IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 June 2024, accepted 22 July 2024, date of publication 30 July 2024, date of current version 29 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3435916

RESEARCH ARTICLE

Advancing E-Commerce Authenticity: A Novel Fusion Approach Based on Deep Learning and Aspect Features for Detecting False Reviews

SAMIA M. ABD-ALHALEM[1], HESHAM ARAFAT ALI[1,2], NAGLAA F. SOLIMAN[3], ABEER D. ALGARNI[3], AND HANAA SALEM MARIE[1]

[1]Faculty of Artificial Intelligence, Delta University for Science and Technology, Gamasa 35712, Egypt
[2]Computers and Control Systems Engineering Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
[3]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

Corresponding author: Samia M. Abd-Alhalem (samia.mohammed@deltauniv.edu.eg)

This work was funded by the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R51), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT In the contemporary digital marketplace, the proliferation of online consumer reviews has a pivotal influence on purchasing decisions. Concurrently, the prevalence of spurious reviews poses a substantial risk to the integrity of e-commerce, misleading consumers, and detrimentally impacting businesses. This paper delineates a pioneering methodology for the identification of counterfeit reviews, which is based on the combination of deep learning attributes and aspect-based analytical features. The main contribution of this research is (1) proposing an aspect fusion network based on the hierarchical attention mechanism to address the problem of multiple aspects of representing review content. The aspect fusion network can help select important aspect words and fuse aspect dictionaries with word-level attention weights. (2) We build a cardinality fusion model so that the heuristic can mitigate the negative impact of random weights and intervals on the auxiliary model. The methodology integrates advanced deep learning paradigms with aspect-based sentiment analysis to detect fraudulent reviews. Specifically, the approach encompasses a dual-method strategy: initially utilizing a Convolutional Neural Network (CNN) for the extraction of profound characteristics from review texts, followed by employing aspect-based sentiment analysis tools, including Part-of-Speech (PoS) tagging and GloVe embedding, for the distillation of aspectual features. Subsequently, these split sets of features are synergized and applied in the training of various classifier layers. Eextensive experiments have been conducted on six public review datasets contrasting the previous work on authenticity and aspect analysis. The effectiveness and performance of the proposed authenticity fusion model have been verified by the detailed analyses. The proposed model outperforms the competitors with remarkable improvement on both review authenticity and aspect analysis. This innovative approach was rigorously evaluated using a dataset of Amazon reviews that encompassed both authentic and counterfeit reviews. The empirical results demonstrate that our proposed method attains a remarkable accuracy rate of 97.73%, substantially surpassing existing state-of-the-art methodologies. The study posits that the strategic fusion of deep learning attributes and aspect-based features significantly enhances the efficacy of counterfeit review detection systems, presenting a formidable tool in the arsenal against e-commerce fraud.

INDEX TERMS E-commerce authenticity, fake review detection, CNN, fraud detection in digital marketplaces, PoS, GloVe, sentiment analysis techniques.

I. INTRODUCTION

Studying and addressing the validity of online reviews plays an important role in the field of e-commerce. In recent years,

The associate editor coordinating the review of this manuscript and approving it for publication was Fu Lee Wang.

some research methods have been proposed and a few real detection models have also been established. These traditional originality detection methods usually aim to learn the representation of the fused features and calculate the validity of the review through a feature fusion process based on traditional machine learning methods. Most of these methods fail to capture features that are very useful for originality detection, so they often perform poorly at this task.

In the realm of online commerce, consumers are often confronted with a deluge of information, marked by discrepancies between the product descriptions available online and the actual attributes of products procured offline. This phenomenon necessitates reliance on product reviews for informed decision-making. Consequently, these reviews hold significant sway over both consumer purchasing tendencies and business interests, as delineated in the existing literature [1]. The influence of reviews is two-fold: while negative reviews can deter potential customers, positive reviews tend to attract them.

In this competitive landscape, some business entities engage in deceptive practices. Specifically, they commission skilled writers to fabricate positive reviews for their products, thereby artificially inflating their appeal and market presence. Simultaneously, these entities may also generate false negative reviews targeting their competitors to undermine them [2]. These nefarious activities not only profoundly mislead prospective customers but also impede the organic growth and trustworthiness of e-commerce platforms [3].

Research indicates a significant challenge faced by customers in distinguishing authentic reviews from counterfeit ones. This situation underscores the pressing need for effective methodologies to identify fraudulent reviews. The development of such methods is imperative for purifying the online shopping milieu, ensuring a wholesome purchasing experience for consumers, and fostering the acquisition of legitimate and constructive feedback.

Online reviews, a cornerstone of the digital commerce ecosystem, are generated for a myriad of reasons. To foster business growth, online retailers and service providers often encourage the dissemination, replication, and unrestricted use of these reviews across various platforms. Such practices mandate appropriate attribution of the original authors and sources, acknowledgment of any modifications made, and clear delineation of associated licensing agreements. Moreover, service providers frequently solicit customer feedback on their product or service experiences. This feedback mechanism serves as a barometer of customer satisfaction, offering insights into consumer perception post-purchase.

In recent years, deep learning representations, including CNNs, have shown superior performance in image processing and pattern recognition tasks. Specifically, the deep features learned by CNNs showed superior performance in the e-commerce authenticity detection task. Notwithstanding, the reliance on customer reviews, predicated on individual experiences, introduces a quandary. The uncritical acceptance of these reviews can be perilous for both service providers and consumers. The notion of 'fake review detection' emerged within this context, initially coined as 'opinion spam detection' [3]. The criticality and scientific ramifications of detecting counterfeit reviews have garnered considerable attention recently. Pioneering investigations in this domain have primarily focused on the manual curation of features in tandem with machine learning algorithms. The key features are the length of the review text [4], [5], its lexical characteristics [6], and its affective polarity [7], all of which constitute essential semantic elements of the text.

The burgeoning expansion of user-generated content on the internet, particularly in the form of product reviews, has precipitated a notable concern regarding the genesis of counterfeit reviews. The imperative to discern and intercept these spurious reviews is paramount in maintaining the integrity of online review platforms, thereby enabling consumers to execute judicious purchasing decisions. This study elucidates a novel methodology for the detection of counterfeit reviews, an approach that combines the prowess of deep learning techniques with aspect-based analytical features.

At the outset, our method employs advanced learning models, notably Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), to meticulously extract intricate patterns and representations from the textual corpus of review data. These models are adept at discerning the nuanced meanings and contextual subtleties embedded within reviews, thereby augmenting our capacity to differentiate between authentic and counterfeit reviews.

Complementing these deep learning techniques, the research integrates aspect-based features that home in on specific product attributes referenced in reviews. These features are derived through meticulous aspect-based sentiment analysis, which enables the capture of sentiments expressed toward various facets of a product. We posit that these aspectual elements provide critical insights for the identification of counterfeit reviews because they frequently exhibit divergent sentiment polarities compared to their genuine counterparts.

A. MOTIVATION

In the modern era, e-commerce has witnessed rapid and steady development, attracting many online consumers. Developing effective trust solutions to combat false and misleading content, and enhance e-commerce authenticity, has become increasingly important. Multi-faceted e-commerce authenticity is offered based on a variety of individual aspects, whether it includes genuine proposition, maintaining themes, optimal advertising, confirmation, and ensuring sufficient quality, among other factors. Based on this multi-faceted authentication, the low implementation ability to detect e-commerce authenticity makes the task difficult due to the mixed nature of different visual objects and features.

The motivation for this research is rooted in the evolving dynamics of user behavior in the digital review space. This behavior can be exemplified by factors such as the volume of positive or negative reviews submitted by users [4] and the frequency of their review submissions [8]. Profit-driven spammers are continually refining their strategies, employing increasingly sophisticated methods to circumvent detection by existing review verification systems.

In response to these challenges, a spectrum of fake review detection techniques leveraging deep learning algorithms has emerged, paralleling the advancements in the field of deep learning [9], [10], [11], [18]. These novel approaches exhibit enhanced domain adaptability and efficacy compared with traditional feature-based methods. They are particularly adept at autonomously extracting semantic information embedded in textual data, thereby circumventing the limitations inherent in manually crafted feature designs.

Despite the successes achieved by these current methodologies, most exhibit a unidimensional focus, either concentrating exclusively on the textual content of reviews or the characteristics of the reviewers. Such an approach often overlooks the multifaceted and implicit expression patterns exhibited by users as well as the intricate interactions that occur among users, products, and the textual content of reviews [11], [12]. Furthermore, it has been observed that genuine reviews, regardless of their positive or negative nature, often contain specific details (e.g., the flavor profile of a dish in a restaurant review) that substantiate the reviewer's emotional expressions. This observation underscores the need for a more holistic and nuanced approach to the detection of counterfeit reviews.

In the rapidly evolving landscape of e-commerce, the integrity of online reviews has become a paramount concern for both consumers and businesses. Recognizing the urgent need for more effective tools in the battle against fraudulent reviews, this study embarks on a quest to develop advanced methodologies for detecting fake reviews. Our study is grounded in the understanding that the current landscape of review authenticity verification faces significant challenges, including the lack of standardized datasets and inadequate feature analysis. Against this backdrop, our research contributes to the field in several key ways.

B. CONTRIBUTION

The main contribution of this work is to develop an authenticity fusion model for e-commerce reviews, considering the authenticity and authenticity aspects of reviews together. With the increasing number of fake reviews being discovered by both customers and platforms, the authenticity of reviews has become one of the top concerns for consumers because their purchasing decisions rely heavily on the information provided by online reviews. The integrity of e-commerce platforms will also be harmed by the spread of fake reviews. Therefore, detecting the authenticity of reviews is a necessary prior step for managing e-commerce platforms and recommendation systems.

Before making detections based on different classifier layers, we need to understand the data better. That is why we devote this paper to extracting features of text data to help us make sense of all these data. in addition, it helps us to process huge amounts of data efficiently and cost-effectively.

- Stage 1: Feature extraction, data-driven, and text preprocessing: This stage involves the extraction of informative features from raw data through data analysis techniques. These features must capture certain characteristics pertinent to the sentiment analysis task that follows it and also entail the preparation of text data.
- Stage 2: Aspect-Specific Analysis with Deep Learning Feature Extraction: Because of this, the present paper proposes a novel approach for integrating deep learning paradigms with aspect-based sentiment analysis, i.e., the sentiment captured according to different aspects or entities detailed in a review. The proposed methodology focuses on aspect-specific analysis with deep learning feature extraction to detect fraudulent reviews. This method is employed to automatically extract high-level features from the preprocessed text data. These features capture sentiment information specific to the aspects (entities or topics) discussed in the text.
- The suggested approach is predicated on employing a fusion network to integrate the retrieved deep-learned features with the aspect-based analytical features. then used to train the classifier layers. These classifiers will learn how to map the extracted features into sentiment labels (e.g., original and computer-generated) or aspect-specific sentiment labels.
- The proposed methodology greatly enhances the detection performance which has been a long-time problem for previous studies.

The main core is to safeguard consumers using a novel approach that prevents employers from manipulating them without compromising detection capabilities. This methodology emphasizes lexical features (i.e., properties obtained from text) such as keywords or -phrases, n-grams, punctuation, semantic similarity, latent subjects, and markers of linguistic styles. Its primary focus is on reviews as textual data. The suggested approach is predicated on employing a fusion network to integrate the retrieved deep-learned features with the aspect-based analytical features. In particular, we suggest a deep feature extraction strategy based on convolutional neural networks (CNNs). Furthermore, multiple methods are utilized to extract aspect-based analytical features, including GloVe embedding and Part-of-Speech (PoS) tagging. See section c of the Proposed Detection Model for additional information on this methodology.

In this paper, we present a structured exploration of fake review detection in e-commerce, Section I represents an introduction and then begins with a comprehensive review of the literature in Section II, where we compare and contrast recent advancements in this field. Following this, Section III delves into a detailed proposed architectural framework. Section IV introduces our experimental and results details. Section V culminates our research with an analytical presentation of the experimental results, evaluating the effectiveness and functional characteristics of our proposed fake review detection system. Finally, Section VI introduces the conclusion.

II. BACKGROUND AND RELATED WORK

The phenomenon of counterfeit reviews has garnered significant attention within the research community, as evidenced by numerous studies [13]. A critical aspect of this research is feature engineering, a process in which salient features are extracted or constructed from data. Crawford et al. [8] provided an insightful overview of the various types of features that can be derived from review texts. A prevalent method among these is the ''bag of words'' approach, in which the most commonly occurring words within the review texts are used as features, either individually or in small clusters.

Further advancing this field, Christopher et al. [14] employed a multifaceted methodology to identify fraudulent reviews on the Yelp platform, specifically targeting restaurant reviews. Their approach encompasses the utilization of a bidirectional long short-term memory (LSTM) network, a form of the recurrent neural network, to leverage the positional relevance of comments within reviews. LSTMs, known for their efficacy in analyzing linguistic patterns in text, enhance the precision of their proposed model by examining multiple segments within each review. In addition, they incorporated the Kullback-Leibler (K-L) divergence technique to assess the variation in word distributions between genuine and spurious Yelp reviews. This ensemble approach, while demonstrating average precision, makes a significant contribution to the methodologies employed in fake review detection.

Wang et al. [15] addressed the complex challenge of identifying loosely affiliated spammer groups. These groups are characterized by the absence of a requirement for each member to review every target product. The study employs a bipartite graph projection technique to effectively tackle this problem. The authors have innovatively developed an algorithm that adopts a divide-and-conquer strategy to detect highly suspicious loose spammer groups. In addition, they introduce a set of group spam indicators to quantitatively assess the spam activity of these groups. The method proposed by Wang et al. is particularly notable for its potential as a preprocessing tool for Frequent Itemset Mining (FIM)-based approaches. The experimental results of their study demonstrate its capability to identify loose spammer groups with high levels of precision and recall and to construct more accurate candidate fake reviewer groups compared with traditional FIM methodologies.

In a separate study, Xiao and Qiu [16] delved into the realm of opinion spammers on social media platforms. These users, a distinct subset of social media participants, are noted for their organized efforts to post comments with the deliberate intention of manipulating public opinion, thereby amplifying the influence of their sponsors. The primary objective of Xiao and Qiu's research is to conduct a quantitative analysis to understand the unique characteristics of opinion spam. This study examines various aspects, including behavioural patterns, network structures, and psycholinguistic idioms associated with opinion spam. A significant finding of their investigation is the effectiveness of context-based collective classification in detecting opinion spam, achieving an F1 score of 91%. However, the study's focus is limited to the Twitter platform, suggesting a need for further research to extend its application to other social media sites such as Facebook and LinkedIn.

Fazzolari et al. [17] re-engineered a collection of useful features for recognizing bogus reviews by considering the Cumulative Relative Frequency Distribution of each feature. Not by making a few tweaks to the tried-and-true state-of-the-art methods, but rather most effectively by changing the input utilized to build supervised classifiers during the training phase. The research shows that the use of distributional characteristics might improve the performance of the classifier through an experimental evaluation performed on real data from Yelp.com. The drawback noted in this study is that it only identifies individual spammers rather than user groups acting in concert and synchronicity with the intent to either promote or denigrate a product.

Additionally, Liu et al. [18] provided a novel method based on a partially supervised model and referred to the process of identifying opinion relations as alignment. Next, a graph-based co-ranking technique was investigated to gauge each candidate's level of confidence. Finally, candidates with a higher degree of confidence as opinion targets or opinion terms were extracted. The author's approach captures opinion relations more precisely than existing methods that are entirely dependent on nearest-neighbor principles, especially for long-span ties. The drawback of this approach is that to co-extract opinion targets and opinion words, relation graphs containing additional types of links between words, such as topical relations, must be considered.

Jindal and Liu [9] undertake a comprehensive examination of issues surrounding product reviews, which are extensively used by both consumers and manufacturers and are often laden with subjective opinions. Their research coincides with the emergence of several startups dedicated to aggregating product opinion reviews in recent years. This study underscores the pressing need for research into spam within these reviews. It notes a gap in the existing literature, as before this study, there had been substantial research on Web spam and email spam, but none specifically addressed spam in product reviews. This study not only scrutinizes various spam behaviors within the context of product reviews but also proposes effective methods for their detection. This study highlights the necessity for enhanced detection techniques and advocates for expanded research into spam across diverse media platforms, including forums and blogs. This approach acknowledges the evolving nature of spam tactics and the need for a multifaceted strategy to counter them effectively.

Asghar et al. [19] presented an innovative extension to a baseline spam detection methodology by integrating additional features pertinent to opinion spam, such as opinion spammers and item spam. Utilizing a dataset comprising Amazon customer reviews and phrases specifically labeled for spam detection, the authors rigorously evaluate the impact of variables related to spam activity in the accurate identification and categorization of spurious (false) indicators, thereby distinguishing them from authentic reviews. The authors propose a novel method for classifying review sentences as spam or non-spam, incorporating a rule-based feature weighting scheme to achieve this objective. This approach results in a notable enhancement of accuracy, increasing from 93% to 96%.

Moreover, the study explores a hybrid combination of features that, significantly elevates accuracy, recall, and F-measure values in the context of fake review detection. This research underscores the potential for even rudimentary spam detection methods to achieve superior performance when supplemented with spam-related features and a rule-based weighting strategy. However, this research is limited by the use of a relatively narrow feature set. Expansion of this feature set across various domains could yield more robust outcomes. In addition, the proposed feature weighting method currently addresses a limited spectrum of spam-related features. Broadening this scope could lead to more conclusive results. Although feature selection in this study is manually conducted, future research could explore the integration of automated feature selection using deep learning models to further enhance detection efficacy.

Lau et al. [20] introduced an advanced approach for detecting fraudulent reviews by, employing semantic language modeling and computational models grounded in text mining. This approach is particularly effective in identifying counterfeit reviews, even in scenarios where the fake reviewer employs sophisticated techniques. The research highlights the efficacy of their supervised model, specifically a Support Vector Machine (SVM), which demonstrates a true positive rate exceeding 95% in the detection of false reviews, particularly when applied to the Amazon review dataset. This performance notably surpasses that of other well-established baseline models. The authors also suggest that the effectiveness of the fake review detection process could be further enhanced by exploring more complex language modeling techniques, such as n-gram language models.

Dixit et al. [21] proposed a spam detection methodology that integrates Naive Bayesian classification with conceptual and semantic similarity measures. This study employs benchmark datasets, including PU1, Linkspam, Spam base, and the Enron corpus, to evaluate the efficacy of their analysis methods. The experimental results demonstrate a remarkable accuracy rate of 98.89%, outperforming the metrics achieved with existing data. However, the study acknowledges a potential limitation: a model trained on a dataset with a skewed balance of fewer emails and more documents may risk overfitting, which could consequently diminish the accuracy of the classifier.

In addition to the primary focus on reviews, there has been considerable research into the characteristics of both reviewers and products. Shojaee et al. [22] explored the use of syntactical and lexical features in their classification models. They employed a Support Vector Machine (SVM) and naive Bayes classifier on their AMT dataset, which comprised 400 deceptive reviews. The SVM model achieved an accuracy of 84%, whereas the naive Bayes model attained an accuracy of 74%. In their study, features were categorized into two groups: review-centric and reviewer-centric [23]. Review-centric features are derived from individual reviews, whereas reviewer-centric features utilize all reviews authored by a specific individual, along with additional information about the reviewer. The latter approach, which focuses on identifying fraudulent reviewers, is more effective in preventing multiple fake reviews, thus yielding superior detection outcomes. Consequently, this study places greater emphasis on reviewer-centric attributes.

Further supporting this approach, Jindal and Liu [9], [24] conducted a comprehensive analysis using a dataset of 5.8 million reviews sourced from Amazon. They employed a Logistic Regression classifier and found that a combination of various feature types generally leads to enhanced performance compared with relying on a single feature type. This is illustrated in Table 1, which presents a comparative analysis of recent studies focusing on the critical task of detecting fake reviews. Table 1 provides a comprehensive overview and summarizes the various approaches, features, classifiers, datasets, limitations, and future scopes of recent studies in the field of fake review detection.

As delineated in Table 1, the field of fake review detection research is confronted with many challenges:

Predominantly, classification algorithms within a fully supervised framework are employed to detect counterfeit reviews. However, this approach is hampered by the substantial requirement for labeled data, which is arduous to procure. Manual labeling of data not only demands extensive material resources and manpower but is also susceptible to inaccuracies due to subjective biases, thereby constraining the progress of fully supervised learning methodologies. Consequently, the accuracy of unsupervised learning approaches, which categorize data through cluster analysis in the absence of supervision, remains suboptimal for complex detection tasks.

Semi-supervised learning emerges as a viable alternative, adept at mitigating the primary shortcomings associated with both fully supervised and unsupervised learning. Nevertheless, the present detection models predominantly utilize basic features, such as part-of-speech tags or n-grams, for modeling. This approach overlooks critical aspects like the interplay between various features, which in turn diminishes the effectiveness of classification.

Furthermore, while current deep learning models demonstrate proficiency in classifying straightforward text, they are less adept at discerning fraudulent reviews. The core challenge lies in the inherent complexity of fake reviews, which often lack identifiable fixed elements in plain text. A multifaceted analysis that incorporates diverse perspectives, including user and company information, is crucial. Therefore, a multidimensional approach to identifying counterfeit reviews is imperative to effectively address these complexities.

In response to identified challenges, including the scarcity of standardized datasets, suboptimal feature selection, and lackluster detection performance, this study proposes a novel detection methodology. This method is meticulously designed to address the prevalent issues in the domain of fake review detection. It encompasses a comprehensive set of procedures for the extraction and quantification of features, rigorously examining the interrelations among these features. Furthermore, the methodology establishes an intricate system dedicated to assessing the veracity of reviews. This system is engineered to enhance detection accuracy and reliability, offering a robust solution to the complexities inherent in identifying counterfeit reviews.

III. PROPOSED DETECTION MODEL

In the proposed methodology, we assume that there are modern e-commerce architectures to facilitate information exchange between users and application providers to deliver services or offers to end consumers. It is important to identify the application or offering providers and users, which represent the companies and users on whose behalf the information is generated. To prove the truth on the ground, users must be matched with companies. The main goal here is to focus only on users without restricting ourselves to data provided by publicly listed companies only. However, the results of this effort limit the generalizability of our findings to the field of e-commerce.

Central to this paper is the optimal selection of features, encompassing text-dependent, text-independent, or a synergistic combination of both, to comprehensively understand the nuances of review content. As illustrated in Figure 1, we propose an integrated system designed to enhance the accuracy and thoroughness of data analysis in the detection of counterfeit reviews. This system innovatively incorporates unstructured reviews and employs advanced text-mining techniques. By extracting and learning novel patterns from the review text, the proposed model represents a significant advancement in fake review detection methodologies. The subsequent subsections meticulously delineate the foundational elements essential for a detailed description of the proposed model.

A. DATASET DESCRIPTION

First, we must first collect a suitable data set of product reviews that can be used to fine-tune the model. In general, data is a vital element in any machine learning model, because it directly affects the quality of the results. The dataset employed in this study comprises review texts sourced from the Open Science Framework (OSF) site [34]. This dataset is characterized by an equal distribution of counterfeit and authentic reviews, encompassing a total of 40,432 reviews across ten categories from Amazon. It includes four primary features: category, rating, label, and review text, with the rating feature ranging on a scale from 1 to 5. Figure 2 provides a representative snapshot of the dataset.

After data preprocessing, some data will typically be deleted. As a result, some features won't be accessible. We will examine the various possibilities of these attributes with the aid of illustrations utilizing several Python tools, including NumPy, Pandas, Matplotlib, Seaborn, WordCloud, and NLTK. We can start by determining whether or not the data is balanced. Figure 3 (a) illustrates that the data is balanced. This is a crucial topic to remember because there are differences between balanced and unbalanced model performance evaluations when generating predictions. The distribution of reviews by rating and category is displayed in Figure 3 (b, c). For every category, the majority of reviews have a score between three and five. Figure 3 (d) illustrates the association between the sentence length and number of reviews in each category [for each label OR and CG].

Average sentence length is a crucial variable that we created. It has two values:

1) Review length $>=$ average sentence length therefore classified as original review
2) Review length $<$ average sentence length therefore classified as original or generated. these reviews feed to the next step (data preprocessing).

Word clouds are an excellent tool for checking the most frequently used terms. It works well to convey information. The most common 100 words in the reviews are shown in Figure 4. The fact that nice words predominated in the dataset is not surprising.

Additionally, descriptive statistical analysis offers insights into the numerical aspects of the dataset. For instance, as shown in Table 2, there is a noticeable skewness towards a rating of 5.0, and the review length presents some outliers.

B. DATA PREPROCESSING

Data preprocessing constitutes a pivotal step in analysing text data, particularly in interpreting and extracting the semantic essence of review texts. To facilitate effective analysis, it is essential to transform these textual reviews into a vectorized format using methodologies capable of comprehending and processing the text. The preprocessing steps implemented in this study, as illustrated in Figure 5,. (a) General overview of data preprocessing methods. (b) Detail of preprocessing steps, the steps include the removal of punctuation, tokenization, elimination of stop words, and lemmatization [9], [35]. Each of these steps plays a critical role in refining the data, ensuring that the subsequent analysis is based on relevant and contextually enriched text.

C. FEATURE EXTRACTION

Feature extraction is an integral process in natural language processing (NLP), which involves the identification and extraction of pertinent features from textual data [36]. The advent of deep learning techniques in recent years has significantly advanced the capabilities of feature extraction

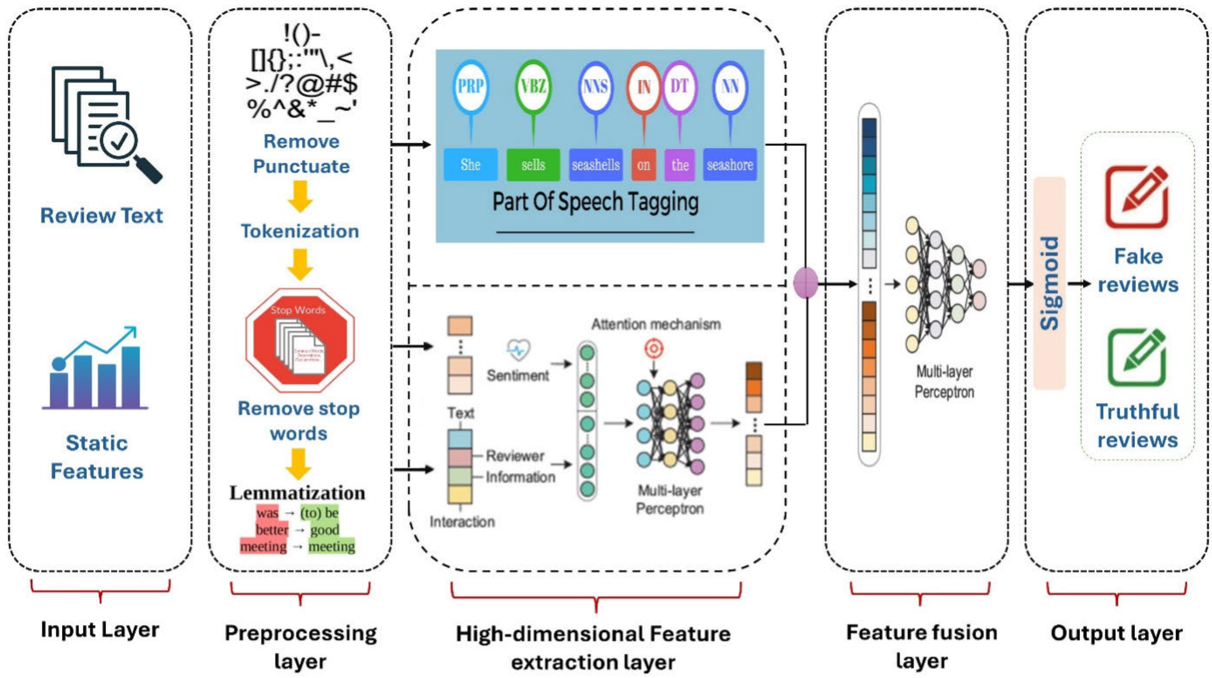


FIGURE 1. The architecture of the proposed detection model.

index	category	rating	label	text_
1	Home_and_Kitchen_5	5.0	CG	love it, a great upgrade from the original. I've had mine for a couple of years
4	Home_and_Kitchen_5	5.0	CG	Very nice set. Good quality. We have had the set for two months now and have not been
11	Home_and_Kitchen_5	5.0	CG	Very handy for one of my kids and the tools are included in the package. I have one in
2	Home_and_Kitchen_5	5.0	CG	This pillow saved my back. I love the look and feel of this pillow.
6	Home_and_Kitchen_5	5.0	CG	They are the perfect touch for me and the only thing I wish they had a little more space.
17	Home_and_Kitchen_5	5.0	CG	These look beautiful and so nice. The only problem is that it's not really a mesh one.
7	Home_and_Kitchen_5	3.0	CG	These done fit well and look great. I love the smoothness of the edges and the extra
13	Home_and_Kitchen_5	1.0	CG	These are so flimsy! They are not the quality you would expect from a piece of furniture.
9	Home_and_Kitchen_5	5.0	CG	My son loves this comforter and it is very well made. We also have a baby
3	Home_and_Kitchen_5	1.0	CG	Missing information on how to use it, but it is a great product for the price! I
14	Home_and_Kitchen_5	5.0	CG	Makes may tea with out stirring. The only problem is that it's kind of hard to put
0	Home_and_Kitchen_5	5.0	CG	Love this! Well made, sturdy, and very comfortable. I love it!Very pretty
16	Home_and_Kitchen_5	5.0	CG	Love this! Perfect size for an entire family!Very good quality.
5	Home_and_Kitchen_5	3.0	CG	I WANTED DIFFERENT FLAVORS BUT THEY ARE NOT.
8	Home_and_Kitchen_5	5.0	CG	Great big numbers & easy to read, the only thing I didn't like is the size of the
18	Home_and_Kitchen_5	5.0	CG	Exactly what you would expect. I love the look and feel of this pillow.
12	Home_and_Kitchen_5	5.0	CG	Did someone say, "Oriental for $60"? It is a great product for the
10	Home_and_Kitchen_5	5.0	CG	As advertised. 5th one I've had. The only problem is that it's not really a
15	Home_and_Kitchen_5	5.0	CG	Absolutely adorable! And excellent price. We have had the wooden ones for a few months now and they
19	Home_and_Kitchen_5	5.0	CG	10 Stars, I would highly recommend this item. We love this blanket.

FIGURE 2. Snapshot from the review dataset.

tasks within NLP, yielding promising results [36], [37], [38], [39]. This advancement has motivated the integration of deep learning methodologies for feature extraction in this study.

This study explores a variety of traditional techniques for feature extraction, including word embedding, Part-of-Speech (POS) tagging, and CNN. These methods were chosen for their effectiveness in extracting meaningful and relevant features from text data. The proposed approach in this research is to train a model using a combination of these features, which is expected to enhance the overall

TABLE 1. Comparative analysis of related studies on fake review detection.

Publication Year	Ref.	Approach	Features	Classifier	Dataset	Limitations/Future Scope
2023	[25]	Supervised Learning	Behavioral and linguistic features	DNN, Bi-LSTM, and CNN	Yelp NYC, Yelp ZIP	Text data only; does not detect new fake reviews. Future: Lexicon-based classifiers
2022	[26]	Semi-Supervised	Behavioral association and linguistic features	SVM	Yelp, Amazon, and IMDB	Feature quality dependent; Imbalance in review types. Future: Incorporating user-profiles and temporal information
2022	[27]	Supervised, Unsupervised, Semi-Supervised	Document vectors, semantic, emotional, user, and product information	SVM, NB, DT, RF, CNN, RNN, and GCN	Multiple datasets including YelpCHI, and Amazon Reviews	Difficulties in labeled dataset acquisition; Limited in identifying fake reviewers. Future: Effective feature extraction
2021	[28]	Supervised, and Semi-Supervised	Textual, and behavioral features	SVM, NB, DT, RF, and etc.	Dataset with 1144 fake reviews, and, 4709 real reviews	Focus on English; Limited to Amazon.
2020	[29]	Supervised and Semi-Supervised	Text features (semantic, lexical, sentiment, character) and behavioral features	SVM, RF, and DT	YelpCHI dataset	Ineffective without labeled data; Not addressing cross-domain analysis
2018	[30]	Semi-Supervised	-	DT, NB, RF, SVM, Logistic regression, and KNN	Chicago Hotel Reviews	Required positive instance (labeled data). Future: Extension to unsupervised learning
2017	[31]	Supervised	Linguistic clues, N-gram	DT, RF, NB, and SVM	Chicago Hotel Review	-
2016	[32]	Supervised and Unsupervised	Sentiment Score, Linguistic feature, Unigram; Reviewer data, Review Data, Product Information	SVM, NB, and DT	20 Chicago Hotel Reviews, and Amazon datasets	A limited number of features
2015	[33]	Supervised	Linguistic clues (writing style, details level, word structure, cognition indicators)	RF, SVM, and NB	15 Asia Hotel Reviews	Only for labeled data; Not suitable for Unlabeled dataset.

performance of the model in detecting fake reviews. The process and workflow of the combined feature extraction approach are depicted in Figure 6.

1) CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Networks (CNNs) are primarily recognized for their efficacy in image recognition tasks [40], [41], [42]. However, their application extends to text classification, particularly in identifying counterfeit reviews [41]. This study employs a CNN model trained on a dataset comprising both authentic and fraudulent reviews. The model's objective is to discern patterns and features indicative of fake reviews, such as specific keywords or phrases.

The architecture of the CNN model encompasses multiple layers, including convolutional, pooling, and fully

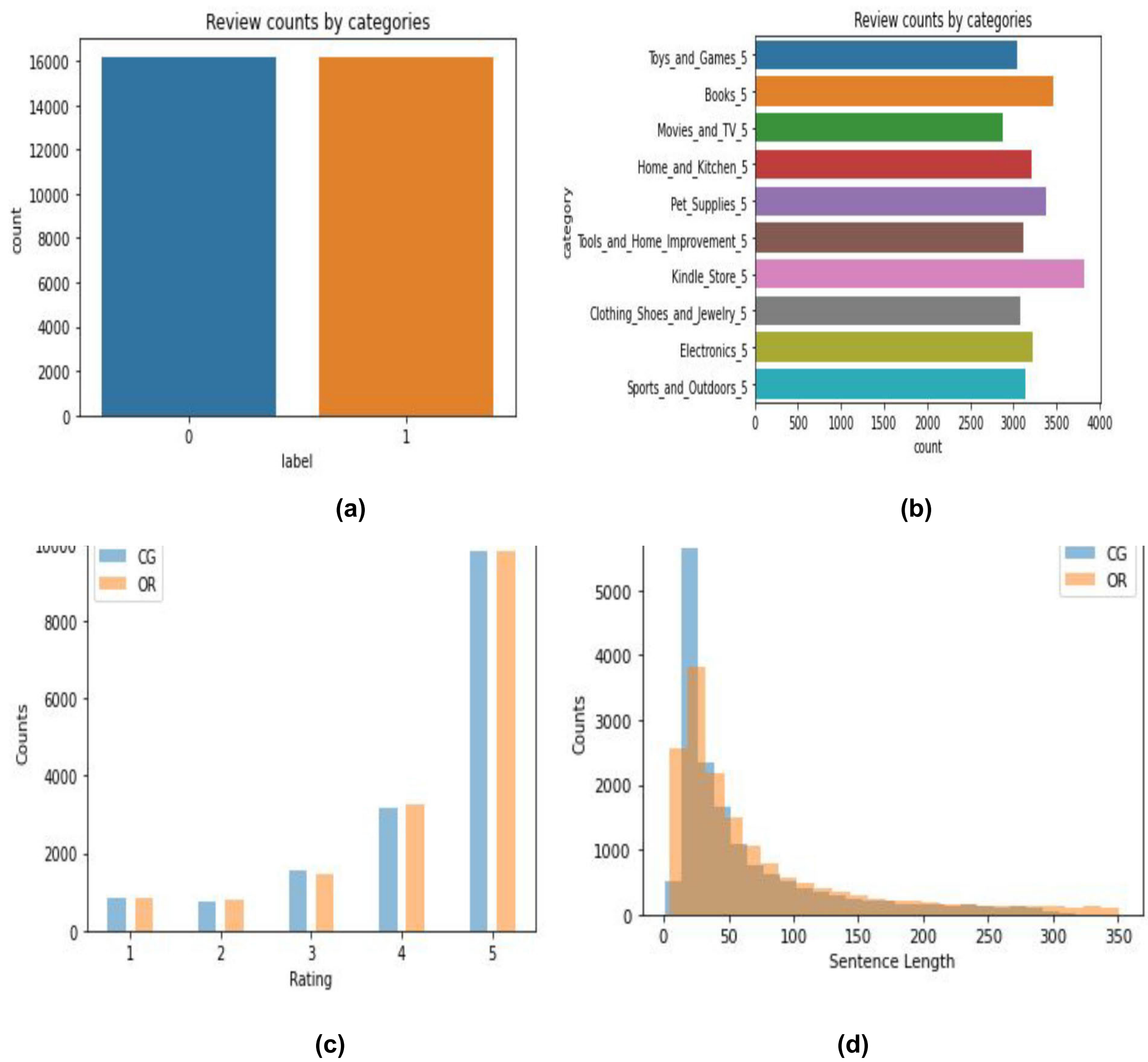


FIGURE 3. Exploratory data analysis - a) Distribution of reviews, b) Category per class visualization, c) Rating per class visualization, d) Sentence length per class.

TABLE 2. Descriptive statistics of dataset features.

	Count	Mean	Std	Min	25%	50%	75%	Max
Rating	32336.6	4.257144	1.141251	1.0	4.0	5.0	5.0	5.0
Label	32336.6	0.500557	0.500007	0.0	0.0	1.0	1.0	1.0
Sentence Length	32336.6	68.531513	70.474200	1.0	21.0	39.0	86.0	351.0
Stop words	32336.6	32.0	32.0	0.0	11.0	25.0	49.0	1296.0
punctuation	32336.6	17.0	23.0	0.0	6.0	11.0	21.0	2087.0
Hashtags	32336.6	0.0	0.0	0.0	0.0	0.0	0.0	32.0
Upper	32336.6	3	7.0	0.0	1.0	2.0	4.0	342.0
avg word	32336.6	6	45.0	0.0	5.0	5.0	6.0	10.0

connected layers. The model's input consists of word embeddings derived, from pre-trained models like GloVe [43] or Word2Vec [44]. Throughout the training phase, the CNN model is designed to extract salient features from the text data

FIGURE 4. Word cloud analysis for a) Computer-generated reviews and b) Original reviews.

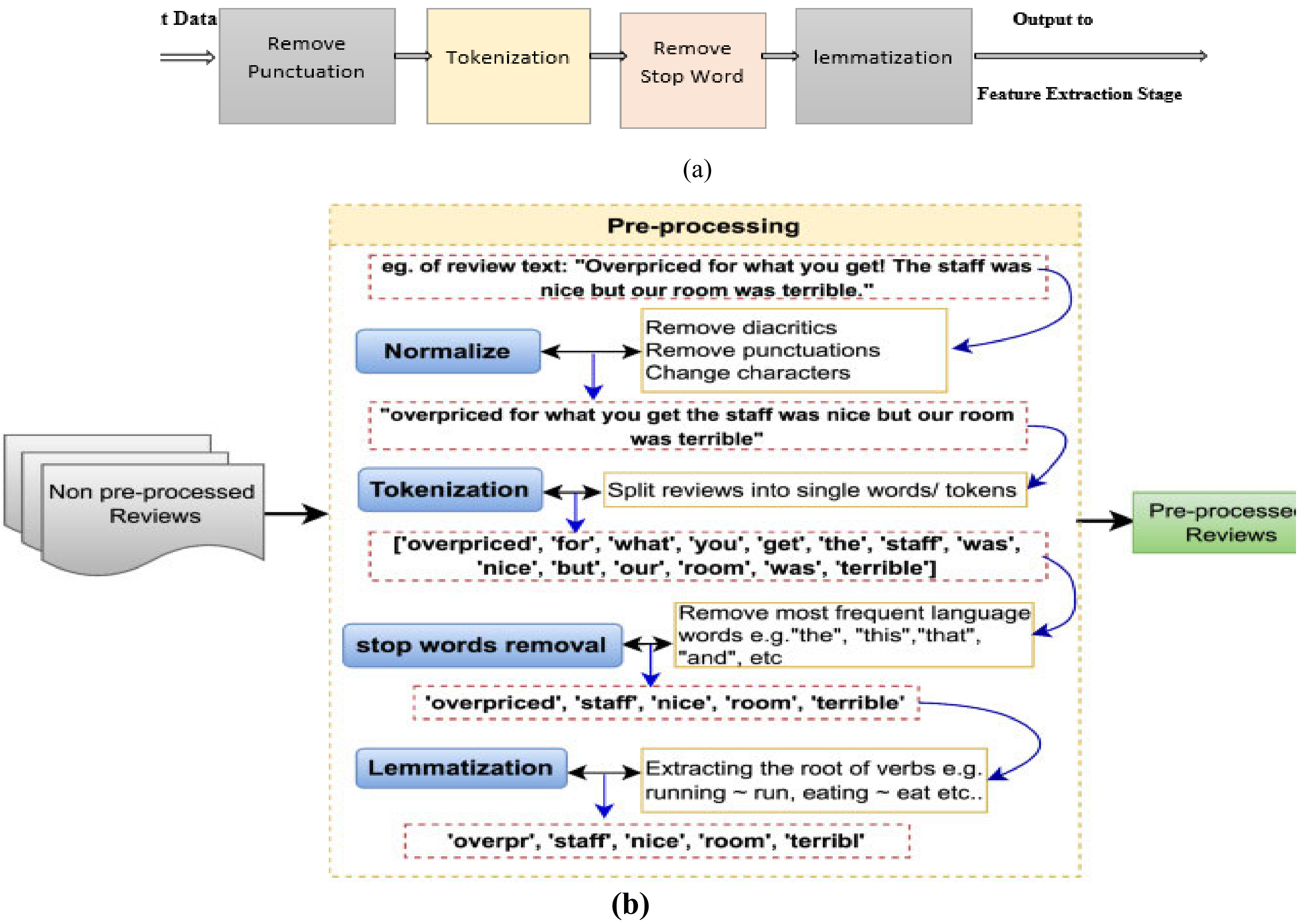


FIGURE 5. (a) General overview of data preprocessing methods. (b) Detail of preprocessing steps.

and correlate them with their respective labels. Subsequently, the trained model is used to extract features from the new text data for classification.

The key functions of the CNN in this context are detailed as follows:

- **Input Layer:** This layer represents the text as a sequence of word embeddings, denoted as $A_i = A_1 \varphi A_2 \varphi A_3 \ldots . \varphi A_n$.
- **Convolutional Layer:** This layer extracts features based on the filter, with the function, $B_i = f(\sum_{i=0}^{n} W_i A_i + b_i)$,

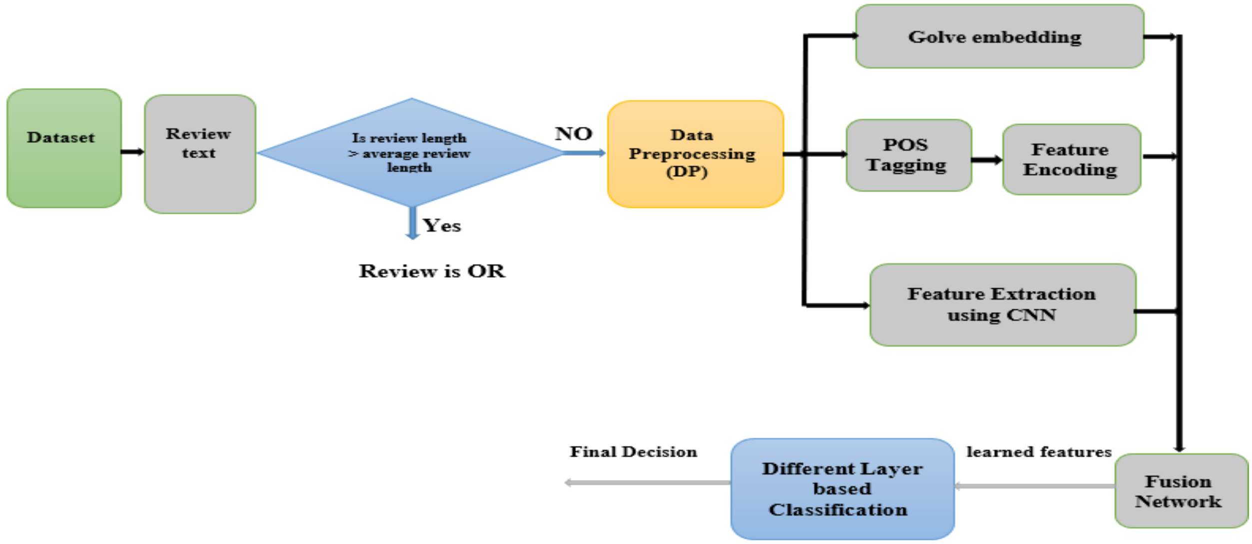


FIGURE 6. Flow diagram of the proposed feature extraction approach.

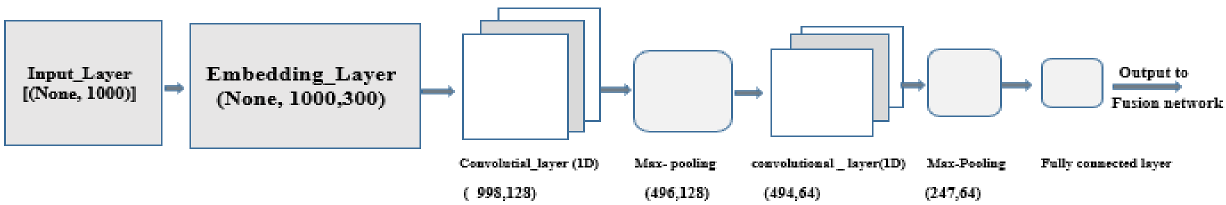


FIGURE 7. Deep-learned features based on the CNN strategy.

where f is a nonlinear activation function, W_i represents the weights, A_i the inputs, and b_i the bias.

- **Average Pooling**: This function calculates the average value for each patch on the feature map, expressed as $\vec{B}$ = Max (B).

The proposed CNN model comprises six layers. The initial four layers consist of convolutional layers, each equipped with 32 filters of kernel size 393, stride 1, and a ReLU activation function, followed by a max-pooling layer and a dropout rate of 0.25. The subsequent layers, 5 and 6, are convolutional layers with 64 filters, a kernel size of 292, and a ReLU activation function. The deep-learned features extracted by these layers are then integrated into the fusion network for further analysis.

2) WORD EMBEDDING

Word embedding is a seminal technique in natural language processing (NLP), renowned for its efficacy in feature extraction from text data. This technique involves the transformation of words into vectors of reduced dimensionality, facilitating efficient representation and processing of textual information. Among various word embedding methodologies, GloVe (Global Vectors for Word Representation) stands out. GloVe employs a co-occurrence matrix to encapsulate the statistical relationships between words [24], [45], [46], [47]. This method has demonstrated exceptional performance across a range of NLP tasks, including sentiment analysis, text classification, and machine translation. By converting words into vector representations, GloVe embeddings enable more nuanced and accurate text classifications grounded, in the deeper semantic meanings of the words.

3) PART-OF-SPEECH (POS) TAGGING

Part-of-Speech (POS) tagging is a critical process in NLP, involving the categorization of each word in a text according to its corresponding POS Linguistic analysis is instrumental in deconstructing and comprehending the grammatical structure of sentences, thereby aiding various NLP applications such as text classification, sentiment analysis, and machine translation. POS tagging employs either statistical models or rule-based algorithms to analyze the contextual usage of a word within a sentence, subsequently assigning an appropriate part of speech tag. These tags encompass a broad spectrum of grammatical categories, including nouns, verbs, adjectives, adverbs, pronouns, prepositions, conjunctions, and interjections [48].

4) FUSION NETWORK

To augment the distributed representations of text data, this study introduces a sophisticated fusion network. This network is designed to integrate various extracted features to enhance text data analysis. The architecture of the fusion network comprises two principal components: the local layer and the fusion layer.

The local layer is composed of two parallel Convolutional Neural Networks (CNNs). Assuming $C^{(i)}$ denotes the features extracted from the i^{th} CNN, the output of the fusion layer can be mathematically represented as

$$\text{Final distributed representations} = \left(\sum_{i=0}^{n} W_c^{(f)}.C^{(i)} + b_c\right) \quad (1)$$

In this equation, b_c represents the bias and $W_c^{(f)}$ denotes the weights of the fusion layer, with n equal to 2. This formulation underlines the method by which the fusion network combines the feature sets from each CNN, thereby synthesizing a more robust and comprehensive representation of the textual data.

IV. EXPERIMENTAL RESULTS

This section presents a comprehensive evaluation of the Fake Review Detection Approach (FRDA). The FRDA is structured into four distinct stages: (i) The Data Preprocessing Stage (DPS), (ii) The Feature Extraction Stage (FES), (iii) The Fusion Network Stage (FNS), and (iv) The Classification Stage (CS), as depicted in Figure 7.

During FES, we introduce and implement a novel technique termed Hybrid Feature Extraction (HFR). This technique integrates word embedding based on GloVe, POS tagging, and a multilayer CNN to meticulously extract both deep-learned and aspect features from the data. Subsequently, a fusion network is employed to combine these features to bolster the overall performance of the detection model. The final stage, CS, involves the application of various classifier layers designed for the rapid and precise detection of fake reviews, leveraging the fused deep-learned and aspect features.

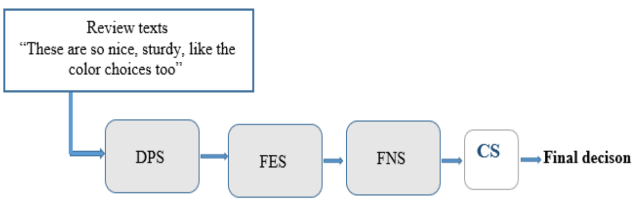


FIGURE 8. Basic stages of the proposed FRDA model.

Our experimental framework comprises two types of evaluations. The first experiment assesses the efficiency of fake review prediction using different layers-based classifiers, considering the number of categories. In the second experiment, the efficacy of the proposed FRDA, particularly its implementation based on Long Short-Term Memory (LSTM), is tested against other contemporary detection models.

FIGURE 9. Accuracy of the proposed FRDA based on several classifier methods considering five categories.

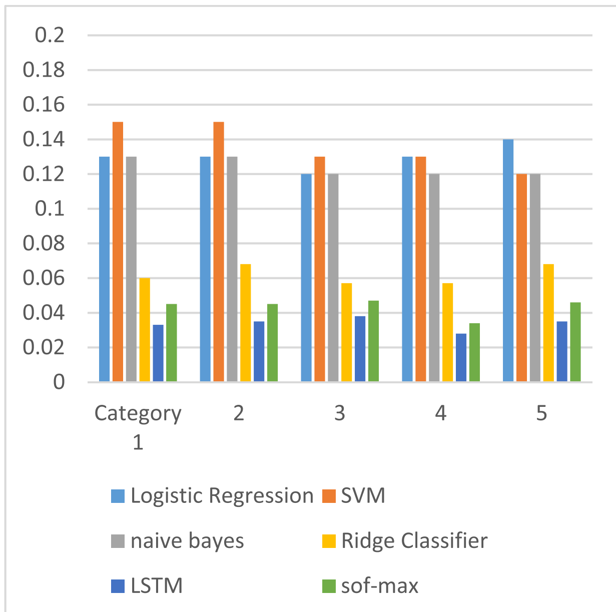


FIGURE 10. Error in the proposed FRDA based on several classifier methods considering five categories.

A. COMPARATIVE ANALYSIS OF CLASSIFIERS USING FRDA

Figures 9-12 exhibit the performance metrics – accuracy, error rate, precision, and recall – of various classifier methods as applied within the Fake Review

Detection Approach (FRDA), considering five distinct categories. These graphical representations elucidate that FRDA when based on Long Short-Term Memory (LSTM), surpasses other comparative methods across all evaluated performance indicators. Tables 3-5 present the empirical outcomes of these

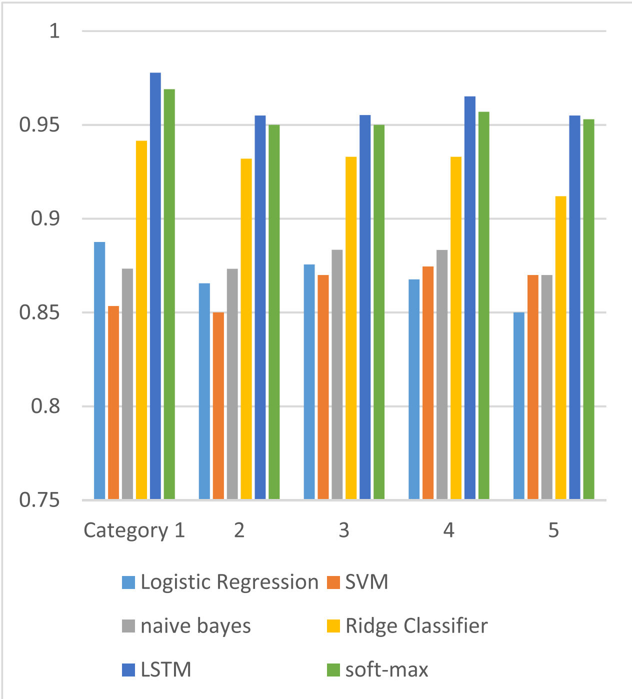


FIGURE 11. Precision of the proposed FRDA based on several classifier methods considering five categories.

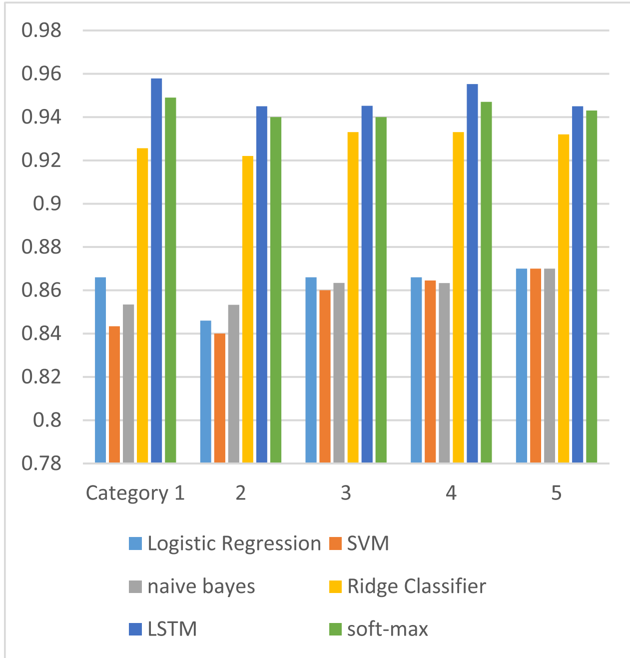


FIGURE 12. Recall of the proposed FRDA based on several classifier methods considering five categories.

TABLE 3. The experimental results of the proposed FDRA for five categories.

Classifier used	Accuracy	Recall	Precision
Logistic Regression	0.8756	0.866	0.8633
SVM	0.8733	0.8645	0.8656
NB	0.876	0.8633	0.8743
RC	0.9356	0.9256	0.9334
LSTM	0.9673	0.9543	0.9652
Soft-max	0.9612	0.943	0.9556

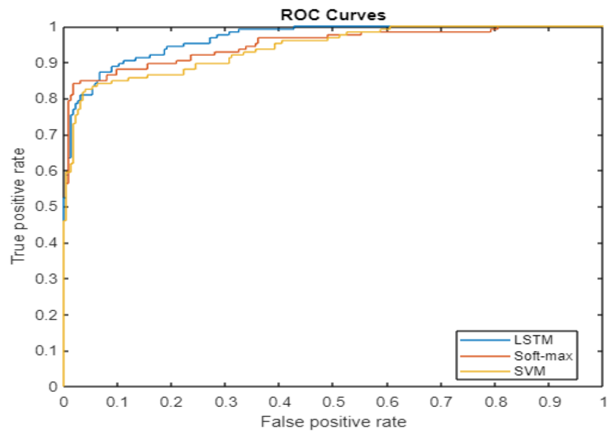


FIGURE 13. ROC plots of the proposed FDRA based on different classifier layers (SVM, Soft-max, and LSTM) for fake review recognition.

TABLE 4. Experimental results of the proposed models for ten categories.

Classifier used	Accuracy	Recall	Precision
Logistic Regression	0.8843	0.866	0.8756
SVM	0.8841	0.8645	0.87
NB	0.8887	0.8733	0.8833
RC	0.9424	0.9356	0.943
LSTM	0.9773	0.9652	0.9752
Soft-max	0.9712	0.9543	0.967

classifiers under the FRDA framework, with a focus on both Category 5 and Category 10. A notable observation is the enhancement in accuracy, recall, and precision concurrent with the increase in the volume of training data. Figure 13 illustrates the Receiver Operating Characteristic (ROC) curve for the proposed methods, providing a visual representation of their classification performance.

B. COMPARATIVE EVALUATION OF THE FRDA BASED ON THE LSTM AND STATE-OF-THE-ART MODELS

In this subsection, a comparative analysis is conducted between the proposed FRDA, specifically its LSTM-based implementation, and several state-of-the-art models. This comparison was conducted across diverse datasets, including Amazon [34], AMT [49], Yelp Chi [10], and the Gold Standard [50], to evaluate the effectiveness of these models in

TABLE 5. Comparison of results between the proposed FDRA based on LSTM and the state-of-the-art models for different datasets.

Dataset	Model	Accuracy	Precision	Recall
Amazon	Logistic Regression [41]	0.8843	0.866	0.8756
	The proposed Model	**0.9773**	**0.9652**	**0.9752**
AMT	SVM [42]	0.912	0.9034	0.8876
	Word order : Preserving CNN [43]	0.9002	0.8956	0.8929
	Recurrent Convolutional Neural Network and Word Context [44]	0.929	0.9138	0.9148
	The proposed Model	**0.94**	**0.9356**	**0.9489**
Yelp chi	Unsupervised Neural Network [45]	0.654	0.6334	0.6402
	The proposed Model	**0.8580**	**0.8313**	**0.8565**
Gold standard	Sentence Weight Neural Network [26]	0.795	0.761	0.898
	The proposed Model	**0.8544**	**0.8335**	**0.8633**

detecting fake reviews across the three most popular corpora. The detection performance of these models is summarized in Table 4. The results demonstrate that the proposed FRDA model, based on the LSTM, exhibits superior performance, achieving an accuracy of 0.9773%. This is in contrast to the accuracies of 0.94%, 0.8580%, and 0.8544% achieved by the models evaluated on the Amazon, AMT, Yelp Chi, and Gold Standard datasets, respectively.

V. DISCUSSION

The following research examined the novel Fake Review Detection Approach (FRDA) that uses LSTM networks for sentiment analysis. The fusion approach allows for more robust detection of false reviews by addressing both the syntactic and semantic elements of the reviews. This dual-layer analysis enhances the detection capability, particularly for sophisticated and subtle fake reviews.

The key discussion is focused on two aspects of sentiment:

1. Performance of Classifiers within FRDA.

2. Comparative Evaluation of LSTM-based FRDA using State-of-the-Art Models

The proposed LSTM-based FRDA was compared in depth with established models across diverse datasets: Amazon, AMT, Yelp Chi, and Gold Standard. It further summarizes the detection performance, indicating that the LSTM-based FRDA demonstrates significantly higher accuracy (0.9773%) than the other models (0.94%, 0.8580%, 0.8544%).

The combination of deep learning features and aspect-based sentiment analysis significantly improved the detection of fake reviews. The FRDA based on the LSTM outperformed all existing state-of-the-art models for fraudulent review detection across different datasets. Other approaches, such as sentence weight networks and recurrent convolutional neural networks, are promising; however, their feature extraction and ability to handle the complexities of natural language are somewhat restricted.

VI. CONCLUSION

In the domain of fake review detection, the integration of deep features and aspect-based analysis has significant potential. This innovative approach harnesses the collective prowess of deep learning and sentiment analysis based on various aspects, leading to the highly accurate identification of fraudulent reviews. When pitted against existing state-of-the-art techniques, this fusion of deep and aspect features has demonstrated exceptional performance in the detection of counterfeit reviews. It is worth noting that alternative methods such as sentence weight networks, recurrent convolutional neural networks, and word context approach, while commendable, are encumbered by inherent limitations stemming from their reliance on predefined features and their constrained ability to grapple with the nuances of natural language.

The success of the FRDA based on LSTM underscores the potential of deep learning models in extracting semantically meaningful features from textual data. This achievement not only exemplifies the effectiveness of our approach but also opens doors to exciting avenues for future research in this field. To further advance the state of fake review detection, several promising directions can be pursued:

- Image analysis and audio sentiment recognition can be incorporated to have a more comprehensive assessment of a review.
- A promising direction in this respect is the exploitation of transfer learning techniques to leverage knowledge from other domains for improved performance.
- The development of techniques to explain the model's classification decisions improves the trust and usability of the model.

In conclusion, the combination of deep and aspect-based features in fake review detection has marked a significant milestone. However, the journey toward more robust and versatile detection methods is far from over. Embracing these future research directions can further strengthen our ability to combat the pervasive issue of fake reviews in an ever-evolving online landscape.

ACKNOWLEDGMENT

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R51), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

REFERENCES

[1] A. Mewada and R. K. Dewang, "Research on false review detection methods: A state-of-the-art review," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7530–7546, Oct. 2022.

[2] D. S. Kostyra, J. Reiner, M. Natter, and D. Klapper, "Decomposing the effects of online customer reviews on brand, price, and product attributes," *Int. J. Res. Marketing*, vol. 33, no. 1, pp. 11–26, Mar. 2016.

[3] S. Ullrich and C. B. Brunner, "Negative online consumer reviews: Effects of different responses," *J. Product Brand Manage.*, vol. 24, no. 1, pp. 66–77, Mar. 2015.

[4] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What yelp fake review filter might be doing?" *Proc. Int. AAAI Conf. Web Social Media*, vol. 7, no. 1, pp. 409–418, Aug. 2021.

[5] S. Rayana and L. Akoglu, "Collective opinion spam detection: Bridging review networks and metadata," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 985–994.

[6] H. Deng, L. Zhao, N. Luo, Y. Liu, G. Guo, X. Wang, Z. Tan, S. Wang, and F. Zhou, "Semi-supervised learning based fake review detection," in *Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl. IEEE Int. Conf. Ubiquitous Comput. Commun. (ISPA/IUCC)*, Dec. 2017, pp. 1278–1280.

[7] X. Hu, J. Tang, H. Gao, and H. Liu, "Social spammer detection with sentiment information," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2014, pp. 180–189.

[8] M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, and H. Al Najada, "Survey of review spam detection using machine learning techniques," *J. Big Data*, vol. 2, no. 1, p. 23, Dec. 2015.

[9] N. Jindal and B. Liu, "Opinion spam and analysis," in *Proc. Int. Conf. Web Search Web Data Mining (WSDM)*, Stanford, CA, USA, 2008, pp. 219–230.

[10] X. Wang, K. Liu, and J. Zhao, "Handling cold-start problem in review spam detection by jointly embedding texts and behaviors," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, vol. 1, 2017, pp. 366–376.

[11] C. Yuan, W. Zhou, Q. Ma, S. Lv, J. Han, and S. Hu, "Learning review representations from user and product level information for spam detection," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2019, pp. 1444–1449.

[12] Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, "Automatic construction of a context-aware sentiment lexicon: An optimization approach," in *Proc. 20th Int. Conf. World Wide Web*, Mar. 2011, pp. 347–356.

[13] D. Zhang, W. Li, B. Niu, and C. Wu, "A deep learning approach for detecting fake reviewers: Exploiting reviewing behavior and textual information," *Decis. Support Syst.*, vol. 166, Mar. 2023, Art. no. 113911.

[14] C. G. Harris, "Detecting fake yelp reviews using a positional LSTM/K-L divergence ensemble approach," in *Proc. 1st Int. Conf. Inf. Syst. Inf. Technol. (ICISIT)*, Jul. 2022, pp. 61–66.

[15] Z. Wang, T. Hou, D. Song, Z. Li, and T. Kong, "Detecting review spammer groups via bipartite graph projection," *Comput. J.*, vol. 59, no. 6, pp. 861–874, Jun. 2016.

[16] Y. Xiao and J. Qiu, "Exploring and detecting opinion spam on social media," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2020, pp. 7570–7573.

[17] M. Fazzolari, F. Buccafurri, G. Lax, and M. Petrocchi, "Experience: Improving opinion spam detection by cumulative relative frequency distribution," *J. Data Inf. Qual.*, vol. 13, no. 1, pp. 1–16, Jan. 2021.

[18] K. Liu, L. Xu, and J. Zhao, "Co-extracting opinion targets and opinion words from online reviews based on the word alignment model," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 3, pp. 636–650, Mar. 2015.

[19] M. Z. Asghar, A. Ullah, S. Ahmad, and A. Khan, "Opinion spam detection framework using hybrid classification scheme," *Soft Comput.*, vol. 24, no. 5, pp. 3475–3498, Jun. 2019.

[20] R. Y. K. Lau, S. Y. Liao, R. C.-W. Kwok, K. Xu, Y. Xia, and Y. Li, "Text mining and probabilistic language modeling for online review spam detection," *ACM Trans. Manage. Inf. Syst.*, vol. 2, no. 4, pp. 1–30, Dec. 2011.

[21] S. Dixit and A. J. Agrawal, "Survey on review spam detection," *Int. J. Comput. Commun. Technol.*, vol. 4, pp. 0975–7449, 2013.

[22] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Human Language Technol.*, 2011, pp. 309–319.

[23] S. Shojaee, M. A. A. Murad, A. B. Azman, N. M. Sharef, and S. Nadali, "Detecting deceptive reviews using lexical and syntactic features," in *Proc. 13th Int. Conf. Intellient Syst. Design Appl.*, Serdang, Malaysia, Dec. 2013, pp. 53–58.

[24] N. Jindal and B. Liu, "Review spam detection," in *Proc. 16th Int. Conf. World Wide Web*, Lyon, France, May 2007, pp. 1189–1190.

[25] R. Mohawesh, S. Xu, M. Springer, Y. Jararweh, M. Al-Hawawreh, and S. Maqsood, "An explainable ensemble of multi-view deep learning model for fake review detection," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 8, Sep. 2023, Art. no. 101644.

[26] L. Xiang, H. You, G. Guo, and Q. Li, "Deep feature fusion for cold-start spam review detection," *J. Supercomput.*, vol. 79, no. 1, pp. 419–434, Jan. 2023.

[27] S. Yu, J. Ren, S. Li, M. Naseriparsa, and F. Xia, "Graph learning for fake review detection," *Frontiers Artif. Intell.*, vol. 5, Jun. 2022, Art. no. 922589.

[28] N. C. S. Ram, G. Vakati, J. V. Nadimpall, Y. Sah, and S. K. Datla, "Fake reviews detection using supervised machine learning," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 5, pp. 3718–3727, May 2022.

[29] J. Wang, H. Kan, F. Meng, Q. Mu, G. Shi, and X. Xiao, "Fake review detection based on multiple feature fusion and rolling collaborative training," *IEEE Access*, vol. 8, pp. 182625–182639, 2020.

[30] S. Banerjee, A. Y. Chua, and J. J. Kim, "Using supervised learning to classify authentic and fake online reviews," in *Proc. 9th Int. Conf. Ubiquitous Inf. Manag. Commun.*, Jan. 2015, pp. 1–7.

[31] W. Etaiwi and G. Naymat, "The impact of applying different preprocessing steps on review spam detection," *Proc. Comput. Sci.*, vol. 113, pp. 273–279, Jan. 2017.

[32] J. K. Rout, S. Singh, S. K. Jena, and S. Bakshi, "Deceptive review detection using labeled and unlabeled data," *Multimedia Tools Appl.*, vol. 76, no. 3, pp. 3187–3211, Feb. 2017.

[33] S. Banerjee, A. Y. K. Chua, and J.-J. Kim, "Using supervised learning to classify authentic and fake online reviews," in *Proc. 9th Int. Conf. Ubiquitous Inf. Manage. Commun.*, Jan. 2015.

[34] Accessed: Dec. 11, 2023. [Online]. Available: https://osf.io/tyue9/files/osfst

[35] J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice-Hall, 2009.

[36] R. Mohawesh, S. Xu, S. N. Tran, R. Ollington, M. Springer, Y. Jararweh, and S. Maqsood, "Fake reviews detection: A survey," *IEEE Access*, vol. 9, pp. 65771–65802, 2021.

[37] G. Bathla, P. Singh, R. K. Singh, E. Cambria, and R. Tiwari, "Intelligent fake reviews detection based on aspect extraction and analysis using deep learning," *Neural Comput. Appl.*, vol. 34, no. 22, pp. 20213–20229, Nov. 2022, doi: 10.1007/S00521-022-07531-8.

[38] F. Anass, R. Jamal, M. A. Mahraz, Y. Ali, and H. Tairi, "Deceptive opinion spam based on deep learning," in *Proc. 4th Int. Conf. Intell. Comput. Data Sci. (ICDS)*, Oct. 2020, pp. 1–5.

[39] M. S. Javed, H. Majeed, H. Mujtaba, and M. O. Beg, "Fake reviews classification using deep learning ensemble of shallow convolutions," *J. Comput. Soc. Sci.*, vol. 4, no. 2, pp. 883–902, 2021.

[40] N. F. Soliman, S. M. Abd-Alhalem, W. El-Shafai, S. E. S. E. Abdulrahman, N. Ismaiel, M. El-S. El-Rabaie, A. D. Algarni, F. Algarni, A. A. Alhussan, and. E. A. El-Samie, "Hybrid approach for taxonomic classification based on deep learning," *Intell. Autom. Soft Comput.*, vol. 32, no. 3, pp. 1881–1891, 1891.

[41] N. F. Soliman, S. M. A. Alhalem, W. El-Shafai, S. E. S. E. Abdulrahman, N. Ismaiel, E.-S. M. El-Rabaie, A. D. Algarni, F. Algarni, and F. E. A. El-Samie, "Bidirectional long short-term memory network for taxonomic classification," *Intell. Autom. Soft Comput.*, vol. 33, no. 1, pp. 103–116, 2022.

[42] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543.

[43] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowl.-Based Syst.*, vol. 108, pp. 42–49, Sep. 2016.

[44] H. Aghakhani, A. Machiry, S. Nilizadeh, C. Kruegel, and G. Vigna, "Detecting deceptive reviews using generative adversarial networks," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2018, pp. 89–95.

[45] F. H. Li, M. Huang, Y. Yang, and X. Zhu, "Learning to identify review spam," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, 2011, pp. 1–6.

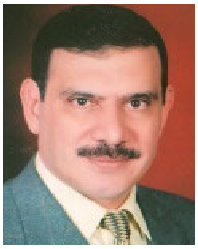
[46] E. F. Cardoso, R. M. Silva, and T. A. Almeida, "Towards automatic filtering of fake reviews," *Neurocomputing*, vol. 309, pp. 106–116, Oct. 2018.

[47] S. Mani, S. Kumari, A. Jain, and P. Kumar, "Spam review detection using ensemble machine learning," in *Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit.* New York, NY, USA: Springer, 2018, pp. 198–209.

[48] S. Zhao, Z. Xu, L. Liu, and M. Guo, "Towards accurate deceptive opinion spam detection based on word order-preserving CNN," 2017, *arXiv:1711.09181*.

[49] W. Zhang, Y. Du, T. Yoshida, and Q. Wang, "DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network," *Inf. Process. Manage.*, vol. 54, no. 4, pp. 576–592, Jul. 2018.
[50] L. Li,W. Ren, B. Qin, and T. Liu, "Learning document representation for deceptive opinion spam detection," in *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*. Nanjing, China: Springer, 2015.

SAMIA M. ABD-ALHALEM received the bachelor's degree in electronics engineering from the Faculty of Engineering, Mansoura University, the master's degree in electronics and communications engineering from the Faculty of Engineering, Zagazig University, and the Ph.D. degree in electronics and electrical communications engineering from the Signal and Image Processing Department, Electronic Engineering Faculty, Minufia University. She is currently an Assistant Professor with the Faculty of Artificial Intelligence, Delta University for Science and Technology. Her research interests include data science, big-data analysis, image processing, and machine learning.



HESHAM ARAFAT ALI received the B.Sc. degree in electrical engineering (electronics) and the M.Sc. and Ph.D. degrees in computer engineering and automatic control from the Faculty of Engineering, Mansoura University, in 1986, 1991, and 1997, respectively. He was the Dean of Artificial Intelligent with Delta University. He has been a Professor in computer engineering and systems, since 2007, and an Associate Professor in information systems and in computer engineering. He was an Assistant Professor with the Faculty of Computer Science, University of Mansoura, from 1997 to 1999. From January 2000 to September 2001, he was joined as a Visiting Professor with the Department of Computer Science, University of Connecticut, Storrs. From 2002 to to 2004, he was the Vice Dean of Student Affair with the Faculty of Computer Science and Information, University of Mansoura. He was awarded with the Highly Commended Award From Emerald Literati Club, in 2002, for his research on network security. Since 2003, he has been an Associate Professor with the Computer Engineering Department, Faculty of Engineering, University of Mansoura. He is a Founder Member of the IEEE SMC Society Technical Committee on Enterprise Information Systems (EIS). He has many book chapters published by international press and about 320 published papers in international (conferences and journals). He (supervised and co-supervised) more than 80 students for (M.Sc. and Ph.D. degrees). He has served as a reviewer for many high quality journals. His research interests include network security, mobile agent, network management, search engine, pattern recognition, distributed databases, and performance analysis.

NAGLAA F. SOLIMAN received the B.Sc., M.Sc., and Ph.D. degrees from the Faculty of Engineering, Zagazig University, Egypt, in 1999, 2004, and 2011, respectively. She has been with the Faculty of Computer Science, PNU, Saudi Arabia, since 2015. She has been a Teaching Staff Member with the Department of Electronics and Communications Engineering, Faculty of Engineering, Zagazig University. Her current research interests include digital image processing, information security, multimedia communications, medical image processing, optical signal processing, big data, and cloud computing.

ABEER D. ALGARNI received the B.Sc. degree (Hons.) in computer science from King Saud University, Riyadh, Saudi Arabia, in 2007, and the M.Sc. and Ph.D. degrees from the School of Engineering and Computer Sciences, Durham University, U.K., in 2010 and 2015, respectively. She has been an Assistant Professor with the College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, since 2008. Her current research interests include networking and communication systems, digital image processing, digital communications, and cyber security.

HANAA SALEM MARIE received the bachelor's degree in electronics engineering and the master's degree in automatic control system from the Faculty of Engineering, Mansoura University, and the Ph.D. degree in artificial intelligence and image processing from the Computer Science and Engineering Department, Electronic Engineering Faculty, Minufia University. She is currently an Associate Professor and the Vice Dean of Education and Student Affairs with the Faculty of Artificial Intelligence, Delta University for Science and Technology. Her research interests include data science, big-data analysis, image processing, and machine learning.