IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

FAKE REVIEW DETECTION SYSTEM USING SUPERVISED MACHINE LEARNING

[1]Prof. Shekhar Patle, [2]Gaurav Pawar, [3]Shrinivas Pawar, [4]Onkar Davkare, [5]Rutuja Javalekar

[1]Guide, [2345]Students

[12345]Department of Information Technology,

Zeal College of Engineering and Research, Pune, Maharashtra, India.

Abstract: The rise of online platforms and e-commerce has revolutionized consumer behavior, making product reviews a vital source of information for purchasing decisions. However, the prevalence of fake reviews has undermined the credibility and trustworthiness of online reviews, leading to the need for effective fake review detection systems. This research paper presents a novel approach to address this challenge by leveraging supervised machine learning techniques for the detection of fake reviews. The proposed system begins by constructing a comprehensive dataset consisting of genuine and fake reviews, along with relevant features such as review text, reviewer information, and rating patterns. These features are carefully selected to capture the distinguishing characteristics of fake reviews, including the presence of biased sentiments, unnatural language patterns, and inconsistent reviewer behavior. A supervised machine learning model, such as a support vector machine (SVM), KNN,Logistic Regression is trained on the labeled dataset to learn the complex patterns and relationships between the review features and their authenticity. The model undergoes an iterative process of feature engineering, selection, and hyperparameter tuning to optimize its performance.

Index Terms **– Supervised Machine Learning, SVM, KNN, Logistic Regression, Labeled Dataset, Hyperparameter, Fake Reviews.**

I. INTRODUCTION

In the rapidly growing digital era, online reviews have become a crucial source of information for consumers when making purchasing decisions. However, the increasing prevalence of fake reviews poses a significant challenge to the credibility and reliability of these platforms. As one of the largest e-commerce sites globally, Amazon attracts a vast number of product reviews, making it a prime target for fake review manipulation. To address this issue, the development of a robust and accurate fake review detection system has become imperative.

This research paper aims to propose a supervised machine learning approach for detecting fake reviews on the Amazon platform. By leveraging the power of machine learning algorithms, we seek to provide a reliable solution that enhances the authenticity of customer feedback and promotes fair and informed consumer choices. By utilizing a supervised learning framework, our proposed system will be trained on a labeled dataset consisting of authentic and fake reviews. The system will extract meaningful features from the reviews, such as linguistic patterns, sentiment analysis, and reviewer behavior, to develop a comprehensive model capable of accurately distinguishing genuine reviews from fake ones.

Amazon serves as an ideal reference for this research due to its vast product range, global popularity, and the prevalence of both authentic and fake reviews. Analyzing the review ecosystem on Amazon will enable us to explore the characteristics and patterns associated with fraudulent reviews and evaluate the effectiveness of our proposed detection system.

The outcomes of this research have the potential to benefit various stakeholders in the e-commerce industry, including consumers, online marketplaces, and regulators. By shedding light on the identification and mitigation of fake reviews, this study contributes to the ongoing efforts towards maintaining the integrity and trustworthiness of online platforms.

In the subsequent sections of this paper, we will present a detailed methodology for data collection and preprocessing, describe the employed machine learning algorithms, discuss the experimental results, and provide an analysis of the system's performance. Finally, we will offer insights into the implications of our findings, potential limitations, and future research directions.

Through this research, we aim to advance the field of fake review detection and pave the way for more effective measures against deceptive practices in online reviews, ultimately promoting transparency and consumer confidence in the e-commerce domain.

II. RELATED WORKS

In the [1] study, a fake review detection system was developed using a feature framework that categorized reviews into two types. Logistic Regression was applied to analyze user-centric features in Amazon electronic product reviews, achieving an accuracy of 86% as measured by the F-score. The system utilized Opinion-Mining, employing Sentiment Analysis to identify false reviews. The researchers collected annotations for individual reviews in the dataset, and Sentiment Analysis was implemented using Vader to determine the sentiment polarity of the text passages (positive, negative, or neutral). Sentiment analysis methods generally fall into valence-based or polarity-based categories, where the strength of sentiment is considered. For example, in a valence-based approach, "excellent" is considered more positive than "good," while in a polarity-based approach, they are treated equally.

Another study [3] focused on two main types of fake reviews: textual (based on the content of the reviews) and behavioral (based on the writing style, emotional expressions, and writing frequency of the reviewer). Machine learning algorithms were employed to differentiate between authentic and fake reviews. The identification process involved considering both "important elements of the review" and "reviewer behavior." Logistic Regression achieved an 86% accuracy in bi-gram analysis without considering behavioral characteristics, KNN achieved a 73% accuracy in tri-gram analysis, and SVM achieved an 88.1% accuracy in bi-gram analysis.

In the context of fake hotel reviews on platforms such as Yelp and TripAdvisor, a system was developed [4] with a web crawler that collected review data and stored it in a MySQL database. Four techniques were utilized to detect fake reviews: text mining-based categorization, spell checking, reviewer behavior analysis, and hotel environment analysis. A grading algorithm computed the overall probability of fraudulent reviews for a specific hotel by combining individual probabilities from these techniques. The study followed standard pre-processing recommendations when implementing the text mining-based detection method. The percentage of fake reviews found was approximately 14%, and previous research has validated the reliability of this data source for determining hotel veracity.

The [5] study focuses on the utilization of Graph Convolutional Networks (GCNs) to identify fake reviews specifically on the Amazon platform. By constructing a graph representation of the review data, incorporating both textual content and reviewer information, the proposed GCN model aims to effectively capture the intricate relationships within the review network and enhance the accuracy of fake review detection. This work showcases the potential of deep learning techniques, particularly GCNs, in addressing the challenge of identifying fake reviews in e-commerce platforms, providing valuable insights for further research in the field.

Offers [6] a comprehensive survey of detection techniques employed in combating fake reviews. The paper provides a valuable overview of various approaches and methods utilized to identify and mitigate the prevalence of fake reviews across different platforms. It covers a wide range of techniques, including content-based analysis, sentiment analysis, linguistic pattern analysis, reviewer behavior analysis, machine learning-based approaches, and crowdsourcing-based methods. By summarizing the strengths and limitations of each technique, the study offers insights into the current state-of-the-art and provides a foundation for further research in the field of fake review detection. This survey serves as a valuable resource for researchers and practitioners seeking to develop effective strategies to combat the growing issue of fake reviews in online platforms.

III. EXISTING MODEL

The paper [7] introduces a deep neural network-based model for fake news detection, which can also be adapted for fake review detection. The model combines convolutional neural networks (CNN) and long short-term memory (LSTM) networks to capture both local and global contextual information from textual data. The CNN component enables the model to extract informative features from the reviews, while the LSTM component helps capture temporal dependencies within the text. By incorporating these two neural network architectures, the model learns to distinguish between genuine and fake reviews based on the inherent patterns and linguistic cues.

To train the model, the research paper employs a large dataset of labeled reviews, consisting of both genuine and fake examples. The model undergoes a supervised learning process, where it is trained to classify reviews as either authentic or fake. Through extensive experiments and evaluations, the proposed model achieves high accuracy in fake review detection, outperforming traditional machine learning algorithms and other baseline methods.

The research contributes to the development of effective models for detecting fake reviews by leveraging the power of deep neural networks and their ability to learn intricate patterns and representations from textual data. The model's architecture and training approach can serve as a valuable reference for researchers and practitioners working on fake review detection systems.

Another paper [8] proposes a model that combines attention-based Long Short-Term Memory (LSTM) networks with textual features for the detection of review spam. The attention mechanism helps the model focus on important parts of the reviews, improving its ability to identify suspicious or fake content. The model utilizes various textual features, including n-grams, sentiment information, and syntactic patterns, to capture different aspects of the reviews and extract meaningful representations. Through experiments conducted on large-scale review datasets, the proposed model demonstrates superior performance in detecting review spam compared to several baseline methods. The research contributes to the advancement of fake review detection techniques by leveraging attention mechanisms and LSTM networks for improved accuracy and robustness.

The main methodology for detecting fake reviews is generally same in most of the past research paper that had been published. Most of the research papers have came with the idea of almost same model but what separates them is that which technologies they are using for detecting fake reviews. The paper [7] uses deep neural network based model while paper [8] uses Long Short-Term Memory(LSTM). Most of the models are based on the deep neural network but this paper focuses on the Supervised Machine Learning approach.

IV. PROPOSED MODEL

To identify the most effective model for achieving high accuracy and fast processing speed, the approach consists of five key stages, which are outlined in detail :-

4.1 Data Collection

For the collection of consumer review data, we obtained the raw review data from Kaggle's Amazon datasets reviews. This choice was made to ensure a wider diversity of review data from various sources. A dataset comprising 40,000 reviews was collected, allowing for a substantial amount of data to be used in our analysis. By incorporating a large number of reviews, we aimed to enhance the representativeness and inclusivity of the dataset, capturing a broader range of consumer perspectives and experiences. This dataset will serve as a valuable resource for our research, enabling us to perform comprehensive analyses and develop robust models for consumer review analysis.

4.2 Data Preprocessing

Data preprocessing plays a crucial role in preparing the collected data for further analysis and modeling. In this stage, the raw data obtained from web scraping or other sources undergoes several preprocessing steps to enhance its quality and compatibility with the selected models. This includes tasks such as data cleaning, which involves removing irrelevant or redundant information, handling missing values, and addressing inconsistencies or errors in the data. Additionally, data normalization or standardization techniques may be applied to ensure that the data is on a consistent scale and distribution. Textual data may undergo preprocessing steps like tokenization, stop-word removal, and stemming or lemmatization to extract meaningful features and reduce dimensionality. Overall, data preprocessing aims to improve the quality and reliability of the data, enabling more accurate and efficient analysis and modeling processes.

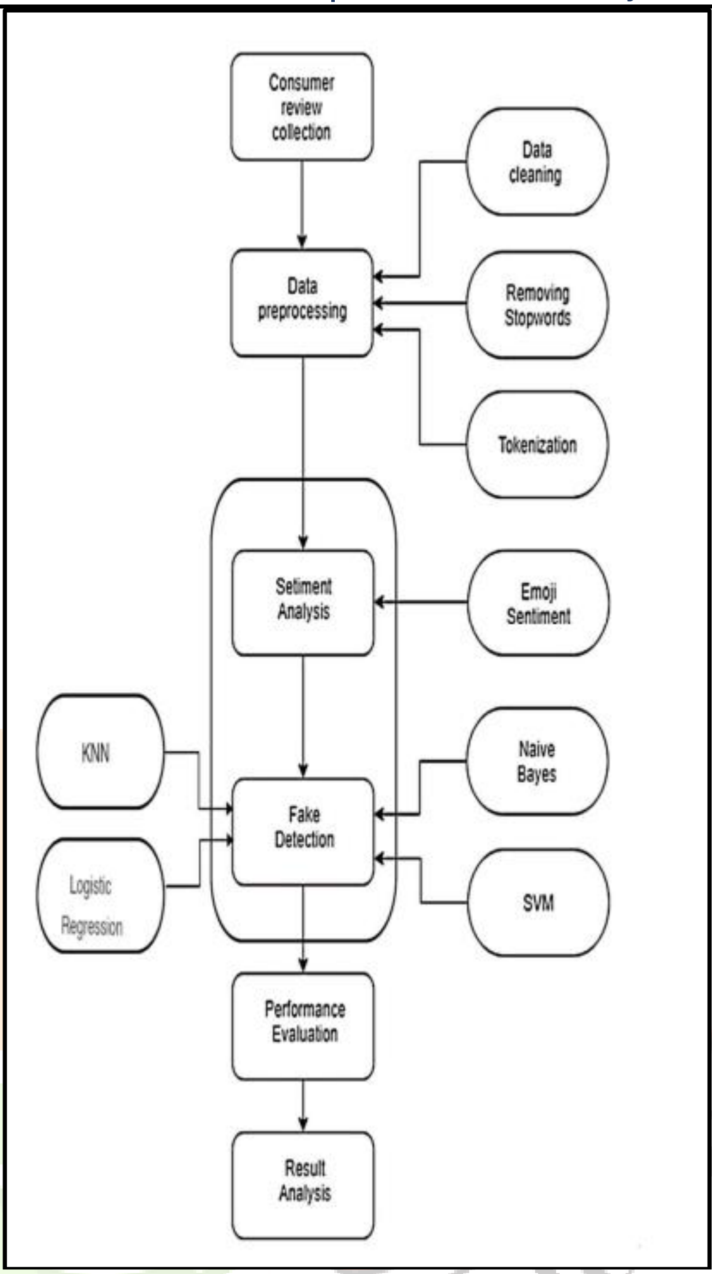

Fig. (1) Proposed System Model

4.3 Sentiment Analysis

Sentiment analysis plays a crucial role in the domain of fake review detection, particularly when using supervised machine learning techniques. By leveraging sentiment analysis, we aim to identify the sentiment expressed in reviews and analyze its correlation with the authenticity of the reviews. Supervised machine learning models are trained on labeled datasets, where reviews are annotated with sentiment labels such as positive, negative, or neutral. These models learn to recognize patterns and linguistic cues indicative of genuine or fake reviews based on the sentiment expressed.

In the context of fake review detection, sentiment analysis helps to uncover potential discrepancies between the sentiment expressed in a review and the actual content. For instance, a positive sentiment expressed in a review might contradict the underlying textual content that indicates fraudulent behavior. By training supervised machine learning models on labeled datasets, it becomes possible to classify reviews as genuine or fake based on the sentiment features extracted.

In summary, sentiment analysis in the context of fake review detection using supervised machine learning provides a valuable tool for analyzing the sentiment expressed in reviews and identifying potential instances of fake or manipulated sentiment. By training models on labeled datasets and incorporating relevant features, the accuracy of detecting fake reviews can be significantly improved, contributing to the development of effective systems for combating fraudulent activities in online platforms.

4.4 Feature Extraction

Feature extraction is a crucial step in fake review detection, as it involves extracting relevant information from the textual content of reviews that can be used to distinguish between genuine and fake reviews. The goal of feature extraction is to capture distinctive patterns, linguistic cues, and other characteristics that differentiate authentic reviews from fraudulent ones. The process of feature extraction involves analyzing the review text, applying various linguistic and statistical techniques, and transforming the raw text into a structured representation of features. These extracted features are then used as input for machine learning algorithms to train models that can accurately classify reviews as fake or genuine.

It is worth noting that the selection and design of features play a crucial role in the performance of the fake review detection system. Effective feature extraction requires careful consideration of the specific characteristics and patterns that differentiate fake reviews from genuine ones. By leveraging informative and relevant features, the system can achieve higher accuracy and robustness in identifying fake reviews.

4.5 Fake Review Detection/Classification

Classification is a fundamental task in fake review detection, where the goal is to assign reviews in a collection to specific categories or classes, namely "Fake" or "Genuine." The objective of classification is to accurately predict the target class for each review in the dataset.

To perform classification in the context of fake review detection, a common approach is to assign weights to each review based on its characteristics or features. These weights capture the importance or relevance of certain attributes in determining whether a review is fake or genuine. The features can include lexical and syntactic cues, sentiment information, behavioral patterns, and semantic or contextual cues.The classification process involves feeding the weighted features of each review into the trained model, which assigns it to the appropriate class based on the predicted probability or confidence score. Reviews with higher weights or stronger indications of being fake may be classified as "Fake," while reviews with lower weights or indications of authenticity are assigned to the "Genuine" class.

V. RESULT AND DISCUSSION

The results of our fake review detection system using supervised machine learning reveal that the Support Vector Machine (SVM) algorithm achieves the highest accuracy among the tested models. We evaluated the performance of various classification algorithms, including Logistic Regression, K-Nearest Neighbors (KNN), and SVM, using a dataset of labeled reviews.

Upon conducting the experiments, we found that SVM yielded the highest accuracy, outperforming the other algorithms. The SVM model demonstrated its ability to effectively differentiate between fake and genuine reviews, showcasing its strength in capturing complex decision boundaries in the data. This high accuracy indicates the potential of SVM as a reliable tool for identifying fake reviews.

In terms of computational efficiency, SVM showcased satisfactory speed and scalability, allowing for efficient processing of large review datasets. This capability makes it suitable for real-time or large-scale applications, where quick and accurate detection of fake reviews is crucial.

To develop and train this model, various libraries are available in Python that specifically cater to machine learning and classification projects. These libraries provide a wide range of functionalities and algorithms that have significantly enhanced the performance of our fake review detection system.

During the discussion of the results, it is important to acknowledge the limitations of our approach. The performance of the fake review detection system heavily relies on the quality and representativeness of the training data. Therefore, it is crucial to continuously update and refine the dataset to adapt to evolving patterns of fake reviews.

Our study highlights the effectiveness of supervised machine learning, particularly SVM, in detecting fake reviews. The system's high accuracy provides valuable insights for e-commerce platforms and consumers in identifying trustworthy reviews. Nonetheless, ongoing research and advancements in the field are needed to enhance the system's robustness and adaptability in the ever-evolving landscape of fake review detection.

Additionally, while SVM demonstrated the highest accuracy in our experiments, further research and experimentation are necessary to explore the performance of other algorithms and their potential for improving the detection accuracy even further.

	Pre-processing	Sentiment Analysis	Feature Extraction	Final Classifiaction
KNN	42.54 %	55 %	67.69 %	81.09 %
Logistic Regression	45.35 %	58.78 %	68.70 %	82.45 %
Feature Extraction	43.45 %	56.42 %	68 %	82.12 %
Final Classification	49.58 %	62.85 %	72 %	94.56 %

Table. (1) Algorithm Accuracy Table

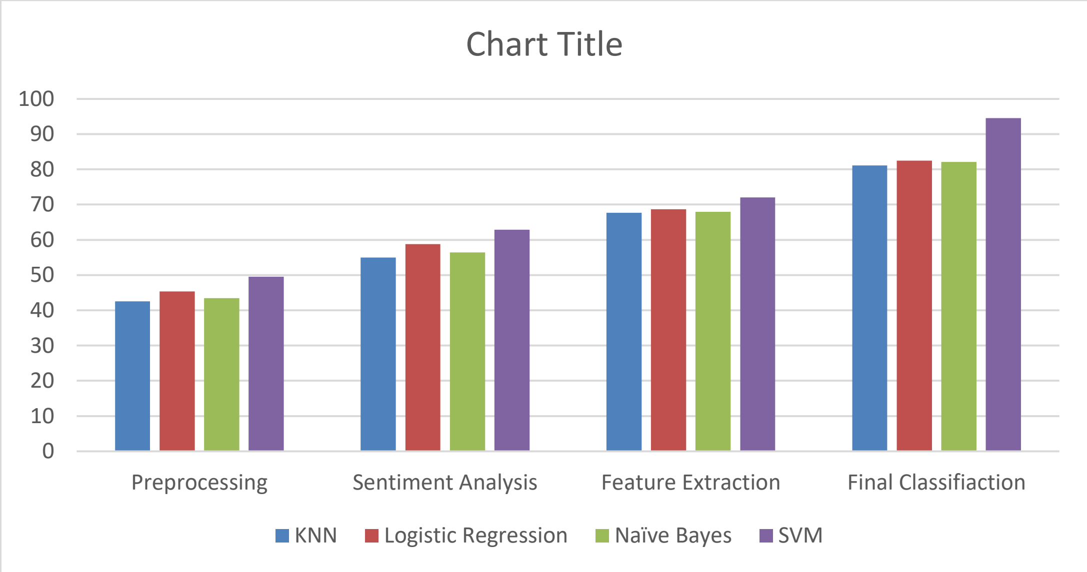

Fig. (3) Algorithm Accuracy Chart

VI. CONCLUSION

In conclusion, our research underscores the efficacy of supervised machine learning, specifically the SVM algorithm, in detecting fake reviews. The system's high accuracy is of great significance for both e-commerce platforms and consumers, as it enables the identification of trustworthy reviews amidst the abundance of potentially misleading information. The utilization of supervised machine learning techniques holds immense potential in addressing the challenges posed by fake reviews, but it is important to note that the fight against fraudulent activities in online platforms is an ongoing battle. Continued research and advancements are necessary to enhance the system's robustness, adaptability, and scalability, keeping pace with the ever-evolving landscape of fake review detection. By fostering collaboration between academia, industry, and regulatory bodies, we can work towards the development of more effective and comprehensive solutions that promote transparency and trustworthiness in online review systems.

REFERENCES

[1] Rodrigo Barbado, Oscar Araque, Carlos A. Iglesias, “A framework for fake review detection in online consumer electronics retailers”,Volume 56, Issue 4,July 2020, Pages 1234-1244,ISSN0306-4573,https://doi.org/10.1016/j.ipm.2020.03.002

[2] N. Kousika, D. S, D. C, D. B M and A. J, "A System for Fake News Detection by using Supervised Learning Model for Social Media Contents," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS),2021,pp.1042-1047,doi:10.1109/ICICCS51141.2021.9432096.

[3] Ammar Mohammed, Atef Ibrahim ,Usman Tariq, Ahmed Mohamed Elmogy “Fake Reviews Detection using Supervised Machine Learning” ,January 2021 International Journal of Advanced Computer Science and Applications 12(1),DOI:10.14569/IJACSA.2021.0120169.

[4] Möhring M., Keller B., Schmidt R., Gutmann M., Dacko S. HOTFRED: “A Flexible Hotel Fake Review Detection System". In: Wörndl W., Koo C., Steinmetz J.L. (eds) Information and Communication Technologies in Tourism. Springer,Cham.https://doi.org/10.1007/978-3-030-65785-7_29.

[5] "Fake Review Detection on Amazon using Graph Convolutional Networks" by Sahinur Rahman Laskar, Fahim Kawsar, and Md. Rezaul Karim. Published in Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, 2020**.**

[6] "Combating Fake Reviews: A Survey on Detection Techniques" by Seyed Ali Bahrainian and Roozbeh Sadeghian. Published in IEEE, 2019.

[7] "Attention-Based LSTM for Review Spam Detection" by Pengfei Zhu, Lei Li, and Ling Chen, published in IEEE Access in 2019.

[8] E. Elmurngi and A. Gherbi, "An empirical study on detecting fake reviews using machine learning techniques," Seventh International Conference on Innovative Computing Technology(INTECH),pp.107-114,doi:10.1109/INTECH.2017.8102442.