

A Comprehensive Survey of Fake Review Detection Technology

TayybaQuyyam^{1,1}, Dr.QichengYU²,

¹London Metropolitan University, Digital Media and Computing, Holloway Road London N7 8DB, United Kingdom

Email: TDQ0018@my.londonmet.ac.uk

²London Metropolitan University, Digital Media and Computing, Holloway Road London N7 8DB, United Kingdom

Email:q.yu,@londonmet.ac.uk

Abstract. Fake reviews (FR) can damage a company's reputation and cause consumers to purchase low-value products and services. With the advancement of artificial intelligence, FRD technology and its detection accuracy have improved significantly. To seek state-of-the-art FRD technology, this paper will conduct a systematic literature survey to explore the solutions of scholars working on methods to effectively detect FR, unsolved problems in this field, and future research directions. The survey covered 30 recent research papers between 2019 and 2023. Our findings are categorized as machine learning, deep learning, and hybrid methods to provide researchers and experts with an outlook on proposed solutions and their limitations. The survey also offers future research directions and an easy way to find datasets, carry out pre-processing, and extract multiple features. One of the main directions in the future is to combine review content with business, product, and reviewer behavior to improve the efficiency of FRD.

Keywords: Fake Review Detection, Sentiment Analysis, Natural Language Processing, Machine Learning, Deep Learning, and Hybrid Methods.

1. Introduction

In recent years, E-commerce platforms have been growing exponentially. Some businesses allow the public to leave their comments, sentiments, and attitudes, which support future customers who want to acquire the same products & services that are already verified and approved as worth purchasing. FRD is critical in the world of E-commerce platforms because fake reviews are a big cause of customers being misled about the specific product. These FRs are the main source of information for investigating uncertainty about the value of goods, mainly in situations where the customer has not seen the goods physically before. FR mislead the clients against the product's value and trust in online shopping is also hampered. In other words, when purchasers view the reviews on an e-commerce website, they as customers, decide whether to purchase the items or turn down their buying plan. FR can have big implications and consequences across several sectors of e-commerce. Fake reviews erode the customer's trust in the products as well as influence the business's reputation. Fake reviews also pose additional problems in that they mislead consumers into spending money with a business or on a product they may not have otherwise chosen.

Many studies have investigated the FRD problem and provided possible solutions. In the early internet era, online review platforms started gaining popularity in the late 1990s and early 2000s when the issue of FR began to appear on the surface. However, at this stage, detection methods were relatively basic. Throughout the period spreading from the mid-2000s to the early 2010s, online businesses had the first automatic systems including algorithms among the review platforms. They analyzed elements such as word patterns, language, timing and other such things to detect suspicious reviews. Vendors in the late 2010s through nowadays have done justice to sentence filtering mechanisms with machine learning (ML), artificial intelligence (AI), and more enhanced algorithms to trademark and eliminate fraudulent reviews. They also use moderators to assist the software systems. Furthermore, consumer advocacy organizations and researchers have been assessing the amount of FR in

the market through studies of this type. Those engaging in FR practices also continuously evolving their tactics to bypass detection systems. As a result, platforms engage in a continual race to improve their detection methods to stay away from fraudulent activities. Some cyber's jurisdictions have introduced laws against FR, which can lead to penalties for businesses involved in fraudulent practices. The FRDs are an ongoing process that will become more sophisticated with the passage of time.

The previous surveys [1][2] [3] provide a summary and analysis of FR, Machine Learning, and Deep Learning Methods from 2007 to 2021. Rami Mohawesh et al [2] provide an extensive survey of FRDs by analyzing and summarizing various approaches and detailing the tasks associated with FRD. These surveys provide valuable guidelines for new researchers in their relevant fields. However, they only covered developments up to 2001, while the field of FRDs continued to grow rapidly with various new methods and techniques were being proposed. One survey-reported work's claims, the FRD accuracy rate already reached 99%. However, according to BBC News reports in October 2023, FRDs continue to damage the reputation of products and services, costing consumers in the United Kingdom approximately 312 million pounds last year. Therefore, it is important to carry on addressing the challenges proposed by FRDs.

Review platforms continuously refine their strategies to combat FR and maintain the credibility of their platforms among consumers. FRs affect the business's reputation and its products & services too. This survey is important for making product development better, gaining user's trust, and improve customer's experience with the products. This survey covers the following matrixes of FRD technologies: Sentiment analysis, Natural Language Processing, Machine Learning, Deep Learning, and Hybrid Methods. They provide awareness for the future direction to researchers for those involved in FRDs.

The consumers' opinions are obtained from the e-commerce platforms and other sources that aim to increase the traffic on the e-commerce sites. However, user trust may often be distracted as there is growing evidence that FR misguides customers' product choices. The different tools and Deep Learning and Machine Learning methods used for the FRDs. This review study is organized as follows Section 2: Methodology of Fake Reviews Survey. Section 3: Literature Reviews. Section 4: Discusses and Limitations. In section 5: Conclusion.

2. Survey of Fake Review Methodology

This paper will conduct a systematic literature survey to explore the solutions of professionals working on methods to effectively detect FR, unsolved problems in this field, and future study directions. Many researchers provide effective methodologies and Khalid Saeed Khan et al. [4] is one of the best in this area of FRD. This section presents the following condensed Survey of Fake Review Methodology.

2.1 Survey Protocol Design

In this survey, we used several Fake reviews detection to search related articles. In this study, we follow the five steps for designing the methodology. The survey of fake review methodology is organized as follows step:

- Listed the Research Question.
- Inclusion Exclusion Criteria. Step
- How many papers are included?
- Summarize the evidence.
- Interpret the findings of the work.

2.1.1 Research Questions

RQ1:How FRD technology has been advanced?
RQ2:How pre-processing of the text reviews impact the FRD results?
RQ3:How data set used for FRD impact its effectiveness?
RQ4:How the different machine learning techniques impact the accuracy.
RQ5:What are the gaps for detecting fake reviews research?

2.1.2 Inclusion exclusion criteria

- Included studies in the range from 2019 to 2023.
- Research Studies Detect Fake Reviews.
- The research study excluded all duplicate articles.
- The research papers are not linked to the research questions.

2.1.3 How many papers are encompassed of FRs

In this comprehensive survey of Fake Reviews detection, approximately 30 research articles have been included between 2019 to 2023 after reviewing more than 100 FRD research articles. Fig (1) of the pie chart shows the class distribution of the research article. The blue class column of the pie chart shows 13 journal IEEE, the orange class column shows 8 Springer journal and the gray class column shows 6 of the other nature of the journal.

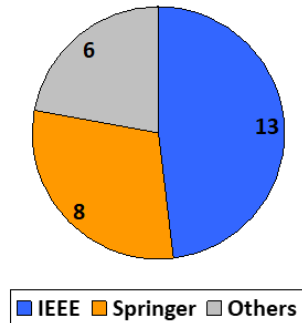


Figure 1 Selected Papers

2.1.4 Summarizing Evidence

This survey covers the Matrix: Methods, Background, Discussion, Strengths and Weaknesses, and Identify the Gaps. This survey results are categorized according to Machine Learning, Deep Learning, and Hybrid methods to provide researchers and experts with an outlook on proposed solutions and their limitations. Each Paper is based on the Fake Reviews Detection Matrix which includes aim and objectives, pre-processing, methods, methodology, research gap, background, datasets, results and evaluation.

2.1.5 Interpreting the findings of the study

This survey includes the research results of the high-accuracy ML, DL, and Hybrid. Table 1 shows the Machine-Learning high results of FRD. Table 2 shows that Deep Learning provides a high accuracy of FRD. Table 3 shows that Hybrid Methods provided a high accuracy of FRD. This survey provides the ranking criteria in Table 4. Table 4 includes the

following Matrix: Descriptions/Methodology, Pre-processing/Features, Datasets and Gaps. In simple preprocessing, I gave the ranking (1) Tokenization, Lowercase, Stemming, punctuation, stop word, and POS Tagging. For multiple preprocessing, I give the ranking complexity of the Preprocessing (2) Simple and Multiple (3). I also gave the Ranking of the datasets. For one dataset use on one source, I give the (1) Two datasets use Two different sources (2) and 3 datasets use (3) More than uses of resources (4). In the features Simple feature Text, Temporal(1), Multiple features like Behavior, Product, linguistic (2), Hybrid features like categorical, Text, Sentiment Temporal features(3).

3. Literature Review

In the past literature study, several Deep Learning, Machine Learning, and Hybrid Methods have been used for E-commerce customers' fake Reviews detection and other domains. Each method follows the matrix of flow. In this chapter, we will review the specific research work of FRD.

3.1 Machine Learning

Afraz Moqueem et al [5][6][7] priorly this studied to helps to address the FRD online shopping platforms and also proposed several ML techniques. These techniques analyze the textual data and classify whether it is genuine or fake and Create a strong system for detecting phoney reviews. Hence, they enhanced the reliability of online products on marketplaces. They investigated preprocessing, ML classifiers, addressed imbalanced data, analyzed results, and compared performances. Faranak Abri et al [8] [9][10] [13] their study aimed to design and assess an (NLP) method for detecting FRD on online platforms. It classified the linguistic features among fake and trustful reviews. Additionally, the study's objective was to analyze 15 linguistic features and several ML methods. One of these methods was to detect FRD based on leveraging multiple features and rolling collaborative methods. The method of categorization used in this work by Ander Sihombing and A .C .M Fong[14][12] [11]. They implemented ML learning methods on the Yelp dataset for the FRD. In this work, the main goal was to detect and filter out FRD to enhance the reliability of the online platform. However, it did not clearly define the matrix evaluation for ML performance. One of the aims of the study develop the ML framework for the FRD on an online platform. The objective of the study is to apply feature engineering evaluation and consider review behavior. The objectives of the study review group concept, collusion relation, and evaluation.

3.1.1 Discussion Table 1. Machine Learning and Other Methods of Accuracy.

Table 1, it shows the popular six Machine learning Methods (NB, SVM, LR, RF, DT, K-NN) that provide accuracies. However, these accuracies are based on authors' results which vary according to their preprocess methods and testing data. Hence, it is not possible to determine which ML algorithm is better suited for FRD simply based on the accuracy figure. To find the answer, we use a democratic voting approach in which each author will have an equal vote to determine the most popular and accurate ML algorithm since every author has compared multiple ML algorithms.

Based on survey details shown in Table, since the SVM algorithm has been used by [5] [6] [7] [9] [8] [10] [11] [12] It is the most popular ML algorithm from the survey and NB is in second place. DT and K-NN are the least popular algorithms.

Since SVM has the highest accuracy in [9] [10] [11] results, it won the most accurate ML algorithm from the survey. NB, RF, and, LR all received two highest accuracy from authors,

but LR has two second higher accuracy from [9] and [11]. Therefore, LR won the second place. The method does not have a very scientific approach, but it could give us useful insight. For more scientific results, we will need to compare all algorithms in the same benchmark environment.

Table 1. Machine learning and Other Methods of accuracy.

<i>Author</i>	<i>NB</i>	<i>SVM</i>	<i>LR</i>	<i>RF</i>	<i>DT</i>	<i>K-NN</i>
[5]	79	78		75		
[6]	84	76				
[7]		89	89	87		
[8]	73	77	73	77	76	
[9]		94	85			85
[10]		84			81	
[11]	84	88	86	83	73	57
[12]	65	77	78			
[13]	95			97		86
[14]	91	88			92	

Naïve Bayes (NB) Sport Vector Machine (SVM) Linear Regression (LR) Random Forest (RF) Decision Tree (DT) K-nearest neighbors.

3.2 Deep Learning

N. Deshail and B. Bhaskara Rao [15][16][17] the study aimed to address the FRD online platform by leveraging CNN and the objective of the study was to implement the novel adoptive-based CNN technique. The main purpose of these study to detect the online FRD. Pratiksha Shetgaonkar et al [18][19][20] the study introduced for FRD aimed to analyze the inconsistencies in the Sentiment Analysis of Rating. The other objective of the study is sentiment analysis, preprocessing, Rating Sentiment, feature extraction, and DL model implementation which was used to detect the FRD. According to Shahbaz Ashraf et al [21][22] the study distinguished the online FRD, as the major impact of customer decision-making. The focus of goals was to create an ensemble model, bag-of N-gram, Leverage N-gram embedding, and consider textual data. the Aim and main objective of the study detect FR online e-commerce platform by using CNN model and Detect the FR online Platform using the model of RHGNN. On the other hand, detect the online FRD on online plate form and the main objective of this work multi-view DL methods were implemented. Jun Zhao et al [23] the Purpose of the study was to detect FR online Platforms using the model of RHGNN.

3.2.1 Discussion Table 2. Deep Learning and Other Methods

Table 2 shows the popular eight Deep learning Methods (CNN, RNN, DNN, LSTM, ANN, Bi-LSTM, RHGNN, BERT) provided the accuracies. However, these accuracies are based on authors' results which vary according to their preprocess methods and testing data. Hence, it was not possible to determine which DL algorithm is better suited for FRD simply based on the accuracy figure. To find the answer, we use a democratic voting approach.

Based on survey details shown in Table 2., Since the CNN algorithm has been using five References such as [15][18][22][16][20]. It is the most popular DL algorithm from the survey. CNN algorithm listed high accuracy [15][16][22] used by authors in the first place.

The LSTM algorithm has been using these five References [15] [17] [18] [19] [21]. It is the most popular DL algorithm from the survey. LSTM algorithm used high accuracy by [17] [18] are abundantly used by authors in the second place. The LSTM provided the High accuracy of [17].

Table 2. Deep Learning and Other Methods of Accuracy.

<i>Author</i>	<i>CNN</i>	<i>RNN</i>	<i>DNN</i>	<i>LSTM</i>	<i>ANN</i>	<i>Bi-LSTM</i>	<i>RHGNN</i>	<i>BERT</i>
[15]	99	98		96	97			
[17]				89				
[18]	81			84				
[19]				91		91		
[22]	85							
[21]		75		67				84
[16]	97							
[23]							85	
[20]	72		81			69		

Convolutional neural network (CNN), Recurrent Neural Network (RNN), Social media Optimization (SMO), Long Short-term Memory (LSTM), Artificial Neural Network(ANN), Deep feed-Forward Neural Network (DFFNN), Bidirectional.

3.3 Hybrid Methods

Gourav Bathla et al [24] predicted FRD on social networks and online platforms. Therefore, the main purpose was to implement the deep learning LSTM methods for the FRD. Xinyue Wang et al [25] the main goal of the study was to create innovative features and machine learning methods were also implemented for FRD. Bo Guo and Zhi-bin Jiang [26] the study aimed to develop a novel framework for FRD on the e-commerce platform. The work introduced by Ge Wang et al [27] to identify FRD on an electronic platform. Priyanka Gupta et al [29] [28] in this study proposed the framework for the FRD online e-commerce platform while using the BERT model. [30] [31] the study aimed to help the customers on the e-commerce platforms. The objective of the study was ML and deep learning methods were implemented.

3.3.1 Discussion Table 3: Hybrid Methods

Table 3 shows that the seven Hybrid Methods (SVM, NB, RF, CNN, LR, K-NN, BERT) provide accuracies. However, these accuracies are based on authors' results which vary according to their preprocess methods and testing data. Hence, it is notable to determine which Hybrid algorithm is better suited for FRD which is simply based on the accuracy figure. To find the answer, we use a democratic voting approach in which each author will have an equal vote to determine the most popular and accurate Hybrid algorithm since every author has compared multiple DL algorithms.

Table 3., Since the SVM algorithm listed by [25] [26] [27] [28] [29] [30] [31]. It is the most popular Hybrid algorithm from the survey. The SVM algorithm used by [25] [27] [29] [31] authors stand in the first place. The RF algorithm has been used by [26] [30] [31] RF

algorithm listed by [26] [30] authors are in second place. The BERT algorithm has been used by [28] [29] and also abundantly used by authors an in the second place.

Table 3. Hybrid Methods of Accuracy.

<i>Author</i>	<i>SVM</i>	<i>NB</i>	<i>RF</i>	<i>CNN</i>	<i>LR</i>	<i>K-NN</i>	<i>BERT</i>
[24]		81		92			
[25]	97	72			97	96	
[26]	77	81	87		84		
[27]	79	72					
[28]	83	90					97
[29]	67						67
[30]	94	88	95				
[31]	94	91	94		94		

Sport Vector Machine (SVM), Nive Bayes (NB), Deep feed Forward Neural Network (DFFNN), Random Forest (RF), Convolutional Neural Network (CNN) Linear Regression (LR), Artificial Neural Network (ANN), Bidirectional Encoder Representations from Transformers(BERF).

4.1 Discussion about Table 4: Machine learning, Deep Learning, Hybrid Preprocessing, Datasets and Labels.

[6][7] [9]the Kaggle and Yelp datasets were used with predicted labels of genuine and fake. Preprocessing like Stemming, Stop Word, and Post Tagging were implemented. The Features such as length, language, and emotive expression also used in this study. [10][32][22]preprocessing techniques were implemented such as text cleaning and handling missing data. Multiple features were also applied like text, sentiment, behavior features, and fusion features.[12] [13] 8] [21] the Restaurant dataset with predicted other label N/Y. The preprocessing and multiple features also applied complex feature review length, etc. [14] [15] [17] the Amazon predicted dataset with other labels fake or not, genuine and fake. The preprocessing and multiple features also applied text, graph, language, collusion /textual features like N-gram and review feature.[16] [23] [18]the Amazon dataset with the label of T- Positive, F-Positive, F-Negative, and T- Negative. Preprocessing and multiple features were implemented on the product and review.[20] [24] [25] [29] the primary features used for FRD such as user-based temporal metadata and semantic and behavioral features.[26] [27] [5]the Amazon dataset was used with fake or genuine, other labels and the preprocessing techniques, other multiple behaviors, statical, Product, opinion, and contextual features were implemented. In this work, traditional features were used and word frequency.[30] [31] [19] the hotel fake review dataset fake or genuine predicted labels were implemented. Preprocessing techniques were implemented such as text cleaning and handling missing data. Multiple features were applied such as semantic and Behavioral features.

Table4. Summary of Machine Learning, DeepLearning, and hybrid

<i>Author</i>	<i>Descriptions/ Methodology</i>	<i>Processing</i>	<i>Features</i>	<i>Data sets</i>	<i>Gaps</i>
[5]	Detect the FR on an online shopping Platform based on ML techniques.	1	1	1	Combined additional features and the Creation of real-time detection capabilities to detect FRD.This model could optimize the speed and

					scalability to handle the FR and enhance the performance of ML for FRD on E-commerce.
[6]	FRD based on NiveBayes Classifier.	1	3	2	The NB does not provide a comprehensive evaluation. ML or ensemble methods may provide high results.
[7]	Resampling imbalanced data for FRD based on ML text features.	2	1	4	Investigate imbalanced datasets on behavioral Features improvements with novel ensemble method.
[8]	Linguistic features for FRD.	1	2	3	In future work, identify linguistic features that can be used to detect phishing attacks.
[9]	FRD involves using text Classification based on KL methods.	2	2	2	We can solve complex real-world problems using this PU Learning approach that will give accurate and well-performing frameworks.
[10]	Multiple Feature fusion and rolling collaboration for FRD.	2	1	1	Improve the accuracy of FRD information across various domains of social networks.
[32]	FRD Based on ML Methods.	2	2	2	Scalability, adversarial attacks, and ML-based solutions for real-world scenarios are potential areas for future work although they may not be addressed within the current scope of this work.
[12]	FRDYelp data set classification techniques of ML.	1	2	1	In the future work can select and identify the features that can enhance the performance of MLforFRD.
[13]	FR detection based on ML Models.	1	2	1	Future work can be the add low-level behavior Features that can be detected of user profile time and date.
[14]	Fake review Detection using group model.	2	1	2	In the future can work Over time, review groups are formed, changed, and influenced by the shared characteristics of their members, such as interests, connections, and behavior.
[15]	CNN and daptive PSO approach for FRD.	2	3	4	These findings do not talk about ethical factors or strategies for mitigating biases in FRD. Future research could explore ethical frameworks and bias mitigation methods.
[17]	FRD with Hybrid Features Using Time-Sequential based on DL.	2	4	4	In this study utilize the text and behavior feature for classification and in the future, we can utilize other features like metadata, sentiment analysis, and linguistic features.
[18]	FRD using sentiment inconsistency.	1	2	1	In future work, combine Ratting as a feature embedding over the embedding features then enhance the effectiveness of the spam detection.
[19]	FRD using sentiment analysis of DL.	2	3	1	In this study use of the Sentiment analysis and NLP for FRD. In future work, advanced NLP techniques will be utilized such as topic modeling, and word embedding for the enhancement of accuracy for FRD.

[22]	FRD classification using DL methods.	2	3	2	future research includes exploring the use of attention mechanisms, graph neural networks, or transformer-based architectures to capture more nuanced patterns and relationships in review.
[21]	Property and political data based on DL methods.	2	4	2	This study utilizes the multiple modalities data and It could be explored in the future for the fusion of multiple modalities and leveraging complementary from unique modalities.
[16]	Sentiment analysis of Amazon FRD with CNN.	2	4	1	Future work can extend our studies by benchmarking different detection models based on CNNs on different datasets from other dominating e-commerce platforms and domains.
[23]	FRD-based reinforced heterogeneous graph NN.	3	1	1	It can be implemented adynamic graph that evolves FRD.
[20]	Ensemble learning techniques Compared with conventional ML	2	4	1	In the future, heuristics hyperparameters and analyze interactions for time series
[24]	Intelligent system FRD using DL.	3	4	1	In the future, stability optimizes the aspect replication based on CNN to enhance the quality.
[25]	FRD semantics and behavior features	3	2	1	In the future can work on advanced ML methods.
[26]	A practical feature-based large-scale.	2	3	2	In future work, behavior patterns are used by paid posters.
[27]	Multidimensional Feature Engineering	2	3	4	In the future, stability Algorithms on several datasets.
[28]	Online FRD based on Bert model.	2	4	1	In the future work can be focused on supervised, semi-supervised, and transformed learning handle FRD.
[29]	NN base Leveraging Transfer BERT.	2	1	1	In future work add a Bi-LSTM layer instead of a dense layer.
[30]	FRDusingthe supervised learning.	3	3	4	In future work, a can be built large dataset covering the e-commerce platform.
[31]	FRD hybrid utilization of fuzzy-based NN.	2	4	4	In the future can work on the state of art for FRD for the benefit of the customer.

Simple Processing (1), Multiple Preprocessing (2) simple and multiple (3) Features, (1) Multiple Features,(2) Hybrid feature (3) Other (4). One dataset 1 source (1) Two dataset2 sources (2) Three datasets 3 sources(3) and more than 3 datasets multiple sources (4).

4. Discussion and Limitations

4.1 Discussion

Recent years have observed significant advancements in FRD technology, marked by the development of sophisticated ML algorithms and Natural Language Processing NLP techniques. Overall preprocessing techniques were applied to the text quality and the

effectiveness of FRD results. A proper preprocessing method can help clean the data and extract features from the text information, leading to a more accurate FRD. Generally, the largest data allow the model to capture the different linguistic and behavioral features associated with FR, leading to good detection results. Consequently, the literature review has focused extensively on ML, DL, and Hybrid techniques for FRD, with popular ML methods like RF, SVM, NB, and various neural network architectures, transfer learning and the pre-train technique to be fine-tuned significantly impacting FRD accuracies. Therefore, SVM, RF, K-NN, and LDA, can be considered the best algorithms for FRD. Deep learning provides the most popular methods CNN and RNN for FRD. Hybrid provides the popular methods NB and SVM for FRD. They provided the SVM with high accuracy. Despite the exhaustive review of 100 articles on FRD, it is important to acknowledge the limitation of space constraints, which precludes the inclusion of all discussions from these papers in this study. Nevertheless, this study endeavors to highlight key insights and trends gleaned from the literature, offering a comprehensive overview of the landscape of FRD methodologies.

4.2 Research Gaps

The use of ML, DL, and Hybrid techniques will help the FRD in the e-commerce platform. Some gaps are also missing in this literature study, the author says that this performance of the classifier is estimated, implemented to support network-based streaming, and implemented Graph neural network. Other possibilities for future work, the author says can be to optimize the algorithm and resources to minimize the computational cost when detecting the accuracy. Additionally, leveraging ML approaches to enhance the accuracy of FRD frameworks on e-commerce platforms. Moreover, exploring different models such as Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM), can significantly improve accuracy in detecting FRD. There are several gaps in FRD Research, therefore, handles missing data, imbalanced data, and unstructured data and enhances the model's interpretability. The quality and quantity such as a very limited dataset can hinder the efficiency of FRD. Thus, one of the main future directions is that accounting for further behavior boosts the efficiency of this fake review detection method by this performance of the estimated classifier.

5. Conclusion

E-commerce platforms have been growing exponentially and most of them allow the public to support their customers by leaving reviews as reviews can damage a company's reputation and cause consumers to purchase low-value products and services. On the contrary, it is difficult to detect them due to the features of fake reviews. This article particularly emphasized on giving a summary of the several approaches that have been used in the e-commerce platforms to identify fake reviews through Machine Learning, Deep Learning, and Hybrid. For this purpose, a systematic literature survey was performed to classify the different methodologies that were used in the fake reviews e-commerce platforms. This literature review helps to identify fake reviews in the previous five years. Therefore, the findings of this literature study are greatly beneficial to our finding limited set of features and contextual understanding, and the unsupervised learning method requires innovative methods to identify patterns without labeled datasets. Furthermore, existing Literature surveys generally consider machine learning, Deep learning, and Hybrid methods which provide the best methods SVM emerged as a popular choice, offering a sharp accuracy of SVM, LR, RF, CNN, RNN, and BERT the sharp accuracy of FRD. Accordingly, we concluded that fake reviews detection is a complex procedure. This research study provides the research direction of the researcher. However, further research study is needed to directly explore how to do FRD accurately.

REFERENCES

- [1] R. A. Duma, Z. Niu, A. S. Nyamawe, J. Tchaye-Kondi, and A. A. Yusuf, "A Deep Hybrid Model for fake review detection by jointly leveraging review text, overall ratings, and aspect ratings," *Soft Comput.*, vol. 27, no. 10, pp. 6281–6296, 2023.
- [2] R. Mohawesh *et al.*, "Fake Reviews Detection: A Survey," *IEEE Access*, vol. 9, pp. 65771–65802, 2021.
- [3] N. A. Patel and R. Patel, "A survey on fake review detection using machine learning techniques," *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, pp. 1–6, 2018.
- [4] K. S. Khan, R. Kunz, J. Kleijnen, and G. Antes, "Five steps to conducting a systematic review," *J. R. Soc. Med.*, vol. 96, no. 3, pp. 118–121, 2003.
- [5] A. Moqueem, F. Moqueem, C. V. Reddy, and D. Jayanth, *Online Shopping Fake Reviews Detection Using Machine Learning*, vol. 4. Springer Nature Switzerland, 2023.
- [6] P. Kalaivani, V. D. Raj, R. Madhavan, and A. P. N. Kumar, "Fake Review Detection using Naive Bayesian Classifier," *2023 Int. Conf. Sustain. Comput. Smart Syst.*, no. Icsccs, pp. 705–709, 2023.
- [7] G. S. Budhi, R. Chiong, and Z. Wang, "Resampling imbalanced data to detect fake reviews using machine learning classifiers and textual-based features," pp. 13079–13097, 2021.
- [8] F. Abri, L. F. Gutierrez, A. S. Namin, K. S. Jones, and D. R. W. Sears, "Linguistic Features for Detecting Fake Reviews 1," pp. 352–359, 2020.
- [9] H. M. Aljahdali, "The Effect of Fake Reviews on e-Commerce During and After Covid-19 Pandemic: SKL-Based Fake Reviews Detection," *IEEE Access*, vol. 10, pp. 25555–25564, 2022.
- [10] 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.
- [11] P. M. Kumar, S. S. Harsha, K. Abhiram, M. Kavitha, and M. Kalyani, "Role of Machine Learning in Fake Review Detection," *6th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2022 - Proc.*, no. Iceca, pp. 1026–1030, 2022.
- [12] A. F. A. Sihombing, "Fake review detection on Yelp dataset using classification techniques in machine learning," *International Conference on Contemporary Computing and Informatics (IC3I)*, 2019.
- [13] M. M. Hassan Sohan, M. M. Khan, I. Nanda, and R. Dey, "Fake Product Review Detection Using Machine Learning," *2022 IEEE World AI IoT Congr. AllIoT 2022*, pp. 527–532, 2022.
- [14] Y. Li, F. Wang, S. Zhang, and X. Niu, "Detection of Fake Reviews Using Group Model," *Mob. Networks Appl.*, vol. 26, no. 1, pp. 91–103, 2021.
- [15] N. D. B. B. Rao, "Unmasking deception: a CNN and adaptive PSO approach to detecting fake online reviews," *Soft Comput.*, vol. 0123456789, pp. 11357–11378, 2023.
- [16] C. Sasikala, S. Ramya, S. Rajesh Kumar, K. B. Ramachandramoorthy, S. Vijayakumar, and K. Umapathi, "Fake Review Detection and Classification Using Improved Convolutional Neural Network on Amazon Dataset," *Proc. - 2023 3rd Int. Conf. Pervasive Comput. Soc. Networking, ICPCSN 2023*, pp. 398–403, 2023.
- [17] Y. L. Weng, S. Y. Chien, and S. Y. Lin, "Fake Reviews Detection with Hybrid Features Using Time-Sequential Deep Learning Model," *Proc. 2022 IEEE Int. Conf. Human-Machine Syst. ICHMS 2022*, pp. 1–3, 2022.
- [18] K. Sharifpour and S. Lahmiri, "Review Detection Using Fake Inconsistency," *2023 Int. Conf. Mach. Learn. Appl.*, pp. 926–931, 2023.
- [19] P. Shetgaonkar, J. T. Rodrigues, S. Aswale, V. L. K. Gonsalves, J. C. Rodrigues, and A. Naik, "Fake Review Detection Using Sentiment Analysis and Deep Learning," *Proc. Int. Conf. Technol. Adv. Innov. ICTAI 2021*, pp. 140–145, 2021.
- [20] R. Mohawesh, S. Xu, M. Springer, Y. Jararweh, and M. Al-hawawreh, "An explainable ensemble of multi-view deep learning model for fake review detection," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 8, p. 101644, 2023.
- [21] A. V. S and G. R. Ramya, "ScienceDirect Identifying Fake Reviews in Relation with Property and Political Data Using Deep Learning Identifying Fake Reviews in Relation with Property and Political Data Using Deep Learning," *Procedia Comput. Sci.*, vol. 218, pp. 1742–1751, 2023.
- [22] S. Ashraf, F. Rehman, H. Sharif, H. Kim, H. Arshad, and H. Manzoor, "Fake Reviews Classification using Deep Learning," *2023 Int. Multi-Disciplinary Conf. Emerg. Res. Trends, IMCERT 2023*, 2023.

- [23] J. Zhao, M. Shao, H. Tang, J. Liu, L. Du, and H. Wang, "Knowledge-Based Systems RHGNN : Fake reviewer detection based on reinforced heterogeneous graph neural networks," *Knowledge-Based Syst.*, vol. 280, no. June, p. 111029, 2023.
- [24] 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, 2022.
- [25] X. Wang, X. Zhang, C. Jiang, and H. Liu, "Identification of fake reviews using semantic and behavioral features," *2018 4th Int. Conf. Inf. Manag. ICIM 2018*, pp. 92–97, 2018.
- [26] B. Guo and Z. Jiang, "What Is the Internet Water Army ? A Practical Feature-Based Detection of Large-Scale Fake Reviews," vol. 2023, no. ii, 2023.
- [27] G. Wang, G. Shang, P. Pu, X. Li, and H. Peng, "Fake Review Identification Methods Based on Multidimensional Feature Engineering," *Mob. Inf. Syst.*, vol. 2022, 2022.
- [28] C. Babi, M. S. Roshini, P. Manoj, and K. S. Kumar, "Fake Online Reviews Detection and Analysis Using Bert Model," *J. Surv. Fish. Sci.*, vol. 10, no. 2S, pp. 2748–2756, 2023.
- [29] P. Gupta, S. Gandhi, and B. R. Chakravarthi, "Leveraging Transfer learning techniques- BERT , RoBERTa , ALBERT and DistilBERT for Fake Review Detection," pp. 75–82, 2021.
- [30] S. N. Alsubari *et al.*, "Data analytics for the identification of fake reviews using supervised learning," *Comput. Mater. Contin.*, vol. 70, no. 2, pp. 3189–3204, 2022.
- [31] S. U. Maheswari, "Detection of Fake and Genuine Reviews with Hybridization of Fuzzy and Neural Networks Techniques Detection of Fake and Genuine Reviews with Hybridization of Fuzzy and Neural Networks Techniques I . Introduction," no. June, 2021.
- [32] P. M. Kumar, S. S. Harrsha, K. Abhiram, M. Kavitha, and M. Kalyani, "Role of Machine Learning in Fake Review Detection," *2022 6th Int. Conf. Electron. Commun. Aerosp. Technol.*, no. Iceca, pp. 1026–1030, 2022.