational intelligence and neuroscience, 2015.
[21] https://www.kaggle.com/PromptCloudHQ/amazon-reviews-unlocked-mobile-phones

Authors Profile

R.Bharathi, B.Tech, M.Tech, Ph.D (Pursuing) is a Research Scholar doing Research Work in the Department of Computer Science & Engineering, Annamalai University, Chidambaram, Tamil Nadu, India. Area of Interest: Data Mining, Machine Learning.

Dr.R.Bhavani, B.E, M.E, Ph.D working as a Professor in the Department of Computer Science & Engineering, Annamalai University, Chidambaram, Tamil Nadu, India. She has 28 years of Teaching Experience. She has published papers in 58 International Journals,19 National and 57 International Conferences and 1 Book has been published. Area of Interest: Image Processing and Computer Graphics.

Dr.R.Priya, B.E, M.E, Ph.D working as a Professor in the Department of Computer Science & Engineering, Annamalai University, Chidambaram, ,Tamil Nadu, India. She has 23 years of Teaching Experience. She has published papers in 38 International and 1 National Journals,10 National and 18 International Conferences and 5 Popular Article has been published. Area of Interest: Image Processing and Pattern Classification Techniques.