Journal of Global Research in Electronics and Communication (JGREC)
Volume 2 | Issue 11 | November 2026 | Pages 08–15
DOI: https://doi.org/xxxxxxxxxxxxx
e-ISSN: 2321-3175
Research Article OPEN ACCESS Available online: https://jgrec.info/index.php/jgrec/index

Artificial Intelligence-Driven Sentiment Analysis for Product Reviews in Online Retailing Platforms

Dr. Bal Krishna Sharma
Professor
Department of Computer Science and Application
Mandsaur University, Mandsaur
drbks7426@gmail.com

Corresponding Author: * Bal Krishna Sharma (drbks7426@gmail.com)

Received: 5 November 2026 | **Revised:** 7 November 2026 | **Accepted:** 12 November 2026

***Abstract*—Machine learning and large language models (LLMs) are two examples of AI methods used for sentiment analysis that can efficiently categories texts as positive, negative, or neutral. Its capacity to automatically detect emotions in texts is utilized by numerous campaigns related to marketing, social media, product reviews, and hate speech. In order to categories the tone of product reviews taken from the Amazon Customer Review dataset, this study suggests a deep learning model. A combination of CNNs and LSTMs (Long Short-Term Memory) networks make up the model. Tokenization, normalization, stop word deletion, and sequence padding are all part of the suggested systematic approach to data preprocessing in the given method. For feature extraction, it recommends the Bag-of-Words model. You may divide the data for training and testing 80/20 with this dataset, which contains 400,000 ratings across five product categories. The LSTM layer use long-term dependency modelling to enhance classification accuracy, whereas the CNN module seeks out local textual patterns. Testing shows that the CNN-LSTM model beats more traditional models like Random Forest, Naive Bayes, and independent CNN in terms of accuracy (98%), precision (98.52%), recall (98.49), and F1-score (98.49). The findings demonstrate that deep hybrid architecture is effective for sentiment analysis of large e-commerce platforms.**

Keywords*—Sentiment analysis, artificial intelligence, machine learning, CNN-LSTM, Amazon product reviews, text classification, emotion detection, class imbalance.

I. INTRODUCTION

The growth of online marketplaces in the last decades has been a sharp vertical and is now a part of everyday business. The increase of online platforms has also resulted in the flourishing of user-generated content, especially product reviews, that is an essential feedback system to the businesses and an informative tool to future customers [1]. Online customer reviews (OCRs)—peer-generated evaluations of products published on a business's or an outside platform's website—are a major factor in shaping consumer behavior and product perceptions [2][3]. Findings indicate that consumers rely heavily on these reviews to make purchases which influences the sales of the products in a great magnitude [4].

The mainstream of the platforms like Amazon, Taobao, and eBay indicate the increasing popularity of e-commerce, as the online shopping experience is more convenient to some of the consumers [5][6][7]. Such a shift is also promoted by the popularity of social networks and general digital media usage, particularly among younger generations, the use of which has established a natural online interaction, including reviews and feedback as a regular part of retail experiences. This has also made online retailing to be closely embedded into the social media platform making companies to consider the customer sentiment in their business strategies [8].

Businesses are increasingly relying on sentiment analysis as an AI tool to gauge customer sentiment, service seriousness, social media mentions, and the effectiveness of customer interactions [9][10]. Future predictive systems of sentiment analysis will provide even more insight about human emotions, which will enable computers to gain a greater insight into context, tone and sentiment in communications as the field of AI keeps improving [11]. This is necessary in modeling more precise and understanding replies to customer data.

There is a greater need than ever before to have reliable and large-scale sentiment analysis systems [12][13]. These systems assist in many business processes, including marketing, customer service and product research by giving live measurements of consumer opinion. Specialists in various spheres, such as politics, sports, and business, turn to sentiment analysis tools more and more often to evaluate the opinion of the population concerning the subject of interest. At the same time, there is a change in the retail sphere due to sustainability issues. The current state of sustainability as a competitive advantage is that it especially appeals to Millennials and Generation Z consumers, who are more concerned with ethical and environmentally conscious behavior [14][15][16]. Not only do retailers who are embracing sustainable practices enjoy the added bonus of consumer trust, but they also experience efficiency in their operations and cost savings. The fashion industry, specifically, has reacted by shifting to circular fashion systems as a response to fast fashion criticisms

Copyright © 2026 The Author(s).
This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author(s) and the source are properly credited.

J. Glob. Res. Elec. Comm., 2 (11):08-15
DOI: https://doi.org/xxxxxxxxxxxx

Nevertheless, the overwhelming amount and unpredictability of consumer views on the internet create confusion and ambiguity regarding customer choice [17][18]. Different viewpoints and ratings on a single product usually cause confusion and indecision when purchasing an item. In order to guarantee the correct use of this data and provide more tailored consumer experiences, businesses are increasingly turning to AI and ML technology. In order to automatically assess the emotional condition of customers expressed in textual feedback, organizations can benefit from sentiment analysis that utilizes artificial intelligence.

A. Motivation and Contribution of Paper

The blistering growth of e-commerce platforms and the sheer number of reviews left by customers are present online. These reviews are vital critic to both customers and companies as they do affect their purchase decisions and product development strategies. However, manually analyzing such a vast corpus of textual data is impractical. Traditional ML techniques often fall short in capturing the nuanced sentiment patterns and contextual dependencies within unstructured text data. Therefore, there is a pressing need for an intelligent, automated, and accurate sentiment analysis system that can efficiently process large-scale review datasets. This led to the creation of a hybrid DL technique that makes use of LSTM networks' advantages for sequential dependency modelling and CNN's advantages for feature extraction. The key contributions of the study are discussed below:

- A large selection of 400,000 evaluations (80,000 from each of five product categories) were the subject of an analysis of a publicly accessible Amazon review dataset from Kaggle.
- Implemented a detailed pre-processing pipeline including data cleaning, normalization of repeated characters, tokenization, padding for uniform length, and removal of hyperlinks and hashtags to refine text data.
- Introduced a robust CNN-LSTM hybrid DL architecture that effectively combines spatial feature extraction and temporal sequence modeling for sentiment classification.
- Performed a thorough performance analysis utilizing the standard metrics F1-Score, Accuracy, Precision, and Recall.

B. Novelty & Justification of the Study

The justification for adopting a hybrid CNN-LSTM architecture lies in the two DL models' complementary strengths: While LSTMs are excellent at collecting contextual sequences and long-range relationships, CNNs are skilled at extracting local and spatial information from text. Traditional ML models and standalone deep networks often struggle to balance these aspects effectively, especially when applied to complex and large-scale review data. The novelty of this study stems from its integration of a sequential preprocessing pipeline with a dual-layered DL model, specifically tailored for sentiment analysis of e-commerce reviews. Unlike prior works that rely solely on either statistical methods or individual deep learning models, this research leverages the synergy between CNN and LSTM to enhance classification accuracy and robustness. Moreover, the model is validated on a sizable and diverse real-world dataset, which strengthens its practical relevance and demonstrates its superiority over baseline approaches in handling heterogeneous textual data in online retail platforms.

C. Structure of Paper

This study is organized into main components, which allows for a complete and consistent presentation. An outline of the existing literature on sentiment analysis of online product reviews is provided in Section II. Methods for collecting data, cleaning it up, and putting it all together are detailed in Section III. The experimental results of the model are covered in Section IV, which also addresses the model's performance and key findings. Section V provides a summary of the study, discusses its shortcomings, and suggests areas for future research.

II. LITERATURE REVIEW

This study of the literature discusses sentiment analysis of online product evaluations and provides a detailed analysis of recent research on the subject. These studies are summarized in Table I presenting the methodologies, the performance outcomes, main findings, identified limitations, and the suggested future directions of study.

Rongala et al. (2025) The availability of product reviews online including on Amazon has grown in recent years as the trend of using online shopping has grown. Sentiment analysis or opinion mining is an idea whose concept is built on the identification of sentiment or opinion of a certain piece of text and its components based on reviews, tweets or customer reactions using eLearning, computational linguistics, ML, and NLP. This paper takes a closer look at how ML and NLP techniques can be applied to product reviews on Amazon and extract influential attitudes, tastes, and trends. The F1-Score, recall, accuracy, and precision are additional metrics considered in model evaluations, such as the BERT. With 89% scores in accuracy, precision, recall, and F1-score, BERT (Bidirectional Encoder Representations from Transformers) outperformed other approaches for sentiment categorization and review analysis [7].

Mane, Dongre and Madankar (2025) customer all the reviews are genuine. But nowadays many of the businesses manipulated the reviews of their products by inserting the fake reviews to misleads the consumer to buy wrong product. There is a need of the effective mechanism which can identify the fake reviews and increases the trustworthiness of the online reviews. Many researchers working on the different techniques available in different domains such as ML, NLP, Opinion Mining, etc. In this paper the fake reviews are identified using the random forest classifier along with different challenges identified for improvement. The amazon reviews dataset has been tested for 90000 reviews to achieves accuracy up to 79.47% [19].

Bagastio, Suakanto and Fakhrurroja (2024) provides a thorough literature analysis that contrasts DL techniques with conventional ML methods for detecting phony reviews across a variety of areas. The study charts the progression from DL techniques like LSTM, CNN, and BERT to ML techniques like SVM and NB. Although ML techniques have been successful in spotting simple patterns, DL models have outperformed them, especially when it comes to spotting more intricate and flexible bogus reviews. The ability of these DL algorithms to identify intricate and flexible bogus reviews has significantly improved [20].

DOI: https://doi.org/xxxxxxxxxxxx

Sridhar and Nagasundaram (2024) proposes a unique method for emotional analysis based on product evaluations by extracting features from internet reviews and classifying them using a ML model. Data was normalized and noise was removed from the input, which was collected from an online review. Then, this data features are extracted and classified using graph recurrent component analysis with support Bayesian fuzzy neural network (GRCA-SBFNN) for detection of the user sentiment for the reviewed product. Recall, F-1 score, AUC, average precision, and training accuracy are some of the metrics that have been experimentally studied for different online product review datasets. The strategy that is suggested F-1 score 92%, training accuracy 98%, average precision 97%, AUC 91%, and recall 96% [21].

N et al. (2024) Fake_reviews_dataset collected by Joni Salminen and have a license :The CC-By Attribution 4.0 International experimentation includes a wide range of evaluations classified as either authentic or fraudulent. Algorithms are able to efficiently process the data since textual reviews are represented as numerical vectors using feature extraction techniques. Several performance measures are utilised in the process of creating, training, and assessing algorithms. These include F1score, recall, precision, and accuracy. Three different algorithms—RF, SVM, and KNN—attained accuracies of 80.56%, 81%, and 71%, respectively, in detecting false reviews. In the fake review identification procedure, the SGD (Stochastic Gradient Descent) algorithm has an accuracy of 86.7%, which is 15.7% higher than KNN, 5.7% higher than SVM, and 6.14% higher than RF [22].

Mukit et al. (2023) monitor reviews given by the customer after buying some product from an online market, where easily find out what are the companies or shops are making frauds with the people. Therefore, produced our own "PR" (Product Review) dataset, which was made up of 4,000 reviews from YouTube videos, social media platforms, online shop pages, and e-commerce websites in both Bangla and Phonetic Bangla. manually classified the gathered data as "positive" or "negative." In order to train our model using supervised ML methods, first extract features from the data using TF-IDF. In both the Bangla and Phonetic Bangla datasets, SVM demonstrated the maximum accuracy, with 82% and 94%, respectively, according to the results [23].

Gope et al. (2022) Analyze the tone of Amazon customer reviews and ratings using their dataset. Machine learning algorithms such as Linear Support Vector Machine, Logistic Regression, Random Forest, Multinomial Naive Bayes, and Bernoulli Naive Bayes were employed. achieved a success rate of 91.90% using the Random Forest classifier. Along with other deep learning techniques, our research utilized RNN with LSTM and achieved an impressive maximum accuracy of 97.52%. When it comes to our model, RNN-LSTM is the way to go [24].

Jain, Kumar and Goyal (2022) The supervised learning model that was created is used when the website is launched. The user must first provide the URL of the website containing the merchandise. Next, Python web scraping techniques are used to extract the dataset from the provided URL. Sound characteristics are then extracted from the data through analysis and dissection utilizing Natural Language Processing algorithms. Finally, use the dataset to train more machine learning models. The experimental findings demonstrate that the model attains a dataset accuracy of 89.12%. In order to prevent companies from losing money at an exponential rate and to provide consumers with more trustworthy review data, this study primarily aims to design a fake review screening system [25].

TABLE I. SUMMARY OF LITERATURE OVERVIEW AND REVIEW ON SENTIMENT ANALYSIS IN ONLINE PRODUCT REVIEWS.

Study	Methodology	Datasets	Key Findings	Limitations/Future Scope
Rongala et al. (2025)	ML and NLP techniques, sentiment analysis using BERT	Amazon product reviews	BERT achieved 89% accuracy, 88% precision and recall, 89% F1-score	Limited to BERT; scope to evaluate other LLMs like RoBERTa, DeBERTa
Mane, Dongre & Madankar (2025)	Random Forest classifier for fake review detection	Amazon reviews (90,000 reviews)	Accuracy of 79.47% for fake review detection	Model may not detect complex fake patterns; future work can include deep learning
Bagastio, Suakanto & Fakhrurroja (2024)	Literature review comparing ML vs. DL methods	Multiple datasets across domains	Deep learning methods like BERT, LSTM outperformed traditional ML	Lacks experimentation; future scope includes hybrid or ensemble models
Sridhar & Nagasundaram (2024)	GRCA-SBFNN model with feature extraction & sentiment classification	Multiple product review datasets	Achieved 98% training accuracy, 97% precision, 91% AUC, 92% F1-score	Model complexity may reduce real-time applicability; future work on simplification
N et al. (2024)	Feature extraction with ML classifiers (RF, SVM, KNN, SGD)	Fake_Reviews_ Dataset (Joni Salminen)	SGD performed best (86.7% accuracy); outperforming SVM, RF, KNN	Evaluation limited to classic ML models; future scope includes BERT-based models
Mukit et al. (2023)	TF-IDF + supervised ML (SVM) on Bangla/Phonetic Bangla	Custom PR dataset (4,000 Bangla reviews)	SVM achieved 82% (Bangla) and 94% (Phonetic Bangla) accuracy	Dataset size is limited; potential to expand and apply DL models
Gope et al. (2022)	ML models (RF, NB, SVM, LR) and RNN-LSTM	Amazon reviews	LSTM achieved 97.52% accuracy, RF: 91.9%	Model tailored to sentiment only; future research could include fake review detection
Jain, Kumar & Goyal (2022)	Web scraping + NLP + ML for fake review filtering	Real-time web-scraped product reviews	Model achieved 89.12% accuracy in identifying fake reviews	Limited dataset scope; automation for real-time detection can be enhanced

III. METHODOLOGY

The process used in online shopping to figure out how people feel about product reviews (Figure. 1) starts with getting the Amazon Customer Review Dataset, which is made up of users' raw text feedback [26]. Data cleaning, tokenization, normalization, hashtag deletion, data cleansing, and sequence padding are all part of the structured preparation phase that this data goes through. Feature extraction is the next step after preprocessing in order to convert textual input

J. Glob. Res. Elec. Comm., 2 (11):08-15
DOI: https://doi.org/xxxxxxxxxxxx

to numerical representations. The following step is to divide the processed dataset into two parts, one for training and one for testing, by dividing it in half, 80:20. We use these subsets to train and evaluate a CNN-LSTM hybrid model that can understand the review texts' long-term dependencies as well as their local semantic information. When assessing the model's efficacy, popular measures including F1-score, recall, accuracy, and precision are employed. Lastly, the accuracy of the e-commerce platform industry's AI-driven sentiment analysis system is assessed by thoroughly analyzing the results to understand the classification outputs.

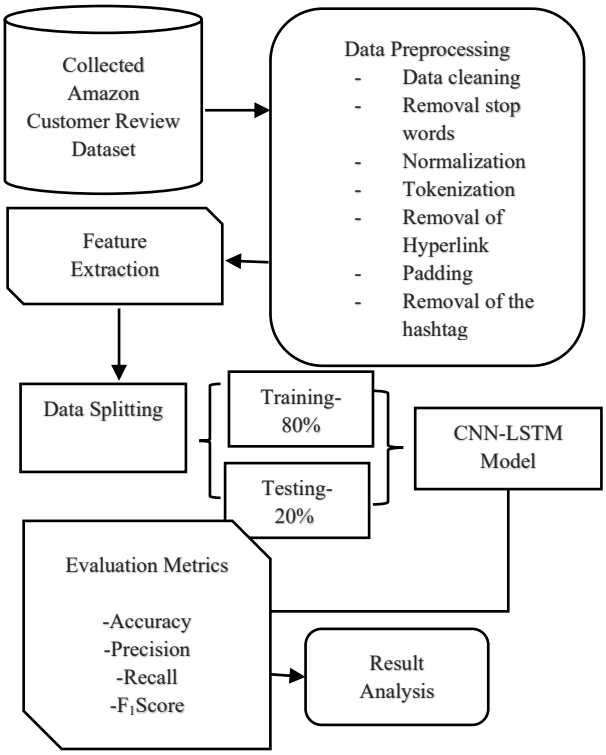

Fig. 1. Flowchart Diagram of the Sentiment Analysis in Online Product Reviews

Here is a comprehensive, step-by-step breakdown of the processes illustrated in the flowchart:

A. Data Collection

The Kaggle platform makes the customer review dataset, which is publicly available and obtained from Amazon, available for sentiment analysis (SA). Reviews from five different product categories are included in the dataset: watches, mobile phones, furniture, cameras, and groceries. These Amazon reviews were available to the general public on Kaggle. With a total of over 100 million reviews, the dataset is large. however, decided to concentrate on 80,000 evaluations from each of these categories, yielding a total dataset of 400,000 entries. In below visualization graphs and word clouds are:

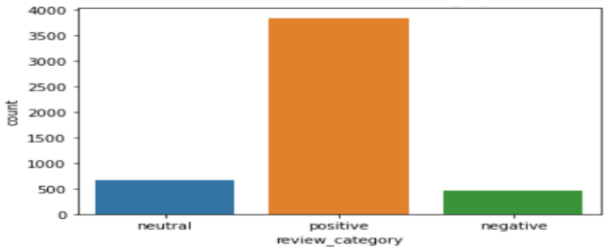

Fig. 2. Distribution of Reviews' Categories

Figure 2 is a bar chart Distribution of reviews categories. The number of reviews classified as neutral, positive, and unfavorable is shown. The review category is represented by the x-axis, which has three labels: "neutral," "positive," and "negative." The y-axis, which ranges from 0 to 4000, shows the number of reviews. The bar for "positive" reviews is the tallest, indicating approximately 3800 reviews. The "neutral" review bar shows a count of around 650, while the "negative" review bar is the shortest, with a count of about 500.



Fig. 3. Most Common Words used in Positive Reviews

The most often used terms for positive reviews are shown in Figure 3. The word cloud shows the frequency of words graphically; higher frequency is shown by larger letter sizes. Prominently displayed in large fonts are words such as "look," "watch," and "good," suggesting their high occurrence in positive reviews. Other frequently appearing words, though in smaller fonts, include "nice," "product," "money," "value," "great," "love," "best," and "amazon." This graphic gives a brief synopsis of the main ideas and phrases found in favorable client reviews.



Fig. 4. Most Common Words in Negative Reviews

The most frequently used terms in bad evaluations are shown in Figure 4. Each word's magnitude in the cloud represents how frequently it appears in unfavorable reviews; bigger terms are more prevalent. Prominent words, indicating high frequency, include "product," "watch," "disappoint," "bad," "quality," "return," and "defect." Other visible words, though smaller in size, that appear in negative reviews are "receive," "strap," "damage," "time," "month," "day," "broken," "poor," and "amazon." This visualization effectively highlights the recurring themes and issues mentioned in critical customer feedback.

B. Data Preparation

An essential step in text data analysis is data pretreatment. [27]. The complexity of text data is increased by the presence of repetitions and redundancies in textual materials like reviews, blogs, and tweets. Data preprocessing steps that are involved in the study are data cleaning, normalization, tokenization, removing hyperlink, padding, removing hashtag and feature extraction. The pre-processing steps discuss below:

DOI: https://doi.org/xxxxxxxxxxxxx

- **Data Cleaning:** Data cleaning was done in a way that was specific to the analytic goals, thus everything that wasn't necessary for the study was removed. Facts, URLs, @user tags, numerals, subjects, punctuation, hashtags, and additional whitespace were all part of the items. The text was further improved by removing special characters in addition to letters and spaces.
- **Removal stop words:** Many words and phrases appear several times in text files. Eliminating the stop words is, hence, crucial. As a matter of fact, stop words never enrich written expression. Frequently, the text contains a lot of these terms.
- **Normalization:** Normalizing repeated letters inside words involved eliminating redundancies, streamlining the language, and reducing successive identical characters to a single instance.[28].
- **Tokenization of the Text:** The method of tokenization involves dividing a text sequence into discrete parts called tokens. Single words, phrases, or even complete sentences may be included in these tokens. After that, these tokens are used as inputs for a number of procedures, including text mining and parsing. The models are able to zero in on specific parts, rather than the entire text, which is essential.
- **Removal of Hyperlink:** Hyperlinks in databases are obsolete. The links serve no purpose other than to facilitate usage. However, in order to change the text's polarity based on the data collect, exclusively use reviews, comments, and tweets to relay thoughts and opinions. Deleting the links from the datasets is therefore essential.
- **Padding:** The classifier has issues during sentiment analysis since the consumer review databases contain both very short and very long reviews. When the network evaluates the review, the number of pixels added is referred to as CNN-related padding [29]. To make sure that all of our customers' reviews have the same length, use padding, which is basically wrapping up our input reviews with zeros.
- **Removal of the hashtag:** These days, hash tagging is also very common. Hash tagging is often utilized in customer feedback. A lot of memory is used up by hashtags. Hashtags provide no purpose in sentiment analysis. The classifiers' uncertainty is only heightened by these. Therefore, hashtag removal is essential. By removing the hashtags from the datasets, the training data becomes more concise and lucid.

C. *Feature Extraction*

Machine learning algorithms decipher data that is in a specific format. The transformation of textual input into numerical feature vectors is known as vectorization. Bag of Words is an example of a method that uses normalization, counting, and tokenization. In this work, Count Vectorizer was used to represent words in terms of the Bow. The N-gram range, a tuple that includes the minimum and maximum length of the word sequence to be considered as features, must be specified in order for Count Vectorizer to function.

D. *Data Splitting*

The dataset is split in half using an 80:20 ratio, with 80% going towards training the model and the remaining 20% put aside for testing to evaluate its overall performance and predicted accuracy.

E. *Model Classifcation of CNN-LSTM Model*

The CNN model has been used for mood analysis problems in the past and has been shown to work. The dynamic Convolutional neural networks for semantic modelling of sentences. The dynamic Performing max pooling over a linear sequence is accomplished by CNN using dynamic K_maxpooling. Word2Vec [30], which airs on CNN- This employs a seven-layer model to enhance the suggested model's performance on the IMDB dataset. Using Word2Vec, the parametric rectified linear unit, and the dropout and CNN models in conjunction, the authors' suggested approach outperforms rival classification techniques.

The CNN is used to analyze sentiment in the dataset of rotting tomatoes. To manage the long-term dependencies needed for sentences, LSTM networks feature a long-term memory in addition to their short-term memory. A word's meaning is determined by its order of appearance. The sentence's sentiment is altered by a word's placement. Various datasets have been trained using CNN and LSTM combinations to improve results [31]. A number of experiments proved that word vectors play an essential role in DL NLP. Using fewer convolution layers with the Stanford sentiment tree bank dataset, an NLP architecture at the character level [32] and a high-level phrase representation achieved notable results in movie reviews. Interactions among an LSTM cell's input, hidden states, and different gates are addressed in Equations (1) through (5):

$$\mathrm{f_t} = \sigma\left(\mathrm{W^{(f)}x_t} + \mathrm{U^{f}h_{t-1}} + \mathrm{b^{(f)}}\right) \quad (1)$$

$$i_t = \sigma\left(W^{(i)}x_t + U^{i}h_{t-1} + b^{(i)}\right) \quad (2)$$

$$o_t = \sigma\left(W^{(o)}x_t U^{o}h_{t-1} + b^{(o)}\right) \quad (3)$$

$$c_t = \sigma\left(f_t \bullet c_{t-1} + i_t \bullet \tanh\left(W^{c}x_t + U^{(c)}h_{t-1} + b^{(ct)}\right)\right) \quad (4)$$

$$h_t = o_t \bullet \tanh(c_t) \quad (5)$$

where W and U stand for the recurrent connections and input weights, respectively. Every little letter is a vector; at time t, x represents input ($x_t \in R^d$), and d indicates each word's feature dimension; σ indicates the sigmoid function and denotes the element-wise product. The memory cell is designed to minimize vanishing, and $c_t(c_t \in R^h)$ indicates the exploding gradient, which aids in long-term dependency learning. With the traditional recurrent network, this was not possible; the input and cell output are controlled by the input gate i_t ($i_t \in R_h$) and the output gate o_t ($o_t \in R^h$) [33]. The memory cell is reset using the forget gate, denoted by f_t ($f_t \in R^h$), while ht ($h_t \in R^h$) describes the hidden state vector.

F. *Evaluation Parameters*

A number of metrics can be employed to evaluate classification systems, including recall, accuracy, precision, and the F1-score. Utilizing these factors, one can gauge the efficacy of supervised machine learning algorithms. They provide an all-encompassing evaluation of the model's performance using information from a confusion matrix or contingency table, which classifies outcomes as either true positives, true negatives, false positives, or false negatives [34]. Defining the confusion matrix parameters is covered in the section that follows:

J. Glob. Res. Elec. Comm., 2 (11):08-15
DOI: https://doi.org/xxxxxxxxxxxxx

- **True Positives (TP):** The quantity of product evaluations that the sentiment analysis algorithm accurately classifies as favorable and that are genuinely positive.
- **True Negatives (TN):** The model correctly labels the TN as unfavorable product reviews.
- **False Positives (FP):** the quantity of TN product reviews that the algorithm mistakenly classifies as positive. This is sometimes referred to as a false alarm or a Type I error.
- **False Negatives (FN):** The quantity of TP product reviews that the model mistakenly classifies as negative. This is sometimes referred to as a missed positive or a Type II mistake.

These parameters are incorporated in several performance metrics listed below:

Accuracy: The percentage of reviews that were correctly categorized out of all the reviews. Derived from Equation (6):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision: The precision shows how accurate the classification is Both high precision and poor accuracy lead to fewer false positives and lower accuracy. Equation. (7) illustrates the precision:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

Recall: The percentage of all real positive evaluations that were accurately predicted to be positive. Equation. (8):

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

F 1-score: F1-Measure is a combination of sensitivity and accuracy. This is how to be sensitive and precise with weighted harmonics. It is mathematically demonstrated in Equation (9):

$$\text{F} - \text{Score} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

IV. Results Analysis and Discussions

The test system utilized a Dell PC outfitted with a 2.90 GHz Intel® Core i7-10700 CPU, 16 GB of DDR4 RAM, and a 512 GB SSD for data storage. Operating on 64-bit Windows 11 Pro, the system offered a contemporary and reliable base for implementing the suggested concept. The CNN-LSTM model's performance results for sentiment analysis in online product reviews are shown in Table II. In every evaluation indicator, the model shows a high degree of efficacy. In particular, it attains a 98% accuracy rate, meaning that the model accurately categorizes the great majority of evaluations. The model's 98.49% recall and 98.52% accuracy show that it can accurately recognize positive and negative emotions with little to no false positives or negatives.

TABLE II. Results of CNN-LSTM Model in Sentiment Analysis in Online Product Reviews.

Metrics	CNN-LSTM
Accuracy	98
Precision	98.52
Recall	98.49
F_1-Score	98.49

The robustness and stability of the hybrid CNN-LSTM model in sentiment categorization for e-commerce review analysis is underscored by the high (98.49%) F1-score, another measure that evaluates the accuracy between recall and precision.

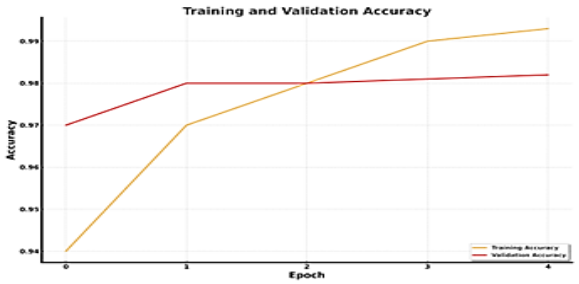

Fig. 5. CNN-LSTM Model for Training and Validation Accuracy Graph

Figure 5 illustrates how a model's accuracy may be trained across five epochs and then verified using the same data. The model is doing well while training using the training data, as indicated by the training accuracy (orange line), which shows a steady upward trend from little over 94% to over 99%. The validation accuracy (red line), on the other hand, demonstrates good generalization to unseen data with no overfitting. The initial value is nearly 97% and it quickly stabilizes at around 98%. A lack of a lean learning process is evident from the model's high accuracy rate and excellent generalization performance, which are caused by the fact that the training and validation curves throughout the epochs are comparable.



Fig. 6. Visual Results of CNN-LSTM Model for Training and Validation Loss

In Figure 6, shows the patterns of training and validation losses throughout the course of five evaluations. In a steady decline from approximately 0.16 to less than 0.02, the training loss (orange line) shows that the model is effectively minimizing error on the training data. The validation loss (red line), on the other hand, first declines before varying little and staying mostly constant between 0.07 and 0.08. This implies that although the model keeps getting better on training data, it does not perform noticeably worse on unknown data, suggesting that there is little overfitting and decent generalization. Overall, the loss trends demonstrate efficient and stable model training.

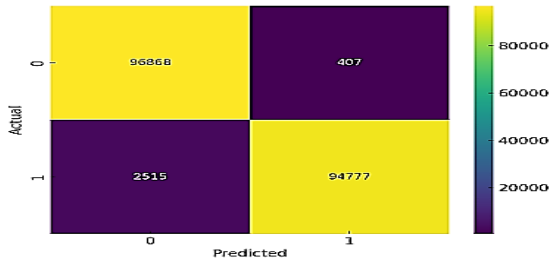

Fig. 7. Confusion Matrix For LSTM + CNN 2 Scores

The LSTM + CNN model's performance in binary classification is seen in Figure 7's confusion matrix. High reliability in both classes was indicated by the right identification of 96,868 TN and 94,777 TP out of the total predictions. In comparison to the overall number of cases, there were only 407 FP and 2,515 FN, which is a little amount. These findings provide a model that is very precise and efficient in reducing classification mistakes, especially when it comes to detecting positive cases while keeping the FPR low.

A. Comparative Analysis

This section provides a comprehensive analysis of the proposed CNN-LSTM model in contrast to popular algorithms such as CNN, NB, and RF. You can see a comparison of the different ML models used to analyses the sentiment of online product reviews in Table III. Similarly, the Random Forest (RF) model failed miserably at sentiment categorization, with a fair accuracy of 68%, precision of 59%, and F1-score of 60%. The Naive Bayes (NB) model, which fared the worst, had an F1-score of 34.7 and an accuracy of 44.9, showing that it struggles to handle complex sentiment patterns. Alternatively, the Convolutional Neural Network (CNN) proved to be highly effective at modelling contextual sentiment qualities, with impressive results of 93% accuracy, 94.02% precision, and a 94% F1-score. Outperforming all other measures, the suggested CNN-LSTM hybrid model achieved a 98% accuracy rate, 98.52 precision, 98.49 recall, and 98.49 F1-score, proving its better ability to grasp the spatial features and sequential dependencies of text input.

TABLE III. VARIOUS ML MODELS' COMPARISON ON SENTIMENT ANALYSIS IN ONLINE PRODUCT REVIEWS.

Models	Accuracy	Precision	Recall	F1-Score
RF [35]	68	59	68	60
NB [36]	44.9	47.1	44.6	34.7
CNN[37]	93	93.02	95	94
CNN-LSTM	98	98.52	98.49	98.49

An impressive 98% accuracy rate in analyzing online product reviews shows that the proposed CNN-LSTM model is the best option. This impressive performance compared to conventional ML algorithms that include RF and NB, as well as single-dimensional DL algorithms such as CNN indicates the power of a combination of convolutional layers and LSTM networks. CNN component is more effective in extracting local textual features, and LSTM component is more effective in long-range and contextual dependencies that allow an accurate sentiment classification. The combination architecture increases the generalization power and resilience of the model, which is good to deal with the sequential and complexity of the textual data in real-world review dataset.

V. CONCLUSION AND FUTURE WORK

In fact, sentiment analysis is an essential method of text data mining in e-commerce that publishes a massive number of reviews, suggestions, and comments every day. Such reviews tend to capture the views of the customers and they make informed decisions. This research paper suggests a hybrid deep learning model consisting of CNN and LSTM network in sentiment analysis of online product reviews. To a great extent, the CNN-LSTM model can capture local semantic patterns and sequential dependencies in text data. The experiment demonstrates that the model suggested works better than traditional ML classifiers RF and NB and pure CNN models. It represented an accuracy of 98%, a precision of 98.52, recall of 98.49, and an F1-score of 98.49, which illustrates its robustness and the effectiveness of its performance on complicated tasks related to the sentiment analysis domain.

Future research can be performed despite the good outcomes. To use the model in international e-commerce platforms, it would be helpful to expand its functionality to work with multilingual datasets. The investigation of domain adaptation methods may enhance the performance in different types of products and emerging language patterns. In addition, the fact that the model could be implemented into real-time systems with the functionality of explainable AI (XAI) provides the opportunity to increase transparency and user confidence in the sentiment-based recommendation systems in online retail environment.

- **Funding:** This research received no external funding.
- **Institutional Review Board Statement:** Not applicable.
- **Informed Consent Statement:** Not applicable.
- **Data Availability Statement:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.
- **Conflicts of Interest:** The author declares no conflict of interest.

How to cite this article:
B. K. Sharma, "Artificial Intelligence-Driven Sentiment Analysis for Product Reviews in Online Retailing Platforms," *Journal of Global Research in Electronics and Communication*, vol. 2, no. 11, pp. 08–15, Nov. 2026. Doi: XXXX

REFERENCES

[1] M. A. Kausar, S. O. Fageeri, and A. Soosaimanickam, "Sentiment Classification based on Machine Learning Approaches in Amazon Product Reviews," *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 3, pp. 10849–10855, Jun. 2023, doi: 10.48084/etasr.5854.

[2] Y. Yiran and S. Srivastava, "Aspect-based Sentiment Analysis on mobile phone reviews with LDA," in *Proceedings of the 2019 4th International Conference on Machine Learning Technologies*, New York, NY, USA: ACM, Jun. 2019, pp. 101–105. doi: 10.1145/3340997.3341012.

[3] S. Pandya, "Comparative Analysis of Large Language Models and Traditional Methods for Sentiment Analysis of Tweets Dataset," *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 12, pp. 1647–1657, 2024, doi: 10.5281/zenodo.14575886.

[4] K. Dixit, "Predictive Analytics in Business Intelligence for Sales Forecasting," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 03, no. 01, August, p. 981, Sep. 2023, doi: 10.48175/IJARSCT-12750G.

[5] D. Patel, "Integrating Price Elasticity and Reinforcement Learning: A Data-Driven Framework for Strategic E-commerce Pricing," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, Houston, TX, USA: IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395747.

[6] M. Alzate, M. Arce-Urriza, and J. Cebollada, "Online reviews and product sales: The role of review visibility," *J. Theor. Appl. Electron. Commer. Res.*, 2021, doi: 10.3390/jtaer16040038.

[7] S. Rongala, S. A. Pahune, H. Velu, and S. Mathur, "Leveraging Natural Language Processing and Machine Learning for Consumer Insights from Amazon Product Reviews," in *2025 3rd International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/ICSSES64899.2025.11009528.

[8] C. Patel, "Effect of Digital Transformation on Customer Engagement in Retail Industries : A Comparative Review," *J. Technol.*, vol. 10, no. 1, pp. 165–174, 2022.

[9] H. Ravilla, *Predictive Analytics for Customer Churn in Salesforce Service Cloud*. Springer Nature Switzerland, 2026.

[10] T. Hou, B. Yannou, Y. Leroy, and E. Poirson, "Mining customer product reviews for product development: A summarization process," *Expert Syst. Appl.*, vol. 132, pp. 141–150, Oct. 2019, doi:

10.1016/j.eswa.2019.04.069.

[11] A. Benlahbib and E. H. Nfaoui, "Aggregating Customer Review Attributes for Online Reputation Generation," *IEEE Access*, vol. 8, pp. 96550–96564, 2020, doi: 10.1109/ACCESS.2020.2996805.

[12] S. Chaturvedi, V. Mishra, and N. Mishra, "Sentiment analysis using machine learning for business intelligence," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, IEEE, Sep. 2017, pp. 2162–2166. doi: 10.1109/ICPCSI.2017.8392100.

[13] J. Kumar Chaudhary, S. Tyagi, H. Prapan Sharma, S. Vaseem Akram, D. R. Sisodia, and D. Kapila, "Machine Learning Model-Based Financial Market Sentiment Prediction and Application," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, May 2023, pp. 1456–1459. doi: 10.1109/ICACITE57410.2023.10183344.

[14] G. Bonifazi *et al.*, "A framework for investigating the dynamics of user and community sentiments in a social platform," *Data Knowl. Eng.*, vol. 146, p. 102183, Jul. 2023, doi: 10.1016/j.datak.2023.102183.

[15] N. Aslam, I. Ullah Khan, F. S. Alotaibi, L. A. Aldaej, and A. K. Aldubaikil, "Fake Detect: A Deep Learning Ensemble Model for Fake News Detection," *Complexity*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/5557784.

[16] S. Pahune and M. Chandrasekharan, "Several Categories of Large Language Models (LLMs): A Short Survey," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 7, pp. 615–633, 2023, doi: 10.22214/ijraset.2023.54677.

[17] R. V. Prabhu, S. Amsa, and C. J. Murthy, "Integrating IoT Data in Retail: Challenges and Opportunities for Enhancing Customer Engagement," *J. Artif. Intell. Res.*, vol. 3, no. 2, pp. 59–68, July, 2023.

[18] H. Taherdoost and M. Madanchian, "Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research," *Computers*, vol. 12, no. 2, p. 37, Feb. 2023, doi: 10.3390/computers12020037.

[19] K. Mane, S. Dongre, and M. Madankar, "Fake Review Detection using Random Forest Classifier," in *2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, IEEE, Jan. 2025, pp. 1–6. doi: 10.1109/SCEECS64059.2025.10940605.

[20] S. J. Bagastio, S. Suakanto, and H. Fakhrurroja, "A Systematic Analysis of Machine Learning and Deep Learning Strategies for Identifying Fake Reviews," in *2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, IEEE, Dec. 2024, pp. 239–244. doi: 10.1109/ICICYTA64807.2024.10913161.

[21] A. S. Sridhar and S. Nagasundaram, "Product Reviews Based Sentimental Analysis by Feature Extraction and Classification Using Machine Learning Algorithm," in *2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, Aug. 2024, pp. 891–897. doi: 10.1109/ICESC60852.2024.10689776.

[22] B. N, C. S, K. sheriff K, and P. T, "Decoding Authenticity in Online Review using Machine Learning," in *2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS)*, IEEE, Apr. 2024, pp. 421–424. doi: 10.1109/ICC-ROBINS60238.2024.10533938.

[23] M. Mukit, M. M. H. Rabbi, M. M. Ahmed, S. Latif, and S. Saha, "Sentiment Analysis on Bangla and Phonetic Bangla Reviews: A Product Rating Procedure using NLP and Machine Learning," in *2023 IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, IEEE, Nov. 2023, pp. 433–437. doi: 10.1109/WIECON-ECE60392.2023.10456378.

[24] J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu, and M. Arifuzzaman, "Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models," in *2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, IEEE, Feb. 2022, pp. 1–6. doi: 10.1109/ICAEEE54957.2022.9836420.

[25] D. Jain, S. Kumar, and Y. Goyal, "Fake Reviews Filtering System Using Supervised Machine Learning," in *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, IEEE, Jul. 2022, pp. 1–4. doi: 10.1109/ICDSIS55133.2022.9915878.

[26] C. Patel, "AI-Driven Recommendation Systems for Improving Online Customer Journey," *Int. J. Curr. Eng. Technol.*, vol. 14, no. 06, pp. 549–556, Nov/Dec, 2024, doi: 10.14741/ijcet/v.14.6.18.

[27] A. Shoukry and A. Rafea, "Preprocessing Egyptian Dialect Tweets for Sentiment Mining," *Fourth Work. Comput. ...*, 2012.

[28] Z. Nassr, F. Benabbou, N. Sael, and T. Hamim, "Improving Sentiment Analysis Performance on Imbalanced Moroccan Dialect Datasets Using Resample and Feature Extraction Techniques," *Information*, vol. 16, no. 1, p. 39, Jan. 2025, doi: 10.3390/info16010039.

[29] A. Iqbal, R. Amin, J. Iqbal, R. Alroobaea, A. Binmahfoudh, and M. Hussain, "Sentiment Analysis of Consumer Reviews Using Deep Learning," *Sustainability*, vol. 14, no. 17, p. 10844, Aug. 2022, doi: 10.3390/su141710844.

[30] B. Jang, I. Kim, and J. W. Kim, "Word2vec convolutional neural networks for classification of news articles and tweets," *PLoS One*, vol. 14, no. 8, p. e0220976, Aug. 2019, doi: 10.1371/journal.pone.0220976.

[31] A. C. M. V. Srinivas, C. Satyanarayana, C. Divakar, and K. P. Sirisha, "Sentiment Analysis using Neural Network and LSTM," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1074, no. 1, p. 012007, Feb. 2021, doi: 10.1088/1757-899X/1074/1/012007.

[32] I. Al-Jarrah, A. M. Mustafa, and H. Najadat, "Aspect-Based Sentiment Analysis for Arabic Food Delivery Reviews," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 7, pp. 1–18, Jul. 2023, doi: 10.1145/3605146.

[33] P. K. Jain, V. Saravanan, and R. Pamula, "A Hybrid CNN-LSTM: A Deep Learning Approach for Consumer Sentiment Analysis Using Qualitative User-Generated Contents," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 5, pp. 1–15, Sep. 2021, doi: 10.1145/3457206.

[34] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," *Expert Syst. Appl.*, vol. 57, pp. 117–126, Sep. 2016, doi: 10.1016/j.eswa.2016.03.028.

[35] H. Ali, E. Hashmi, S. Yayilgan Yildirim, and S. Shaikh, "Analyzing Amazon Products Sentiment: A Comparative Study of Machine and Deep Learning, and Transformer-Based Techniques," *Electronics*, vol. 13, no. 7, p. 1305, Mar. 2024, doi: 10.3390/electronics13071305.

[36] P. S. Ghatora, S. E. Hosseini, S. Pervez, M. J. Iqbal, and N. Shaukat, "Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM," *Big Data Cogn. Comput.*, vol. 8, no. 12, p. 199, Dec. 2024, doi: 10.3390/bdcc8120199.

[37] L. Ashbaugh and Y. Zhang, "A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning," *Computers*, vol. 13, no. 12, p. 340, Dec. 2024, doi: 10.3390/computers13120340.