Journal of Hunan University

(Natural Sciences）

Vol. 51 No. 8

August 2024

Available online at

http://jonuns.com/index.php/journal/index

Open Access Article https://doi.org/10.55463/issn.1674-2974.51.8.5

A Hybrid CNN-LSTM Based Natural Language Processing Model for Sentiment Analysis of Customer Product Reviews: A Case Study from Ghana

Nouhaila El Koufi[1]*, Yaw Marfo Missah[2], Abdessamad Belangour[1]

([1] Department of Mathematics and Computer Science, Hassan II University, Casablanca, Morocco

[2] Department of Computer Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana)

* Corresponding author: Elkoufi.nouhaila1@gmail.com

Article History:

Received: July 5, 2024
Revised: August 1, 2024
Accepted: August 6, 2024
Published: August 31, 2024

Abstract: As new technologies rapidly evolve and consumer habits and lifestyles change, there is a noticeable shift toward online shopping and services among customers. Consequently, there has been a consistent surge in the volume of customer data. A significant portion of these data revolves around consumers' perceptions and opinions regarding organizations' products or services. This study aims to develop a novel hybrid CNN-LSTM model to analyze customer sentiment and satisfaction with e-commerce platforms. By integrating deep learning techniques with sentiment analysis and natural language processing, this approach offers a comprehensive system capable of understanding consumer feedback with greater accuracy. The purpose is to create a robust tool for precision marketing that not only captures customer sentiments, but also enhances decision-making in e-commerce. These data are of considerable importance for market intelligence collectors operating in areas such as marketing, customer relationship management, and customer retention. Sentiment analysis is employed to scrutinize customer sentiment, marketing campaigns, and product evaluations, thus assisting e-commerce companies in acquiring a more profound understanding of their customers' viewpoints and satisfaction with a product or service. This valuable insight can help managers improve their decision-making regarding future products and services, marketing strategies, promotional channels, and customer service improvements. By harnessing artificial intelligence techniques, such as deep learning, sentiment analysis, and natural language processing, it becomes possible to create and implement systems capable of analyzing consumer satisfaction and feedback on e-commerce platforms. The novelty of this study lies in its innovative combination of CNN and LSTM architectures, which allows the model to effectively capture both spatial and temporal patterns in textual data, providing deeper insights into consumer behavior than traditional methods. A case study was conducted using authentic data from an e-commerce enterprise in Ghana to demonstrate the practical application of our approach. The research findings revealed that the suggested model yielded positive results. The proposed hybrid CNN-LSTM model is applied to target marketing to obtain precise consumer profiles and enhance decision-making, with the aim of increasing enterprise revenue.

Keywords: artificial intelligence; decision-making; deep learning; natural language processing; e-commerce

基于深度学习的客户满意度评价系统：来自加纳的案例研究

摘要：随着新技术的快速发展以及消费者习惯和生活方式的改变，客户明显转向在线购物和服务。因此，客户数据量持续激增。这些数据中很大一部分围绕消费者对组织产品或服务的看法和意见。本研究旨在开发一种新颖的混合 CNN-LSTM 模型来分析客户对电子商务平台的情绪和满意度。通过将深度学习技术与情绪分析和自然语言处理相结合，这种方法提供了一个能够更准确地理解消费者反馈的综合系统。目的是创建一个强大的精准营销工具，不仅可以捕捉客户情绪，还可以增强电子商务中的决策能力。这些数据对于在营销、客户关系管理和客户保留等领域运营的市场情报收集者来说非常重要。情绪分析用于审查客户情绪、营销活动和产品评估，从而帮助电子商务公司更深入地了解客户对产品或服务的观点和满意度。这种宝贵的见解可以帮助管理人员改善有关未来产品和服务、营销策略、促销渠道和客户服务改进的决策。通过利用深度学习、情感分析和自然语言处理等人工智能技术，可以创建和实施能够分析消费者对电子商务平台的满意度和反馈的系统。这项研究的新颖之处在于它创新地结合了 CNN 和 LSTM 架构，这使模型能够有效地捕捉文本数据中的空间和时间模式，比传统方法更深入地洞察消费者行为。我们使用来自加纳一家电子商务企业的真实数据进行了案例研究，以展示我们的方法的实际应用。研究结果表明，建议的模型产生了积极的结果。提出的混合 CNN-LSTM 模型应用于目标营销，以获得精确的消费者资料并增强决策能力，目的是增加企业收入。

关键词：人工智能；决策；深度学习；自然语言处理；电子商务

1. Introduction

In recent years, the consumer market has undergone a notable transformation owing to the rapid expansion of digital networks and the proliferation of information technology [1]. This transformation has led to the widespread adoption and advancement of various technologies, including the internet, mobile devices, logistics solutions, and secure payment systems. Concurrently, there has been a notable evolution in consumer behavior and preferences, with an increasing number of people embracing online shopping as their primary method for purchasing goods and services. The development of e-commerce has been propelled by technological advancements and shifting consumer trends [2]. What started as a niche market has now become a global phenomenon, with online shopping emerging as a preferred mode of commerce for consumers worldwide. This transition has been facilitated by the convenience, accessibility, and variety offered by e-Commerce platforms. Furthermore, the rise of e-commerce has been significantly propelled by the COVID-19 pandemic, as the imposition of lockdowns and social distancing measures has led more consumers to rely on online platforms for their shopping needs [3]. This surge in online shopping has underscored the resilience and adaptability of e-commerce businesses, further solidifying their position as key drivers of economic growth. As e-commerce continues to mature and innovate, it is poised to play a more significant role in shaping the future of retail and commerce. Investments in emerging technologies such as artificial intelligence, augmented reality, and blockchain are expected to further enhance the online shopping experience, driving continued growth and expansion of the e-commerce sector.

The emergence of online services and e-commerce platforms has led to the generation of vast amounts of data on customer behavior, preferences, and sentiments [4]. Through every interaction, a user makes valuable data on these platforms, whether it is browsing products, making purchases, leaving reviews, or engaging with customer support. These data encompass a wide range of information including demographic details, browsing habits, purchase history, feedback, and social media interactions. With the advent of sophisticated tracking tools and analytics platforms, businesses can effectively gather, process, and analyze data to gain deep insights into customer preferences and trends [5]. Through the use of sophisticated algorithms and machine learning methods, businesses can uncover patterns, correlations, and predictive models that help them understand what drives customer behavior and satisfaction. This wealth of data not only allows businesses to customize their products and marketing approaches according to the unique preferences of each customer but also facilitates the development of personalized recommendations, targeted promotions, and seamless user experiences. Moreover, the continuous accumulation of customer data allows companies to iteratively refine their products and services, adapt to evolving market trends, and stay ahead of competition in the fast-paced world of online commerce.

Client comments and feedback on e-commerce platforms are invaluable resources for understanding customer needs and measuring user satisfaction, which in turn enhances decision-making in marketing [6]. These comments offer direct insights into the experiences, inclinations, and challenges faced by customers, providing businesses with a clear understanding of what drives customer satisfaction [7]. By systematically analyzing feedback, companies can recognize patterns, trends, and opportunities for enhancement, allowing them to adjust their products, services, and marketing strategies to meet customer expectations more effectively. Moreover, monitoring sentiment in client feedback allows businesses to gauge overall user satisfaction levels and prioritize initiatives that most effectively strengthen customer loyalty. By incorporating insights from client feedback into marketing decision-making processes, businesses can develop more targeted and effective campaigns, address customer concerns, and ultimately achieve sustained success in a competitive e-commerce landscape. Sentiment analysis is a branch of computer science that examines sentiments expressed in text with a focus on extracting opinion data. Also known as opinion mining or emotion AI, this technique employs natural language processing (NLP) to assess whether text data conveys positive, negative, or neutral sentiments. It is widely utilized to analyze textual data, particularly to assist companies in monitoring customer feedback for brand and product insights. Understanding customer preferences, including the products that they are likely to recommend, poses a significant challenge for e-commerce businesses.

The integration of deep learning methods with sentiment analysis plays a pivotal role in analyzing client satisfaction and enhancing decision-making in precision marketing. Deep learning algorithms, with their ability to process large volumes of complex data coupled with sentiment analysis techniques, offer a powerful tool for extracting valuable insights from client feedback and sentiment-expressed text. This combined approach allows businesses to attain a more profound comprehension of client sentiments, including subtle nuances, such as sarcasm or implicit emotions, which traditional methods might miss. By uncovering hidden patterns and correlations within client feedback data, businesses can personalize their marketing strategies more effectively, ensuring that messages resonate with individual clients at a deeper level. Ultimately, the integration of deep learning and sentiment analysis empowers businesses to make more precise and informed decisions for accurate marketing, leading to improved client satisfaction and loyalty.

This study aims to enhance decision-making in accurate marketing. To achieve this goal, this study introduces a hybrid (convolutional neural network - long short-term memory) CNN-LSTM model based on sentiment analysis to analyze client satisfaction and comprehend client needs. By leveraging the advantages of both the CNN and LSTM networks, the proposed model excels in extracting nuanced sentiments from textual data and capturing temporal dependencies within the sequence of sentiments. This innovative approach allows for a comprehensive understanding of consumer sentiment, allowing businesses to acquire profound insights into customer preferences and behavioral trends. By applying the proposed method to precision marketing, businesses can generate precise portraits for consumers and identify their preferences, interests, and pain points with high accuracy. Based on these insights, businesses can effectively tailor their marketing strategies, delivering personalized experiences and targeted offerings that resonate with individual consumers. Ultimately, the proposed method empowers businesses to make well-informed decisions, allocate resources efficiently, and improve customer satisfaction in the dynamic landscape of accurate marketing.

2. Literature Review

With the rapid rise in competition within the online market, enterprises face a pressing need to develop targeted and precise marketing strategies. Precision marketing aims to help companies create effective strategies that meet customer desires while maintaining market competitiveness through low cost, rapid implementation, and optimized resource use. To address this challenge, artificial intelligence methods are specifically selected for their powerful capabilities to enhance accurate marketing efforts. In the ever-changing online market, where competition is fierce and consumer behavior complex, deep learning excels in recognizing and analyzing intricate patterns. These algorithms can adapt to and learn from new data, ensuring that marketing strategies remain flexible and responsive to shifting market trends. By utilizing deep learning and sentiment analysis, enterprises can efficiently analyze client satisfaction and improve decision-making in accurate marketing. Deep learning algorithms can process large volumes of complex data and, when combined with sentiment analysis techniques, provide a robust toolset for extracting valuable insights from client feedback and sentiment-expressed text. This combined approach allows businesses to gain a deeper understanding of client sentiments, including subtle nuances, such as sarcasm or implicit emotions that traditional methods might overlook. By uncovering hidden patterns and correlations within client feedback data, businesses can personalize their marketing strategies more effectively, ensuring that messages resonate with individual clients at a deeper level. Ultimately, integrating deep learning and sentiment analysis enables businesses to make more precise and informed decisions in precision marketing, resulting in improved client satisfaction and loyalty.

Research on predicting customer preferences by analyzing feedback to enhance marketing decisions has resulted in the creation of various techniques, with a particular emphasis on artificial intelligence methods. An extensive study on Arabic sentiment analysis models was presented in [8] with an emphasis on deep learning applications in e-marketing. This study evaluated current methodologies and performance metrics while exploring potential enhancements in preprocessing techniques. To detect polarity-based sentiment features and calculate sentiment scores for products, a novel automatic recommender system was introduced that employs hybrid deep-learning techniques for sentiment analysis across different levels [9]. This system enables product ranking based on positive attributes and utilizes visualization tools such as dashboards to support decision-making for marketers and consumers. Validated through an analysis of smartphone reviews from various e-commerce platforms, the method demonstrated promising results and significantly improved the effectiveness of recommender systems. Raju et al. [10] addressed the problem of customer evaluation on e-commerce platforms by proposing an advanced sentiment analysis model designed to enhance consumer experience. The model utilizes Fejer Kernel filtering for accurate data estimation, and employs fuzzy dictionary-based semantic word feature extraction. Feature extraction optimization was achieved through Simulated Annealing, and customer opinions were classified using the BERT deep-learning model. Additionaly, to assess customer reviews and assist patrons in making safe dining decisions during the COVID-19 pandemic, a model was proposed that combines Bidirectional LSTM with Simple Embedding and Average Pooling for evaluating customer opinions [11]. A case study involving 112,412 restaurant reviews from Yelp.com in the first half of 2020 was used to simulate the proposed model. A methodology for sentiment analysis and opinion mining aimed at tackling the intricate challenges of automating sentiment detection in texts across multiple languages was introduced in [12]. Their approach employs a machine learning framework that vectorizes textual documents to train a polarity classification model. The experiments were performed on datasets consisting of online user reviews in both Greek and English. Additionally, a comparative analysis of various deep learning models and input features, demonstrated the increasing effectiveness of deep learning in sentiment analysis [13]. According to [14], a proposed novel sales forecasting method combines sentiment analysis with the Bass/Norton model, utilizing both historical sales data and online review content. They validated their approach by using real-world data from the automotive industry and associated online reviews. Some studies have compared various deep learning models, such as LSTM, CNN, and K-NN, to evaluate customer sentiments on social media in the context of English language evaluation [15, 16]. Therefore, according to [17], a hybrid model that combines BiLSTM and CNN was proposed for predicting e-commerce sales by incorporating customer reviews into the forecasting process. This approach illustrates how the combination of product attributes and comment sentiments can accurately forecast sales. The SLCABG model was proposed to extract principal sentiment and contextual features from reviews in a Chinese-language dataset gathered from an e-commerce site [18]. A proposed approach for sentiment analysis [19] focuses on text from various Internet and social media platforms, and recognizing it as a crucial form of unstructured data. Their technique encompasses a thorough process including data collection, preprocessing, feature encoding, and classification using three variations of long short-term memory (LSTM) models. This study aims to address the challenges of accurately detecting the polarity of consumer reviews based on large volumes of data. By applying their method to different textual datasets, they demonstrated that their approach not only produced better or comparable results but also reduced computational complexity. This study highlights the significance of sentiment analysis in extracting meaningful insights from consumer reviews and social media content. According to [20], a deep neural network model is designed to extract meaningful features from the review characteristics matrix, effectively addressing challenges such as sparsity, ambiguity, and redundancy. To manage the massive influx of textual content generated daily from sources, such as social media, e-commerce sites, and news websites, a consumer review summarization model utilizing NLP techniques and LSTM was developed [21]. This model aims to provide businesses with concise summaries of customer feedback, helping them to better understand consumer behavior and preferences. Additionally, a method for evaluating service quality in e-commerce was proposed in [22] by analyzing customer reviews, offering a quicker alternative to traditional sampling methods. They employed sentiment analysis to classify online reviews into positive and negative sentiments using machine learning techniques. Xu et al. a proposed continuous Naïve Bayes learning framework improves sentiment classification of product reviews from e-commerce platforms [23]. The authors introduced methods to fine-tune the learned distributions based on three types of assumptions, thereby enhancing the model adaptability to various domains. A Hindi-based sentiment analysis method was tailored for e-commerce applications in [24] using Twitter as a key data source for customer opinions on products. Recognizing the complexity of sentiment analysis in Hindi, they utilized an LSTM model combined with a continuous bag of words (BoW) approach.

Based on the literature review, we observed that studies combining deep learning and sentiment analysis for the Arabic language to enhance decision-making and develop targeted marketing are rare. Most studies focus on other languages, such as English and French, whereas those addressing Arabic primarily aim to detect customer churn. In this study, we propose a decision-making framework that integrates deep learning with sentiment analysis. This system involves the development of a hybrid CNN-LSTM model applied to accurate marketing to obtain accurate consumer profiles and enhance decision making. The proposed system can analyze consumer satisfaction and feedback on e-commerce platforms. These valuable insights can help managers improve decisions related to future products and services, marketing strategies, promotional channels, and customer service enhancement.

3. Materials and Method

In this section, we describe the components of the proposed system, which include six essential phases, as shown in Figure 1. These steps include data collection, exploration, preprocessing, splitting, modeling, and strategic marketing decision-making. Initially, data were collected from a Ghanaian E-commerce business containing general information about clients and their opinions. The data were examined to extract and analyze customer opinions and gain a clearer understanding of the sentiment frequencies associated with each service provider. Subsequently, these data were pre-processed using advanced methods to obtain precise insights. It was then divided into three classes for training, validation, and testing. Subsequently, a powerful deep-learning model was created to support the training and prediction processes. This was followed by a performance evaluation. Ultimately, a suitable marketing decision was made based on the results of the prior phases.

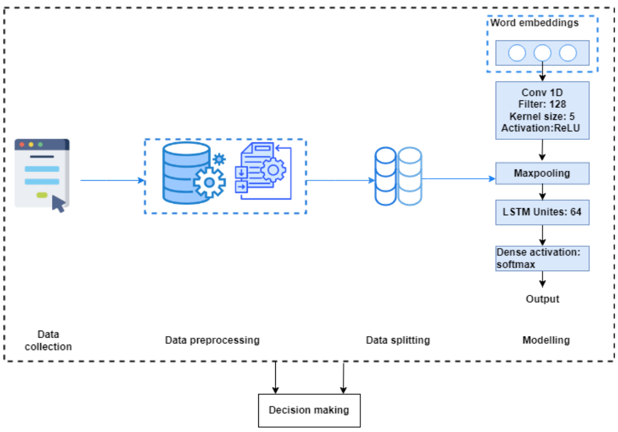

Figure 1. The proposed system architecture (The authors’ elaboration)

3.1. Data Collection

Data collection primarily involves gathering valuable information from the initial transactional data. For this study, data were collected from a real-world omni-channel business in Ghana. This business typically sells products in four categories: Beauty, Organic Food, Health (including gluten-free, dietary supplements, and sugar-free products), and Eco-friendly Household Products. This business operates both a local store in Ghana and an e-commerce platform for product sales. In this study, we used a dataset from an e-commerce platform. Additionally, we developed a form through which we requested customers to provide feedback after purchasing products. This form requests general information, such as Name, Age, City, Gender, and comments about the provided products in the four categories.

3.2. Data Exploration

The aim of data exploration is to analyze client feedback, specifically focusing on understanding sentiment frequencies related to individual service providers across four product categories: Beauty, Organic Food, Health, and Eco-friendly Household Products. By exploring these sentiment frequencies, the authors can improve their capacity to analyze and interpret data more effectively. First, sentiment frequencies for business product categories, including beauty, organic food, health, and eco-friendly household products, were analyzed. Each product category featured three sentiment categories: negative, neutral, and positive, representing varying degrees of customer satisfaction ranging from dissatisfaction to satisfaction. The degree of client satisfaction for the four product categories is shown in Figure 2. As depicted in Figure 2(a), the frequencies of each sentiment category for beauty products were 5063 (positive), 1524 (neutral), and 233 (negative). Similarly, in Figure 2(b), the frequencies of each sentiment category for Eco-friendly Household products are 1400 (positive), 4300 (neutral), and 6270 (negative). In Figure 2(c), the frequencies of each sentiment category for health products are 6100 (positive), 5000 (neutral), and 580 (negative). Finally, in Figure 2(d), the frequencies of each sentiment category for Organic Food products are 3640 (positive), 7020 (neutral), and 2330 (negative).

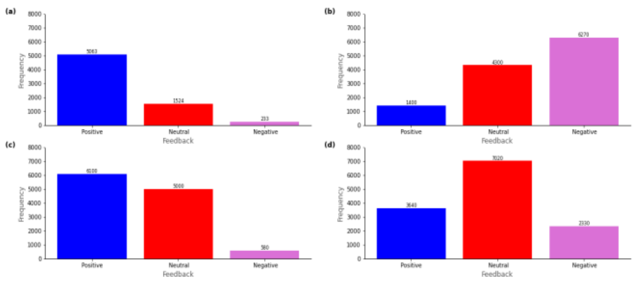

Figure 2. Degree of client satisfaction regarding product categories (The authors' elaboration)

Second, Figure 3 depicts the degree of client satisfaction regarding the four product categories by gender.

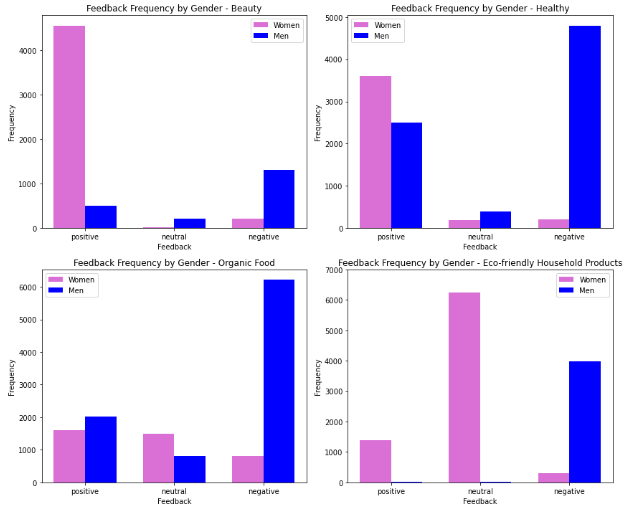

Figure 3. Degree of client satisfaction regarding the four product categories by gender (The authors' elaboration)

Figure 4 illustrates the degree of satisfaction of clients regarding the four product categories by urbanity of residence, which was extracted from the city column.

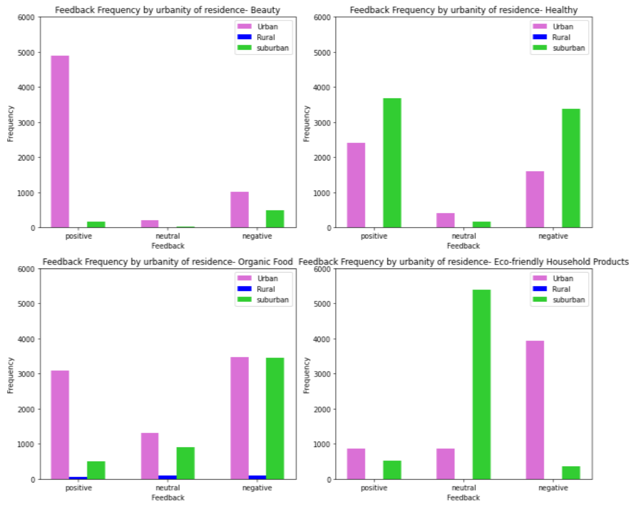

Figure 4. Degree of client satisfaction regarding the four product categories by residenceurbanity (The authors' elaboration)

Figure 5 shows the degree of client satisfaction regarding the four product categories by age class.

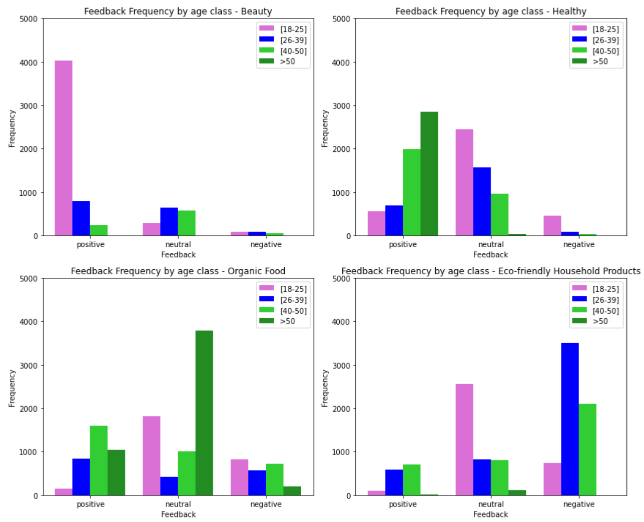

Figure 5. Degree of client satisfaction regarding the four product categories by age class (The authors' elaboration)

The incorporation of bigrams is crucial to bolster Arabic machine translation, linguistic models, and sentiment analysis. Bigrams function as invaluable aids in comprehending Arabic linguistic structures, elucidating the organization and importance of words, and emphasizing the relationships between adjacent terms. They play a pivotal role in text analysis and NLP, enabling the scrutiny of Arabic bigrams through frequency, co-occurrence, and contextual analyses. Various applications, such as sentiment analysis, language modeling, and machine translation, depend on recognizing positive, negative, or neutral sentiment patterns, and understanding language-specific dependencies and word combinations.

3.3. Data Preparation

Data preparation served as the cornerstone of our project, guaranteeing the quality and consistency of the collected data. The efficient processing of raw data is not only fundamental but also critical in any data-related task, particularly in the context of sentiment extraction methods. The extraction of high-quality text is not just a preference, but also an absolute necessity for accurate analysis and meaningful insights. Therefore, meticulous attention to data processing is paramount as it lays the foundation for robust sentiment analysis and the generation of valuable insights across various applications and scenarios. Our dataset comprises two primary types of data: structured data containing essential customer information, and unstructured data consisting of customer comments. The cleaning process begins by eliminating missing values, incorrect records, incomplete entries, outliers, and noise within the data. This initial step ensured that the subsequent analysis was based on reliable and consistent data. For comment data specifically, we employed additional cleaning steps to further refine the text. This included the removal of French words, emojis, hashtags, English words, links, mobile numbers, and symbols. By eliminating these elements, we ensured that the text was clear and free from any extraneous information that could potentially skew the sentiment analysis results. By meticulously preparing and cleaning our data, we established a foundation for precise sentiment analysis that can yield valuable insights into customer opinions and preferences. This thorough data-processing approach improves the reliability and validity of our findings, enabling us to make well-informed decisions and recommendations based on the insights obtained from our analysis.

In our study, we utilized tokenization, an essential step in preprocessing text data that converts unstructured text into a format compatible with machine-learning models. This process entails breaking down the text into smaller tokens to facilitate a more efficient analysis. Each token is assigned a vocabulary or word index by the tokenizer, with common tokens assigned to lower indices. Once the vocabulary is defined, tokenizers transform text sequences into sequences of word indices that are suitable for machine-learning algorithms. The adoption of tokenizers empowers developers to leverage deep-learning models for tasks such as language generation, sentiment analysis, and text classification. By tokenizing words across texts, it becomes easier to recognize patterns and extract insights from the textual data. Furthermore, we employed various preprocessing techniques on our dataset to refine Arabic text to its fundamental linguistic components. This improved dataset provides opportunities for model training and evaluation in areas, such as sentiment analysis, machine translation, and named-entity recognition.

To improve the organization, quality, and compatibility of categorical classes with various applications, this study focused on converting them into numeric values. Data transformation is a key aspect of this process, involving tasks such as reformatting, validating, and restructuring data. We also address issues such as null values, redundant duplicates, inaccurate indexing, and inconsistent formats. By systematically refining and meticulously validating the data, we significantly improved their overall integrity. This process ensures that data can be effectively utilized across different systems and serves as a safeguard against potential issues arising from flawed or incompatible data representations. The data transformation conducted in this study is exemplified as follows (for example, gender: women > 0, men > 1).

3.4. Data Splitting

To evaluate the performance of our prediction model, it was vital to partition the data into two sets: training and testing. Nonetheless, given the complexity of our data, additional considerations were necessary during the evaluation. Various validation techniques can also be employed. The most straightforward method involves dividing the data into two segments: training and testing sets. Sometimes, when refining the hyperparameters of the prediction algorithm, a third set, known as the validation set, is employed. This facilitates fine-tuning and enhances the performance of the model. In this study, data were partitioned into three distinct segments: training, testing, and validation. Specifically, 80% of the data were used for the training phase, whereas the remaining 20% were set aside for validation and testing purposes. This approach was undertaken to assess the performance of the proposed model' on unseen data and ascertain its ability to generalize beyond the training samples. By evaluating the effectiveness of the model on a separate test set, the researcher determined its suitability for real-world applications.

3.5. Prediction Model

After completing the data preparation stage, we turned our attention to developing a prediction model by carefully considering the unique features of our dataset. First, we propose a model that combines two neural networks, CNN and LSTM. We conducted this fusion to explore how well the CNN adapts to LSTM, considering that LSTM is renowned for its effectiveness in sentiment analysis. By merging these two architectures, CNN-LSTM models offer a blend of spatial feature extraction and sequential modeling precision, making them suitable for tasks such as sentiment analysis and text classification that require both spatial and temporal data processing. While CNN focuses on extracting spatial features from the input data, LSTM Excel was used to handle sequential dependencies.

Each component of the model architecture included several key parameters. The model is capable of handling inputs up to the length of the tokenizer's index-word dictionary plus an additional unique word or token. Each input word or token was represented by a dense vector with 64 dimensions, encapsulating its semantic meaning. The Conv1D layer, equipped with 128 filters, functions as a feature detector and analyzes patterns in the input data. It processes blocks of five words or tokens using a kernel size of five, and employs the ReLU activation function to efficiently capture complex data patterns. The LSTM layer, with 64 units, establishes hidden state dimensions and number of memory cells, which are vital for understanding intricate temporal dependencies. Three units in the dense layer produce output probabilities for multiclass classification using the Softmax activation function. These parameters collectively shape the 'architecture of the model, thereby influencing its data processing and learning capabilities. The model is trained with the Adam optimizer and utilizes a categorical cross-entropy loss function to assess the discrepancy between the predicted probabilities and true class labels. was used to assess the model' performance. The training process involved a batch size of 256 samples over 50 epochs specifically designed for a three-class output scenario. Figure 6 illustrates the model's architecture for analyzing sentiments in customer satisfaction from an e-commerce platform.

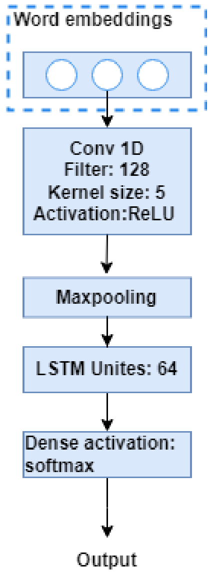

Figure 6. CNN-LSTM model (The authors' elaboration)

To compare and evaluate the performance of the models in the sentiment classification task against our model, we developed four additional models: LSTM, CNN, BiLSTM, and GRU.

The LSTM model was constructed with an architecture comprising the input, hidden, and output layers. The LSTM architecture, which is a variant of RNNs, was created to address the difficulties in capturing long-term dependencies within sequential data. LSTMs use specialized memory cells and gating mechanisms to selectively maintain and modify information over multiple time-steps. Each LSTM cell includes four fundamental components: cell state, input gate, forget gate, and output gate. The LSTM architecture was designed to address the vanishing gradient problem, allowing it to effectively store and retrieve long-term dependencies. The integration of LSTMs has significantly enhanced various NLP tasks, including language modeling, machine translation, sentiment analysis, and speech recognition.

Moreover, the CNN model, which is known for its versatility in processing both text and audio signals, was introduced, making it highly beneficial for NLP applications. One-dimensional CNNs efficiently manage sequence data such as word embedding or character-level encoding. Parsed sentences and documents are transformed into numerical representations that can be interpreted by neural networks. In the convolutional layer, one-dimensional filters or kernels move across the input data and execute convolutional operations along the sequential dimensions. The pooling layer then reduces the feature maps produced by the convolutional layer, while the fully connected layers proceed to process and learn relevant feature combinations. One-dimensional CNNs perform tasks, such as text classification, sentiment analysis, named-entity recognition, and speech recognition. Architectural modifications have enhanced the performance of 1D CNNs in NLP tasks, improving their ability to model one-dimensional sequential data and effectively extract features.

Furthermore, a BiLSTM model was developed. BiLSTM networks incorporate both past and future data, enabling them to process sequential data proficiently with contextual information from both directions. The input sequence is first transformed into numerical representations and then sequentially processed by a forward LSTM layer to capture the historical context, whereas a backward LSTM layer simultaneously processes it to capture the future context. By integrating both forward and backward LSTM layers, bidirectional LSTM networks can gain a comprehensive understanding of the input sequence at any given point. Although BiLSTM networks incur higher training and inference costs, they are frequently preferred for sequence modeling because of their enhanced performance and ability to capture dependencies in both directions.

Finally, a gated recurrent unit (GRU) model is developed. The GRU network, a type of recurrent neural network architecture, is designed to process sequential data in tasks such as NLP and speech recognition. It addresses the limitations of the original RNN structure and offers an improved performance. The GRU nodes serve as both update and reset gates within the network, determining the retention and forgetting of information when computing the current hidden state. GRU networks excel in modeling long-term dependencies owing to their streamlined parameter structure, enabling the efficient processing of longer sequences. GRUs have shown strong performance across a range of fields, including machine translation, speech recognition, sentiment analysis, and language modeling.

3.6. Performance Evaluation Metrics

In this study, we selected a set of performance metrics to assess the effectiveness of the proposed model. These metrics include accuracy, sensitivity, specificity, precision, and the F1-measure.

3.6.1. Accuracy

Accuracy represents the proportion of correct predictions relative to the total number of predictions made by the classification model.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

3.6.2. Sensitivity (Sen)

Sensitivity calculates the proportion of actual positive cases that a classification model correctly identifies [25]. This measures the effectiveness of the model in detecting positive instances.

$$Sen = \frac{TP}{TP + FP} \tag{2}$$

3.6.3. Specificity (Spe)

Specificity represents the accuracy with which a classification model correctly identifies negative instances. This provides insight into the ’ability of the model to avoid false positives.

$$Spe = \frac{TN}{TN + FN} \tag{3}$$

3.6.4. Precision

Precision measures the accuracy of positive predictions made by a classification model by calculating the proportion of true-positive predictions among all positive predictions [26].

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

3.6.5. F1-Measure

The F1-measure, which is derived as the harmonic mean of precision and recall, helps balance these two metrics [27]. This is particularly useful for evaluating a model’s performance in cases where the class distribution is imbalanced.

$$F - measure = \frac{(2 \times Precision \times Recall)}{(Recall + Precision)} \tag{5}$$

4. Results and Discussion

In this investigation, we utilized real-world data collected from an Omni-channel business in Ghana to validate and illustrate the advantages of the proposed system. This business operates both a local store in Ghana and an e-commerce platform for product sales. Our focus is specifically on e-commerce. We designed a feedback form to appear after clients had completed a purchase from any product category offered by the business. The form requested basic information, such as name, age, gender, and city, along with feedback in the form of comments about the purchased products. After meticulously analyzing the provided data, we undertook a thorough data-cleaning process, which involved two main types of data: structured data containing essential customer information and unstructured data consisting of customer comments. This cleaning process included the removal of missing values, outliers, and noise. Furthermore, we employed advanced techniques to refine the text of the customer comments and eliminate irrelevant elements. Our meticulous preparation ensured high-quality data for accurate sentiment analysis, thus providing valuable customer insights to support informed decision making.

An initial analysis of the data was conducted, focusing particularly on three key factors: age, gender, and urbanity of residence, to gain a comprehensive understanding of the studied population. Figures 3, 4, and 5 visually depict the findings. Men emerged as the primary purchasing demographic. Regarding age distribution, the highest proportions were observed within the 26-39 and 18-25 age groups. In terms of urban residence stratification, urban and suburban populations represented the majority. These insights emphasize the importance of precise control over gender, geographic location, and age when implementing accurate marketing strategies.

After preprocessing and splitting the data, we developed the proposed model. The performance evaluation, as shown in Figures 7 and 8, revealed that our model achieved an accuracy of 0.8501 and a mean squared error of 0.1931. These results indicate a promising performance, suggesting that our model effectively captures and predicts the underlying patterns in the data.

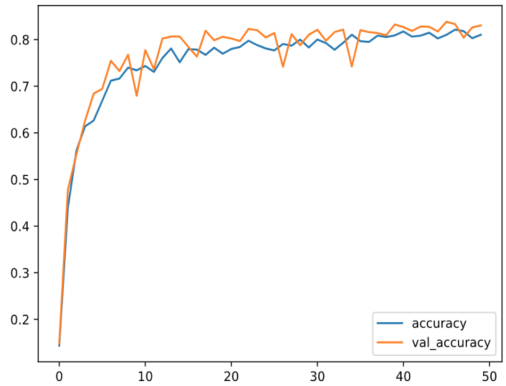

Figure 7. CNN-LSTM accuracy (The authors' elaboration)

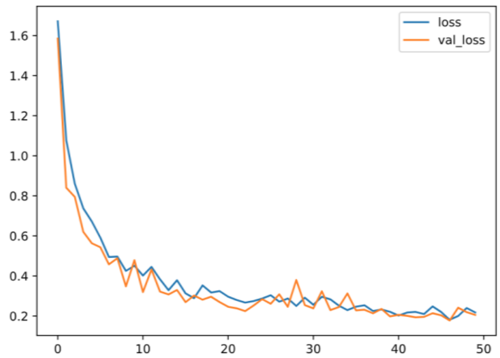

Figure 8. CNN-LSTM loss (The authors' elaboration)

Subsequently, we compared the performance of our proposed model with that of four other models based on selected performance metrics including accuracy, sensitivity, specificity, and F1 score. Table 1 and Figure 9 present the results of the models. Based on the results, our proposed CNN-LSTM model outperforms the other models across multiple metrics. In terms of accuracy, the CNN-LSTM model achieved the highest value of 0.8501, indicating its superior ability and performance compared with the LSTM, CNN, BiLSTM, and GRU models. Similarly, the CNN-LSTM model demonstrated the highest sensitivity score of 0.8132. Moreover, the CNN-LSTM model exhibited the highest specificity score of 0.8674, highlighting its effectiveness. Additionally, the CNN-LSTM model attained the highest F1-score of 0.8378, which serves as a comprehensive metric for model evaluation. Overall, these results suggest that the CNN-LSTM model offers the best overall performance compared to the other models.

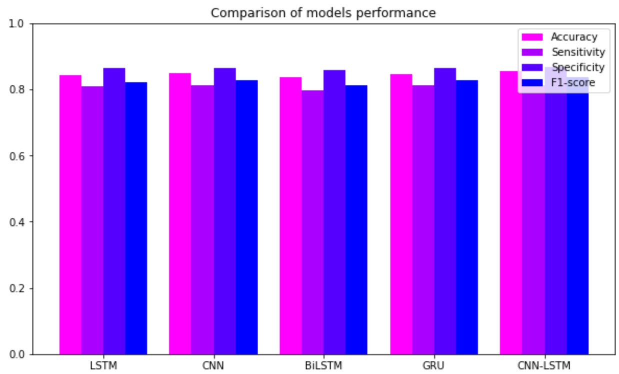

Figure 9. Comparison of model performance (The authors' elaboration)

We compared the profits of Ghanaian e-commerce businesses studied as a case study before and after the implementation of accurate marketing. Our analysis reveals a significant increase in profitability. According to the average annual data, the monthly income from target marketing is 0.852, compared to only 0.621 for traditional marketing strategies.

Table 1. Performance of models (The authors' elaboration)

Models	Acc	Sen	Spe	F1-score
LSTM	0.842	0.810	0.863	0.821
CNN	0.848	0.811	0.865	0.827
BiLSTM	0.836	0.795	0.856	0.811
GRU	0.845	0.810	0.864	0.826
CNN-LSTM	0.850	0.813	0.867	0.837

5. Conclusion

This study introduced a novel hybrid CNN-LSTM model that effectively analyzes customer sentiment and satisfaction in the e-commerce sector, outperforming traditional models such as CNN, LSTM, BiLSTM, and GRU in terms of accuracy, sensitivity, specificity, and F1-score. The main findings highlight the model's superior performance in generating precise consumer profiles, which is crucial for effective precision marketing, particularly in emerging markets such as Ghana. Compared with previous research, this study uniquely combines CNN and LSTM architectures, offering enhanced sentiment analysis capabilities. These implications are significant for e-commerce businesses, providing a powerful tool to improve marketing strategies, customer satisfaction, and loyalty through better decision-making. We recommend further research to validate this model across different languages and markets and to explore the integration of additional techniques to further enhance performance. Future studies could also examine the 'applicability of the model in other industries where customer feedback is critical, such as hospitality, retail, and financial services.

Declarations

Author Contributions

Conceptualization, N.E.K. and Y.M.M.; methodology, N.E.K.; software, A.B.; validation, N.E.K.; formal analysis, Y.M.M.; investigation, N.E.K. and A.B.; resources, Y.M.M.; data curation, N.E.K.; writing—original draft preparation, A.B.; writing—review and editing, N.E.K.; visualization, Y.M.M.; supervision, N.E.K.; project administration, N.E.K. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

Funding information is not available.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

References

[1] EL KOUFI N., & BELANGOUR A. Toward a responsive system for precision marketing based on RFM model, deep learning and features importance ranking: a case study of Morocco. *International Journal of Advanced Computer Science Applications,* 2024, 16(2): 1-10.

[2] BAI X., AW E. C. X., TAN G. W. H., & OOI K. B. Livestreaming as the next frontier of e-commerce: A bibliometric analysis and future research agenda. *Electronic Commerce Research and Applications,* 2024, 110(2): 1-10. https://doi.org/10.1016/j.elerap.2024.101390

[3] CHEN H., LI H., & PIRKKALAINEN H. How extended reality influences e-commerce consumers: A literature review. *Electronic Commerce Research and Applications*, 2024, 10(14): 1-13. https://doi.org/10.1016/j.elerap.2024.101404

[4] EL KOUFI N., BELANGOUR A., & SADIQ M. A systematic literature review of machine learning techniques applied to precision marketing. *Technical and Physical Problems of Engineering*, 2022, 14(7): 1-7.

[5] ZAGHLOUL M., BARAKAT S., & REZK A. Predicting E-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches. *Journal of Retailing and Consumer Services*, 2024, 79(33): 1-17. https://doi.org/10.1016/j.jretconser.2024.103865

[6] EL KOUFI N., BELANGOUR A., & SADIQ M. Research on precision marketing based on big data analysis and machine learning: case study of Morocco. International *Journal of Advanced Computer Science and Applications*, 2022, 13(10): 1-8.

[7] EL KOUFI N., & BELANGOUR A. Research intelligent precision marketing of insurance based on explainable machine learning: a case study of an insurance company. Journal of Theoretical and Applied Information Technology, 2024, 102(6): 1-14. http://dx.doi.org/10.14569/IJACSA.2022.0131008

[8] ALMUTAIRI S. M., & ALOTAIBI F. M. A Comparative Analysis for Arabic Sentiment Analysis Models In E-Marketing Using Deep Learning Techniques. *Journal of Engineering and Applied Sciences*, 2024, 10(1): 1-13. https://doi.org/10.5455/jeas.2023050102

[9] RAVIYA K., & VENNILA M. An approach for recommender system based on multilevel sentiment analysis using hybrid deep learning models. International Conference on Smart Structures and Systems (ICSSS), IEEE, 2022, 13(4):1-06.

[10] RAJU B., SRIRAM M., & GANESAN V. Optimized deep learning model for sentimental analysis to improve consumer experience in ecommerce websites. *Journal of Theoretical and Applied Information Technology*, 2024, 102(7):1-14.

[11] LUO Y., & XU X. Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *International Journal of Hospitality Management*, 2021, 94(34):1-19.

[12] GIATSOGLOU M., VOZALIS M. G., DIAMANTARAS K., VAKALI A., SARIGIANNIDIS G., & CHATZISAVVAS K. C. Sentiment analysis leveraging emotions and word embeddings. *Expert Systems with Applications*, 2017, 69(20): 214-224.

[13] DANG N. C., MORENO-GARCÍA M. N., & DE LA PRIETA F. Sentiment analysis based on deep learning: A comparative study. *Electronics*, 2020, 9(3): 1-8.

[14] FAN Z. P., CHE Y. J., & CHEN Z. Y. Product sales

forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, 2017, 74(30): 1-10.
[15] PRIYADARSHINI I., & COTTON C. A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, 2021, 77(12): 1-20.
[16] ALZAHRANI M. E., ALDHYANI T. H., ALSUBARI S. N., ALTHOBAITI M. M., & FAHAD A. Developing an Intelligent System with Deep Learning Algorithms for Sentiment Analysis of E-Commerce Product Reviews. *Computational Intelligence and Neuroscience*, 2022, 2022(1): 1-21.
[17] ZHU H. A deep learning based hybrid model for sales prediction of e-commerce with sentiment analysis. *International Conference on Computing and Data Science (CDS), IEEE*, 2021, 30(5): 1-9.
[18] YANG L., LI Y., WANG J., & SHERRATT R. S. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access*, 2020, 8(5): 1-20.
[19] IQBAL A., AMIN R., IQBAL J., ALROOBAEA R., BINMAHFOUDH A., & HUSSAIN M. Sentiment analysis of consumer reviews using deep learning. *Sustainability*, 2020, 14(17): 1-9.
[20] SHOJA B. M., & TABRIZI N. Customer reviews analysis with deep neural networks for e-commerce recommender systems. *IEEE Access*, 2019, 7(3): 1-30.
[21] KAUR G., & SHARMA A. A deep learning-based model-using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, 2023, 10(1): 1-5.
[22] SARI P. K., ALAMSYAH A., & WIBOWO S. Measuring e-Commerce service quality from online customer review using sentiment analysis. *Journal of Physics: Conference Series*, 2018, 971(54): 1-20.
[23] XU F., PAN Z., & XIA R. E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Information Processing & Management*, 2020, 57(5): 1-21. https://doi.org/10.1016/j.ipm.2020.102221
[24] YADAV V., VERMA P., & KATIYAR V. Long short term memory (LSTM) model for sentiment analysis in social data for e-commerce products reviews in Hindi languages. *International Journal of Information Technology*, 2023, 15(2): 759-772.
[25] EL KOUFI N., & BELANGOUR A. Artificial Intelligence Techniques in Precision Marketing: A Multi-criteria Analysis and Comparative Study. *International Conference on Computing, Intelligence and Data Analytics*, Springer, 2023, 2(1): 1-7.
[26] EL KOUFI N., & BELANGOUR A. Toward a decision-making system based on artificial intelligence for precision marketing: A case study of Morocco. *Journal of Open Innovation: Technology, Market, and Complexity*, 2024, 10(1): 1-20. https://doi.org/10.1016/j.joitmc.2024.100250
[27] EL KOUFI N., & BELANGOUR A. Artificial intelligence techniques applied in precision marketing: a survey. *International Conference on Electrical, Computer, Communications and Mechatronics Engineering*, IEEE, 2024, 13(6): 1-5. https://doi.org/10.1109/ICECCME57830.2023.10253083

参考文:

[1] EL KOUFI N.，和 BELANGOUR A. 基于 RFM 模型、深度学习和特征重要性排名的精准营销响应系统：以摩洛哥为例。国际高级计算机科学应用杂志，2024，16(2)：1-10。
[2] BAI X.、AW E. C. X.、TAN G. W. H. 和 OOI K. B. 直播作为电子商务的下一个前沿：文献计量分析和未来研究议程。电子商务研究与应用，2024，110(2)：1-10。https://doi.org/10.1016/j.elerap.2024.101390
[3] CHEN H.、LI H. 和 PIRKKALAINEN H. 扩展现实如何影响电子商务消费者：文献综述。电子商务研究与应用，2024，10(14：1-13。https://doi.org/10.1016/j.elerap.2024.101404
[4] EL KOUFI N.、BELANGOUR A. 和 SADIQ M. 应用于精准营销的机器学习技术的系统文献综述。工程技术和物理问题，2022，14(7)：1-7。
[5] ZAGHLOUL M.、BARAKAT S. 和 REZK A. 预测电子商务客户满意度：传统机器学习与深度学习方法。零售和消费者服务杂志，2024，79(33)：1-17。https://doi.org/10.1016/j.jretconser.2024.103865
[6] EL KOUFI N., BELANGOUR A., 和 SADIQ M. 基于大数据分析和机器学习的精准营销研究：以摩洛哥为例。国际先进计算机科学与应用杂志，2022，13(10)：1-8。
[7] EL KOUFI N.，和 BELANGOUR A. 基于可解释机器学习的保险智能精准营销研究：以保险公司为例。理论与应用信息技术杂志，2024，102(6)：1-14。http://dx.doi.org/10.14569/IJACSA.2022.0131008
[8] ALMUTAIRI S. M.，和 ALOTAIBI F. M. 使用深度学习技术对电子营销中的阿拉伯语情感分析模型进行比较分析。《工程与应用科学杂志》，2024，10(1)：1-13。https://doi.org/10.5455/jeas.2023050102
[9] RAVIYA K.，和 VENNILA M. 一种基于混合深度学习模型的多级情感分析推荐系统方法。国际智能结构与系统会议(ICSSS)，IEEE，2022，13(4)：1-06。
[10] RAJU B.、SRIRAM M. 和 GANESAN V. 优化的情感分析深度学习模型，以改善电子商务网站的消费者体验。理论与应用信息技术杂志，2024，102(7): 1-14。
[11] LUO Y.，和 XU X. 新冠疫情时代在线餐厅评论分析深度学习模型的比较研究。国际酒店管理杂志，2021，94(34): 1-19。
[12] GIATSOGLOU M.、VOZALIS M. G.、

DIAMANTARAS K.、VAKALI A.、SARIGIANNIDIS G. 和 CHATZISAVVAS K. C. 利用情感和词嵌入进行情绪分析。应用专家系统，2017，69(20): 214-224。

[13] DANG N. C.、MORENO-GARCÍA M. N. 和 DE LA PRIETA F. 基于深度学习的情绪分析：一项比较研究。电子产品，2020，9(3)：1-8。

[14] FAN Z. P.、CHE Y. J. 和 CHEN Z. Y. 使用在线评论和历史销售数据进行产品销售预测：一种结合低音模型和情绪分析的方法。商业研究杂志，2017，74(30)：1-10

[15] PRIYADARSHINI I. 和 COTTON C. 一种用于情绪分析的新型 LSTM-CNN-网格搜索深度神经网络。这超级计算杂志，2021，77(12)：1-20。

[16] ALZAHRANI M. E. 、 ALDHYANI T. H. 、 ALSUBARI S. N.、ALTHOBAITI M. M. 和 FAHAD A. 开发一种具有深度学习算法的智能系统，用于电子商务产品评论的情绪分析。计算智能与神经科学，2022，2022(1): 1-21。

[17] ZHU H. 一种基于深度学习的混合模型，用于情绪分析的电子商务销售预测。国际计算与数据科学会议（CDS），IEEE，2021，30（5）：1-9。

[18] YANG L.，LI Y.，WANG J.，和 SHERRATT R. S. 基于情感词典和深度学习的中文电子商务产品评论情绪分析。IEEE 访问，2020，8（5）：1-20。

[19] IQBAL A.，AMIN R.，IQBAL J.，ALROOBAEA R.，BINMAHFOUDH A.，和 HUSSAIN M. 使用深度学习对消费者评论进行情绪分析。可持续性，2020，14（17）：1-9。

[20] SHOJA B. M.， 和 TABRIZI N. 使用深度神经网络进行电子商务推荐系统的客户评论分析。IEEE 访问，2019，7(3)：1-30。

[21] KAUR G.，和 SHARMA A. 一种基于深度学习模型的混合特征提取方法，用于消费者情绪分析。大数据杂志，2023，10(1)：1-5。

[22] SARI P. K.、ALAMSYAH A. 和 WIBOWO S. 使用情绪分析从在线客户评论中衡量电子商务服务质量。物理学杂志：会议系列，2018，971(54)：1-20。

[23] XU F.、PAN Z. 和 XIA R. 基于朴素贝叶斯持续学习框架的电子商务产品评论情绪分类。信息处理与管理，2020，57(5)：1-21。https://doi.org/10.1016/j.ipm.2020.102221

[24] YADAV V.、VERMA P. 和 KATIYAR V. 用于印地语电子商务产品评论社交数据情绪分析的长短期记忆(LSTM)模型。国际信息技术杂志，2023，15(2)：759-772。

[25] EL KOUFI N. 和 BELANGOUR A. 精准营销中的人工智能技术：多标准分析和比较研究。国际计算、智能和数据分析会议，施普林格，2023，2(1)：1-7。

[26] EL KOUFI N. 和 BELANGOUR A. 面向精准营销的人工智能决策系统：以摩洛哥为例的研究。《开放式创新：技术、市场和复杂性》杂志，2024，10(1)：1-20。https://doi.org/10.1016/j.joitmc.2024.100250

[27] EL KOUFI N. 和 BELANGOUR A. 人工智能技术在精准营销中的应用：一项调查。国际电气、计算机、通信和机电一体化工程会议，IEEE，2024，13(6)：1-5。https://doi.org/10.1109/ICECCME57830.2023.10253083