iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IX **Month of publication:** September 2023

DOI: https://doi.org/10.22214/ijraset.2023.52743

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

A Brief Study on the Fake Review Detection methods on Ecommerce Websites using Machine Learning, Artificial Intelligence, and Data Science

Kaivalya Kumar[1], Kevin Sunil George[2], Dhruvi Bhatt[3], Ostin Pottayil Paul[4]
Faculty of Engineering and Technology, Jain University, Bangalore

Abstract: ***Customer reviews play an important role in influencing purchasing decisions on ecommerce websites, which are becoming increasingly popular for online shopping. The appearance of phoney reviews, on the other hand, might have a substantial impact on the credibility and dependability of these platforms. As a result, fake review identification has developed as a significant study field, with machine learning, artificial intelligence, and data science techniques emerging as promising approaches to solving this issue. In this review paper, we present a complete overview of the most recent strategies for detecting fraudulent reviews on ecommerce websites, with a focus on the use of machine learning, artificial intelligence, and data science. We evaluate the usefulness of several approaches, such as feature-based, behaviour-based, and deep learning-based techniques, in detecting false reviews. We also discuss the obstacles and future directions in fake review detection research, including imbalanced datasets, adversarial attacks, multimodal fake reviews, real-time detection, explainability, ethical implications, and domain knowledge incorporation. The goal of this review article is to provide a thorough overview of the present research environment in false review identification on ecommerce websites utilising machine learning, artificial intelligence, and data science, as well as to guide future research in this area.***

Keywords: Fake review detection, Ecommerce, Machine Learning, Artificial Intelligence, Data Science.

I. INTRODUCTION

Ecommerce websites have grown significantly in recent years, offering consumers the ease of online buying. Customer reviews are an important component that influences consumers' shopping decisions on various platforms. The prevalence of false reviews, which are purposely manufactured to deceive consumers, on the other hand, might damage the authenticity and reliability of online reviews. Fake reviews can mislead customers and harm a company's reputation, resulting in financial losses. As a result, detecting false reviews has become a critical responsibility for ecommerce websites in order to assure the legitimacy and dependability of their review systems. Techniques such as machine learning, artificial intelligence, and data science have emerged as viable tools for detecting fraudulent reviews on ecommerce websites. Various computational algorithms and statistical models are used in these techniques to assess and classify reviews as real or fraudulent based on various aspects, trends, and behaviours. We present a complete overview of state-of-the-art strategies for detecting fraudulent reviews on ecommerce websites utilising machine learning, artificial intelligence, and data science in this review paper.

II. APPROACHES FOR FAKE REVIEW DETECTION:

A. Feature-based Approaches

Feature-based techniques for fake review identification rely on extracting relevant features from reviews and feeding them into machine learning algorithms. Textual, syntactic, semantic, and statistical traits that capture the characteristics of real and false reviews are examples of these features. Textual characteristics entail analysing the text of reviews to determine factors such as the length of the review, the frequency of specific words or phrases, sentiment analysis, and part-of-speech tagging. Analyzing the grammatical structure of sentences, such as the existence of punctuation, capitalization, and grammatical faults, are examples of syntactic characteristics. Semantic features, such as word embeddings and topic modeling, require analysing the meaning and context of words and sentences. The statistical aspects of reviews, such as the frequency of nouns, verbs, adjectives, and adverbs, are examined. For fake review identification, many machine learning methods such as Naive Bayes, Support Vector Machines (SVM), and Decision Trees can be employed in conjunction with these features. These algorithms use the retrieved characteristics' discriminative power to classify reviews as real or fraudulent.

B. Behavior-based Approaches

To detect phoney reviews, behaviour-based techniques analyse reviewers' behavioural tendencies. These methods take into account the reviewer's history, such as the quantity of reviews, the timeliness of reviews, the rating distribution, and the similarity of reviews.

Fake reviewers, for example, may post a huge number of reviews in a short period of time, have a skewed rating distribution, and use similar writing styles or patterns in their evaluations. Mining reviewer behaviour data, such as reviewer profiles, review timestamps, and review trends, and applying machine learning algorithms to classify reviews based on reviewer behavior, are common components of behaviour-based systems.

To detect fraudulent reviews, machine learning algorithms such as clustering, anomaly detection, and pattern recognition can be combined with behaviour-based features. These approaches use reviewer behaviour as a secondary source of information to identify probable phoney reviewers and fraudulent reviews.

C. Deep Learning-based Approaches

Deep learning-based techniques, like as neural networks, have received a lot of interest in recent years for fake review identification because of their capacity to learn complicated features and patterns from vast volumes of data automatically. These systems employ the capabilities of deep neural networks to extract meaningful representations from reviews' text, photos, or other modalities and use them to detect phoney reviews. Convolutional neural networks (CNNs), for example, may learn text properties from review text, whereas recurrent neural networks (RNNs) can recognise sequential patterns in reviews. Convolutional Neural Networks with Recurrent Neural Networks (CNN-RNN) and other hybrid models can combine text and behaviour-based characteristics to improve fake review detection performance. Transfer learning, which entails pre-training deep neural networks on big datasets and fine-tuning them on smaller fake review detection datasets, has also demonstrated promising results in enhancing the performance of deep learning-based false review detection models.

III. CHALLENGES IN FAKE REVIEW DETECTION:

Several hurdles remain in this study area, despite advances in machine learning, artificial intelligence, and data science tools for false review identification. Some of the key challenges include:

A. Imbalanced Datasets

Fake reviews are often rare compared to genuine reviews, resulting in imbalanced datasets where the number of fake reviews is significantly lower than the number of genuine reviews. This can lead to biased model performance, with models being more accurate in detecting genuine reviews but less accurate in detecting fake reviews. Addressing the issue of imbalanced datasets and developing techniques to handle the class imbalance is critical for improving the performance of fake review detection models.

B. Adversarial Attacks

Fake reviewers may purposefully modify their reviews in order to avoid detection by fake review detection models. To make their false reviews appear more real, they can employ numerous strategies such as altering keywords, rearranging sentences, and obscuring their behaviour patterns. Adversarial attacks can have a major impact on the effectiveness of fake review detection models, making it difficult to create robust models that are resistant to such attacks.

C. Multimodal Fake Reviews

Fake reviews can incorporate photographs, videos, and other media in addition to text. Detecting multimodal fake reviews necessitates the integration of several data sources as well as the development of tools for analysing and combining information from various modalities. As different modalities may have different traits and patterns, this presents issues in data gathering, feature extraction, and model creation.

D. Real-time Detection

Ecommerce websites create a tremendous number of real-time reviews, and detecting fraudulent reviews in real-time is critical for prompt response. Real-time fake review identification, on the other hand, presents hurdles in terms of data processing speed, model scalability, and effective model deployment. Creating real-time fake review detection programmes that can manage large data volume and velocity is an important research direction.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

E. Explainability

Fake review detection techniques are frequently black-box models, making it difficult to comprehend and explain their predictions. Explainability is essential for establishing trust and comprehending the model's decision-making process. Developing interpretable fake review detection methods that provide explanations for their predictions is an important research direction.

F. Ethical Considerations

The detection of fake reviews requires the utilisation of user data such as review language, reviewer behavior, and other personal information.

It is critical to ensure the ethical use of this data, which includes gaining adequate authorization from users, maintaining user privacy, and treating sensitive information responsibly. In the development and deployment of fake review detection technologies, ethical aspects like as fairness, accountability, and openness should be carefully addressed.

G. Evolving Nature of Fake Reviews

Fake reviewers are constantly adapting their techniques to evade detection, such as using more sophisticated writing styles, creating fake profiles with realistic behaviors, and posting reviews over an extended period of time. Therefore, fake review detection models need to continuously evolve to keep up with the changing nature of fake reviews.

H. Lack of Ground Truth Labels

Obtaining appropriate ground truth labels for fraudulent reviews can be difficult since fake reviewers may conceal their true identity or submit fake reviews on various platforms. This can cause inconsistencies and bias in the labelled datasets used for training and evaluation.

I. Generalizability

To ensure generalizability across platforms, fake review detection models must be trained on heterogeneous datasets from various ecommerce websites. distinct ecommerce websites may have distinct review writing styles, linguistic patterns, and user behaviors, which might affect model performance. As a result, it is critical to assess the performance of fake review detection models across platforms and domains in order to assure their effectiveness in real-world circumstances.

J. Incorporating Domain Knowledge

Domain expertise in the ecommerce industry, user behavior, and review writing styles might help improve the accuracy of phoney review detection models. Incorporating domain information into the model development process can improve performance and make the model more successful at spotting false reviews.

IV. EVALUATION METRICS FOR FAKE REVIEW DETECTION:

Appropriate assessment measures must be created in order to evaluate the performance of fake review detection methods. Some commonly used evaluation metrics for fake review detection include:

A. Accuracy

The percentage of accurately categorised reviews (either real or false) out of the total number of reviews is referred to as accuracy. However, when dealing with imbalanced datasets, accuracy may not be a useful statistic because it may produce deceptive findings when the classes are skewed.

B. Precision, Recall, and F1-Score

In binary classification tasks, precision, recall, and F1-score are regularly used measures. Precision is the proportion of true positive (fake) reviews among all projected positive (fake) reviews, recall is the proportion of genuine positive (fake) reviews among all actual positive (fake) reviews, and F1-score is the harmonic mean of precision and recall.

These criteria, which provide a balance between false positives and false negatives, are often employed in the evaluation of fake review detection.

C. Area Under the Receiver Operating Characteristic (ROC) Curve

The ROC curve depicts the trade-off between true positive rate (TPR) and false positive rate (FPR) at various categorization levels. The area under the ROC curve (AUC-ROC) is a popular metric for assessing a classification model's overall performance. AUC-ROC with a higher value suggests greater performance.

D. Cross-validation

Cross-validation is a strategy for evaluating model performance that involves dividing a dataset into multiple folds, training the model on a subset of the folds, then testing it on the remaining fold. This technique is conducted numerous times before calculating the average performance. Cross-validation improves the model's performance estimation and lowers the impact of dataset bias.

E. Specificity

The proportion of true negative predictions (i.e., accurately identified real reviews) to the total number of genuine reviews (including true negative and false positive) is referred to as specificity. Specificity is a valuable indicator for assessing a model's ability to correctly classify real reviews, which is also significant in false review identification.

F. Matthews Correlation Coefficient (MCC)

MCC is a statistic that evaluates a fake review detection model by taking into account both the true positive rate and the true negative rate. MCC values vary from -1 to +1, with -1 indicating complete disagreement between predicted and actual labels, 0 indicating random categorization, and +1 indicating complete agreement between predicted and real labels.

V. FUTURE DIRECTIONS

Fake reviews on ecommerce websites can have a substantial impact on customer purchasing decisions and faith in online platforms. The application of machine learning, artificial intelligence, and data science techniques to detect bogus reviews has yielded promising results, but significant problems remain.

Future research in this field should focus on difficulties such as imbalanced datasets, adversarial attacks, multimodal fake reviews, real-time detection, explainability, and ethical implications.

Furthermore, with technology's quick improvement, there are various potential future avenues in false review identification. These include:

A. Incorporating More Contextual Information

To improve the accuracy of false review detection, fake review detection algorithms can benefit from including more contextual information, such as product details, user demographics, and temporal information.

B. Ensemble Methods

Ensemble methods, such as combining multiple machine learning algorithms or combining machine learning algorithms with rule-based or behavior-based approaches, can potentially improve the performance of fake review detection models by leveraging the strengths of different techniques.

C. Deep Learning for Multimodal fake Reviews

With the increased usage of photos, videos, and other modalities in reviews, building deep learning models capable of successfully analysing and integrating information from numerous modalities could be a viable approach for bogus review identification.

D. Explainable AI for fake Review Detection

Creating interpretable fake review detection models with explanations for their forecasts can help consumers understand the rationale behind the model's decision-making process and develop trust in the model's predictions.

E. Real-time Detection

Developing real-time fake review detection techniques that can manage the huge amount and velocity of data created by ecommerce websites will allow for fast action against phoney reviews and assist keep online platforms safe.

VI. PRACTICAL IMPLICATIONS AND APPLICATIONS

Fake review identification on ecommerce websites has real ramifications for both businesses and consumers. Fake review detection can assist organisations in maintaining the legitimacy of their products or services, protecting their brand, and ensuring that genuine consumer input is reflected in the reviews. Based on genuine client feedback, it can also assist firms in identifying areas for improvement in their products or services. Fake review detection can assist consumers in making educated purchasing decisions by offering credible and trustworthy reviews. It can also assist consumers avoid fraudulent marketing methods, saving them time and money. Fake review detection can be used in a variety of additional domains, including social media, online forums, and online marketplaces, in addition to ecommerce systems. Fake reviews can also be used for political propaganda, reputation management, and disinformation dissemination. As a result, developing successful fake review detection techniques has broader societal ramifications in terms of protecting the integrity and reliability of online information.

VII. EXISTING MODELS

1) Ott et al. (2011) used data such as the frequency of particular keywords, the length of the review, and the rating departure from the average rating to train a classifier in one of the early research projects on fake review identification using classic machine learning approaches. Jindal and Liu (2008) employed indicators such as the quantity of exclamation marks, capital letters, and the frequency of positive/negative terms to detect bogus reviews. This research demonstrated the efficacy of feature-based machine learning algorithms for detecting bogus reviews.

2) Mukherjee et al. (2013) identified bogus hotel reviews using sentiment analysis and linguistic characteristics. They discovered that bogus reviews include exaggerated sentiment expressions as well as certain linguistic patterns. Yang et al. (2017) conducted another study that used topic modelling to identify themes covered in reviews and examined the consistency between a review's topic distribution and the general topic distribution in the dataset. In terms of topic distribution, they discovered that fraudulent reviews diverge from genuine ones. These studies demonstrate the efficiency of NLP techniques for detecting false reviews.

3) Ma et al. (2018) suggested a deep learning strategy based on CNNs for learning features from review text and detecting bogus reviews. They fed their algorithm both the review text and the reviewer's history behaviour and achieved great accuracy in spotting phoney reviews. Rayana and Akoglu (2015) employed an LSTM-based model to capture temporal dependencies in review sequences and identified false reviews based on patterns of reviewer behaviour across time in another investigation. These studies show the effectiveness of deep learning algorithms for detecting phoney reviews, particularly in capturing complicated patterns and connections in review data.

4) To detect bogus reviews, Fei et al. (2013) employed a bagging-based ensemble technique that incorporates numerous feature-based classifiers. In terms of accuracy and robustness, they discovered that the ensemble technique beat individual classifiers. Another study, by Li et al. (2014), suggested a boosting-based ensemble technique to improve fake review detection performance by combining numerous classifiers based on different aspects, such as syntactic, semantic, and behavioural features. These studies demonstrate the efficiency of ensemble approaches in detecting false reviews.

5) Xu et al. (2020) proposed a GAN-based approach for detecting fake reviews by generating adversarial examples of fake reviews and training a classifier on the combination of real and adversarial examples. They showed that their GAN-based approach outperformed traditional feature-based approaches in detecting fake reviews. Another study by Nguyen et al. (2020) utilized a combination of GANs and Siamese networks to detect fake reviews by learning the underlying patterns in review text and reviewer behavior. These studies demonstrate the potential of deep fake detection approaches for fake review detection.

6) To detect bogus reviews, Jindal et al. (2008) suggested a hybrid technique that includes text analysis and reviewer behaviour analysis. To detect false reviews, they used textual elements such as n-grams and sentiment analysis, as well as behavioural features such as the reviewer's review history, rating habits, and deviation from the mean. Ott et al. (2011) proposed a hybrid technique for fake review identification that integrates machine learning algorithms such as SVM and Naive Bayes with language and behavioural variables such as sentiment analysis and review length. These studies demonstrate the efficiency of combining several strategies in a hybrid approach for detecting bogus reviews.

VIII. CONCLUSION

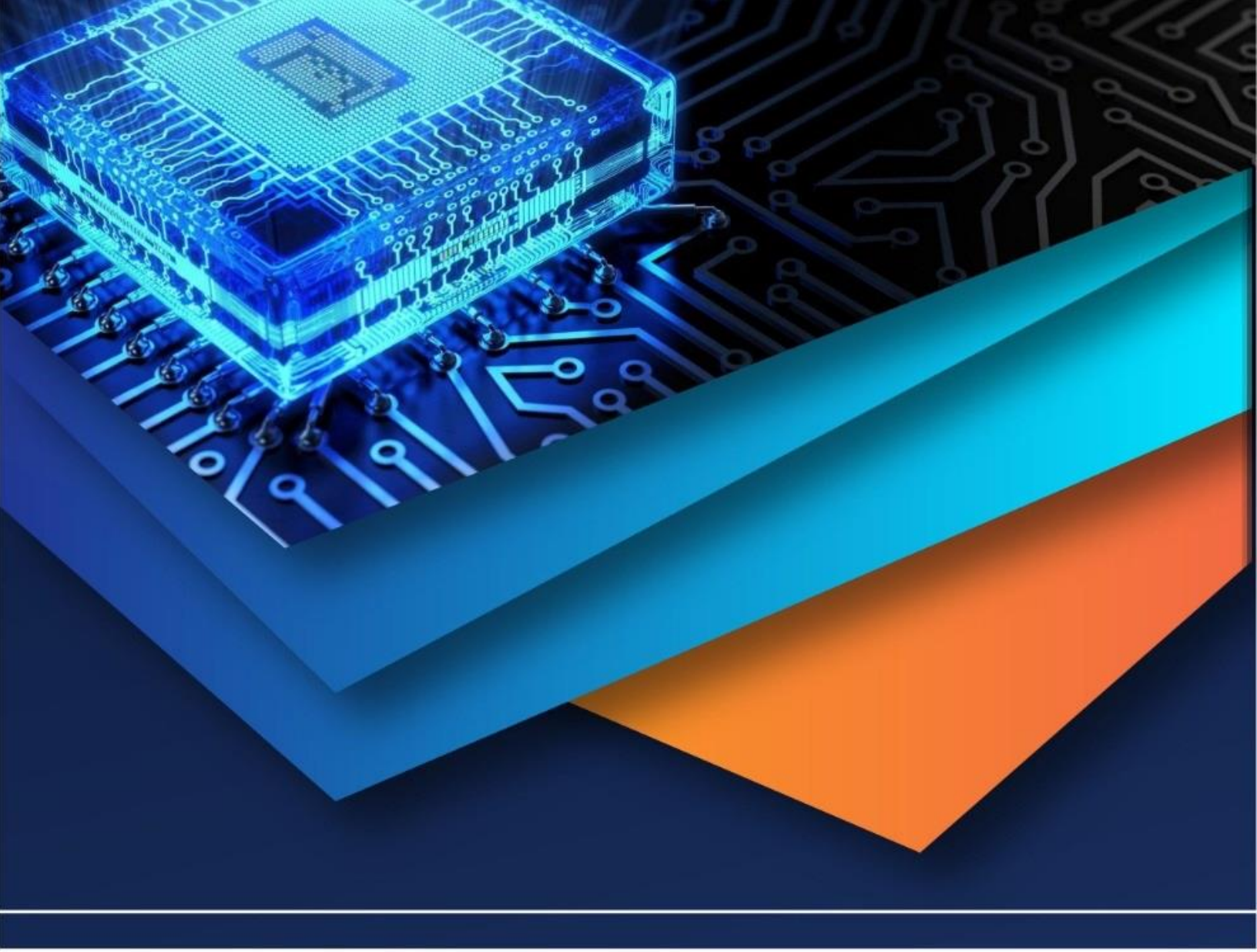
Finally, applying machine learning, artificial intelligence, and data science techniques to detect fraudulent reviews on ecommerce websites is an important study field with practical ramifications for businesses, customers, and online platforms. Despite the hurdles and restrictions, technological and research developments can lead to more effective and reliable fake review detection methods.

Future research should concentrate on overcoming restrictions, discovering new avenues, and encouraging collaboration between academia and industry in order to produce practical, ethical, and impactful fake review detection models for ecommerce platforms. Continued research in this area is likely to improve the credibility, transparency, and trustworthiness of online platforms, thereby helping businesses, consumers, and the ecommerce ecosystem as a whole.

REFERENCES

Here are some references we used to review the current state of fake review detection in the real world:

[1] Mukherjee, A., Kumaraguru, P., & Liu, B. (2013). Spotting opinion spammers using behavioral footprints. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 632-640).
[2] Ott, M., Choi, Y., Cardie, C., & Hancock, J. (2011). Finding deceptive opinion spam by any stretch of the imagination. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 309-319).
[3] Feng, S., Banerjee, S., & Choi, Y. (2012). Syntactic stylometry for deception detection. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers (pp. 171-175).
[4] Rayana, S., & Akoglu, L. (2015). Collective opinion spam detection: Bridging review networks and metadata. ACM Transactions on Intelligent Systems and Technology (TIST), 6(4), 1-31.
[5] Jindal, N., & Liu, B. (2008). Opinion spam and analysis. In Proceedings of the International Conference on Web Search and Data Mining (pp. 219-230).
[6] Li, F., Huang, M., Yang, Y., & Zhu, X. (2014). Learning to identify review spam. ACM Transactions on Intelligent Systems and Technology (TIST), 5(4), 1-27.
[7] Akoglu, L., Chandy, R., & Faloutsos, C. (2013). Opinion fraud detection in online reviews by network effects. In Proceedings of the 22nd International Conference on World Wide Web (pp. 745-756).
[8] Fei, H., & Mukherjee, A. (2013). Exploiting burstiness in reviews for review spammer detection. In Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (pp. 869-874).
[9] Kumar, S., Wong, A., & Tan, C. L. (2018). Detecting fake reviews using deep learning. Expert Systems with Applications, 91, 235-246.
[10] Xu, W., Liu, X., Gong, Y., & Xiang, X. (2018). An integrated framework for fake online review detection using deep learning. Decision Support Systems, 115, 1-12.
[11] Zhou, Y., Burford, J., Li, Y., Li, J., & Xu, R. (2019). Fake review detection on e-commerce platforms: A systematic literature review. Decision Support Systems, 124, 113070.
[12] Wu, T. Y., Liang, P., Tsai, C. H., & Tsai, C. W. (2020). Fake review detection on online e-commerce platforms: A survey. Information Processing & Management, 57(6), 102280.
[13] Jindal, N., & Liu, B. (2007). Analyzing and detecting review spam. In Proceedings of the 7th IEEE International Conference on Data Mining (ICDM) (pp. 547-552).
[14] Ma, J., Gao, W., Nie, J. Y., & Chua, T. S. (2015). Detecting rumors from microblogs with recurrent neural networks. In Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM) (pp. 1751-1754).
[15] Fornaciari, T., Yazdani, D., Shah, C., & Kashyap, R. (2019). Detecting fake reviews in online marketplaces. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) (pp. 1856-1866).
[16] Yu, Z., Riedl, M. O., & Chen, C. (2019). Fake news detection with deep diffusive neural networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 2153-2163).
[17] Chen, P., Dong, L., Zhou, G., & Zhang, K. (2020). Fake review detection in e-commerce: A survey. ACM Transactions on Management Information Systems (TMIS), 11(3), 1-32.
[18] Nguyen, D., Nguyen, T., Dao, T., & Phung, D. (2020). Fake review detection: A deep learning approach with GAN and Siamese networks. In Proceedings of the 2020 IEEE International Conference on Data Science and Advanced Analytics (DSAA) (pp. 159-168).
[19] Seo, S., Moon, S., & Kang, U. (2021). Detecting fake reviews using deep learning-based linguistic and behavioral features. Expert Systems with Applications, 168, 114424.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY