BioGecko

Vol 12 Issue 03 2023
ISSN NO: 2230-5807

Fake Review Detection Using Fuzzy Logic andMachine Learning

Mrs. G. Sharmila [a*], S. Abinesh [b], Athul Dhanesh [b], S. Nitthiyan Annamalai [b]
[a] Assistant Professor, Department of CSE, Manakula Vinayagar Institute of Technology, Puducherry;
[b] UG Students, Manakula Vinayagar Institute of Technology, Puducherry.

Abstract— Making choices for consumers when they shop online is greatly influenced by product reviews in the digital age. Fake reviews are a widespread problem that risks the reliability of these platforms, misleading users and affecting businesses. Fake reviews detection, uses web scraping and machine learning techniques to develop a sentiment analysis system for product evaluations. Users can provide a product URL and the system will retrieve reviews from the associated webpage using a Flask web application. The application scrapes the review content using the BeautifulSoup library. The reviews go through a preprocessing stage when they are acquired, which involves taking out stop words, lemmatizing tokens, and turning them into feature vectors. Each review's sentiment is generated using TextBlob, and fuzzy logic is used to determine the membership values. Following the preprocessing, TFIDF vectorizer is used to integrate and convert the preprocessed data, sentiment scores, and membership values. To forecast the reviews' sentiment, a pre-trained machine learning model loaded from a pickle file is employed. By combining web scraping, natural language processing, and machine learning, the programme successfully demonstrates how to gain insightful information about the tone of product reviews. With a strong tool to fight fake reviews and help consumers make smart choices, our project tries to create a fair and reliable purchasing environment.
Keywords— Fake review, Machine learning, Sentimental analysis, Natural language processing, Fuzzy logic

I. INTRODUCTION

Anyone may freely and without fear of consequences express their thoughts and beliefs. It is simpler to Post honestly and with confidence thanks to social media and online publication. These opinions are given the proper input to connect with the appropriate person who can assist with the problem, occasionally even a fraud, these opinions are valued if they are twisted. As a result, people Maliciousness makes it simple to fool the system with their authenticity and expression of opinion in order to promote their own products or criticise those of competitors without disclosing their own identity or that of the company they are working for. Opinion spam is a term that describes these individuals and the activities they engage in. Opinion spam can take many different forms. A method of being constructive Reviews that are untrue or negative or that are expressed with the intention of influencing customers to purchase a product are harmful to its reputation. When developing models utilising various types of sentiment analysis on data from various sources, there are several studies in the field of sentiment analysis. Sentiment analysis, commonly referred to as opinion mining. Machine learning, a key component of artificial intelligence, is used in automatic opinion mining. Software that can extract knowledge from datasets and integrate some additional data to increase performance can be used to build polling systems. One of the most important applications of opinion mining is in the reviews, comments, and services of consumer items that are posted online and through e- commerce.

Because they are so helpful to both the customer and the merchant, e-commerce websites encourage its users to provide comments and reviews about the products or servicesthey have purchased. Potential customers can use these reviews to gather vital information about the experiences of previous or current users before deciding to buy a product from a specific supplier. The seller or service providers similarly employ this data to identify any shortcomings or problems clients may have with their goods and to comprehend competition data to distinguish their products from those of comparable rivals. In the modern era, this technique of encouraging customers to post product reviews has evolved into an excellent strategy for businesses trying tomarket their products using the voice of the target audience. Such valuable information has been distorted and spammed. Numerous methods have been developed to recognise this kind of spam. Each of the methods produced varying degrees of precision. The main objective of this project in making better judgements about online purchases of items by guiding them in identifying fake reviews on websites and social media. Informing retailers and manufacturers of fake and genuine reviews of their products is also beneficial. The most recent techniques, such fuzzy logic and machine learning, are used to differentiate fake reviews from genuine ones. These techniques evaluate comments from user, which are then separated into fake and genuine reviews.

II. CLASSIFICATION ALGORITHM

Logistic Regression Classifier : It is a machine learning classification algorithm that chooses a result by considering one or more independent variables. Due to the dual nature of the variable being used to determine the outcome, there are only two possible outcomes. The goal of logistic regression is to identify the dependent variable's best-fitting connection to a group of independent variables. In comparison to other binary classification techniques like nearest neighbour, it performs better since it unbiasedly describes the factors guiding classification.

Naive Bayes Classifier: It is a classification method based on the Bayes theorem, which presumes independence between predictors. Simply expressed, a Naive Bayes classifier thinks that the existence of one feature in a class has no bearing on the presence of any more features. Each of these attributes increases the chance on its own even though they are interrelated. The Naive Bayes model is straightforward to build and works best with data sets that are at least somewhat large. Even with a simple methodology, Naive Bayes is known to outperform the majority of machine learning's classification algorithms.

Decision Tree Classifier: The decision tree technique is used to build the classification model's tree structure. It uses the if-then categorization criteria, which are both comprehensive and mutually exclusive. As the process proceeds, the data is incrementally broken down into smaller structures and connected with a decision tree. When the design is finished, it resembles a tree with nodes and leaves. The rules are successively learned using one training batch of data at a time. Following the learning of each rule, the tuples containing the learned rules are removed. The process runs on the training set up until it reaches the termination point. The tree is constructed using the divide-and-conquer method from the top down. A decision node will have two or more branches, and each branch's leaf will represent a category or option. The root node of a decision tree is the topmost node that corresponds to the best predictor; this is its best feature. A decision tree may handle both categorical and numerical data.

Random Forest Classifier: A popular machine learning algorithm frequently used for classification applications is the Random Forest Classifier. It is a member of the ensemble learning family, which combines many decision trees to generate predictions. Using random selections of the training data, Random Forest builds a set of decision trees. Each tree is trained independently, and the combined forecasts of all the trees yield the final prediction. Accuracy and robustness are enhanced by this method, particularly for large and complicated datasets. Random Forest, as opposed to single decision trees, can accommodate both numerical and categorical information with minimal preprocessing. Additionally, it is less likely to overfit. Incorporating "bagging" and "feature randomness," the technique generates a variety of unrelated trees that are different and uncorrelated, hence lowering variance and enhancing generalisation. Additionally, it offers feature significance metrics for identifying key features. Due to its adaptability, precision, and capacity for handling large datasets, Random Forest finds use in a variety of industries, including banking, healthcare, marketing, and image identification.

Support Vector Machine Classifier: An extremely efficient classification algorithm used in machine learning is called the Support Vector Machine (SVM). It is a supervised learning method created especially to deal with challenging classification issues. In a high-dimensional space, SVM creates a hyperplane to divide several classes of data points. Finding the hyperplane with the largest margin—that is, the distance between the hyperplane and the closest data points in each class—is the goal. This enables SVM to handle both linearly and non-linearly separable data and achieve outstanding generalisation performance. The strength of SVM is its effective handling of high-dimensional data. SVM can implicitly shift the input data to a higher-dimensional feature space where linear separation is possible by using the kernel technique. As a result, SVM is appropriate for uses in bioinformatics, image recognition, and text classification. A strong theoretical foundation based on statistical learning theory and optimisation concepts also supports SVM. In comparison to other classifiers, it can handle huge datasets and is less prone to overfitting. However, selecting the right kernel function and fine-tuning its parameters are crucial for SVM's success.

Vol 12 Issue 03 2023
ISSN NO: 2230-5807

Fig 2.1 Fake Review

III. FUZZY LOGIC

a) Sentiment Analysis:

Classifying the evaluations into groups based on whether they are emotional favourable, negative, or neutral. It involves predicting the tone of reviews based on the text' word choice, emoji selection, review rating, and other elements. On the other hand, fake reviews generate more positive or negative feelings than legitimate reviews. It's Vital to Consider The Subjective to Objective Ratio, Advertisers create reviews by including more subjective data and emphasise feelings, such as their happiness, than they do facts about the product or what it does. Determine whether a review is real or false through analysing its emotional content.

b) Fuzzy Logic:

A form of logic known as fuzzy logic enables reasoning with fuzzy or unclear concepts. Traditional logic is based on binary true/false values; fuzzy logic is different. Concepts in fuzzy logic are represented by fuzzy sets and have varying degrees of truth. A fuzzy set is one that enables objects to only partially belong to the set. The definition of "tall" might be expressed, for instance, as a fuzzy set, where an object's membership in the set is determined by how closely it follows to the standard for tallness. Fuzzy logic is helpful when there is a lot of uncertainty or confusion, like when making decisions or processing natural language. It can be used to build systems that can make judgements based on ambiguous inputs, such the attitude of a text message or the tone of a review. In general, fuzzy logic is a potent instrument that may be used in a wide range of sectors when there is a high level of ambiguity or uncertainty. When working with complex data, it enables advanced decision-making and can produce more accurate results.

IV.NECESSITY

The important role that online reviews play in creating brand reputation and influencing customer behaviour underlies the requirement for creating a fake review detection system. Online reviews are an essential information source for prospective clients in the current digital era, assisting them in making selections about goods and services. However, the surge in dishonest behaviours, such writing false reviews, calls into question the reliability and authenticity of online platforms. Fake reviews can harm businesses, confuse customers, and affect market perceptions. Therefore, it has become crucial for businesses to have effective systems that can detect and flag fraudulent reviews.

Businesses may safeguard their reputations, preserve openness, and uphold customer trust by putting fake review detection projects into place. Such a project makes use of cutting-edge technology like sentiment analysis and machine learning algorithms to analyse review patterns, linguistic signals, and user behaviours in order to spot fake or doubtful reviews. In the end, a trustworthy system for identifying fake reviews helps create a trustworthy and fair online market where customers can securely rely on real reviews to make informed purchasing decisions.

V. LITERATURE SURVEY

Saeedreza et al [1] introduces a novel spam detection framework namely NetSpam based on a meta path concept as well as a new graph-based method to label reviews relying on a rank-based labelling approach. The performance of the proposed framework is evaluated by using two real- world labelled datasets of Yelp and Amazon websites. Our observations show that calculated weights by using this meta path concept can be very effective in identifying spam reviews and leads to a better performance. In addition, we found that even without a train set, NetSpam can calculate the importance of each feature and it yields better performance in the features' addition process, and performs better than previous works, with only a small number of features. Moreover, after defining four main categories for features our observations show that the reviews behavioral category performs better than other categories, in terms of AP, AUC as well as in the calculated weights. The results also confirm that using different supervisions, similar to the semi- supervised method, have no noticeable effect on determining most of the weighted features, just as in different datasets. Hina Tufail, M. Usman Ashraf [2], proposed a fake review detection model by using Text Classification and techniques related to Machine Learning. They used classifiers such as Support Vector Machine, K-Nearest Neighbour, and logistic regression (SKL), using a bigram model that detects fraudulent reviews based on the number of pronouns, verbs, and sentiments. Our proposed methodology for detecting fake online reviews outperforms on the yelp dataset and the TripAdvisor dataset compared to other state- of-the-art techniques with 90% and 89.03% accuracy. Jindal and Liu [3] that two duplicate reviews were labelled as fake reviews. The cosine similarity method was used to identify fake reviews and then manually confirmed them. Conversely, the reviews that did not have a cosine similarity above a certain threshold with any other reviews were kept as truthful reviews and not manually reviewed. The dataset from Amazon.com contains 54,618 reviews, of which 6% were labelled as fake. SLM method was used to give each review a spamming score. The experimental results of the proposed model achieved a 0.9987 AUC score, which outperformed SVM. Further, SLM was effective in detecting fake reviews. However, considering duplicate reviews as fake can be unreliable. Huayi Li [4] implemented working with Dianping, the largest Chinese review hosting site, we present the first reported work on fake review detection in Chinese with filtered reviews from Dianping's fake review detection system. Dianping's algorithm has a very high precision, but the recall is hard to know. This means that all fake reviews detected by the system are almost certainly fake but the remaining reviews (unknown set) may not be all genuine. Since the unknown set may contain many fake reviews, it is more appropriate to treat it as an unlabeled set. This calls for the model of learning from positive and unlabeled examples (PU learning). By leveraging the intricate dependencies among reviews, users and IP addresses, we first propose a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extend it to Collective Positive and Unlabeled learning (CPU). Our experiments are conducted on real-life reviews of 500 restaurants in Shanghai, China. Results show that our proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Since our models only use language independent features, they can be easily generalized to other languages. Sedighi et al. [5] proposed a decision tree method todetect fake reviews. They used traditional feature selection techniques to select suitable features and evaluate them. The proposed model can be improved by taking into account the data correlation in choosing the appropriate features. In the study by. Khurshid et al., the authors proposed a supervised machine learning model to detect fake reviews based on content features and primal features. The proposed model used five classifiers to classify the reviews: Naive Bayes, Random Forest, JRip, AdaBoost, and J48. The results on a real-life dataset, showed that the AdaBoost with combined features performed better than other classifiers with an accuracy of 73.4%.

Khurshid et al. [6], the authors proposed a supervised machine learning model to detect fake reviews based on content features and primal features. The proposed model used five classifiers to classify the reviews: Naive Bayes, Random Forest, JRip, AdaBoost, and J48. The results on a real-life dataset [8], showed that the AdaBoost with combined features performed better than other classifiers with an accuracy of 73.4%. Further, using Primal features has a significant impact on improving performance. However, the proposed model did not perform well with an imbalanced dataset. Khurshid et al. extended their previous work and proposed an ensemble learning model to detect fake reviews based on selected features. The proposed model consisted of two tiers: Tier 1 used three classifiers and Tier 2 used Logistic Regression classifier to introduce an accurate result. They also used the following feature selections to extract structural and linguistic features: Particle swarm optimization
, Cuckoo Search, Greedy stepwise, carried out in vector space and Chi-Squared utilized to evaluate the worth of an attribute by calculating the value of Chi-Squared statistic value. The experimental results showed that the chi-squared feature plays a significant role in improving the proposed model's performance with an 84.1% accuracy on the Yelp restaurant dataset and 81.7% semi-real dataset. Yao et al. [7] proposed an ensemble fake review detection model based on review content and reviewer features. The author handled the unbalance data by combining the grid search method and resampling by finding the best sampling ratio for each classifier. Finally, they utilized majority voting and stacking methods to enhance the classification model performance. The experimental results on the Yelp dataset showed that the proposed model did not outperform the state-of-the-art methods. Further, the proposed model suffers from time complexity. Sánchez-Junquera et al. [8] proposed an adaptation model for detecting fake reviews in cross domain. The proposed model frequently used Co-occurring Entropy to find the domain features and then used a

mismatch method to mask them. The gold standard dataset results using naïve Bayes classifiers showed that the proposed model had difficulty in detecting fake reviews in cross-domain. While the authors highlighted a concept drift problem in fake reviews where the characteristics of reviews change-over time. The authors utilized two methods, statistical machine learning technique, and benchmark concept drift detection methods, to investigate and prove their argument. They tested four real life Yelp datasets and they found that the classifier performance dropped significantly due to the changing of fake review characteristics over time. Furthermore, they stated a strong relation between concept drift and classification performance which negatively affects the prediction algorithm performance. This study indicates the importance of developing fake review detection models that can handle this issue. On the other side, the authors proposed a framework to investigate the review inconsistency based on different features (content, language, and rating) in fake reviews detection. The extracted features are fed into different machine learning classifiers (SVM, NB, RF, and MLP) in order to identify whether the review is fake or genuine. They collected datasets from Yelp.com to evaluate the proposed model. The experimental results show that the review inconsistency features can boost the performance in fake review detection. However, the proposed model works on limited data, and using word embedding representation with deep learning can enhance the performance. K Indhuja, Raj P C Reghu [9] The paper proposes to build an application oriented fuzzy based system of sentiment analysis of product reviews. Practical applications of this new area of research include for instance, emotion aware robots, automatic movies classifiers, intelligent computer interfaces or avatars, next generation video games design and automatic marketing surveys. Sentiment relevance is a concept to distinguish content informative for determining the sentiment of a document from uninformative content. In-cooperation of fuzzy logic induces fuzziness to system. Fuzzy logic is used to compute fuzzy score of the given document. The proposed model uses novel fuzzy functions that emulate the effect of different linguistic hedges and incorporated them in the sentiment classification task. Fuzzy logic can virtually deal with any proposition expressed in natural language [2]. An important concept in fuzzy logic lies in the concept of linguistic variables. Fuzzy set theory provides a straight forward way to model the intrinsic fuzziness between sentiment polarity classes. These properties make Fuzzy logic an apt tool for identifying sentiment classification of product reviews. Wenqian Liu [10] proposed a method for the detection of fake reviews based on review records associated with products. We first analyze the characteristics of review data using a crawled Amazon China dataset, which shows that the patterns of review records for products are similar in normal situations. In the proposed method, we first extract the review records of products to a temporal feature vector and then develop an isolation forest algorithm to detect outlier reviews by focusing on the differences between the patterns of product reviews to identify outlier reviews. We will verify the effectiveness of our method and compare it to some existing temporal outlier detection methods using the crawled Amazon China dataset. We will also study the impact caused by the parameter selection of the review records. Our work provides a new perspective of outlier review detection

V. EXISTING WORK

To identify any such spammed fake reviews, the present work advises categorising reviews into fake and genuine. utilising methods for Naive Bayes, Linear SVC, SVM, Random Forest, and Decision Trees that are part of machine learning. In addition to the review specifics, other features are used to boost accuracy, such as comparing the review's sentiment, confirmed purchases, ratings, emoji count, and product category with the overall score. A classifier is developed using the features that were found. Additionally, those traits are assigned a likelihood component or weight based on the categorise training sets. To distinguish between fake and true reviews, this supervised learning method employs a variety of machine learning methods.

To forecast the outcome, the current model follows these steps:
Data collection: Raw review data was gathered from a variety of sources, including Amazon and booking websites.

Data Preprocessing: After removing noisy and unreliable data, as well as redundant and irrelevant data from the review dataset, the data are processed and polished. It involves tokenizing sentences, removing punctuation, tokenizing words, and removing stop words.The technique of separating a string into words or tokens is known as word tokenization. The NLTK package is used to complete it. With the aid of the NLTK software, the full data preprocessing procedure is carried out.

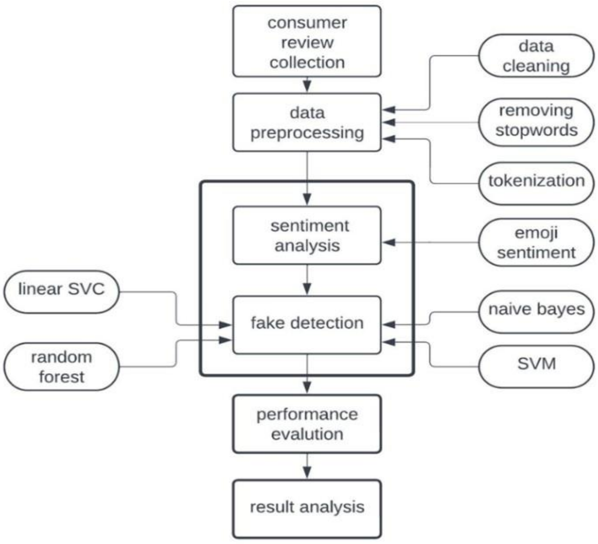

Fig 5.1 existing system architecture diagram

Feature Extraction: Using particular parameters, such as sentiment analysis, confirmed purchase details, reviewer id, and review length, the preprocessed data is transformed into a set of features.

Fake Review Detection: Identifying fake reviews is the main objective of classification, which is to correctly identify the target class for each example in the data. Each piece of information in the review file is given a weight, and based on that weight, it is divided into the corresponding classes of Fake and Genuine. Results of Performance Evaluation comparing the accuracy of different models and classifiers while adding improvements for improved outcomes.

VI. PROPOSED WORK

Our planned methodology entails a number of clearly defined stages, starting with the diligent collection of a varied and precisely labelled dataset, where evaluations are painstakingly classified as either phoney or genuine. Our thorough preprocessing procedures use complex tokenization, lemmatization, and stop-word removal methods together with careful handling of missing values in order to maximise accuracy.
We put a lot of attention on feature engineering and use the well-known term frequency-inverse document frequency (TF-IDF) methodology to turn the unprocessed review text into a representative feature space that accurately represents the subtle value of each word. Our cutting-edge approach, however, goes beyond standard methods by utilising fuzzy logic, which uses carefully created membership functions to assign sentiment scores to each review.
Our system then advances to the training stage, where a variety of state-of-the-art machine learning models, such as Random Forest, Logistic Regression, Naive Bayes, Decision Trees, and Support Vector Machines, are thoroughly trained. These models successfully capture the semantic meaning and sentiment subtleties hidden in the reviews by combining the TF-IDF features with the sentiment ratings. By combining the predictions from various models into a single, strong conclusion using a VotingClassifier, we utilise the potent ensemble learning approach to further improve the precision and performance of our system.
The crowning achievement of our suggested approach manifests as a user-focused, intuitive web application that has been painstakingly created using the flexible Flask framework. This interactive platform gives consumers control by offering a smooth interface to enter a certain product link and seamlessly retrieve the corresponding reviews.
We quickly obtain the targeted reviews using sophisticated web scraping techniques, painstakingly preprocess them, and then use the carefully trained ensemble classifier to correctly forecast each review's genuineness. Users may easily tell phoney reviews from real ones because to the results' persuasive presentation in an attractive and intuitive way.

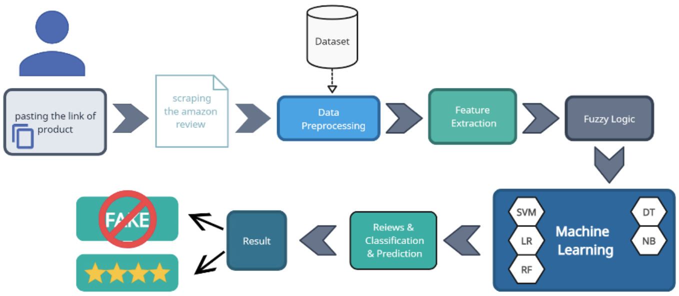

Fig 6.1 proposed system Architecture diagram

VII. RESULT AND DISCUSSION

The web application includes two primary interfaces: the "Review Paste Link" page and the "Fake or Genuine" resultpage.

On the "Review Paste Link" page, users are presented with a user-friendly interface where they can input the Amazon product link. This interface is designed to collect the necessary information needed to initiate the review analysis process. Once the link is submitted, the application proceeds to scrape the product's reviews and conduct sentiment analysis and fake review detection. The complex technical details of these processes are hidden from the user, ensuring a smooth and intuitive experience.

Once the analysis is complete, users are directed to the "Fake or Genuine" result page. This page provides a clear and concise presentation of the analysis outcomes for each review associated with the provided product link. The results are presented in a manner that allows users to quickly distinguish between "Fake" and "Genuine" reviews. This empowers users to make informed decisions based on the classification without being burdened by the intricacies of the underlying algorithms and procedures.

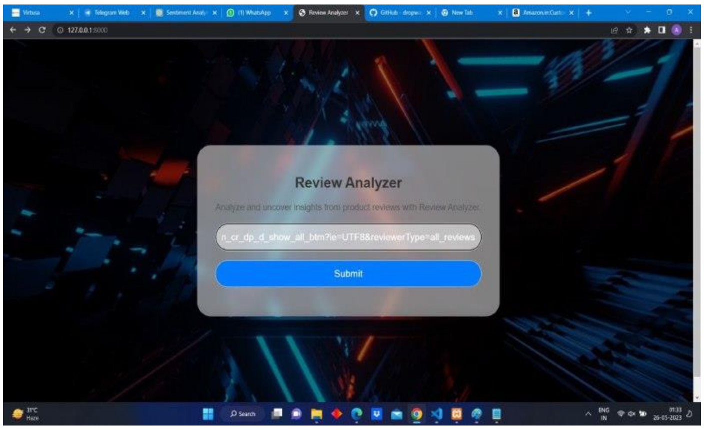

Fig 7.1 Drop link interface

Vol 12 Issue 03 2023
ISSN NO: 2230-5807

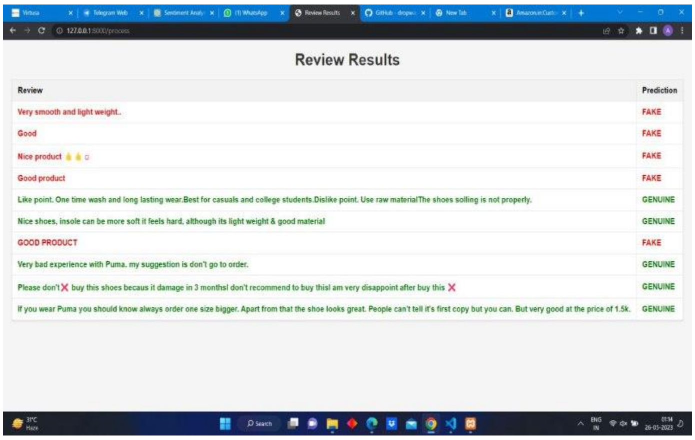

Fig 7.2 Review result interface

IX. CONCLUSION

As a result, our approach for detecting fake reviews successfully integrates fuzzy logic methods with machine learning algorithms to produce a workable and accurate answer. The system provides strong performance in identifying genuine from false reviews by preprocessing the review text, extracting pertinent characteristics, combining sentiment analysis using fuzzy logic, and using an ensemble classifier. Fuzzy logic integration enables a more thorough investigation of review authenticity by allowing a more complex portrayal of sentiment. The performance of the ensemble classifier, which combines several basic classifiers, is enhanced by using the advantages of each classifier in separation. The system is more usable and open to users because it is implemented as a Flask web application. Users may easily enter review text, choose relevant variables, and get fast feedback on the accuracy of the reviews due to the user-friendly web interface. The technology offers trustworthy insights into the accuracy and effectiveness of the fake review identification process due to its extensive evaluation metrics and performance assessment. Our system gives companies and customers a strong tool to spot fake reviews and make wise judgements by combining text analysis, feature extraction, fuzzy logic, and machine learning.

Overall, our method offers a viable solution to the problem of detecting false reviews, giving companies and customers thetools to assess the reliability of user-generated reviews and make better decisions.

FUTURE WORKS

1. Including real-time and time-based data: Including real- time and time-based data in the feature representation is one possible ways for future research. The timestamps of comments or reviews can be used to identify short-term trends and patterns in user sentiment or behaviour.
2. Examining different machine learning algorithms: While the existing system makes use of a voting ensemble of algorithms, there is scope for further study different methods and evaluate their effectiveness. In the future, the ensemble may incorporate more methods like logistic regression, support vector machines, or neural networks. Furthermore, investigating deep learning methods like convolutional neural networks or recurrent neural networks can reveal newinformation on the recognition of complex patterns in reviewdata.
3. Future study can concentrate on developing unsupervised learning methods for detecting fake reviews in addition to the current supervised learning strategy. This is especially helpful if there is a sizable volume of unlabeled data accessible. Abnormal detection algorithms might be used to find reviews that drastically differ from typical or comparable reviews could be grouped together based on various attributes. By using unsupervised learning techniques, the system may handle unlabeled data more skillfully and gain a wider perspective on the existence of fake reviews.

4. Integrating user and product metadata: Adding user and product metadata to the feature representation can improve the context for spotting false reviews. It can be useful to include features like user reviews, review history, or product attributes to spot trends or anomalies that might point to the presence of false reviews. By taking into account other elements that affect review validity, the integration of external data sources—such as social media profiles or reputation systems—can help improve the system's overall accuracy.

REFERENCES

1. Saeedreza Shehnepoor, Mostafa Salehi, Reza Farahbakhsh, Noel Crespi, “Netspam: A Network-Based Spam Detection Framework For Reviews In Online Social Media” In Ieee Transactions On Information Forensics And Security 2017
2. Hina Tufail, M. Usman Ashraf, Khalid Alsubhi, Hani Moaiteq Aljahdali, "The Effect Of Fake Reviews On E-Commerce During And After Covid-19 Pandemic: Skl-Based Fake Reviews Detection," In Ieee 2022.
3. N. Jindal And B. Liu, ‘‘Opinion Spam And Analysis,’’ In Proc. Int. Conf. Web Search Web Data Mining (Wsdm), 2008, Pp. 219–230.
4. Huayi Li, Zhiyuan Chen, Bing Liu, Xiaokai Wei, Jidong Shao, "Spotting Fake Reviews Via Collective Positive-Unlabeled Learning" In Ieee International Conference On Data Mining 2014.
5. Z. Sedighi, H. Ebrahimpour-Komleh, And A. Bagheri, ‘‘Rlosd: Representation Learning Based Opinion Spam Detection,’’ In Proc. 3rd Iranian Conf. Intell. Syst. Signal Process. (Icspis), Dec. 2017, Pp. 74–80.
6. F. Khurshid, Y. Zhu, Z. Xu, M. Ahmad, And M. Ahmad, ‘‘Enactment Of Ensemble Learning For Review Spam Detection On Selected Features,’’ Int. J. Comput. Intell. Syst., Vol. 12, No. 1, Pp. 387–394, 2018.
7. J. Yao, Y. Zheng, And H. Jiang, ‘‘An Ensemble Model For Fake Onlinereview Detection Based On Data Resampling, Feature Pruning, And Parameter Optimization,’’ Ieee Access, Vol. 9, Pp. 16914–16927, 2021.
8. J. Sánchez-Junquera, L. Villaseñor-Pineda, M. Montes-Y-Gómez,P. Rosso, And E. Stamatatos, ‘‘Masking Domain-Specific Information For Cross-Domain Deception Detection,’’ Pattern Recognit. Lett., Vol. 135, Pp. 122–130, Jul. 2020
9. Wenqian Liu, Jingsha He, Song Han, Fangbo Cai, Zhenning Yang, Nafei Zhua, "Method For The Detection Of Fake Reviews Based On Temporal Features Of Reviews And Comments In Ieee Engineering Management Review 2019.
10. K Indhuja, Raj P C Reghu, "Fuzzy Logic Based Sentiment Analysis Of Product Review Documents" In First International Conference On Computational Systems And Communications (Iccsc) 2014.