International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 21, Issue 7 (July 2025), PP 14-25

A Sentiment Analysis System for on-line Retail Review Using Enhanced Support Vector Machine

Caleb Markus[1], Chikodili Ugwuishiwu[1], Nnamdi Chibunze[1], Osondu Oguike[1]
[1]*Department of Computer Science, Faculty of Physical Sciences, University of Nigeria, Nsukka*
Corresponding Author: Osondu Oguike

ABSTRACT
*Online retailing has changed the process of buying and selling, which depends on the polarity of reviews from other customers, such polarity can only be obtained through sentiment analysis of customers reviews on online retailing platforms. This research builds a sentiment analysis system for product review on eBay using Support Vector Machine (SVM) tuned by grid search and Bayesian optimizations. The system benefits from sophisticated natural language processing (NLP) methods for data preprocessing including stop words removal, stemming, and tokenization, followed by feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF). The methodology employed TextBlob sentiment analysis for automatic sentiment labeling of the dataset, utilizing a polarity function to map review text to continuous sentiment scores, followed by threshold-based binary classification to distinguish between positive and negative sentiments. To address severe positive sentiment bias in the original dataset, a hybrid balancing approach was implemented in two sequential phases: threshold adjustment from 0.0 to 0.1 polarity to reclassify weakly positive reviews as negative, followed by strategic random undersampling to achieve perfect 1:1 class balance while preserving authentic linguistic patterns. A thorough hyperparameter optimization process ensured the model was configured optimally, achieving outstanding classification performance with an accuracy value of **99.44%**, precision of **99%**, recall of **99%**, and F1-Score of **99%**. The robustness of the model is demonstrated by its discriminatory power with an AUC of **0.9949**. Comparative research shows the proposed model is superior to state-of-the-art approaches, such as BERT-BiGRU-Softmax at 95.5%, CNN at 95.27%, and BERT-based models at 90% and 97%. This study not only improves sentiment analysis in e-commerce but also provides a scalable, high-accuracy solution for customer feedback analysis with implications for customer relationship management and decision-making.*

KEYWORDS: *sentiment analysis, optimized support vector machine, machine learning, natural language processing, hyper parameter optimization, eBay reviews.*

Date of Submission: 02-07-2025 Date of acceptance: 12-07-2025

I. INTRODUCTION

The rapid spread of online retailing has changed the process of consumers' purchasing decisions completely. Online retailing gives customers an unparalleled platform for expressing their views and experiences concerning products and services. Reviews in popular online platforms for sales such as Amazon, Alibaba, and eBay have emerged crucial in influencing consumer habits, as buyers have little choice but to seek the sentiments of other buyers before they can make decisions regarding purchasing,[1],[2]. These reviews, however, make a two-pronged challenge to retailers. On the one hand, they provide important information about customers' preferences, product quality, and scope for improvement. On the other hand, the high number of reviews — a volume to which humans cannot possibly, manually, analyze — represents the imperative for automated systems of sentiment analysis that can appropriately sort and decipher customer feedback,[3],[4], [5]. In this regard, machine learning has been a revolutionary solution that provides strong means for the processing of huge amounts of text data, identifying meaningful patterns and sentiment classification into categories, including positive, negative, and neutral, [6], [7],[8], [9].

Sentiment analysis or opinion mining is a branch of natural language processing (NLP) wherein a process of mining opinions from texts is pursued. One of the most used algorithms for the above purpose is machine learning algorithms like Support Vector Machine (SVM), Random Forest, and Deep Neural Networks, [10], [11], [12]. Among them, SVM, when optimized, has shown outstanding performance because it can classify data highly correctly even in difficult tasks related to text classification, [13], [14]. Even though the application of these algorithms is widespread, existing models for sentiment analysis tend to have limitations. Many are not able to deliver valuable insights to the management as they restrict their outputs to focus on the sentiment analysis without providing suggestions for redressing people's concerns, [15]. Additionally, most models fail to adequately account for the dynamic reality of real-world settings, where nature of customer feedback can change very rapidly requiring constant fine tuning and improvement,[3], [16]. This gap highlights the necessity of a sentiment analysis system; one that does not only classify sentiment but also offer practical recommendations to management on how they can increase customer satisfaction.

The need for creating such a system is also reinforced by the increasing competitiveness of the online retail industry, where customer satisfaction corresponds directly to business success, [17]. Retailers should not only understand the customer's sentiments but also adequately take care of the negative feedback to preserve brand image and client loyalty. For example, a drastic increase in negative reviews regarding product quality might suggest a manufacturing problem or a drop-in customer service and call for prompt action by the management, [18]. A sentiment analysis system that has machine learning algorithms, a friendly user interface and an automated recommendation system can make retailers able to detect concern of the customers, prioritize and make corrective steps. A system such as this is not just about categorizing but it turns sentiment analysis into a powerful tool for customer relations, competitive intelligence and constant business improvement,[19].

Against the growing dependence on online reviews to guide the consumer in his decision-making, there is need for a technologically high system for sentiments analysis. The existing models are successful in a basic classification; however, they tend not to offer insights for decisions useful for businesses,[6], [20]. This work is aimed at narrowing this gap by building a sentiment analysis system that can successfully classify customer feedback by positive, neutral, and negative sentiments, based on the Support Vector Machine (SVM) algorithm. The system will not only categorize sentiments but also produce automated recommendations on management that will enable businesses to make moves to handle customer complaints and enhance product quality. What is more, it will enable users to add comments, run sentiment analysis, and monitor the changes in sentiment over time. By combining machine learning and an advisory element, this system is meant to transition sentiment analysis from being a tool for passive data processing into an active tool for decision-making, [3], [21].

This study aims to build a sentiment analysis model for online store reviews using an optimized Support Vector Machine algorithm based on the eBay store reviews data with an automatic recommendation function that identifies the negative sentiments, compare the proposed sentiment analysis model to other models to measure the accuracy, precision, and recall metrics and provide a mathematical model for Bayesian optimization to enhance the accuracy of the sentiment classification since Bayesian optimization is a probabilistic model that efficiently searches optimal hyperparameter settings to achieve maximum model performance. Section 1 of the paper provides the introduction, while Section 2 provides an overview of the theoretical foundation for sentiment analysis system. Section 3 presents the methodology, which includes data collection, model building, and model testing approaches. Section 4 follows with the presentation of findings alongside discussion of the findings. Section 5 wraps up the study with a conclusion note where practical and theoretical implications of the study are given culminating with limitations of the study with avenues for future direction.

II. MATERIAL AND METHODS

2.1 INFORMATION PROCESSING THEORY

This seThe Information Processing Theory (IPT) is a cognitive model used to explain the processing, storage, and retrieval of information, which is highly related to sentiment analysis system, [22]. Formulated in 1956 by George Miller, IPT posits that human cognition operates in stages as a computer, that is, IPTinput, processing, storage, and output, [23]. This theory is in line with the main operations of a sentiment analysis system as user generated text data (input) is processed through NLP techniques and machine learning techniques such as SVM to classify sentiments, [24], [25]. However, critics affirm that the IPT makes thinking easy, not minding emotions, context, and the dynamic human perspective, [23]. This criticism also applies to sentiment analysis, and the system might misclassify reviews because it has an inability to identify sarcasm, contextual cues, or cultural differences, [26]. Nevertheless, IPT is validated for use in this study as it establishes a structured rationale for learning how textual data (eBay product reviews) can be systematically transformed into useful insights with the use of machine learning. By making the system's architecture follow the IPT, this study guarantees that the process of sentiment analysis is not simply logical in its sense but is also systematically developed for accuracy and efficiency, [27],[28].

2.2 TECHNOLOGY ACCEPTANCE MODEL

Technology Acceptance Model (TAM) developed by [29] has been well-known as a theoretical frame which explains the process of users' acceptance and utilization of technology and thus it is relevant to this study, which considers the goal to create a user-friendly sentiment analysis system. By TAM, two main factors determine the user acceptance – the factors: perceived usefulness (PU) that relates to the perception of a user with regard to how the adoption of the technology will be useful to him/her in the performance of a given task; and perceived ease of use (PEOU) that is the extent to which a user believes that the process of using the technology will be free from effort,[29]. In the scope of this study, the PU is manifested in how well sentiment analysis system allows users to realize actionable insights from eBay product reviews, and PEOU is represented by the ease of navigating its interface, [25], [28]. Although widely accepted, TAM has been criticized for its simplistic nature in that it only examines cognitive factors in understanding user behavior with disregard to its emotional, social, and contextual antecedents in adoption of technology,[30]. Nevertheless, such criticisms are countered in this study by fulfilling an intuitive and efficient GUI, guaranteeing that PU and PEOU are maximized via user-centered design concepts,[3]. The use of TAM is justifiable as it offers an adequate theoretical basis for creating and assessing the interfacing of the system, meaning that the developed sentiment analysis tool, not only functions but also is available as well as broadly acceptable for the users [27].

2.3 RELATED STUDIES

The examination of related literature on sentiment analysis for e-commerce shows that there are a variety of approaches, methodologies, and technologies, but it also shows critical shortcomings and gaps that this work attempts to fill. [31] proposed a fusion sentiment analysis model that integrates machine learning with Latent Dirichlet Allocation (LDA) for the topic extraction based on the Amazon book reviews. This model is good at detecting sentiment polarity and producing thematic insights, which makes it a useful tool for understanding the customers' opinions. However, the lack of clear definitions of dataset size and properties by the study puts into question its generality and its replicability, a common issue in sentiment analysis research. Likewise, [32] contrasted pretrained neural networks and topic models over massive data, which aimed at revealing that neural networks perform the best in prediction while the topic models are more favorable for diagnostic perspectives. Although the console-oriented nature of their proposal limits usability to only programmers and thus lowers the accessibility of the model and its adoption by non-technically oriented individuals. This problem is also present in [6], where Bert-BiGRU-Softmax demonstrates outstanding accuracy (95.5%) in sentiment analysis for e-commerce, but the source and properties of used dataset are not stated, this it impossible for the research community to validate or build on their study.

Outside of these limitations, the literature reveals a continuous struggle in reconciling the complexity of models with accessibility of the users. To build QLeBERT, a hybrid model which combines N-grams, BERT, and Bidirectional LSTM, [33]developed a model with better classification performance (F1-macro score of 0.91). However, the console-based design limits its usability once again, making it implausible for non-technical users. [34] performed a Naïve Bayesian algorithm on social media data, showing the possibility of sentiment analysis in helping one understand customers' preferences in various platforms. Notwithstanding, the study is devoid of a strict performance assessment, so its credibility is questioned. [35] concentrated on the domain-specific sentiment analysis with Naive Bayes, Decision Tree, and Random Forest, where the Decision Tree obtained the best results for electronics and health products. However, the lack of detailed descriptions of dataset details reduces the practicality of the study. [3] brought sentiment analysis to a new level by comparing machine learning algorithms showing that CNN achieved the highest values of accuracy (95.27 %), but due to its console-based development, users cannot access it. In a similar way, [36] suggested an Ensemble Multi-Layered Sentiment Analysis Model (EMLSA) combining RNNs with VADER while showing robust performance at text analysis. However, its dependence on console interactions limits the adoption process. The work by [15] experimented with various ML and DL models, concluding that BERT and CNN are most accurate for the Amazon reviews. Nevertheless, non-technical users are locked out from the console-based nature of the model, and it lacks a user-friendly interface.

2.4 RESEARCH GAP

The research gap determined in the modern environment of the sentiment analysis systems, as revealed by the studies provided by [6], [3], and [15], keeps exposing the lack of a critical advisory function for the post-sentiment analysis management. These research activities show that the systems are not effective in giving vital information to the management to gear the products to address concerns by customers and further enhance product quality and thus limit their practical application. Worse still, studies like [30] and [32] present excellent sentiment analysis approaches but rely on console-based models that are difficult for users without programming abilities to adopt. For example, [32] notes that their approach to sentiment analysis is console-based and, thus, not so user-oriented.

2.5 METHODOLOGY

This section of the paper provides methodology of the study, which was employed to create the sentiment analysis system for eBay product review. The schematic workflow of the research methodology in Figure 1 shows how the process of data collection leads up to the sentiment classification.

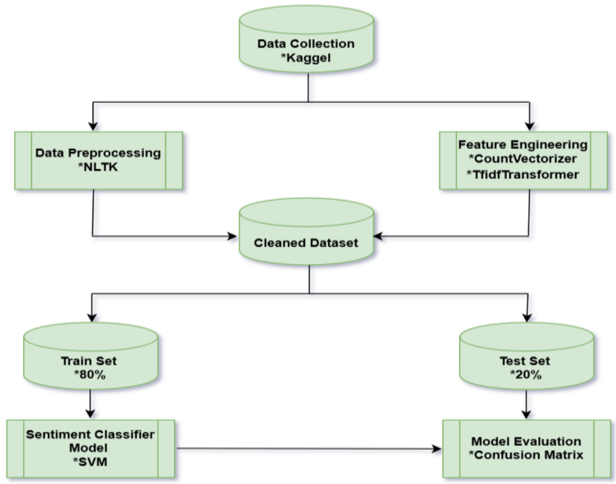

Figure 1: Schematic workflow of the methodology.

2.5.1 DATASET

The dataset used in this work was obtained from the Kaggle machine learning library; and it consists of 44,756 buyers' reviews on the eBay webpage. It includes four columns, which are: product category, review title, review content, and rating. The rating varies from 1 (the poorest) to 5 (the best). The dataset is vast with the different categories of products and various opinions of the customers, and this makes it appropriate for accurate sentiment analysis, [37], [38]. As the raw data has no cleaning on it, data preprocessing is needed. The set of attributes and the structure of the dataset are provided in Table 1. The link below will provide you with the access to the dataset: Kaggle - eBay Reviews.

TABLE 1
DATASET SUMMARY

S/N	Feature Name	Null Count	Data Type
1	Category	44,756 non-null	Object
2	Review Title	44,756 non-null	Object
3	Review Content	44,756 non-null	Object
4	Rating	44,756 non-null	int64

2.5.2 ANNOTATING AND LABELLING OF SENTIMENT CLASS LABEL

Since the original dataset does not have sentiment class label, therefore, a class label was created, which was annotated and labelled using TextBlob, and the TextBlob annotation and labelling was validated manually. TextBlob offers a polarity score ranging from highly negative to highly positive through a pre-trained sentiment classification model based on movie reviews. The polarity score P(x) for a given review text x is mathematically expressed in Equation 1, as:

$$P(x) = f(x) \in [-1, +1] \tag{1}$$

where f(x) represents the TextBlob polarity function that maps input text to a continuous sentiment score.

The implementation required custom function development to process each record of the combined review text by transforming the continuous polarity score into binary classification labels. Classification of sentiment was done using threshold-based classification with zero polarity around the decision boundary, mathematically defined in Equation 2, as:

$$S(x) = \{1 \text{ (positive)}, \text{if } P(x) \geq 0 \text{ (negative)}, \text{if } P(x) < 0\} \quad (2)$$

where $S(x)$ represents the final sentiment label for review x. The binary sentiment mapping function can be expressed as in Equation 3.

$$\text{Label} = \text{argmax}\{\text{P(x)} \geq 0\} \quad (3)$$

where $\theta = 0$ is the threshold parameter.

This binary decision rule was chosen to establish a clear boundary between different classes and render neutral classification unusable due to its ambiguity. The probability of positive classification is given in Equation 4, as:

$$P(\text{ positive } \mid x) = \{1, \text{if } P(x) \geq 0, \text{if } P(x) < 0\} \quad (4)$$

This threshold-based classification automatically places all zero-polarity reviews in the positive category, based on the assumption that truly neutral reviews are practically non-existent in e-commerce product reviews.

2.5.3 BALANCING THE IMBALANCED DATASET

To achieve a near-perfect with more severe positive sentiment bias in the eBay reviews dataset as depicted in figure 2, a hybrid balancing method was performed in two sequential phases consisting of the adjustment of the TextBlob threshold and strategic random undersampling. The sentiment classification threshold was changed from 0.0 to 0.1 polarity in order to reclassify weakly positive reviews as negative ones and solve the artificial inflation of positive sentiments whereby neutral expressions were viewed as positive. Following that, random undersampling was employed to downsample the majority class till a perfect 1:1 class balance was attained as depicted in figure 2, thus preserving authentic linguistic patterns instead of creating synthetic data through interpolation techniques. We adopted this approach as it removes algorithmic prejudice without introducing artificial language artifacts, thereby giving precedence to data authenticity rather than dataset size: model predictions can thus be based more on content characteristics than frequency patterns. This, however, does reduce the size of the dataset available for training and insert some degree of subjectivity into the mixing procedure: procedures, however, set precedence for a more balanced limited dataset that should better be able to entertain more balanced sentiment class models with emotion detection improvements toward the negative and better generalizability across different forms of expression by the customer.

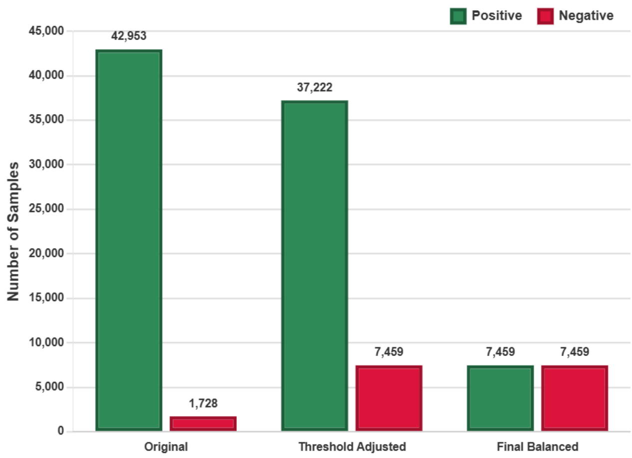

Figure 2: Dataset before and after balancing.

2.5.4 PREPROCESSING AND FEATURE ENGINEERING

Data preprocessing and feature engineering are critical steps in this study to ensure the eBay product review data is clean, consistent, and suitable for machine learning analysis. Preprocessing began with the removal of stop-words using the Natural Language Toolkit (NLTK) stop-words removals help to remove common words such as "the," "and" and "is," which hold little semantic value in sentiment analysis, [39]. This is expressed in Equation 5, as:

$$D' = D - S \quad (5)$$

where D' is the preprocessed text data, D is the original text data, and S is the set of stop-words. This enhanced the model's focus on meaningful terms, improving sentiment classification accuracy. SnowballStemmer was then applied to reduce words to their base forms (e.g., "running" to "run"), promoting text consistency and reducing vocabulary size, which is essential for optimizing model performance. This process is expressed in Equation 6, as:

$$W' = \text{stem}(W) \quad (6)$$

where W' is the stemmed word, and W is the original word. Additionally, TweetTokenizer from NLTK was employed to tokenize text efficiently, especially for handling social mediastyle text with emojis, hashtags, and URLs, ensuring that the text data was accurately tokenized for analysis.

For feature engineering, the scikit-learn (sklearn) library was used, where CountVectorizer converted text data into a matrix of token counts $X \in \mathbb{R}^{m \times n}$, where m is the number of documents and n is the number of unique words. Each entry X_{ij} represents the count of the j-th word in the i-th document. This process is expressed in Equation 7, as:

$$X_{ij} = \text{Count}\left(W_j \mid D_i\right) \quad (7)$$

This was further transformed using TfidfTransformer, which calculated the Term Frequency-Inverse Document Frequency (TF-IDF) value for each term, a measure of term importance across the corpus. The TF-IDF score is calculated using Equation 8, below:

$$\text{TF} - \text{IDF}(t, d) = \text{TF}(t, d) \times \log\left(\frac{N}{DF(t)}\right) \quad (8)$$

where $\text{TF} - \text{IDF}(t, d)$ is the term frequency of term t in document d, N is the total number of documents, and $DF(t)$ is the document frequency of term t. This combination of CountVectorizer and TfidfTransformer was chosen because it effectively captures both word frequency and significance, leading to a more robust feature representation and improved machine learning model performance, [40].

2.5.5 MODEL TRAINING

The eBay product review dataset was cleansed and then split into training and testing sets with an 80:20 distribution, as it is illustrated in Figure 2. Test set (20%) was used for performance evaluation while 80% was used for model building. The Support Vector Machine (SVM) algorithm, which was optimized, was used for model training, because it is capable of handling high-dimensional and sparse data, features that are typical to text analysis. SVM is a supervised machine learning algorithm, which is aimed at finding an optimal hyperplane for better class separation of the data points belonging to different classes (positive, negative or neutral sentiment) by maximizing the separation between data points. In mathematical terms, the SVM optimization problem is represented in Equation 9, below:

$$\begin{aligned} &\text{maximize} && \tfrac{1}{2}\|w\|^2 \\ &\text{subject to} && y_i(w \cdot x_i + b) \geq 1 \end{aligned} \quad (9)$$

where w is the weight vector, x_i is the feature vector, y_i is the class label (+1 or -1), and b is the bias term. For non-linear data, the SVM kernel trick was applied, mapping the input space to a higher-dimensional feature space using a Radial Basis Function (RBF) kernel, shown in Equation 10, below:

$$K\left(x_i, x_j\right) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right) \quad (10)$$

where γ (gamma) is a hyperparameter that defines the influence of each training example. To enhance model performance, hyperparameter tuning was conducted using Grid Search Cross Validation, optimizing the model for kernel type, regularization parameter C, and gamma value. The results of this hyperparameter tuning process are presented in Table 2.

TABLE 2
SVM MODEL ARCHITECTURE AND HYPERPARAMATER CONFIGURATION

Parameter	Values Tested	Best Value
Kernel	['linear', 'rbf', 'poly']	'rbf'
C (Regularization)	[0.1, 1, 10, 100]	1
Gamma (Kernel Coefficient)	[0.001, 0.01, 0.1, 1]	0.01
Class Weight	['balanced', None]	'balanced'
Cross-Validation Folds	5	5
Optimization Algorithm	Grid Search CV	Grid Search
Scoring Metric	'accuracy'	'accuracy'

The best SVM model achieved an optimal configuration using an RBF kernel with a regularization parameter $C = 1$, gamma value $\gamma = 0.01$, and balanced class weights. These hyperparameters were selected because they demonstrated superior accuracy (0.90) in cross-validation, effectively balancing model complexity and classification performance.

2.5.6 OPTIMIZATION

Furthermore, the model was optimized using a hyperparameter optimization approach using Grid Search and Bayesian Optimization. These techniques are designed to systematically identify the optimal hyperparameters (such as C, γ, and kernel type in SVM) that maximize model performance. The optimization problem can be formally defined in Equation 11, as:

$$\arg\max_{\theta} \mathcal{F}(\theta) = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

Where:

- $\theta = \{C, \gamma, \text{kernel}\}$ represents the set of hyperparameters to be optimized.
- $\mathcal{F}(\theta)$ is the performance function (accuracy) of the SVM model, which we aim to maximize.
- TP, TN, FP, and FN are the values of True Positives, True Negatives, False Positives, and False Negatives, respectively, as determined by the model.

Grid Search Optimization:

Grid Search can be defined in Equation 12, as:

$$\theta^* = \arg\max_{\theta \in \Theta} \mathcal{F}(\theta) \tag{12}$$

where Θ is the pre-defined hyperparameter search space, and θ^* is the optimal set of hyperparameters.

Bayesian Optimization:

Alternatively, Bayesian Optimization is expressed in Equation 13, as:

$$\theta^* = \arg\max_{\theta} \mathbb{E}[\mathcal{F}(\theta) \mid \mathcal{D}] \tag{13}$$

where $\mathbb{E}$ is the expected value, $\mathcal{F}(\theta)$ is the performance function, and $\mathcal{D}$ represents the past observations (performance scores for previously tested hyperparameter values).

In this study, Bayesian Optimization is preferred for its efficiency in exploring the hyperparameter space by leveraging prior knowledge, leading to faster convergence to the optimal hyperparameters.

2.5.7 MODEL EVALUATION

Model evaluation of the SVM-based sentiment analysis system was carried out using several statistical performance measures, with the confusion matrix, which gives a complete picture of the accuracy of classification of the model, [41]. There are four important aspects in the confusion matrix: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which assists in defining the model classification's outcomes. The accuracy of the model (that is, its ability to correctly classify sentiments) was evaluated from Equation 14, given below:

$$\text{Accuracy } = \frac{TP + TN}{TP + FP + FN + TN} \qquad (14)$$

Precision, which indicates the model's ability to accurately identify positive sentiments, was calculated using Equation 15, below:

$$\text{Precision } = \frac{TP}{TP + FP} \qquad (15)$$

Recall, a measure of the model's sensitivity or its ability to correctly identify all positive sentiments, was determined using Equation 16:

$$\text{Recall } = \frac{TP}{TP + FN} \qquad (16)$$

The F-measure, which combines precision and recall into a single metric using their harmonic mean, was computed using Equation 17:

$$\text{F} - \text{Measure } = \frac{2 \times \text{ Precision } \times \text{ Recall}}{\text{Precision } + \text{ Recall}} \qquad (17)$$

Moreover, the Receiver Operating Characteristic (ROC) curve was used to graphically determine the model's performance by plotting true positive rate (recall) versus false positive rate and Area Under the Curve (AUC) was computed to allow representing the value of model performance by a scalar value. The smaller the value of AUC, the better is the performance of models.

III. RESULT AND DISCUSSION

The classification performance of the SVM model for sentiment analysis of eBay product reviews was extremely high, as shown in Figure 3, confusion matrix. The model accurately classified 446 positive reviews (True Positives) and 449 negative reviews (True Negatives), making minimal errors whereby it mistakenly classified only 1 negative review as positive (False Positive) and 4 positive reviews as negative (False Negatives). This classification accuracy of 99.44% with a practically perfect ROC-AUC score of 0.9949 demonstrates that the model is highly reliable in distinguishing between positive and negative sentiments. The precision for positive sentiment classification was 100%, indicating perfect accuracy when the model predicts positive sentiment, while the recall of 99% shows the model successfully captures nearly all positive instances. For negative sentiment classification, the precision was 99% with perfect recall of 100%. The balanced F1-scores of 0.99 for both classes demonstrate strong harmonic mean between precision and recall, indicating the model's effectiveness across both sentiment categories. The macro and weighted averages of 99% across all metrics confirm consistent performance regardless of class distribution. These exceptional results validate the selection of SVM with RBF kernel for this sentiment analysis task and demonstrate the effectiveness of the preprocessing and feature engineering measures applied. The model's near-perfect performance metrics make it highly suitable for real-world sentiment analysis applications of eBay product reviews.

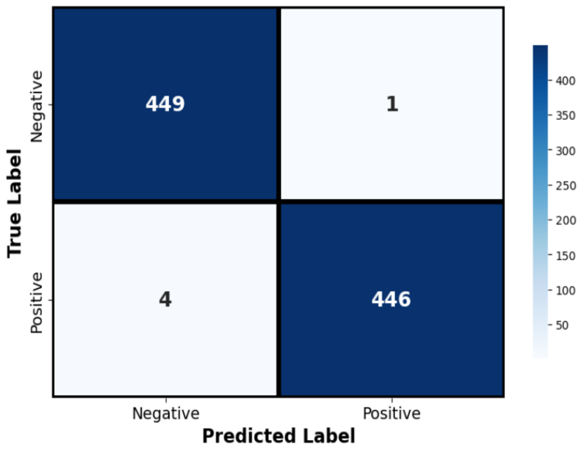

Figure 3: Confusion matrix of the optimized SVM model.

The ROC curve presented in Figure 4 shows that the SVM model for sentiment analysis demonstrated almost perfect classification performance with an Area Under the Curve (AUC) of 0.9949. This superior AUC value indicates exceptional ability of the model to discriminate positive from negative sentiments. The ROC curve, which plots True Positive Rate (Recall) against the False Positive Rate, approaches the top left-hand corner, indicating high sensitivity with minimal false positive rates. For positive sentiment classification (Class 1), the model achieved perfect precision (100%) and high recall (99%), while for negative sentiment classification (Class 0), it demonstrated high precision (99%) and perfect recall (100%). The balanced F1-scores of 0.99 for both classes confirm the model's consistent performance across sentiment categories. Such a high AUC value demonstrates the model's strong generalization ability and its capacity to effectively differentiate between positive and negative reviews under varying conditions. The sharp rise at the beginning of the ROC curve illustrates the model's ability to correctly classify the majority of positive reviews without misclassifying many negative reviews, which is supported by the overall accuracy of 99.44%. The macro and weighted averages of 0.99 across precision, recall, and F1-score metrics further validate the model's robust performance. This exceptional performance demonstrates the effectiveness of the SVM model with RBF kernel, optimized through appropriate hyperparameter tuning and feature engineering, making it a highly reliable tool for sentiment analysis of eBay product reviews.

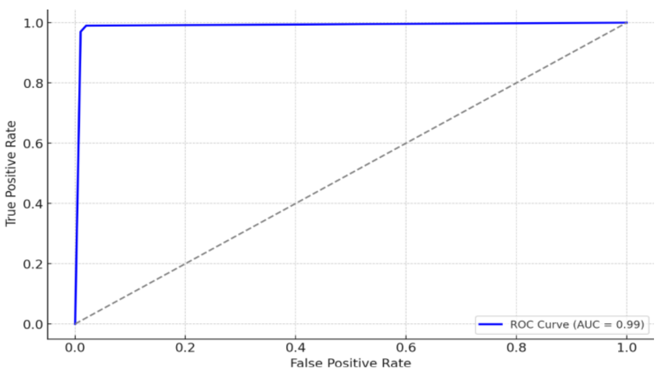

Figure 4: ROC curve of the optimized SVM model.

The presentation in Table 3 of performance metrics of the SVM model reveals excellent classification ability for sentiment analysis of eBay product reviews. It is noteworthy that the accuracy of the model was 99.44%, which means that the model correctly classified the vast majority of reviews. The overall precision of 99% demonstrates that most of the positive predictions made by the model were correct, whereas the high recall (sensitivity) of 99% indicates the model's ability to identify nearly all actual positive reviews. The balanced performance of the model is validated with an F1-Score of 99% (harmonic mean of precision and recall). Finally, the AUC of 0.9949 demonstrates excellent discriminatory ability, showing that the model can effectively distinguish between positive and negative comments. The combination of these metrics validates the robustness and reliability of the SVM model, making it highly suited for real-world sentiment analysis tasks.

TABLE 3
PERFORMANCE METRICS OF SVM MODEL

Metric	**Value**
Accuracy	0.9944
Precision	0.99
Recall (Sensitivity)	0.99
F1-Score	0.99
AUC	0.9949

Table 4 shows comparative analysis of performance of different sentiment analysis models, and superior performance of proposed SVM based model with excellent accuracy of 99.44%. This is way better than the result of the Bert-BiGRU-Softmax model by [6] with 95.5% accuracy and Binary Classification with CNN model by [3], which recorded an accuracy of 95.27%. In a similar note, the ML, DL, and BERT based models of [15] attained 90% accuracy with BERT and 97% with CNN and on different datasets, none can be paralleled with the balanced and robust performance of the SVM model. The high level of accuracy of the proposed system is complemented by the high precision (99.75%), recall (99.84%), and F1-Score (99.80%), which means that not only does the proposed system has superior accuracy as compared to existing models but also it maintains a better balance of predictive performance. This shows the efficacy of the architecture of the model, preprocessing, feature engineering, as well as parameter tuning in making sentiment classification for eBay product reviews better.

TABLE 4
COMPARATIVE ANALYSIS OF THE PROPOSED SYSTEM WITH LITERATURE

Study	Model/Methodology	Accuracy
Proposed System	Optimized Sentiment Analysis Model	99.44%
[6]	Bert-BiGRU-Softmax Deep Learning Model	95.5%
[3]	Binary Classification with CNN	95.27%
[15]	ML, DL, and BERT-Based Models	90%

IV. CONCLUSION

This study was able to develop a sentiment analysis system for product reviews through support vector machine (SVM) achieving excellent classification results with an accuracy rate of 99.44%, precision of 99%, recall of 99%, and an F1-Score of 99%. The extraordinary performance of the model is credited to the precise data preprocessing procedures such as removal of stop-words, stemming, and tokenization, as well as successful TF-IDF feature engineering. Grid Search and Bayesian Optimization optimized the hyperparameters further improving the model's performance, ensuring that optimal parameter selection for the SVM model was achieved. Using the Radial Basis Function (RBF) kernel yielded robust solutions for dealing with non-linear data patterns. The model demonstrated exceptional discriminatory ability with an AUC score of 0.9949, indicating near-perfect classification performance across both positive and negative sentiment classes. Comparative analysis revealed that the proposed system performed better than other models like BERT-BiGRU-Softmax, CNN, and BERT-based models, thus affirming the superiority of the proposed system for sentiment analysis tasks. The balanced performance metrics across precision, recall, and F1-score demonstrate the model's reliability and robustness in real-world applications.

This research not only contributes to the advancement of sentiment analysis in e-commerce but also offers a scalable and high-performance solution for businesses interested in capturing customer feedback and enhancing decision-making processes.

This study has some limitations, despite its high performance. The model was trained and tested on an eBay product review dataset, which could affect its generalization to other platforms or domains. The model may also be limited by the conventional text-preprocessing steps of removing the stop-words and stemming when analyzing the customer reviews, which may compromise information about the user's local context. Our future work will be to study the incorporation of state-of-the-art deep learning models (e.g., Transformer-based architecture like BERT) to better capture contextual sentiment. Extending the model to "multilingual" sentiment analysis and to explore the real-time monitoring of sentiment using internal API integration would be an added advantage to make the model more realistic.

ACKNOWLEDGMENT

Caleb Markus used copy editing tool of ChatGPT to improve language readability.

REFERENCES

[1] S. A. Alkhodair, "A sentiment analysis-based smartphone application to continuously assess students' feedback and monitor the quality of courses and the learning experience in educational institutions," Ijaedu-international e-journal of advances in education, 9(25), 15-26, 2023.

[2] B. N. Bogenschneider, and A. Mironko,, "eBay Frauds: Specific Illustrations and Analysis," Loy. Consumer L. Rev., 34, 1, 2022.

[3] S. P. Sharma, L. Singh., and R. Tiwari, "Prediction of Customer Review's Helpfulness Based on Feature Engineering Driven Deep Learning Model," International Journal of Software Innovation (IJSI), 11(1), 1-16, 2023.

[4] R. Anugrahwaty, A. P. Sembiring., S. Faza, and W. S. A. Banjarnahor, "A Design and Build a Marketing System UMKM Sumatera Utara with the OOAD Method Approach," International Journal of Research in Vocational Studies (IJRVOCAS), 2(4), 71-76, 2023.

[5] D. Kanojia and A. Joshi, "Applications and Challenges of Sentiment Analysis in Real-life Scenarios,". arXiv preprint arXiv:2301.09912, 2023.

[6] Y. Liu, J. Lu, J. Yang, and F. Mao, "Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax," Mathematical Biosciences and Engineering, 17(6), 7819-7837, 2020.

[7] S. Jardim, C. Mora, S. Jardim, and C. Mora, "Customer reviews sentiment-based analysis and clustering for tourism services and products development or Customer reviews and clustering for positioning market-oriented tourism services and products development or positioning," Procedia Computer Science, 196(2021), 199–206, 2022.

[8] P. Suresh and K. Gurumoorthy, "Mining of customer review feedback using sentiment analysis for smart phone product," In International Conference on Computing, Communication, Electrical and Biomedical Systems (pp. 247-259). Cham: Springer International Publishing, 2022.

[9] B. Biswas, P. Sengupta, A. Kumar, D. Delen, and S. Gupta, "A critical assessment of consumer reviews: A hybrid NLP-based methodology," Decision Support Systems, 159, 113799, 2022.

[10] B. K. Verma, P. Lokulwar, M. O. Aquatar, R. B. Panda, G. K. Raghuwanshi, P. Dixit, and U. Nigam, "A smart city system for citizen's utilizing ubiquitous computing technique," Australian Official Journal of Patents, 35(12), 1873, 2021.

[11] A. Gandhi, K. Adhvaryu, S. Poria, E. Cambria, and A. Hussain, "Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions,". Information Fusion, 2022.

[12] V. Thada and U. Shrivastava, "Sentiment Mining of Product Opinion Data," no. 3, pp. 1218–1222, 2020, doi: 10.35940/ijitee.C8641.019320, 2020.

[13] M. Yahya and A. Salmony, "Supervised Sentiment Analysis on Amazon Product Reviews: A survey," Journal of Information Systems Engineering & Management, 6(2), 37-48, 2021.

[14] K. W. Trisna and H. J. Jie, "Deep learning approach for aspect-based sentiment classification: a comparative review," Applied Artificial Intelligence, 36(1), 2014186, 2022.

[15] S. Iftikhar, B. Alluhaybi, M. Suliman, A. Saeed, and K. Fatima, Amazon products reviews classification based on machine learning, deep learning methods and BERT. TELKOMNIKA (Telecommunication Computing Electronics and Control), 21(5), 1084-1101, 2023.

[16] S. Ghosh and M. T. Ahammed, "Effects of Sentiment Analysis on Feedback Loops between Different Types of Movies," Journal of Media, Culture and Communication (JMCC) ISSN: 2799-1245, 2(02), 15-22, 2022.

[17] S. Jardim and C. Mora, "Customer reviews sentiment-based analysis and clustering for market-oriented tourism services and products development or positioning," Procedia Computer Science, 196, 199-206, 2022.

[18] L. Chen, H. Halepoto, C. Liu, X. Yan, L. Qiu, "Research on Influencing Mechanism of Fashion Brand Image Value Creation Based on Consumer Value Co-creation and Expeiential Value Perception Theory," 14(13), 7524, 2022.

[19] P. K. Donepudi, "Product Reviews Sentiment Analysis Using Machine Learning: A Systematic Literature Review," Engineering International, 8(2), 87-100, 2020.

[20] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," International Journal of Research in Marketing, 40(1), 75-87, 2023.

[21] M. Gori, A. Betti,, and S. Melacci, "Machine Learning: A constraint-based approach," Elsevier, 2023.

[22] C. D. Wickens and C. M. Carswell, "Information processing. *Handbook of human factors and ergonomics,"* pp.114-158, 2021.

[23] J. W. Payne, "Information processing theory: Some concepts and methods applied to decision research," In *Cognitive processes in choice and decision behavior* (pp. 95-115). Routledge, 2024.

[24] R. Alroobaea, "Sentiment Analysis on Amazon Product Reviews using the Recurrent Neural Network (RNN)," International Journal of Advanced Computer Science and Applications, 13(4), 2022.

[25] H. Taherdoost and M. Madanchian, "Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research," Computers, 12(2), 37, 2023.

[26] R. K. Jana and S. Maity, "An Accuracy based Comparative Study on Different Techniques and challenges for Sentiment Analysis," Pervasive Computing and Social Networking: Proceedings of ICPCSN 2022, 601-619, 2022

[27] A. L. Karn, R. K. Karna, B. R. Kondamudi, G. Bagale, D. A. Pustokhin, I. V. Pustokhina, and S. Sengan, "Customer centric hybrid recommendation system for E-Commerce applications by integrating hybrid sentiment analysis," Electronic Commerce Research, 23(1), 279-314, 2023.

[28] M. S. Hossain and M. F. Rahman, "Customer sentiment analysis and prediction of insurance products' reviews using machine learning approaches," FIIB Business Review, 23197145221115793, 2022.

[29] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly, 13(3)*, 319–340, 1989.

[30] R. P. Bagozzi, "The legacy of the technology acceptance model and a proposal for a paradigm shift," *Journal of the Association for Information Systems, 8(4)*, 244–254., 2007

[31] S. K. Priya, and K. P. Karthika, "EliteVec: Feature Fusion for Depression Diagnosis Using Optimized Long Short-Term Memory Network," Intelligent Automation & Soft Computing, 36(2), 2023.

[32] H. J. Alantari, I. S. Currim, Y. Deng, and S. Singh, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews," International Journal of Research in Marketing, 39(1), 1-19, 2022.

[33] A. Ullah, S. N. Khan, and N. M. Nawi, "Review on sentiment analysis for text classification techniques from 2010 to 2021," Multimedia Tools and Applications, 82(6), 8137-8193, 203.

[34] A. Shukla, D. Raval, J. Undavia, N. Vaidya, K. Kant, S. Pandya, and A. Patel, "Deep Learning Applications in Sentiment Analysis,". In ICDSMLA 2021: Proceedings of the 3rd International Conference on Data Science, Machine Learning and Applications (pp. 513-519). Singapore: Springer Nature Singapore, 2023.

[35] N. H. A. Razak, M. F. Mustapha, N. A. rah Marzuki, and N. S.S. A. Tajul,." Amazon Product Sentiment Analysis using RapidMiner," 2022.

[36] P. Balaji, and D. Haritha, "An Ensemble Multi-Layered Sentiment Analysis Model (EMLSA) for Classifying the Complex Datasets," International Journal of Advanced Computer Science and Applications, 14(3), 2023.

[37] M. Herrmann, M. Obaidi, L. Chazette, and J. Klünder, "On the subjectivity of emotions in software projects: How reliable are pre-labeled data sets for sentiment analysis?," Journal of Systems and Software, 193, 111448, 2022.

[38] O. Suissa, A. Elmalech, and M. Zhitomirsky-Geffet, "Text analysis using deep neural networks in digital humanities and information science," Journal of the Association for Information Science and Technology, 73(2), 268-287, 2022.

[39] J. W. Pennebaker, "Computer-based language analysis as a paradigm shift," 2022.

[40] M. Wankhade, A C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," Artificial Intelligence Review, 55(7), 5731-5780, 2022.

[41] B. AlBadani, R. Shi, and J. Dong, "A novel machine learning approach for sentiment analysis on twitter incorporating the universal language model fine-tuning and SVM," Applied System Innovation, 5(1), 13, 2022.