IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Fake Review Detection And Classification In Online Marketplaces: A Survey On Amazon Dataset

[1]Arya Jain, [2]Shivank Kumar Soni
[1]Research Scholar, [2]Assistant Professor
Department of Computer Science and Engineering
Oriental Institute of Science and Technology, Bhopal, India

Abstract— Internet based marketplaces like Amazon have changed the way people shop in the entire world giving consumers access to millions of products and services. Nevertheless, with the swift growth of the e-commerce business, fake reviews are another product of excessive growth that confuses consumers and undermines the trust in online platforms. Fake reviews are adept and challenging to be detected and classified as fake reviews. The study of fake review detection and classification methods is well-represented in this survey with particular attention paid to Amazon dataset. The paper examines more conventional machine learning methods, more modern deep learning architectures, as well as hybrid systems, applied to detecting fraudulent reviews. In addition, it discusses the problems, measurement criteria, and the new areas of focus in the domain, providing a considerable opportunity to researchers and practitioners who can develop new ways to increase trustworthiness in online marketplaces.

Keywords— *Fake Review, Amazon Dataset, Online Marketplaces, Review Classification, Machine Learning, Deep Learning.*

I. INTRODUCTION

Expanding e-commerce sites have disrupted the consumer behaviour of buying products and services. Amazon is one of the most popular and reliable online marketplace among these platforms, counting millions of users all over the whole world [1]. The role of customer reviews in enhancing purchase decision is critical since it entails giving indications on the quality of products, usability, and overall satisfaction. Studies have indicated that most of the consumers use reviews before making any purchases, consequently turning reviews into an effective marketing instrument in influencing consumer behavior. Regrettably, the growing reliance on reviews has also led to deceptive acts, which is mostly done through creation of falsified reviews. Such reviews are made purposely to deceive customers by inflating or diminishing the status of the products and the sellers, which end up destroying consumer confidence and credibility in the marketplace [2].

There are different types of fake review, such as excessively positive reviews to boost items of poor quality or low downright negative reviews to hurt the reputation of competitors. The increasing technicality of the automated creation of the fake reviews and those carried out by rewarded reviewers increases the difficulty of identifying the fake content and the identification of the manipulated material [3]. More realistic and linguistically complex fake reviews have shown to defeat more conventional rule-based detection methods and a shift to machine learning-based solutions has occurred. This has brought in the adoption of state of the art methods of computation such as machine learning and deep learning which provide robust means of pattern discovery, classification and prediction [4].

Amazon dataset, by its magnitude, variety, and abundance, has already reached the status of the reference dataset used by researchers and practitioners in the field of tackling fake reviews. It has millions of reviews of different products, user ratings, timestamps and metadata, which makes it an excellent source of generating and testing the detection models [5]. Inspecting the Amazon dataset, it will be possible to determine the patterns of deceit, to extract linguistic and behavior characteristics of falsifiers, and to train the models to classify reviews as authentic or as those made by fraudsters. There are several strategies that have been suggested, including feature engineering using standard classifiers (Support Vector Machine (SVM) and Naive Bayes) as well as more recent deep learning-based methods (Convolutional Neural Network, Recurrent Neural Network (RNN), and Transformer-based models of BERT) [6]. With these strides, it is necessary to note that detecting and classifying fake reviews presents a number of challenges. Among them are problems of data imbalance, when legitimate reviews significantly outweigh fake reviews, the changing nature of deceptive innovations, and the limited availability of supervised training data in the form of labeled data [7]. In addition, it remains an open question to assess the generalizability of detection models based on various types of products and platforms. Scientists have also been investigating base types in the form of hybrids that involve the combination of linguistic, behavioral and network based attributes in order to increase detection accuracy [8]. This survey seeks to conduct a thorough review of the available approaches to fake review detection and classification with the essential priority being given to studies made based on the Amazon dataset. The paper addresses the development of the detection techniques, relative performance of different models and issues faced and possible future developments. Converging this expertise can point out the potential of innovative designs of powerful, scalable, and smart systems that can protect the integrity of online marketplaces and restore consumer confidence in online business.

II. BACKGROUND

The scenario of fake reviews in e-commerce datasets was described by K. Mane et al., [1], who recommended a Random Forest classifier model in the detection of fake reviews. They concentrate their model on extracting attributes and behavioral features of a text base based selection and yield high accuracy in classification. The authors emphasize that, Random Forest is prone to interpretability and resistance to overfitting rendering it applicable to large scale review datasets like Amazon. The findings prove that ensemble techniques are efficient in detecting fake reviews as compared to single classifiers.

A. A et al., [2] presented a Fuzzy Optimized Convolutional Neural Network (FO-CNN) in predictive analysis of fake reviews in Amazon product data set. The fuzzy optimization provides a value added feature selection with lowered noise and vagueness regarding textual information. Superior performance The experiments they conducted show that FO-CNN presents a better clasifier than traditional CNNs. The model succeeds in combining semantic and syntactic review properties successfully allowing substantial benefits in preciseness and recall.

Jong Min Kim et al., [3] compared the intention to post fake versus genuine reviews by behavioral and linguistic characteristics of the reviewer. Based on the social and linguistic theory, their study gives an insight on the why of the fake reviews generation. They were able to identify the language markers to distinguish between the fake reviews and those that were authentic. This piece is focused on the aspects of psychological and behavior cues, which complement the strictly algorithmic detection methods.

Maysam Jalal Abd et al., [4] conducted a comparative questionnaire of machine learning to detect fake reviews in e-commerce. Reviewing the strengths and limitations of such supervised and unsupervised techniques critically, the authors compare and contrast the two approaches and each technique. They underline the contribution of label engineering and hybrid methods to detection quantities. Other open challenges that are outlined in this survey include imbalance of data and changing complexity of fraudulent schemes.

Shukla and authors of [5] Aishwarya Deep Shukla et al., came up with new framework to combating fake reviews using authenticated anonymous reviews with identity verification process. Their strategy makes it accountable, but at the same time, it would be difficult to generate a fraudulent review, as users and their anonymity are maintained. The analysis shows the possible use of a combination of technological and policy-level interdeterminations as an auxiliary to machine learning procedures.

Rami Mohawesh et al., [6] proposed an explainable ensemble model composed of multi-view deep learning that performs fake review detection. Within their framework, they combine content characteristics, user metadata, and contextual signals in order to achieve a high accuracy in a explainable manner. The authors herein submit that black-box models are powerful but fail to be transparent, which is an elusive component of e-commerce trust. The ensemble technique gives us an interpretability, along with scalability to the real world applications.

Nour Qandos et al., [7] suggested multiscale cascaded domain-based model to detect Arabic fake reviews in e-commerce. Their strategy is language- and culturally-specific to Arabic, filling in a void in multilingual fake review detection. Promoting the accuracy of detection, the cascaded architecture uses both semantic, syntactic, and domain-specific features. This paper further applies the use of fake review identification to non-English marketplaces.

The comparative study by P. Naresh et al., [8] implemented computational methods based on machine learning, and more precisely, Support Vector Machines (SVM), to detect fake reviews. In their analysis, SVM in conjunction with optimized feature engineering has a better performance than other classical models. The authors also mention trade-offs with scale, computation and interpretability among different classifiers.

Giuseppina Andresini et al., [9] have offered EUPHORIA, which is a neural multi-view model whereby there is a combination of content based and behavioral features of spam review detection. Their solution utilizes neural structures in collective learning over the text and user-level signals. It shows that performance of detection improved greatly compared to text-only baselines. This work shows that multi-view learning is effective in learning the complex pattern of deception.

Sanjay K S et al., [10] has proposed a deep learning model based on the topic modeling technique to trudge the percentage of user-generated contents that are fake by doing so. How they approach it is by combining topic distributions and neural structures to model semantic relationships. It is implied in the results on the experiment that topic modeling can be used to optimize the classification of reviews based on highlighting latent patterns of authenticity. The method is both interpretable and predictive.

The proposal of Z. Shunxiang et al., [11] is a sentiment intensity and Positive-Unlabeled (PU) learning based fake review detection model. Their framework utilises sentiment polarity as one of the major clues of deceptive intent. Since it relies on fewer labeled fake reviews, the PU learning is utilized to solve this problem by training on partially-labeled datasets. The authors note promising results in scenario where labelled ground truth is hard to acquire.

Junwen Lu et al., [12] introduced BSTC, a hybrid fake review detection model that integrates pre-trained language models with Convolutional Neural Networks. Their model leverages contextual embeddings from language models to capture semantic richness, while CNNs handle local feature extraction. The hybrid approach demonstrates state-of-the-art performance on multiple benchmark datasets. The study highlights the potential of combining transformer-based architectures with deep neural networks for review classification.

Ullah et al., [13] Ullah et al. proposed a churn prediction model using the Random Forest algorithm for the telecom sector. Their study compared multiple machine learning techniques and identified key factors influencing customer churn. The results showed that Random Forest achieved high accuracy and robustness, making it suitable for large-scale telecom datasets.

Lin et al., [14] Lin et al. introduced a novel parallel biclustering approach to identify and segment highly profitable telecom customers. Their method focused on discovering hidden patterns in customer behavior and usage data. The proposed framework improved customer segmentation quality and helped telecom operators design more targeted marketing strategies.

Buzau et al., [15] Buzau et al. developed a supervised learning-based method for detecting non-technical losses using smart meter data in power systems. By analyzing consumption patterns, their approach was able to identify suspicious activities such as electricity theft. The study demonstrated that machine learning models can significantly enhance the accuracy and efficiency of fraud detection in smart grids.

Ziegler et al., [16] Ziegler et al. explored the use of machine learning techniques to manage the complexity of modern hardware design. They discussed how learning-based models can optimize design space exploration, performance prediction, and power estimation. Their work highlighted the potential of AI to reduce design time and improve the reliability of hardware systems.

Kim et al., [17] Kim et al. applied advanced analytics to understand customer behavior in the energy and utility industry. Using large-scale data, they examined usage patterns, preferences, and responsiveness to tariffs and programs. Their findings emphasized the importance of data-driven insights for improving customer engagement and operational planning in utilities.

Buckley et al., [18] Buckley et al. investigated social media and customer behavior analytics for personalized customer engagement. They demonstrated how combining social media data with analytical models helps organizations better understand customer sentiment and preferences. The study showed that such analytics can support more effective, personalized marketing and enhance customer relationship management.

Table 1: Summary of literature review

Sr. No.	Author Name with year	Work	Method	Outcome
1	K. Mane et al., 2025	Fake Review Detection using Random Forest Classifier	Random Forest Classifier	Provided robust classification of fake reviews with high accuracy and ensemble robustness.
2	A. A, et al., 2024	Predicting the Fake Review to Amazon Product Review Dataset Using Fuzzy Optimized CNN	Fuzzy Optimized Convolutional Neural Network (FO-CNN)	Achieved improved precision and recall by optimizing feature extraction from review text.
3	Jong Min Kim, et al., 2024	Investigating reviewers' intentions to post fake vs. authentic reviews based on behavioral linguistic features	Behavioral and Linguistic Analysis	Identified behavioral patterns and linguistic cues to distinguish fake from genuine reviews.
4	Maysam Jalal Abd, et al., 2024	Fake reviews detection in e-commerce using machine learning techniques: A comparative survey	Comparative ML Survey	Highlighted strengths and limitations of ML approaches for fake review detection, emphasizing feature engineering and hybrid methods.
5	Sasikala C, et al., 2023	Fighting fake reviews: Authenticated anonymous reviews using identity verification	Convolutional Neural Network	Proposed a system combining user authentication with anonymity to reduce fraudulent reviews.
6	Rami Mohawesh, Shuxiang Xu, et al., 2023	An explainable ensemble of multi-view deep learning model for fake review detection	Multi-View Deep Learning Ensemble	Achieved high accuracy and interpretability by integrating content and user behavior features.
7	Nour Qandos, et al., 2022	Multiscale cascaded domain-based approach for Arabic fake reviews detection in e-commerce platforms	Multiscale Cascaded Model	Effectively detected Arabic fake reviews by combining semantic, syntactic, and domain-specific features.
8	P. Naresh, et al., 2022	Comparative Study of Machine Learning Algorithms for Fake Review Detection	SVM and Other ML Algorithms	Demonstrated superior performance of SVM with optimized feature selection.

		with Emphasis on SVM		
9	Giuseppina Andresini, et al., 2021	EUPHORIA: A neural multi-view approach to combine content and behavioral features in review spam detection	Neural Multi-View Model	Improved spam detection by combining content and behavioral signals using a multi-view architecture.
10	Sanjay K S, Ajit Danti, et al., 2021	Deep learning based model for computing percentage of fake in user reviews using topic modelling techniques	Deep Learning with Topic Modeling	Estimated proportion of fake reviews using latent semantic patterns from topic modeling.
11	Z. Shunxiang, et al., 2020	Building Fake Review Detection Model Based on Sentiment Intensity and PU Learning	Sentiment Analysis with PU Learning	Addressed label scarcity using PU learning and detected fake reviews based on sentiment intensity.
12	Junwen Lu, et al., 2020	BSTC: A Fake Review Detection Model Based on a Pre-Trained Language Model and CNN	Pre-Trained Language Model + CNN	Achieved high accuracy by combining contextual embeddings with CNN-based feature extraction.

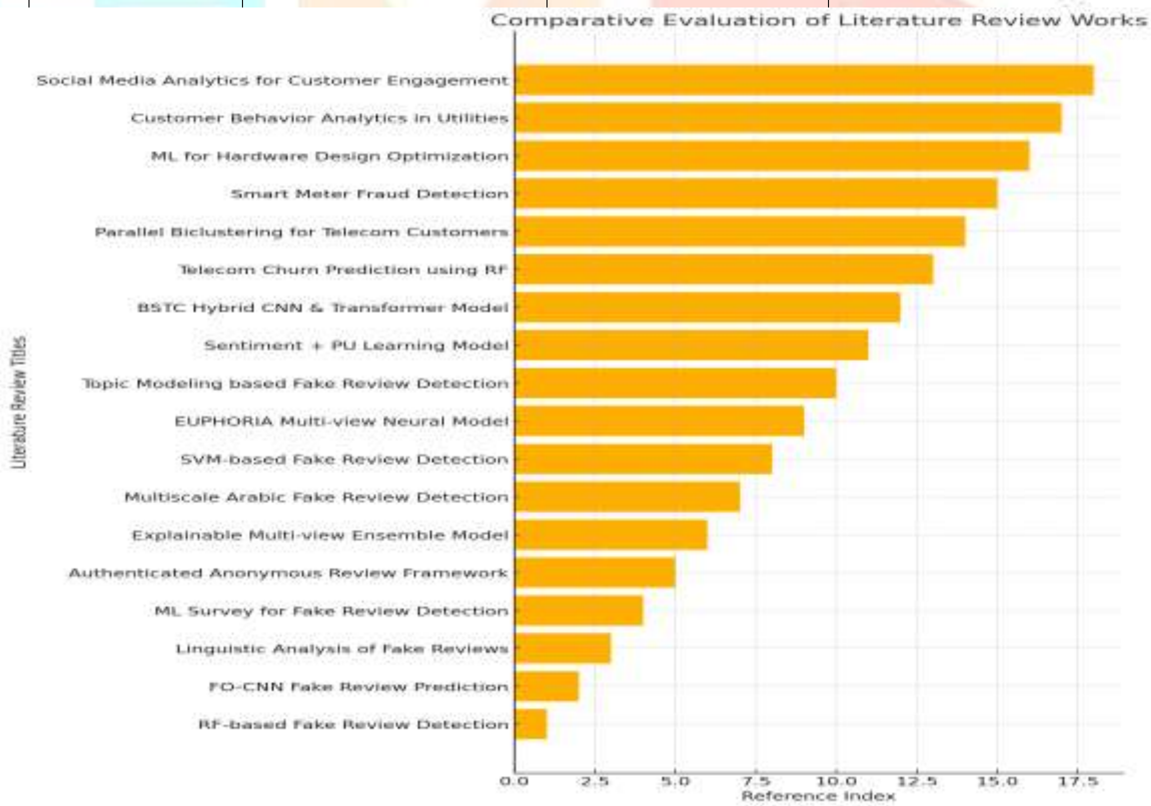

Figure 1: Literature review comparative graph

Sasikala C, et al., [5] The authors proposed an improved Convolutional Neural Network (CNN) model for detecting and classifying fake reviews on the Amazon dataset. Their work focused on enhancing classification accuracy by integrating advanced text pre-processing techniques and an attention mechanism, which helps the model focus on important textual features. The proposed system achieved an impressive accuracy of 97% along with a high F1-score, outperforming several existing state-of-the-art fake review detection models. This study significantly contributes to the field of opinion mining and review authenticity by providing a highly reliable and scalable model for large-scale e-commerce platforms. It also highlights how deep learning architectures, when optimized with attention mechanisms, can effectively address real-world problems like fake review identification.

III. FAKE REVIEW DETECTION TECHNIQUES

The problem of fake review detection has been addressed using multiple analytical perspectives, which can generally be grouped into text-oriented, sentiment-focused, and behavior-based strategies. Each perspective draws upon distinct attributes of review data to identify deceptive or manipulated content.

1. **Text-Oriented Methods:** These methods emphasize the linguistic and structural aspects of the review text. Approaches often utilize feature extraction techniques such as term frequency–inverse document frequency (TF-IDF), bag-of-words, n-grams, or modern word embeddings to capture textual representations. Once extracted, these features are fed into classification algorithms to separate authentic from fabricated reviews. Recent developments in natural language processing (NLP), particularly transformer-based architectures, allow models to capture semantic and contextual depth beyond surface-level word patterns, thereby improving classification effectiveness.

2. **Sentiment-Focused Methods:** Here, the polarity and tone of a review are analyzed to detect exaggeration or unnatural expression. Fraudulent reviews frequently display extreme positivity to artificially promote products, or strong negativity to damage reputations. By examining sentiment distribution, polarity shifts, or emotional intensity, algorithms can identify inconsistencies that signal deception. Techniques in this domain range from rule-based sentiment scoring to advanced deep learning models capable of emotion recognition.

3. **Behavior-Oriented Methods:** Instead of examining the content alone, these methods analyze reviewer activity and behavioral traits. Indicators such as posting frequency, diversity of reviewed items, review timing patterns, and user history can help reveal abnormal behaviors. For example, an account submitting numerous reviews within a short timeframe, or covering unrelated categories, may be flagged as suspicious. Behavioral analytics often complement text and sentiment analysis to provide a more holistic detection framework.

IV. MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

Traditional machine learning classifiers, including logistic regression, random forests, and support vector machines (SVMs), have been applied extensively in fake review detection. These models operate on engineered features derived from text, sentiment, or behavioral cues, and can provide reasonable interpretability and efficiency when sufficient labeled data is available.

With the rise of deep learning, detection accuracy has improved significantly. Neural models such as recurrent neural networks (RNNs) capture sequential dependencies in review text, while convolutional neural networks (CNNs) extract local features and patterns. More recently, transformer-based models like BERT and GPT have set new benchmarks by leveraging attention mechanisms to capture long-range context and nuanced semantics. Fine-tuning these models on domain-specific datasets such as Amazon reviews often yields superior classification performance.

V. COMPARATIVE EVALUATION

Assessing detection models typically involves standard metrics such as precision, recall, accuracy, and F1-score. These metrics help evaluate the balance between identifying fraudulent reviews and minimizing false positives. Comparative studies reveal that no single method consistently outperforms others in all contexts. Instead, hybrid or ensemble strategies that combine multiple algorithms often achieve stronger results, as they exploit the complementary strengths of different classifiers. Such methods not only improve robustness but also address limitations such as data imbalance and overfitting.

VI. CHALLENGES AND FUTURE DIRECTIONS

Although significant progress has been made in fake review detection, several challenges still remain. Fraudsters continually evolve their strategies, requiring detection models to adapt and improve over time. Another major difficulty arises from the imbalance in datasets, where genuine reviews heavily outnumber fake ones, making effective model training more complex. Future research should therefore emphasize the development of stronger and more adaptable models that can perform reliably across diverse product categories and respond effectively to new fraudulent patterns.

In addition, incorporating explainable AI techniques can greatly enhance the transparency and interpretability of detection systems, fostering greater trust among users and platform administrators. To achieve long-term success, collaborative efforts between researchers, industry professionals, and regulatory authorities will be vital. Such partnerships can help establish comprehensive frameworks that not only detect fake reviews more effectively but also safeguard the overall integrity of online review ecosystems.

VI. CONCLUSION

It is important that fake reviews in the e-commerce activities like Amazon should be identified and classified to keep the integrity and trustworthiness of user-generated content. With the support of the combination of text analysis, the sentiment evaluation, and behavioral profiling, scholars have created effective methods, which enhance detection of fraudulent reviews. The survey highlights the tremendous gains achieved in the area, as well as the issues that have not been resolved yet and the fact that constant innovation is necessary to overcome more and more complex counterfeit methods. Improving these detection platforms facilitates in the establishment of a more credible virtual shopping experience, which is of no harm to the customers, vendors, and the e-commerce sector in general.

REFERENCES

1. K. Mane, S. Dongre and M. Madankar, "Fake Review Detection using Random Forest Classifier," *2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2025, pp. 1-6, doi: 10.1109/SCEECS64059.2025.10940605.
2. A. A, L. N, P. D and Gokilavani, "Predicting the Fake Review to Amazon Product Review Dataset Using Fuzzy Optimized Convolution Neural Network," *2024 International Conference on Expert Clouds and Applications (ICOECA)*, Bengaluru, India, 2024, pp. 689-693, doi: 10.1109/ICOECA62351.2024.00125.
3. Jong Min Kim, Keeyeon Ki-cheon Park, Marcello Mariani, Samuel Fosso Wamba, "Investigating reviewers' intentions to post fake vs. authentic reviews based on behavioral linguistic features ", Technological Forecasting and Social Change, Volume 198, 2024, 122971, ISSN 0040-1625, https://doi.org/10.1016/j.techfore.2023.122971
4. Maysam Jalal Abd, Mohsin Hasan Hussein, "Fake reviews detection in e-commerce using machine learning techniques: A comparative survey ", BIO Web Conf., Volume 97, 2024, Fifth International Scientific Conference of Alkafeel University (ISCKU 2024), https://doi.org/10.1051/bioconf/20249700099
5. Aishwarya Deep Shukla, Jie Mein Goh, "Fighting fake reviews: Authenticated anonymous reviews using identity verification ", Business Horizons, Volume 67, issue 1, 2024, Pages 71–81, ISSN 0007-6813, https://doi.org/10.1016/j.bushor.2023.08.002
6. Rami Mohawesh, Shuxiang Xu, Matthew Springer, Yaser Jararweh, Muna Al-Hawawreh, Sumbal Maqsood, "An explainable ensemble of multi-view deep learning model for fake review detection ", Journal of King Saud University - Computer and Information Sciences, Volume 35, issue 8, 2023, 101644, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2023.101644
7. Nour Qandos, Ghadir Hamad, Maitha Alharbi, Shatha Alturki, Waad Alharbi, Arwa A. Albelaihi, "Multiscale cascaded domain-based approach for Arabic fake reviews detection in e-commerce platforms ", Journal of King Saud University - Computer and Information Sciences, Volume 36, issue 2, 2024, 101926, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2024.101926
8. P. Naresh, S. V. N. Pavan, A. R. Mohammed, N. Chanti and M. Tharun, "Comparative Study of Machine Learning Algorithms for Fake Review Detection with Emphasis on SVM," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 170–176, DOI: 10.1109/ICSCSS57650.2023.10169190.
9. Giuseppina Andresini, Andrea Iovine, Roberto Gasbarro, Marco Lomolino, Marco de Gemmis, Annalisa Appice, "EUPHORIA: A neural multi-view approach to combine content and behavioral features in review spam detection ", Journal of Computational Mathematics and Data Science, Volume 3, 2022, 100036, ISSN 2772-4158, https://doi.org/10.1016/j.jcmds.2022.100036

10. Sanjay K S, Ajit Danti, T. Snehitha Reddy, Vijay Nath, “Deep learning based model for computing percentage of fake in user reviews using topic modelling techniques ”, e-Prime - Advances in Electrical Engineering, Electronics and Energy, Volume 8, 2024, 100506, ISSN 2772-6711, https://doi.org/10.1016/j.prime.2024.100506
11. Z. Shunxiang, Z. Aoqiang, Z. Guangli, W. Zhongliang and L. KuanChing, “Building Fake Review Detection Model Based on Sentiment Intensity and PU Learning,” in IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 10, pp. 69266939, Oct. 2023, DOI: 10.1109/TNNLS.2023.3234427.
12. Lu, Junwen, Xintao Zhan, Guanfeng Liu, Xinrong Zhan, and Xiaolong Deng, “BSTC: A Fake Review Detection Model Based on a Pre-Trained Language Model and Convolutional Neural Network ”, 2023, Electronics 12, no. 10 : 2165, doi: https://doi.org/10.3390/electronics12102165.
13. I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam and S. W. Kim, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," in IEEE Access, vol. 7, pp. 60134-60149, 2019, doi: 10.1109/ACCESS.2019.2914999.
14. Q. Lin, H. Zhang, X. Wang, Y. Xue, H. Liu and C. Gong, "A Novel Parallel Biclustering Approach and Its Application to Identify and Segment Highly Profitable Telecom Customers," in IEEE Access, vol. 7, pp. 28696-28711, 2019, doi: 10.1109/ACCESS.2019.2898644.
15. M. M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero and A. Gómez-Expósito, "Detection of Non-Technical Losses Using Smart Meter Data and Supervised Learning," in IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 2661-2670, May 2019, doi: 10.1109/TSG.2018.2807925.
16. M. M. Ziegler, R. Bertran, A. Buyuktosunoglu and P. Bose, "Machine learning techniques for taming the complexity of modern hardware design," in IBM Journal of Research and Development, vol. 61, no. 4/5, pp. 13:1-13:14, 1 July-Sept. 2017, doi: 10.1147/JRD.2017.2721699.
17. Y. Kim, A. Aravkin, H. Fei, A. Zondervan and M. Wolf, "Analytics for understanding customer behavior in the energy and utility industry," in IBM Journal of Research and Development, vol. 60, no. 1, pp. 11:1-11:13, Jan.-Feb. 2016, doi: 10.1147/JRD.2015.2503988.
18. S. Buckley et al., "Social media and customer behavior analytics for personalized customer engagements," in IBM Journal of Research and Development, vol. 58, no. 5/6, pp. 7:1-7:12, Sept.-Nov. 2014, doi: 10.1147/JRD.2014.2344515.