WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Fake Review Monitoring System

Dr. B. Karthikeyan[1], Mr T. Guru Prakash[2]

[1]Assistant Professor [SG], Department of Computer Science, Bishop Heber, College (Autonomous), Afflicted to Bharathidasan University Trichy-17, Tamilnadu, India

[2]Department of Computer Science, Bishop Heber College (Autonomous), Afflicted to Bharathidasan University Trichy 17,Tamilnadu, India

ABSTRACT:

The rising trend of online product reviews has had a tremendous impact on customer decision-making within ecommerce. While authentic reviews contribute towards building the reputation of a product, deceptive reviews such as false negative and positive reviews can confuse the potential purchaser and manipulate product ratings The project intends to create a machine learning model that is able to distinguish between real and fake news headlines. The training and testing data set consists of headlines and their respective results. In order to model the headlines, several preprocessing steps such as text cleaning, stemming, and vectorization are applied. Different classification techniques such as Naive Bayes classifiers, Multilayer Perceptron (MLP), Decision Trees, Support Vector Machines (SVM), Logistic Regression, and Stochastic Gradient Descent (SGD) are utilized to develop individual models. Accuracy, mean squared error, confusion matrix, and classification report are the metrics utilized for training and evaluation of each model. Ensemble learning techniques are also investigated in the project using a Voting Classifier for combining various models. The results of the evaluation indicate the capacity of the models in classifying news headlines precisely and demonstrate the utility of ensemble learning in enhancing the accuracy of predictions.

KEYWORD: Fake Review , Detection, Naive Bayes classifiers, Multilayer Perceptron (MLP), Decision Trees, SupportVector Machines (SVM), Logistic Regression, and Stochastic Gradient Descent (SGD)

Introduction: -

This Python program classifies text according to headlines and their outcomes as fake or real news. The text data is processed in advance with TF-IDF for vectorization and stemming and a Stochastic Gradient Descent classifier trained with the data. Other classifying models such as Decision Trees, Support Vector Machine, Naive Bayes, Logistic

Regression, and Multi-Layer Perceptron are discussed. Accuracy, mean squared error, confusion matrix, and classification report are performance metrics used to measure the performance of each model. The script also illustrates utilizing Voting Classifier in order to form ensemble models using soft and hard voting strategies.

Utilizing a set of review, this technique tries to classify text by evaluating whether the review are real or fake review. The objective of this coding is to apply various machine learning algorithms to develop and evaluate models which are able to differentiate between a true and a false review. The code seeks to examine a number of approaches for handling the fake review detection issue through the use of methods

such as text preprocessing, TFIDF feature extraction, and model training such as Support Vector Machines, Decision Trees, Multilayer Perceptron, and Naive Bayes. Moreover, to combine the strengths of single models and perhaps enhance overall prediction accuracy, ensemble methods such as Voting Classifier are utilized. The necessity to combat fake review and promote the dissemination of authentic review in society is what drives our project.

Literature Review: -

Syed Mohammed Anas.et al [2021] discusses the problem of counterfeit reviews in the ecommerce sector through the creation of a monitoring and deletion system based on Natural Language Processing (NLP) and Machine Learning (ML) algorithms. Using the Amazon Yelp dataset, the models—Naïve Bayes and Random Forest—were tested for their performance in identifying reviews as authentic or counterfeit. Random Forest performed better than Naïve Bayes with improved accuracy (89.487%), precision (85.577%), recall (94.389%), and F1 score (89.768%), and hence is more appropriate for real-world use cases. The model demonstrates the possibility of scalability and implementation on platforms such as Amazon and Flipkart with the goal of restoring consumer confidence by removing false reviews. However, limitations such as small dataset size and the growing sophistication of fake reviews indicate the need for advanced techniques and larger-scale implementations.

Manasi Bansode,[2021] The work is to detect and categorize fake online reviews with different machine learning models. The platform filters hotel reviews into truthful and fake, positive and negative ones, to support consumers in taking smart decisions and companies in analyzing customers' emotions. Five models of machine learning— Stochastic Gradient Descent (accuracy: 77.81%), Logistic Regression (77.58%), Support Vector Machine (76.25%) Multinomial Naïve Bayes (75.38%), and KNearest Neighbor (60.54%) were trained on 104,000 reviews and pickled for web app deployment. A voting approach aggregates the outputs of all the models to produce the final classification. The final output, which is saved as a CSV file, facilitates accurate fake review identification and helps users assess services through genuine feedback.

Myasar Tabany, Meriem Gueffal et al. [2024] This project tries to perform sentiment analysis of short and long Amazon reviews and report their impact on the supervised learning Support Vector Machines (SVM) model, to bridge for fake reviews classification. To begin, the SVM model was tested by ranking its performance against Naïve Bayes, Logistic Regression, and Random Forest models and turned out to be better (second assumption) according to accuracy (70%), precision (63%), recall (70%), and F1-score (62%). Hyperparameter tuning enhanced the SVM model for sentiment analysis (accuracy of 93%), then changing the review length impacted the performance of the model, which confirmed that review length impacts the classifier (first assumption). Secondly, performed fake reviews classification on the dataset of fake reviews achieved 88% accuracy, whereas merged subsets of the two datasets achieved 84% accuracy. Wael et al. employ supervised learning algorithm to detect fake review. Prior to the application of the classification approach, various preprocessing steps are done; these involve stemming, deletion of punctuation symbols and stop word elimination.They employ linguistic feature to detect fake reviews. Linguistic feature has POS and bag-of-words. Bag-ofwords features are single word or set of words that occur in specified text. Then various classification algorithms are used such as decision tree, random forest, support vector machine, naive bayes and gradient boosted trees. In these naive bayes and support vector machine provide better result. Jitendra et al. used various features on the basis of sentiment polarity and content similarity for detection of fake and real reviews. In this authors utilized sentiment score on the basis of sentiment polarity i.e. positive and negative reviews, linguistic and unigram as feature. They then utilized three algorithms 1) support vector machine, 2) naive bayes and 3) decision tree.

Rohit et al., employ PU-learning algorithm with diverse classifiers. Author employs six types of classifiers here to identify spoofed reviews. These are decision tree, naive bayes, random forest, support vector machine, logistic regression and k-nearest neighbour classifiers. Here, logistic regression classifier provides best performance than all the six different algorithms. Hernandez et al. compared modified PUlearning algorithm with traditional PUlearning. With modified PUlearning method, author analysed that it can detect fewer instances from the unlabeled set.In each iteration, only the new negative instances are taken into account that are created by output of last iteration and classifier is imposed on that only new negative instances.Therefore in each iteration, negative examples are minimized and end examples are properly identified as fake or real reviews.Authors also identify positive and negative imposter reviews in this paper. They employed naive bayes and support vector machine classifier with both unigrams and bigrams features and reviews are labeled as fake and non-fake reviews

Methodology: -

The approach involves several steps of text classification with the aid of machine learning algorithms. The first step is loading and preprocessing the dataset, including stemming where words are reduced to the root word and stop words removal. The text is transformed into numerical features with the use of TF-IDF vectorization. The dataset is split into training and testing sets. Various classifiers, such as Stochastic Gradient Descent, Logistic Regression, Support Vector Machine, Decision Trees,

MLPClassifier, BernoulliNB, and MultinomialNB, are trained and evaluated using metrics such as accuracy, mean squared error, confusion matrix, and classification report. Additionally, ensemble approaches suchsince Voting Classifier are utilized for voting aggregations of different models, both soft and hard voting. This approach spans data preprocessing, feature engineering, model training, evaluation, and ensemble learning methods for text classification.

Step 1-Research and Survey: It is important to do an intensive analysis of present research and practices regarding spam filtering in product reviews. Analyze the weaknesses in current approaches and understand how the same can be solved or overcomed in this present research work for detecting better fake reviews.

Step 2- Data Acquisition: Obtain real-world review datasets from online shopping websites like Amazon and Flipkart. Both fake and real reviews are included in the datasets. The data is randomly sampled across different product categories, where 80% is used to train the model and 20% is held out for testing.

Step 3- Data Integration: Integrate multiple datasets obtained from various e-commerce platforms into one unified dataset. This combined dataset enables wider analysis and avoids any bias of individual platforms.

Step 4- Spam Identification Labelin Manually tag the reviews in the dataset as spam or non-spam according to pre-defined criteria. This process entails reviewing the content of the reviews and their sources to identify whether they display features of fake reviews, including excessively positive wording, repetitive phrases, or suspicious patterns.

Step 5-Pre-processing: Use several data cleaning methods to make sure that the dataset is clean and ready for analysis. This involves missing value handling, noisy data handling (irrelevant or wrong data), and eliminating inconsistencies. Pre-processing can also include eliminating duplicates and fixing formatting problems.

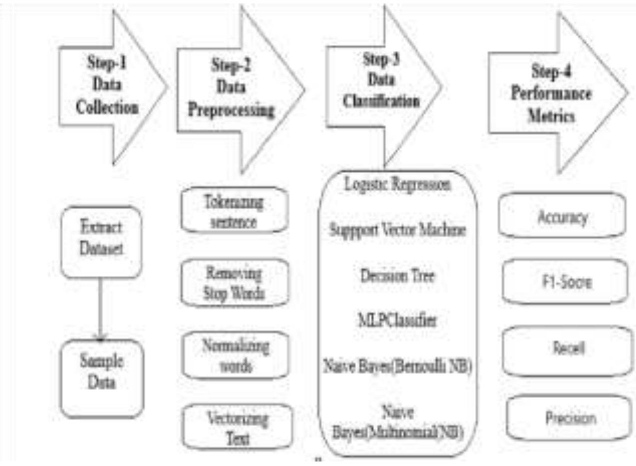
Step 6-Tokenization: Divide the text of each review into individual components (tokens), for example, words or phrases. Tokenization simplifies the process for machine learning models to work with and analyze the text by considering single units, which can be tested for relevance or sentiment afterward.

Step 7-Elimination of Stop-words**: Eliminate frequent words (e.g., "the," "and," "is") from the dataset that do not hold significant information for analysis. These words are generally regarded as insignificant in text mining since they occur frequently and do not indicate anything about the content of the review.

Step 8-Bag-of-Words Model: Represent the tokenized reviews as numerical data through the Bag-of-Words model. This entails having a vocabulary of distinctive words in the text corpus and expressing each review as a vector, where every dimension relates to the existence or frequency of a word in the vocabulary.

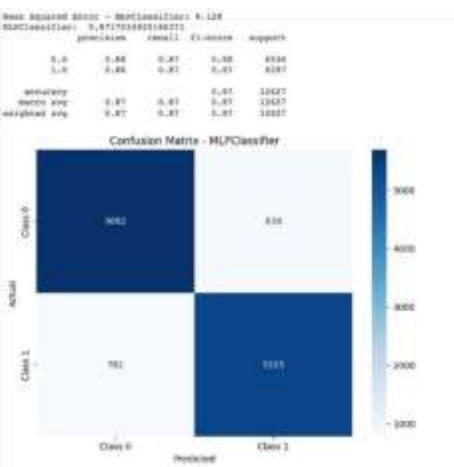
Step 9-Classifier Training**: Train machine learning classifiers using the preprocessed data and Bag-of-Words features. The classifier is trained on the training set (80% of the data), which will learn patterns from the labeled reviews to predict whether new reviews are fake or real.

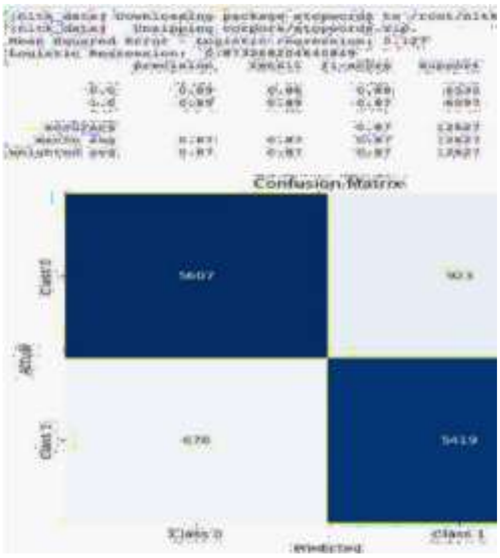
Step 10-Sentiment Analysis: Conduct sentiment analysis on the reviews with various classifiers (e.g., Naive Bayes, SVM, Logistic Regression). This step classifies reviews into positive or negative sentiments, serving to eliminate fake reviews. Also, the system detects and deletes duplicate reviews from the same user, displaying only the latest review for each product.



Result and Discussion: -

On the basis of performance measurement of some of the models over the provided data, the ensemble of models, which incorporated both Decision Tree Classifier and MLP Classifier in equal proportions, gave the highest accuracy.This ensemble approach, using a soft voting method, aggregates the prediction of the various classifiers to effectively classify news articles as authentic or fabricated. In addition, the model demonstrates resilience by having a range of methods that combine the features of decision trees and neural networks. With a higher ensemble accuracy score compared to the performance of single models, this method enhances overall prediction performance, thereby implying its suitability for fake vs. real review discrimination





Conclusion :-

The process involves data pre processing steps like text cleaning, stemming, and vectorization using TF-IDF. These algorithms are then trained and validated with accuracy metrics, mean squared errors, confusion matrices, and classification reports. Ensemble methods like Voting Classifier are also utilized to combine predictions from various models. This code concludes that ensemble methods like Voting Classifier, especially soft voting, enhance classification accuracy compared to single models. Nevertheless, additional fine-tuning and testing with other algorithms and hyper parameters can enhance the overall performance of the classification task. Additionally, exploring more advanced natural language processing methods and deep learning architectures can yield even better results, especially for applications using text data.

REFERENCES:

[1] Fake Product Review Monitoring System" Kunisetti Sudha Manikanta Dr. B.Raja Sarath Kumar , International Journal of Engineering Innovations in Advanced Technology ISSN: 2582- 1431 (Online), Volume-4 Issue-4, December 2022

[2] Intelligent Interface for Fake Product ReviewMonitoring and Removal - Ata-Ur-Rehman1, Nazir M. Danish2, Sarfraz M. Tanzeel3, Nasir Usama4, AslamMuhammad5, Martinez- Enriquez A. M.6, Adrees Muhammad7 - 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) Mexico City, Mexico. September 11-13,

[3] Ensemble Machine Learning Model for Classification of Spam Product Reviews Muhammad Fayaz , 1 Atif Khan,2 Javid Ur Rahman,3 Abdullah Alharbi,4 M. Irfan Uddin,5 and Bader Alouffi

[4] Fake Product Review Monitoring and Removal for Genuine Online Product eview Using IP Address Tracking M.Abi Priya1 , S.Hema2 , R.Dhivya Praba

[5] Fake Product Review Monitoring and Removal for Proper Ratings Amit Sawant, Shraddha Bhange, Shruti Akash Pandey

[6] FAKE PRODUCT REVIEW MONITORING SYSTEM Prof. Archana Burujwale, Sakshi Govind, Nikhil Kadam, Bhavesh Jadhav, Prasad Patil