EAI Endorsed Transactions
on AI and Robotics
Research Article EAI.EU

# Cutting-Edge Techniques for Detecting Fake Reviews

Kuldeep Vayadande[1*], Amit Mishra[2], Gajanan R. Patil[3], Yogesh Bodhe[4], Pavitha Nooji[5], Ninad Kale[1], Anish Katariya[1], Amey Kharade[1], Parth Supekar[1], Lalit Patil[1]

[1]Vishwakarma Institute of Technology, Pune, India
[2]Dr. Vishwanath Karad MIT World Peace University, Pune, India
[3]Army Institute of Technology, Pune, India
[4]Government Polytechnic, Pune, India
[5]Vishwakarma University, Pune, India

## Abstract

The paper reviews various approaches for detecting fake reviews using different machine learning techniques, each with distinct strengths and limitations. It examines existing literature on supervised learning methods, unsupervised techniques, graph-based models, and hybrid approaches. Among these, unsupervised models rely on pattern recognition, while supervised methods, including SVM and transformer-based models like BERT, offer high accuracy but struggle with class imbalance and computational efficiency. Unsupervised and graph-based models serve as effective alternatives when labeled data is scarce or when complex relationships between reviews and users must be analyzed. Additionally, hybrid approaches that integrate multiple techniques are gaining traction, as they enhance feature selection and model performance. In this paper, we explore different methodologies for fake review classification, analyze their advantages and drawbacks, and highlight key challenges in the field.

**Keywords:** Review Classification Techniques, Deep Learning, Hybrid Models

Received on 20 March 2025, accepted on 05 July 2025, published on 09 July 2025



doi: 10.4108/airo.8945

[*]Corresponding author. Email: amey.kharade231@vit.edu

## 1. Introduction

The rise of e-commerce has fundamentally altered consumer purchasing behavior, with online reviews playing a crucial role in decision-making. Reviews provide insights into product quality and reliability, shaping customer perceptions. However, the increasing dependence on online reviews has led to a significant rise in fraudulent reviews, which can mislead customers, manipulate product reputations, and cause financial losses. This has necessitated the development of robust fraud detection mechanisms. Earlier studies primarily relied on heuristic and rule-based systems, analyzing features like sentiment polarity, review length, and reviewer behavior to detect fraudulent patterns. Statistical classifiers such as Naïve Bayes, Decision Trees, and Support Vector Machines (SVM) were commonly used for fake review classification. While these approaches demonstrated moderate accuracy, they struggled with evolving fraud patterns and lacked adaptability. Machine learning models, particularly supervised learning techniques like Random Forest, Logistic Regression, and Gradient Boosting Machines, have improved fraud detection accuracy by incorporating engineered textual and behavioral features. Research [7] combines human expertise with data-driven methodologies by using two datasets labeled through different methods, achieving 82% accuracy in differentiating fake and genuine reviews. However, these methods depend heavily on feature engineering, which may not generalize well across datasets with different linguistic and structural properties. Additionally, classical ML models have struggled to keep up with the continuously evolving fraudulent strategies employed by fraudsters.

Deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been explored for detecting fake reviews by learning hierarchical text representations. More recently, transformer-based architectures like BERT and Ro BERT have gained popularity due to their ability to capture contextual dependencies in text data. Research in [28] highlights the lack of diverse linguistic datasets in fake review detection research. To address this, a dataset in the Italian language was developed using BERT and ELECTRA, achieving 95% accuracy. Despite these advancements, transformer models remain computationally expensive and require large labeled datasets for training. Moreover, adversarial attacks on these models can lead to decreased robustness in real-world applications. Graph-based machine learning models, particularly Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT), have shown promise in detecting fraudulent reviews by modeling relationships between users, products, and reviews. These models construct graphs where nodes represent entities such as users, products, and reviews, and edges capture interactions among them. The convolutional operations in GCNs enable the aggregation and propagation of information across the network, allowing the detection of anomalies indicative of fraud. However, optimizing these models for real-time fraud detection and large-scale deployment remains a challenge due to their computational complexity. Hybrid approaches combining GCNs and transformers have demonstrated potential for addressing the shortcomings of individual models. While GCNs excel at capturing relational dependencies among users and products, transformers effectively analyze textual cues in reviews. This dual approach enhances fraud detection by leveraging both structural and textual information. However, existing research lacks studies on hyperparameter optimization techniques that can fine-tune these hybrid models for optimal performance. With generative AI platforms now able to produce pseudonymous content at scale, there is an ever-growing need to create adaptive fake review detection technologies. The impact of such manipulation cascades downstream of a single transaction, extending to brand reputation, trust online, and equitable market competition.

The Whale Optimization Algorithm (WOA), a nature-inspired metaheuristic technique based on humpback whale foraging behaviour, has shown promise in optimizing high-dimensional search spaces. WOA can be employed to fine-tune parameters of GCN and transformer-based models, balancing complexity and performance while mitigating overfitting. Research in [45] demonstrates that WOA improves neural network training efficiency by enhancing model generalization on unseen data. However, existing research does not explore the full integration of WOA into fraud detection pipelines, limiting its practical application in real-time fraud prevention. Despite advancements in fraud detection models, several research gaps remain unaddressed. While many studies focus on offline batch processing, limited research explores real-time fraudulent review detection for dynamic e-commerce platforms. The high computational cost of transformer models and GCNs presents a challenge in large-scale deployment, requiring more research on optimization for practical applications. Many existing studies focus on high-resource languages, ignoring the need for fraud detection models that work across diverse linguistic contexts. Similarly, multi-modal approaches integrating textual, visual, and behavioral data remain underexplored. Fraudulent review generators continuously evolve, making it crucial to develop fraud detection models resistant to adversarial attacks. While hybrid GCN-transformer approaches are promising, research on optimal hyperparameter tuning techniques, such as WOA, remains limited. By addressing these gaps, our research aims to enhance the effectiveness of fraudulent review detection through advanced hybrid models and optimization techniques, ultimately improving trust and transparency in e-commerce platforms.

Table 1 Performance comparison of various methods in %

| Ref no. | Technique | Classifier | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|---|---|
| [2] | CNN | LSTM | 97.73 | 96.52 | 97.52 | 97.05 |
| [8] | NLP | BERT & ELECTRA | 95.00 | - | - | 95.00 |
| [22] | Supervised Learning | XGB, LSVC, SGD | 58.33 | 93.66 | 57.66 | 77.66 |
| [25] | Deep Learning | RoBERTa | 91.02 | 92.50 | 90.00 | 90.50 |
| [29] | GNN | TF-IDF + BERT | 87.60 | 86.81 | 88.67 | 87.74 |

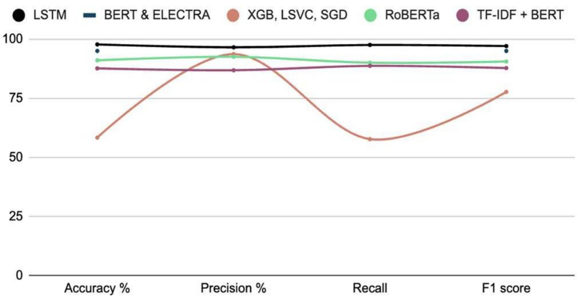


**Figure 1** Comparisons of metrics of different Classifiers

## 2. Techniques for Fake Review Detection

From 2017 to 2024, research on Fake Review Classification methods, the field can be divided into following directions: Supervised, Unsupervised, Graph- Based, Neural Networks, Hybrid Approaches. It can be seen that deep learning and neural networks methods account for the largest proportions among the current fake review classification methods. Fig. 2 shows Distribution of the researches of Classification Techniques whereas Fig. 3 shows Distribution of the researches of various techniques and Table 2 gives Distribution of the researches of various techniques.

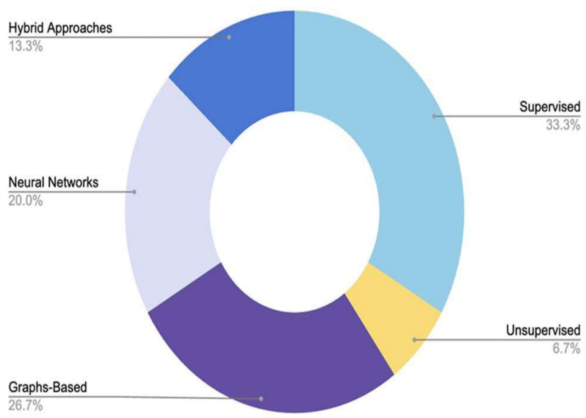


**Figure 2** Distribution of research across

classification techniques. This visual roadmap reinforces the structured categorization and highlights areas of emerging interest.

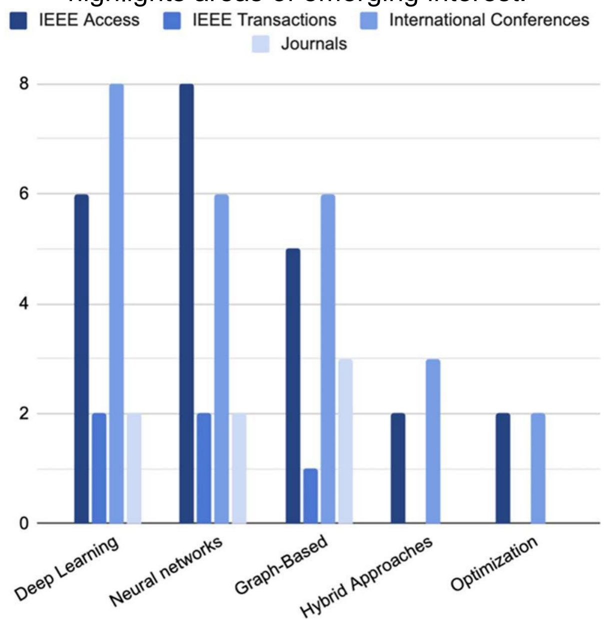


**Figure 3** Distribution of techniques by publication source. It provides insight into the popularity and research intensity within each category, aiding bibliometric understanding.

Table 2 Distribution of the researches of various techniques

| Technique | IEEE Access | IEEE Transaction | International Conferences | Journals |
|---|---|---|---|---|
| Deep Learning | 6 | 2 | 8 | 2 |
| Neural Networks | 8 | 2 | 6 | 2 |
| Graph-Based | 5 | 1 | 6 | 3 |
| Hybrid Approaches | 2 | 0 | 3 | 0 |
| Optimization | 2 | 0 | 3 | 0 |
| Other Techniques | 3 | 0 | 2 | 1 |

## 2.1 Supervised Learning

Supervised learning methods are dependent on labeled datasets, where each review is marked as either genuine or fake. These algorithms train models using known examples and apply this training to unseen data for classification. Supervised approaches for classification of reviews are extensively employed due to its simplicity, and ease of implementation. Decision trees are used for the classification of reviews, as they capture many patterns in the data. A decision tree recursively splits the dataset into feature values and constructs a model capable of classifying whether a review is real or fake. Combination of decision trees is a random forest, to improve the robustness of an individual tree by training different subsets. Also, it reduces overfitting, improves the model's accuracy. Various research in the last few years resulted in some salient features in enhancing fake review detection models using algorithms based on decision trees. Further, several researchers have explored the deficiencies of single decision trees and have come up with ensemble techniques such as random forests, which can be considered to yield much better generalization on unseen data with reduced problems of overfitting compared to their simpler counterparts. Figure 4 shows the Decision trees concept.

### 2.1.1 Decision tree

Decision trees are widely used for fake review detection as they effectively capture patterns in data. A decision tree splits the dataset based on feature values, while random forests, an ensemble of multiple decision trees, enhance robustness, reduce overfitting, and improve accuracy. Beyond technical aspects, fake review detection has economic, social, and has technological implications. Economically, it prevents market distortion and revenue loss; socially, it preserves trust in e-commerce; technologically, it advances AI-driven fraud detection. Adaptability across different platforms is crucial, as fraud tactics vary. Models must be fine-tuned to handle platform-specific nuances using real-time data and contextual analysis. Additionally, data-driven decision-making strengthens fraud detection by leveraging large datasets, real-time analytics, and anomaly detection to counter evolving fraudulent strategies.

### 2.1.2 Support Vector Machine

SVMs are widely used for textual classification tasks, including classification of reviews. It constructs a subspace that separates reviews into two categories: genuine and fake. SVM works effectively for high dimensional spaces text reviews, suitable for fake review detection [7],[24]. For example, SVM models are paired with Word embeddings and TF-IDF, to represent text reviews in numerical form, capable of training.

### 2.1.3 Outlier detection

Some of the supervised models take fake reviews as outliers in the data. The methods make use of traditional classifiers in the identification of abnormal patterns in review data, taking the extremely positive, negative reviews as a strong indication of fraud [27]. They embed statistical outlier detection with supervised learning. The result is a very robust approach to identifying reviews that are deceptive.

## 2.2 Unsupervised Learning

Unsupervised learning methods do not need labeled data, so unsupervised learning is of great use in fake review classification where typically labeled examples are very few or not available. This method looks for patterns and anomalies in textual data without any predefined categories. Based on the references, we summarize the main unsupervised learning approaches reviewed for the fake review's detection as under:

### 2.2.1 Clustering Algorithms

These clustering methods segment the reviews into groups of similar items, helping identify unusual patterns or anomalies in those segments that may raise a flag for fake reviews. K-means and hierarchical clustering are generally implemented in segmenting reviews into clusters and analyzing those reviews that are much different from typical reviews. These techniques can help explore hidden patterns in the review data by grouping similar reviews together and finding outliers.

### 2.2.2 Anomaly Detection

Anomaly detection identifies reviews that cannot be found as usual and may show fraud instances. The Isolation Forest and One Class SVM methods model the normal reviews distribution to flag reviews that fall out of this distribution in textual data. This kind of approach works pretty well to identify fake reviews because they do not fit into the regular patterns seen by typical reviews. Figure 5 shows Outliers

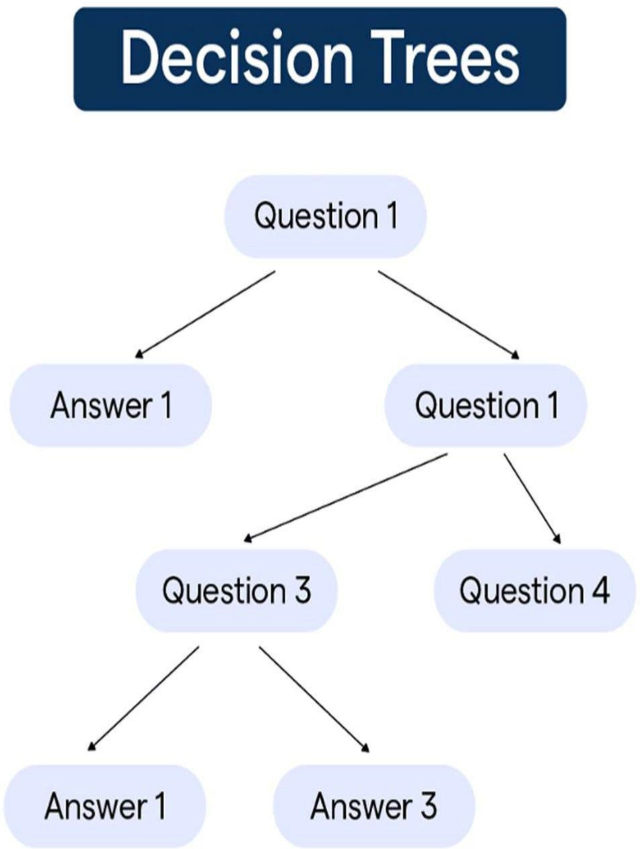


**Figure 4** Decision Tree

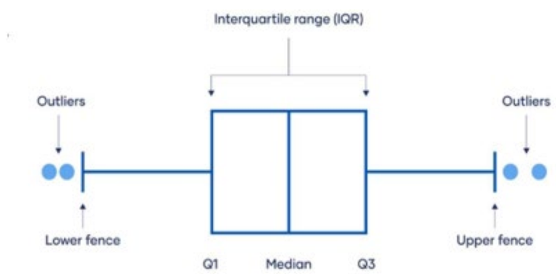


**Figure 5** Outliers

## 2.3 Graph-Based Techniques

Graph-based methods leverage the relationships between entities in data reviews, and products, to classify reviews. By implementing the structure of these relationships, these methods identify patterns and interpret fake reviews as components of a deceptive network. The graph-based approaches discussed below are derived from the references.

### 2.3.1 GNN

The recent breakthroughs have resulted in the use of Graph neural networks, which apply deep learning techniques directly on the graph-structured data. GNN is able to learn dependencies between various entities in the graph and find patterns of behavior. These networks learn to classify activities by training on the node features and structural information of the graph, so they fit well for the identification of intricate fake review networks.

## 2.4 Neural Networks

Neural networks are among the best approaches for classifying reviews due to their ability to automatically build up higher-order and complex patterns from textual data. Contrasting traditional ML approaches, which rely on manually designed features, a neural network learns the representation of review texts by capturing, in an efficient way, subtle differences between genuine and fake contents. Various neural network-based approaches are drawn from among 57 references for addressing these challenges in fake review detection.

### 2.4.1 RNN

RNNs are a sub-concept of neural networks, they are designed to handle sequential data. In the task of fake review detection, RNNs model sequential dependencies between words, capturing the flow of the language and unnatural patterns that may mark a review as being fake. For example, fraudulent reviewers often follow certain templates, which can be detected by word sequence analysis. However, RNNs have been prone to the problem of vanishing gradients, a factor that limits their full capacity in capturing long term dependencies within text. That limitation has led to the adoption of LSTM networks-advanced RNN architecture, capable of storing more information over longer sequences.

### 2.4.2 LSTM

LSTM networks improve conventional RNNs by adding memory cells, enabling retention of information across long sequences. This way, LSTMs can model dependencies in texts of reviews to identify repeating deceptive patterns or unusual linguistic construction across longer passages. Therefore, LSTMs are applicable when the text is long form or when fraudulent reviewers subtly modify their language over time to avoid detection, as observed in [10], [29], and [33].

### 2.4.3 CNN

Although CNNs are proposed for image processing tasks, recently they have been applied to several NLP tasks with success, including the detection of fake reviews. CNNs apply convolution processes on texts to get the local features of word n-grams. By learning the local patterns of words, the CNN learns typical linguistic cues or writing styles that define a fake review. They used max- pooling layers to concentrate on the most important features. Diminishing the dimensionality results in better performance for large review datasets, as reported in [6], [19], and [36].

## 2.5 Hybrid Approaches

Combining deep learning with feature-based sentiment analysis are widely employed hybrid techniques in classification of review. Feature-based sentiment analysis focuses on extracting features from review texts, then calculating the sentiment towards these features. When integrated with deep learning models, this approach enhances the ability to capture linguistic patterns/sentiments that fake reviews often manipulate [4], [17], [25].

### 2.5.1 Metaheuristic and graph-based methods

Metaheuristic algorithms namely Genetic and Particle swarm optimization algorithms have been combined with deep learning to advance the selection of features and hyper-parameters in fake review detection systems. These optimization algorithms enhance models by tuning the parameters more efficiently than traditional methods. Metaheuristics guide the model towards finding the best performing combinations of features, avoiding the trap of local minima, which could limit the effectiveness of deep learning in large datasets [2], [31], [41].

Table 3 Recommended Parameter Defaults

| Optimizer | Population Size | Max Iterations | Common Use |
|---|---|---|---|
| WOA | 30-50 | 100-300 | Tuning BERT |
| GWO | 20-30 | 100-200 | Feature Selection |
| Bat Algorithm | 25-40 | 200-500 | Hyper parameter search |

## 3. Datasets

## 3.1 YelpCHI dataset

This dataset is sourced from Yelp, consists of about 67,000 reviews for the same set of restaurants, and hotels in Chicago, USA. Each entry includes info- user-related, product, a timestamp, ratings, and a text review. The data contains 201 restaurants, and hotels that are reviewed by 38K reviewers. Yelp co. does have an algorithm to classify reviews, further categorizing them on a filtered list. On Yelp, filtered reviews are publicly available, recommended reviews are featured on a business's page. Yelp's anti-fraud filter is far from perfect; it was found to be catching the accurate results; hence the near ground truth. This dataset contains both recommended and filtered reviews. In this dataset, 13.2 filtered and have been authored by 20.3 spammers. Table 3 Comparison table of datasets

Table 4 Comparison of datasets

| Dataset | Description | Number of Reviews | Additional Details |
|---|---|---|---|
| YelpCHI | Reviews from Yelp for hotels and restaurants in Chicago. | 67,395 | 201 establishments, 38,063 reviews; 13.23% filtered, 20.35% spammers. |
| Amazon Reviews | Reviews from Amazon spanning 18 years. | 31,686,770 | 1,300,000 training, 200,000 testing samples per sentiment class. |
| Maxwell Fake Reviews | Collection of labeled fake vs. genuine reviews. | 20,000 fake, 20,000 genuine | Useful for machine learning model testing; balanced distribution. |
| Italian Cultural Heritage | Reviews for 20 hotels in Naples. | 1,600 | 800 reviews; 10 positive and 10 negative per hotel; average length 61 words. |

## 3.2 Amazon reviews' dataset

The dataset consists of reviews on Amazon, there are approximately 34M amazon reviews from 6.6M users on 2.4M products. A subset contains 1.8M training and 200K testing samples in each polarity sentiment. The dataset is of reviews on Amazon, span of those reviews' dataset is of 18 years (till Mar. 2013);it has about 35M reviews. Amazon reviews' polarity dataset is created by considering review scores 1 and 2 as negative, 4 and 5 as positive, and samples of score 3 are ignored. In the dataset, class 2 is the positive whereas class 1 is the negative. Each class has 200K test and 1.8M training samples.

## 3.3 Maxwell Fake reviews dataset

This is a collection of reviews labeled as either fake or genuine. Fake Reviews Dataset contributed and hosted on Kaggle by Maxwell. The dataset consists of 5050 towards the development and testing of ML models useful in the classification of reviews. Each review features a number of important columns: textual content of review, rating, timestamp, and information about the reviewer. The dataset is structured in a way that textual and metadata insights could be derived, that are helpful in the integration of gamut techniques for the classification of Reviews, not limited to sentiment and behavioral analysis, and text-based modeling. The dataset provides substantial balance in the distribution of real and fake reviews; hence, the model will have stronger training. This makes it a useful resource for advanced algorithm testing, namely deep learning and graph neural networks, due to its diversity in structure regarding reviews.

### 3.4 Italian cultural heritage dataset Datasets

It consists of reviews of 20 hotels in Chicago, totaling 1600 reviews with labels- true and fake, negative and positive sentiment. It also describes the dimension of Italian cultural heritage in Naples. The data consists of 800 reviews. Namely, 10 positive, 10 negative reviews were collected for each one of the selected 20 places within the city. The average length of reviews is around 61 words/review.

## 4. Performance Comparison

With the rising need to detect fake reviews, researchers have explored various methods, including machine learning, deep learning, and graph-based techniques. A structured methodology is essential, covering data collection, preprocessing, model selection, training, and evaluation. Ensuring error-free calculations and validating statistical techniques enhance model reliability. Key metrics like accuracy, precision, recall, and AUC-ROC should be used for proper assessment. This review analyzes existing research, compares models across platforms, and evaluates real-time detection effectiveness. Findings are clearly presented using tables and graphs for logical interpretation of results.

The research [1] discusses use of GNNs for analyzing graph data, highlighting their effectiveness in capturing complex relations within heterogeneous networks and compares MP-GT with existing models like SA-GCN and CAP. The authors of [2] address the significance of online consumer reviews on purchasing decisions, highlighting the challenge posed by spurious reviews in e-commerce. In [3], the paper addresses the challenges of edge detection in colour images by UAVs, highlighting issues such as noise and distortion proposes an improved whale optimization algorithm (WOA) that utilizes quaternion representation for better edge detection results. The authors of [5] introduced a novel fuzzy optimized convolutional neural network aimed at enhancing the accuracy of user opinion predictions. In [6], the authors highlighted the limitations of existing datasets, which are predominantly in English. Also, previous research has explored Support Vector Machines, Naive Bayes as ML techniques, but authors in this research aims to create a baseline using modern language models like BERT and ELECTRA. Traffic networks are complex and exhibit spatial-temporal dependencies, making accurate prediction challenging and Recent advancements in deep learning have led to various models focusing on spatial-temporal dependency modeling, primarily using convolutions to separate spatial and temporal correlations but real time traffic state forecasting remains difficult due to the intricate nature of traffic networks. So, to capture dynamic dependencies simultaneously, review [9] proposes a hybrid DL framework for traffic prediction.

The authors of [10] found out the problem in traditional methods for node embeddings- graphs require all nodes to be present during training, which limits their ability to generalize to unseen nodes. So, they introduced GraphSAGE, an inductive framework that generates embeddings for unseen nodes by using local neighborhood features of the node enhancing scalability and adaptability in dynamic graph scenarios. In literature [11], the authors gave the result that the whale bionic algorithm reduced prediction error and gave greater accuracy when compared to the prediction results of models-RNN, LSTM, WOA-LSTM. In [12], the authors addressed the problem of increase in nonlinear loads in power systems which led to significant harmonic distortion in voltage and current, necessitating effective solutions. So, the authors used Whale optimization algorithm (WOA) to employ Selective Harmonic Elimination method to optimize inverter output voltage waveforms and eliminate low-order harmonics more efficiently. The paper [13] addresses the challenges of managing grid stability due to the variability of renewable energy sources like wind and solar, which Table 5 Performance comparison of Supervised learning methods with graph Fig. 6 Comparison graph that can cause transient overvoltage and other disturbances.

Table 5 Performance comparison of Supervised learning methods with graph.

| Methods | Metrics | Performance | Dataset | Reference |
|---|---|---|---|---|
| SVM | F-measure | 78.10% | Yelp, etc. | [51] |
| LIBSVM | Accuracy | 89.60% | Yelp, etc. | [52] |
| NB-SVM | Accuracy | 91.90% | IMDB | [53] |
| WMUSVM | Recall | 82.50% | TripAdvisor | [54] |
| BERT | Accuracy | 90.50% | Yelp | [55] |
| XGBoost | Precision | 99.00% | Yelp | [56] |

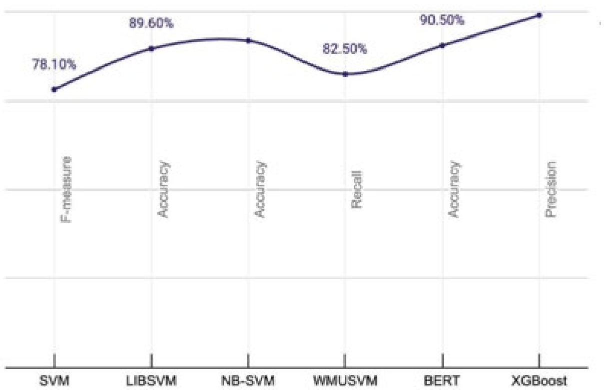


**Figure 6** Comparison graph.

Hence, The whale optimization algorithm is proposed as a solution optimizing the allocation of resources and control parameters and improving frequency regulation. The paper [14] discusses Author Gender Detection as a critical issue in communication under internet security. So, research employs ML and meta- heuristic algorithms, specifically an artificial neural network (ANN) combined with Whale optimization algorithm (WOA) to improve accuracy of classification. In [15], the authors aim to enhance fraud detection by employing a hybrid approach combining supervised and unsupervised algorithms, specifically Light Gradient Boosting Machine (LGBM) and Kernel Principal

Component Analysis (KPCA) to improve the Receiver Operating Characteristic-Area Under Curve. Credit card fraud is a significant issue in digital transactions, highlighting the need for effective detection methods as traditional methods often fail to capture complex fraudulent patterns. The study [16] focuses on outlier detection as a promising method to identify fraudulent activities within transactional data.

The paper [17] addresses the challenge of understanding disease associations through biological data, emphasizing the need for effective models to analyze complex genetic interactions. GraphSAGE and GCN are used to perform convolution operations on graph data, to capture the relationships between nodes and their features. In study [18], the authors have introduced a method Traditional Image Steganography for embedding messages in the images so that the messages will not be detected by steganalysis techniques. The development of GANs has enabled the implementation of this method enhancing the effectiveness of steganography and effectively resisting both CS-steganalysis and RF-steganalysis.

The paper [19] investigates the challenges of AI-based fraud detection systems, focusing on real-time transaction surveillance and detection to enhance digital payment security. The study employs a comparative exploration approach such as integration of AI technologies to continuously monitor transactions and identify fraudulent activities. In [20], the authors introduced a deep learning-based engine that utilizes feature selection attention CNN module to adaptively emphasize the importance of relevant prior knowledge to enhance the model's ability to accurately

classify Consumer fraud which involves illegal activities aimed at generating revenue. The paper [32] reviews recent literature

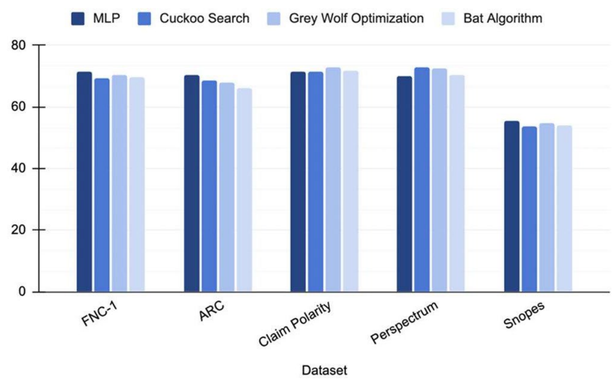


**Figure 7** Comparison of Accuracies (%) of various algorithms on different datasets

on credit card fraud detection (CCFD) using Deep Learning techniques, highlighting their effectiveness compared to traditional machine learning methods. It provides a comprehensive overview of various DL techniques such as, RNN, GRU, LSTM, CNN comparing their performance. The researchers of [33] identify the challenge of timely detection of fraud in the financial sector, where traditional methods are often too slow and propose the use of cloud AI systems, which can process and analyze large datasets rapidly, significantly improving the speed of fraud detection.

In [37], paper discusses the challenges of solving the first Fredholm integral problem related to particle size distribution highlighting the limitations of conventional optimization algorithms, whale optimization algorithm (WOA) is presented as a promising alternative due to its global performance, speed, and ability making it suitable for complex particle size inversion problems.

In [39] , the proposed SAGE- Net address The Limitations of existing methods by integrating semantic and geometric features, enhancing the robustness of global descriptors for place recognition. Experiments on benchmark datasets indicate that SAGE- Net outperforms state-of-the art approaches, demonstrating its effectiveness and ability in unseen scenes. The authors discuss the automated synthesis of analog circuits, highlighting various frameworks and methodologies emphasizing a hierarchical design approach where circuit blocks are built from fundamental analog primitives for creating complex system-level designs [40]. GNN generalizes CNN to graphs, inspired by Fourier transformation Heterogeneous GNN methods transform graphs, apply GNN, and aggregate representations in [41].

The authors of [42] introduced a novel approach using GNNs to extract features from review networks, enhancing the detection of fake reviews. Also, the authors propose improvements to GNNs by integrating TrustRank and a multi-feature detection framework, which aids in better identifying fraudulent activities.

In [43], the researchers introduces a GNN framework to detect fake reviews, emphasizing the importance of social context in understanding user interactions and reviews addressing the challenge of verifying online opinions due to the lack of verification procedures.

The authors of [44] categorized fake reviewer detection methods into three main streams: "Text-based approach", "Behavior-based approach" and "Graph-based approach" which highlighted the evolution and challenges in the field of fake review detection, creating an overview for the proposed RHGN framework.

The paper [46] introduces GraphSAGE, a graph-based approach, enhanced with a covertness model to improve the detection of fake reviews by analyzing user interactions and review content. The proposed model shows competitive performance in Recall and AUC metrics when tested on real-world Amazon datasets, outperforming other models like GCN and GAT.

Researchers of [47] combined rotation forest algorithm with the WOA for feature selection to identify whether emails are spam or not. Remarkably, this hybrid method achieved a 99.9 significantly outperforming previous techniques. The proposed approach in [49] leverages a GCN that integrates semantic similarity. Fig. 7 shows comparison of accuracies (algorithms on different datasets. By combining word-level and document-level information, the researchers construct a GNN.

Table 6 Comparison of Dataset

| **Dataset** | **MLP** | **Cuckoo Search** | **Grey Wolf Optimization** |
|---|---|---|---|
| FNC-1 | 71.29% | 69.17% | 70.54% |
| ARC | 70.43% | 68.54% | 67.99% |
| Claim | 71.33% | 71.43% | 72.85% |

| | | | |
|---|---|---|---|
| Polarity | | | |
| Perspectrum | 70.06% | 72.93% | 72.34% |
| Snopes | 55.49% | 53.79% | 54.86% |

Therefore, authors displayed a comparative study in which Node2vec, GraphSAGE methods were tested that converted the input data into a format which is meaningful and useful for classification by vector representation of node-based graphs. Table 5 compares the performance of some ML optimization techniques in processing some datasets related to the detection of misinformation or fake review.

Comparison optimization methods include MLP, Cuckoo Search, Grey Wolf Optimization, and Bat Algorithm. These algorithms will be evaluated on five different datasets: FNC-1, ARC, Claim Polarity, Perspectrum, and Snopes. In this case, the performance of the MLP acts as the baseline across the datasets. For example, on the FNC-1 dataset, it performs quite well at 71.29%. Cuckoo Search is a population-based metaheuristic optimization method that draws inspiration from the brood parasitism of some cuckoo species. It has somewhat poor performance compared to most datasets, such as that carried out on the FNC-1 dataset with 69.17% surpassed by Grey Wolf Optimization, which is inspired by the hunting strategy of grey wolves, it manages to achieve 72.85%. Claim Polarity, setting a high bar for model performance improvements in certain cases. Another algorithm, Bat Algorithm, inspired by echolocation in bats, can also do comparably well, achieving results comparable to, or very slightly below, those of Grey Wolf Optimization: for example, 69.54% Perspectrum. The table 6, can be referred to by the same datasets for the performance of different optimization algorithms and, thus, may show which methods should perform better for any particular type of data. Especially promising look Grey Wolf Optimization and Bat Algorithm because their performance in many cases is superior or in-line with the results of other methods, such as Cuckoo Search or MLP for different datasets. Table 7 shows some comparisons of a few feature extraction techniques applied in different fake review detection systems. This process is very crucial in the conversion of data into a numerical format from textual format such that various machine learning algorithms can understand. The following five techniques are now discussed and compared: TF-IDF, BERT, Word2Vec, GloVe, Count Vectorization, and FastText.

Table 7 Comparative Analysis

| Year | Ref. | Approach | Classifier | Dataset | Acc./AUC |
|---|---|---|---|---|---|
| 2024 | [7] | Human-driven and data-driven | 1. Deep Learning<br>2. SVM<br>3. KNN | 1. Yelp Filtering Algorithm Dataset<br>2. Crowds Perception Dataset | 80-85% |
| 2023 | [8] | Language model-based approach | 1. BERT<br>2. ELECTRA | Italian Cultural Heritage (ICH) Dataset | 95% |
| 2023 | [25] | Survey approach | Deep learning - RoBERTa | Deception Dataset | 90% to 92.5% |
| 2024 | [29] | Metaheuristic AI-based approach | 1. Random Forest<br>2. Feed-Forward Neural Network<br>3. KNN<br>4. Logic Regression | Amazon Reviews Dataset | 75.92%, 75.92%, 78.87%, 84.07% |
| 2021 | [45] | Improved Whale Optimization Algorithm (IWOA) | 1. MLP<br>2. CS<br>3. GWO<br>4. BA | 1. FNC-1 Dataset<br>2. ARC Dataset<br>3. Claim Polarity Dataset<br>4. Perspectrum Dataset<br>5. Snopes Dataset | 76.53%, 74.72%, 78.45%, 79.85%, 79.02% |

Table 8a Analysis of Feature Extraction Techniques

| Feature Extraction Technique | Advantages | Disadvantages | Time Complexity | Dimensionality | Applicability for Fake Review Detection |
|---|---|---|---|---|---|
| TF-IDF | Fast, good for smaller datasets | Ignores word order, context, and semantics | O(n * m) | High | Useful for capturing important root text |
| BERT | Captures context and semantics from both directions | Computationally expensive, slow for large datasets | $O(n^2 * m)$ | Low | Highly effective in capturing complex linguistic cues |
| Word2Vec | Captures semantic relationships between words | Requires large training data, doesn't handle out-of-vocabulary words | O(n) | Low to medium | Effective for capturing word meaning and context |
| GloVe | Captures global statistics and relationships | Requires pre-training on large corpora | O(n) | Low | Good at representing word relationships, less dynamic in real-time |
| Count Vectorization | Simple, fast, works well with small datasets | Ignores context, creates sparse vectors | O(n * m) | High | Basic and limited but useful in some cases |
| FastText | Handles out-of-vocabulary words by breaking words into subwords | Less accurate for very small datasets | O(n * log n) | Low | Context-based with word variations, useful for rare words or languages |

Table 8b Inference Time and Memory Comparison

| Model | Inference Time (CPU) | Inference Time (GPU) | Memory |
|---|---|---|---|
| TF-IDF | 5 ms | - | ~0.1GB |
| BERT | 200 ms | 30 ms | 2-4 GB |
| FastText | 10 ms | 5 ms | ~0.5GB |
| GCN+GAT | 400 ms | 50 ms | 6+GB |

Each technique is analyzed along the following dimensions, including advantages, disadvantages, and time complexity. Advantages are strengths within each approach, namely, simplicity, capturing semantic relationships, and ability to handle large datasets. Disadvantages refers to a limitation that forms a part of using the techniques, especially the limited context which simple methods like TF-IDF can capture while more advanced models like BERT are computationally too expensive. Time Complexity refers to Approximate time complexity for each technique. In the time complexity estimates, 'n' represents the average number of tokens per review and 'm' denotes the vocabulary size. For BERT, complexity is approximated as $O(n^2 \cdot m)$ due to its self-attention mechanism over tokens. For TF-IDF and Count Vectorization, complexity is $O(n \cdot m)$ as each token in a review is matched against vocabulary terms linearly.

For example, BERT and FastText are spotted to capture context and subword level information and hence suitable for the detection of subtle cues in fake reviews. In short, the table 8 compares these techniques in tabular form regarding showing performance complexity trade-offs and suitability for fake review detection.

## 4.1 Evolution of Fake review techniques

2011-2014: Supervised Machine Learning Era.
2011: Supervised machine learning techniques were introduced by using labeled datasets. For fake review detection, decision trees and random forests were being considered.
2012: Feature engineering improvements review length, user behavior, and temporal features in metadata.
2014: Some features of supervised classification model were mainstreamed for detection of fake reviews. Such as, TF- IDF and bag-of-words.
2015-2017: CNNs and RNNs are deep learning methods which capture semantic and syntactic features of review texts more effectively.
2016: Techniques Capturing long-term dependencies in review data using LSTM networks evolved.
2017: Introduction of attention mechanisms to enhance the focus on more relevant text parts.
2018- 2020: Graph-based and Ensemble Methods introduced.
2018: Graph Neural Networks come into play; for example, GraphSAGE and Graph Attention Networks, to capture relationships among users, products, and reviews.
2019: Ensemble methods that combine multiple models and hybrid systems which integrate DL and traditional ML methods start to rise in popularity.
2020: Increased interest in using Word Optimization Algorithms and sophisticated feature selection methods for better detection.
2021-2024: Advanced Graph Learning and Explainability introduced.
2021: GraphSAGE and GAT find their extensive use in the task of fake review detection, with increased attention towards modeling user-user and user-product interactions from reviews.
2022: Explainability in AI techniques start to be proposed for fake review detection by incorporating techniques into models that will render model predictions more understandable.
2023: Using a whale optimization algorithm in conjunction with graph neural networks to enable optimization for large-scale fake review detection tasks.
2024: GAT, GraphSAGE, WOA hybrid models applied to real-time detection in an interpretable and highly scalable manner.

# 5. Discussion and Future scope

In this section, we discuss the impact of various techniques for review classification, their real-world implications, and future development directions. Fake review detection plays a crucial role in maintaining consumer trust, preventing financial losses for businesses, and ensuring fair competition in e-commerce. Fraudulent reviews mislead customers, influence purchasing behavior, and damage brand reputations, making detection essential for platform integrity. Supervised methods like BERT and Support Vector Machines effectively utilize labeled data to spot fake reviews but face challenges such as class imbalance, high computational costs, and reliance on extensive annotated datasets. Neural networks, including CNNs and LSTMs, can learn higher-order features but require significant computational resources and often lack interpretability, making real-world deployment challenges. Unsupervised learning approaches, such as clustering and topic modeling, provide alternatives when labeled data is scarce; however, they are sensitive to noise, prone to generating false positives, and tend to have lower accuracy. Graph-based models are promising for detecting fake reviews by analyzing relationships between users, products, and reviews, but they require well-structured data and large-scale networks to perform effectively. A major challenge in fake review detection is dataset quality, as biased, imbalanced, or small datasets can reduce model accuracy and generalization. Ensuring diverse, high-quality datasets with real-world review patterns is essential for improving detection reliability. Additionally, many datasets lack multilingual and cross-domain coverage, limiting model adaptability across different platforms. Addressing these challenges with large-scale, unbiased, and domain-specific datasets can significantly enhance detection performance and real-world applicability.

# 6. Conclusion

The paper reviewed the various approaches that have targeted classifying fake reviews using different machine learning techniques, each with different strengths and weaknesses. This paper reviewed the literature on different approaches, including those falling under supervised learning methods, unsupervised techniques, graph-based models, and hybrid methods. Among them, unsupervised models are based on patterns observed; supervised learning methods including SVM and BERT are the most accurate, though problematic to deal with class imbalance issues and are computationally intensive. Unsupervised methods and graph- based models are helpful alternatives in case of limited labeled data or complex relationships among objects.

Hybrid approaches that combine various techniques are promising, as they provide room for optimization regarding both feature selection and the performance of the models themselves. In this paper, we have identified different approaches that exist in the classification of reviews and its impacts. Next, we discussed their advantages, and limitations in order to review these techniques.By spotlighting research gaps and proposing WOA-based optimization for hybrid systems, this paper not only synthesizes existing knowledge but also paves the way for future advancements in scalable, real-time fake review detection.

## References


[1] X. Fang, H. Yang, L. Zhang, et al., (2024). Enhancing App Usage Prediction Accuracy With GCN- Transformer Model and Meta-Path Context. IEEE Access, vol. 12, pp. 53031 53044. https://doi.org/10.1109/ACCESS.2024.337239.

[2] Abd-Alhalem, S. M., Ali, H. A., Soliman, N. F., Algarni, A. D., and Marie, H. S. (2024). Advancing E- Commerce Authenticity: A Novel Fusion Approach Based on Deep Learning and Aspect Features for Detecting False Reviews. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3435916.

[3] Liu, D., Zhou,S., Shen, R., and Luo,X.(2023).Color Image Edge Detection Method Based on the Improved Whale Optimization Algorithm. IEEE Access, vol. 11, pp. 5981 5989. https://doi.org/10.1109/ACCESS.2023.3236761.

[4] Li, X., and Chen, L. (2023). Fake Review Detection Using Deep Neural Networks with Multimodal Feature Fusion Method. 2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS), pp. 2869-2872. https://doi.org/10.1109/ICPADS60453.2023.00411.

[5] A, A., L, N., P, D., and Gokilavani. (2024). Predicting the Fake Review to Amazon Product Review Dataset Using Fuzzy Optimized Convolution Neural Network. 2024 International Conference on Expert Clouds and Applications (ICOECA), pp. 689- 693. https://doi.org/10.1109/ICOECA62351.2024.00125.

[6] Tajrian, M., Rahman, A., Kabir, M. A., and Islam, M. R. (2023). A Review of Methodologies for Fake News Analysis. IEEE Access, vol. 11, pp. 73879-73893. https://doi.org/10.1109/ACCESS.2023.3294989.

[7] Abedin, E.,Mendoza, A.,Akbarighatar, P., and Karunas k era, S. (2024). Predicting Credibility of Online Reviews: An Integrated Approach. IEEE Access, vol. 12, pp. 49050 49061. https://doi.org/10.1109/ACCESS.2024.3383846.

[8] Catelli, R., et al. (2023). A New Italian Cultural Heritage Data Set: Detecting Fake Reviews with BERT and ELECTRA Leveraging the Sentiment. IEEE Access, vol. 11, pp. 52214-52225. https://doi.org/10.1109/ACCESS.2023.3277490.

[9] Liu, T., Jiang, A., Zhou, J., Li, M., and Kwan, H. K. (2023). GraphSAGE-Based Dynamic Spatial– Temporal Graph Convolutional Network for Traffic Prediction. IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 10,pp.11210-11224.https://doi.org/10.1109/TITS.2023.3279929.

[10] Hamilton, W., Ying, R., and Leskovec, J. (2017). Inductive Representation Learning on Large Graphs. Dept. Comp. Sci., Stanford Univ., California. https://doi.org/10.48550/arXiv.1706.02216.

[11] Feng, C., Li, Z., Shao, L., Li, J., Li, C., and Liu, H.(2024). Load Prediction of Composite Curing Equipment Cluster Based on Combined EEMD-WOA LSTM Model. 2024 IEEE International Conference on Mechatronics and Automation (ICMA),pp.183-188. https://doi.org/10.1109/ICMA61710.2024.1063294.

[12] Li, W., and Zhang, W. (2024). Research on Selective Harmonic Elimination Technology Based on Enhanced Whale Optimization Algorithm. 2024 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 1-6. https://doi.org/10.1109/ICMA61710.2024.1063299 0.

[13] Ma, X., Gu, W., Lv, S., and Zheng, C. (2024). Optimal Configuration of Grid-Forming Energy Storage System Based on the Whale Optimization Algorithm. 2024 3rd International Conference on Energy, Power and Electrical Technology (ICEPET), pp. 754- 759. https://doi.org/10.1109/ICEPET61938.2024.10627.

[14] Safara, F., et al. (2020). An Author Gender Detection Method Using Whale Optimization Algorithm and Artificial Neural Network. IEEE Access, vol. 8, pp. 48428 48437. https://doi.org/10.1109/ACCESS.2020.2973509.

[15] Aquino, J. P. J., De Guia, A. P. G., Dela Cruz, D. C., and De Goma, J. C. (2024). Fraud Detection in Online Credit Card Transactions Using Deep Learning. 2024 5th International Conference on Industrial Engineering and Artificial Intelligence (IEAI),pp.85-89. https://doi.org/10.1109/IEAI62569.2024.00023.

[16] Devavarapu, Y., Bedadhala, R. R., Shaik, S. S., Pendela, C. R. K., and Ashesh, K. (2024). Credit Card Fraud Detection Using Outlier Analysis and Detection. 2023 4th International Conference on Intelligent Technologies (CONIT), pp. 1-7. https://doi.org/10.1109/CONIT61985.2024.106264

[17] Riccard, K., and Bandara, D. (2023). Autism Risk Classification Using Graph Neural Networks Applied to Gene Interaction Data. 2023 Congress in Computer Science, Computer Engineering, Applied Computing (CSCE), pp. 1574-1580. https://doi.org/10.1109/CSCE60160.2023.00259.

[18] Pei, G., Tang, Y., Yan, S., and Hu, D. (2024). Is this Real? Image Steganography without Embedding Spoofs Fake Image Detector. 2024 3rd International Conference on Image Processing and Media Computing (ICIPMC), pp. 192-198. https://doi.org/10.1109/ICIPMC62364.2024.105866

[19] Rani,S., and Mittal, A.(2023).Securing Digital Payments a Comprehensive Analysis of AI Driven Fraud Detection with Real Time Transaction Monitoring and Anomaly Detection. 2023 6th International Conference on Contemporary Computing and Informatics(IC3I), pp. 2345-2349. https://doi.org/10.1109/IC3I59117.2023.10397958.

[20] P, A., Bharath, S., Rajendran, N., Devi, S. D., and Saravanakumar, S. (2023). Experimental Evaluation of Smart Credit Card Fraud Detection System using Intelligent Learning Scheme. 2023 International Conference on Innovative Computing, Intelligent Communication and Security (ICIVICS), pp.10-14. https://doi.org/10.1109/ICICICS53744.2023.102827

[21] Mewada, A., and Dewang, R. K. (2022). Research on False Review Detection Methods: A State-of-the- Art Review. Journal of King Saud University- Computer and Information Sciences, vol. 34, no. 9, pp. 7530-7546. https://doi.org/10.1016/j.jksuci.2021.07.021.

[22] Hameed, W., Allami, R., and Ali, Y. (2023). Fake Review Detection Using Machine Learning. Revue d' Intelligence Artificielle, vol. 37. https://doi.org/10.18280/ria.370507.

[23] Yu, S., Ren, J., Li, S., Parsa, M. N., and Xia, F. (2022). Graph Learning for Fake Review Detection. Frontiers in Artificial Intelligence, vol. 5, pp. 922589. https://doi.org/10.3389/frai.2022.922589.

[24] Kumar, K., George, K. S., Bhatt, D., and Paul, O. P. (2023). A Brief Study on the Fake Review Detection Methods on E-

commerce Websites using Machine Learning, Artificial Intelligence, and Data Science. International Journal of Recent Advances in Multidisciplinary Research, vol. 11, no. 4, pp. 52743. https://doi.org/10.22214/ijraset.2023.52743.

[25] Muhawesh, R., Xu, S., Tran, S., Ollington, R., Springer, M., Jararweh, Y., and Maqsood, S. (2021). Fake Reviews Detection: A Survey. IEEEAccess,vol.9,pp.43142-43164. https://doi.org/10.1109/ACCESS.2021.3075573.

[26] Ben Jabeur, S., Ballouk, H., Ben Arfi, W., and Sahut, J.-M. (2023). Artificial Intelligence Applications in Fake Review Detection: Bibliometric Analysis and Future Avenues for Research. Journal of Business Research, vol. 158, pp. 113631. https://doi.org/10.1016/j.jbusres.2022.113631.

[27] Sheikhi, S. (2021). An Effective Fake News Detection Method Using WOA-xgbTree Algorithm and Content Based Features. Applied Soft Computing, vol. 109, pp. 107559. https://doi.org/10.1016/j.asoc.2021.107559.

[28] Oak, R. (2024). Detecting Review Fraud Using Meta heuristic Graph Neural Networks. International Journal of Information Technology, vol. 16, no. 4, pp. 1-10. https://doi.org/10.1007/s41870-024-01896-

[29] Iqbal, A., Rauf, M. A., Zubair, M., and Younis, T. (2023). An Efficient Ensemble Approach for Fake Reviews Detection. 2023 3rd International Conference on Artificial Intelligence (ICAI), pp.7075.https://doi.org/10.1109/ICAI58407.2023.10136652.

[30] Lai, S., Wu, J., Ye, C., and Ma, Z. (2024). UCF PKS: Unforeseen Consumer Fraud Detection with Prior Knowledge and Semantic Features. IEEE Transactions on Computational Social Systems, vol. 11, no. 4, pp. 5454 5467.https://doi.org/10.1109/TCSS.2024.3372519.

[31] Mienye, I. D., and Jere, N. (2024). Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions. IEEE Access, vol. 12, pp. 96893-96910.https://doi.org/10.1109/ACCESS.2024.3426955.

[32] Abitova, G. A., Abalkanov, M., Shuteyeva, G., Aitmukhanbetova, E., and Kulniyazova, K. (2024). Review of Cloud AI for Real-Time Fraud Detection. 2024 10th International Conference on Automation, Robotics and Applications (ICARA), pp.454-460. https://doi.org/10.1109/ICARA60736.2024.105531

[33] Silpa, C., Prasanth, P., Sowmya, S., Bhumika, Y., Pavan, C. H. S., and Naveed, M. (2023). Detection of Fake Online Reviews by Using Machine Learning. 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), pp. 71-77. https://doi.org/10.1109/ICIDCA56705.2023.100997

[34] Shan, Z., Wang, Y., Liu, X., and Wei, C. (2023). Fuzzy Automatic Disturbance Rejection Control of Quadrotor UAV Based on Improved Whale Optimization Algorithm. IEEE Access, vol. 11, pp. 69117-69130. https://doi.org/10.1109/ACCESS.2023.3292265.

[35] Liu, D., Zhou, S., Shen, R., and Luo, X. (2023). Color Image Edge Detection Method Based on the Improved Whale Optimization Algorithm. IEEE Access, vol. 11, pp. 5981-5989. https://doi.org/10.1109/ACCESS.2023.3236761.

[36] Lu, Z., et al. (2024). Inversion of Bubble Size Distribution Based on Whale Optimization Algorithm. IEEE Photonics Journal, vol. 16, no. 4, pp. 1-6. https://doi.org/10.1109/JPHOT.2024.3406886.

[37] Riccard, K., and Bandara, D. (2023). Autism Risk Classification Using Graph Neural Networks Applied to Gene Interaction Data. 2023 Congress in Computer Science, Computer Engineering, Applied Computing (CSCE), pp. 1574-1580. https://doi.org/10.1109/CSCE60160.2023.00259.

[38] Hao, W., Zhang, W., and Jin, H. (2024). SAGE-Net: Employing Spatial Attention and Geometric Encoding for Point Cloud Based Place Recognition. IEEE Robotics and Automation Letters, vol. 9, no. 6, pp. 4958-4965. https://doi.org/10.1109/LRA.2024.3387112.

[39] Wu, Z., and Savidis, I. (2024). Comparative Analysis of Graph Isomorphism and Graph Neural Networks for Analog Hierarchy Labeling. 2024 25th International Symposium on Quality Electronic Design (ISQED), pp. 1 7. https://doi.org/10.1109/ISQED60706.2024.105287

[40] Tang, S., Jin, L., and Cheng, F. (2021). Fraud Detection in Online Product Review System via Heterogeneous Graph Transformer. IEEE Access, vol. 9, pp. 167364-167373, https://doi.org/10.1109/ACCESS.2021.3084924.

[41] Ren, X., Yuan, Z., and Huang, J. (2022). Research on Fake Reviews Detection Based on Graph Neural Network. Proc. SPIE, vol. 12250, International Symposium on Computer Applications and Information Systems (ISCAIS 2022), 1225015. https://doi.org/10.1117/12.2639534.

[42] Cheng, L.-C., Wu, Y. T., Chao, C.-T., and Wang, J. H. (2024). Detecting Fake Reviewers from the Social Context with a Graph Neural Network Method Decision Support Systems, vol. 179, pp. 114150. https://doi.org/10.1016/j.dss.2023.114150.

[43] Zhao, J., Shao, M., Tang, H., Liu, J., Du, L., and Wang, H. (2023). RHGNN: Fake Reviewer Detection Based on Reinforced Heterogeneous Graph Neural Networks. Knowledge-Based Systems, vol. 280, pp.111029. https://doi.org/10.1016/j.knosys.2023.111029.

[44] Pandey, A. C., and Tikkiwal, V. A. (2021). Stance Detec tion Using Improved Whale Optimization Algorithm. Complex Intelligence and Systems, vol. 7, pp. 1649–1672. https://doi.org/10.1007/s40747-021-00294-0.

[45] Narayan, A., Tharun, D., S. D, M. K., and Chacko,A. (2024). Elevating GraphSAGE for Covertness: A Strategic Approach to Unmasking Fake Reviews in E Commerce. Workshop Proceedings of the 18th International AAAI Conference on Web and social media. https://doi.org/10.36190/2024.65.

[46] Shuaib, M., Abdulhamid, S. M., Adebayo, O. S., et al. (2019). Whale Optimization Algorithm-Based Email Spam Feature Selection Method Using Rotation Forest Algorithm for Classification. SN Applied Sciences, vol. 1, pp. 390. https://doi.org/10.1007/s42452-019-0394-7.

[47] H.Yu, W.Liu, N. Zhu, P. Li, and X. Luo, ”IN-GFD: An Interpretable Graph Fraud Detection Model for Spam Reviews,” IEEE Transactions on Artificial Intelligence. https://doi.org/10.1109/TAI.2024.3420262

[48] A. Shivhare and R. Dubey, ”Deep Learning Model for Classifying Spam Review Over Social Media,” 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, pp. 688 694.https://doi.org/10.1109/ICIMIA60377.2023.104 26476

[49] A. A. Ibrahim, M. G. Abdulkareem, and I.Al-Jadir, ”A Comparison of Deep Learning Techniques Used to Identify Social Media Reviews,” 2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT), Al-Muthana, Iraq, pp. 18-22. https://doi.org/10.1109/AICCIT57614.2023.102182

[50] Aritsugi M, et al. ”Exploiting function words feature in classifying deceptive and truthful reviews” In: 2018 Thirteenth International Conference on Digital Information Management (ICDIM). IEEE, 2018, pp. 51–6. https:// 10.1109/ICDIM.2018.8846971

[51] OttM, Cardie C, Hancock J Estimating the prevalence of deception in online review communities. In: Proceedings of

the 21st international conference on World Wide Web, 2012; pp. 201–10. https://doi.org/10.1145/2187836.21878
[52] Meanin G, Mikolov ,Ranzato M ,and Bengio Y. Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews. arXiv preprint; 2014 https://doi.org/10.48550/arXiv.1412.5335
[53] Yang X and Yu X. Recognizing deceptive reviews based on weighted multi-instance unbalanced support vector machine. In: Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science, 2019, pp. 705–8. https://doi.org/10.1145/3349341.3349494
[54] Kennedy S, Walsh N, Sloka K, Mccarren A, and Foster J. Fact or factitious? Contextualized opinion spam detection. In: Proceedings of the 57th Annual Meeting of the association for computational linguistics: student research workshop, 2019. https:// 10.18653/v1/P19-2048
[55] Devlin J, Chang MW, Lee K, and Toutanova K. Bert: pre train- ing of deep bidirectional transformers for language understand- ing. 2018. https:// 10.18653/v1/N19-1423
[56] Sihombing A and Fong ACM. Fake review detection on yelp dataset using classification techniques in machine learning. In: 2019 International conference on contemporary computing and informatics (IC3I). IEEE, 2019, pp. 64–8. https:// 10.1109/IC3I46837.2019.9055644
[57] R. H. Al-Furaiji and H. Abdulkader, "A Comparison of the Performance of Six Machine Learning Algorithms for Fake News", EAI Endorsed Trans AI Robotics, vol. 3, Mar. 2024. https://doi.org/10.4108/airo.4153
[58] U. Tank, S. Arirangan, A. R. Paduri, and N. Darapaneni, "A Study Towards Building Content Aware Models in NLP using Genetic Algorithms", EAI Endorsed Trans AI Robotics, vol. 2, Nov. 2023. https://doi.org/10.4108/airo.4078
[59] M. Tyagi, P. K. Singh, S. K. Yadav, and S. K. Soni, "A Multi-Channel Spam Detection System Utilizing Natural Language Processing and Machine Learning", EAI Endorsed Trans AI Robotics, vol. 4, Mar. 2025. https://doi.org/10.4108/airo.8309