JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Fake Review Detection Using SGD

Amirthavarshini B,Perumal P, Anu Deepthi V, Brinda Iswarya Lakshmi R

Student, Professor, Student, Student

Sri Ramakrishna Engineering College

ABSTRACT

Online customer reviews have a big impact on what consumers decide to buy in the current digital age. But as the significance of these reviews has grown, so too has the prevalence of phony ratings, which are posted by companies hoping to discredit rivals or enhance the reputation of their own goods. Such dishonest tactics are quite dangerous, especially for small businesses, where even one bogus negative review can have a significant negative impact. In response, this work uses machine learning (ML) techniques to propose an innovative way of identifying and categorizing bogus reviews. The Kindle dataset with a focus on book sales was used to test the suggested algorithm. A rigorous preparation workflow that included stemming, tokenization, normalization, and stop word removal was applied to textual data. For the classification, three machine learning methods were applied: BI Long Short Term Memory, Stochastic Gradient Descent and Gradient Boosting. Accuracy was assessed using methods for oversampling and under sampling both balanced and unbalanced datasets. The research's conclusions offer positive prospects for enhancing the credibility of online reviews and shielding businesses from the damaging effects of fraudulent evaluations. By revealing phony reviews, this study safeguards the interests of businesses and customers while also preserving the integrity of online review systems.

INTRODUCTION

Opinion mining using computer methods has become a potent technique in the digital age for obtaining subjective data, opinions, and feelings from large volumes of text. Opinion mining has proved particularly useful in the area of consumer acceptability and perceptions of goods and services. The emergence of social media and internet platforms has made it easier for people to voice their thoughts freely and anonymously without worrying about the consequences. Although these views provide insightful criticism, they also have a drawback in that they are easily swayed.

Opinion spamming has become a more common phenomenon, in which users take advantage of the system to either further their own interests or degrade those of competitors by writing false evaluations. This approach presents a serious risk, especially in social and political contexts where it has the power to skew public discourse and mobilize sizable audiences in favour of immoral objectives. Opinion spam is predicted to become more common and sophisticated, making it harder to identify, as long as opinions expressed on social media continue to impact decisions made in real life.

By applying machine learning approaches to the dataset, this work tackles the crucial problem of detecting false reviews. To be exact, we only develop successful system can tell fake comments from real ones. Our approach was to study how reviewers behave and discern genuine from fake reviews. We used machine learning classifiers such as Bi-LSTM, Extreme Gradient Boosting (XGboost) and Stochastic Gradient Descent(SGD) to analyse this. To ensure that our model is effective, we will test it on both balanced and unbalanced datasets, importing strategies like over sampling into the equations.

Our research intends to improve the resilience and accuracy of fake review detection systems by utilizing the complementary characteristics of both techniques. We suggest a hybrid framework that combines the effectiveness of SGD optimization and the predictive capacity of Gradient Boosting for

classification problems with the deep learning capabilities of Bi LSTM for feature extraction and representation learning.

The effectiveness of using the Bi LSTM, SGD, and Gradient Boosting algorithms for the problem of false review identification is examined in this work. Recurrent neural network variation Because Bi LSTM excels at spotting sequential relationships in textual input, it is a suitable method for analysing reviews that contain complex linguistic patterns and nuanced contextual information. SGD, on the other hand, provides an efficient optimization framework that allows for rapid convergence and scalability, when dealing with imbalanced datasets and extracting high-dimensional feature representations.

The research holds relevance as it has the potential to counteract the spread of fraudulent reviews, protect consumer confidence, and promote an equitable and transparent virtual economy. We want to empirically evaluate our suggested approach on real-world datasets to show its effectiveness in properly detecting fraudulent reviews while reducing false positives. This work advances the field's understanding of false review identification, which is a valuable contribution to the ongoing efforts to protect online review systems and advance consumer welfare in the digital era.

In conclusion, by offering a thorough framework that makes use of the synergies between the Bi LSTM, SGD, and Gradient Boosting algorithms, this research adds to the rapidly developing field of false review detection. Our research highlights how crucial it is to use a variety of approaches to tackle challenging problems in consumer protection and online reputation management.

PROPOSED SYSTEM

The suggested fake review detection system makes use of sophisticated machine learning algorithms like Gradient Boosting, Bi-Long Short-Term Memory (Bi-LSTM), and Stochastic Gradient Descent (SGD) in order to get around the drawbacks of current techniques. These algorithms have a number of benefits, such as the capacity to handle high-dimensional feature spaces, the ability to capture temporal relationships.

The system will perform tokenization, normalisation, and feature extraction on the textual content of reviews. Machine learning models will then be trained and classified using the pre-processed data. We will employ SGD, Bi-LSTM, and Gradient Boosting algorithms to find patterns and relationships in the review data. This will make it possible to identify minor indicators of phoney reviews. The suggested method will aggregate the predictions of several models using ensemble learning techniques, improving overall accuracy and lowering the possibility of false positives. Furthermore, the system will be built with feedback mechanisms to continuously enhance performance and adjust to changing strategies used by bad actors.

The suggested system seeks to achieve state-of-the-art performance in fake review identification by utilising Gradient Boosting, Bi-LSTM, and SGD algorithms. This will help to effectively mitigate the impact of fraudulent activities on online platforms. The system will also be adaptable and expandable, making it possible to integrate it easily into current review platforms used in a variety of sectors and domains.

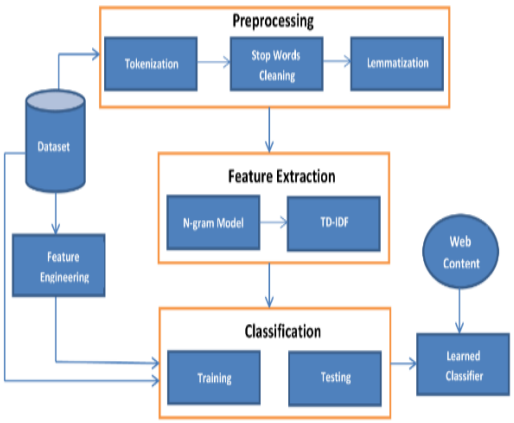

Fig 1 Proposed System

SOFTWARE REQUIREMENTS

The software requirements outline the essential software components and configurations necessary to install, run, and maintain the fake review detection system. These requirements ensure compatibility, stability, and optimal performance of the system's software stack.

Operating System: The fake review detection system should be compatible with popular operating systems such as Windows, Linux (e.g., Ubuntu, CentOS), or macOS. The choice of the operating system depends on the user's preference, development environment, and deployment environment.

Programming Languages: The system is typically developed using programming languages such as

Python, which offers extensive libraries and frameworks for machine learning, natural language processing, and deep learning tasks. Versions compatible with Python 3.x are recommended for compatibility with the latest libraries and dependencies.

Development Tools: Integrated Development Environments (IDEs) such as PyCharm, Jupyter Notebook, or Visual Studio Code are commonly used for developing, debugging, and testing the system's codebase. Version control systems like Git are essential for collaborative development and managing project versions.

Libraries and Frameworks: The system relies on various libraries and frameworks for implementing machine learning algorithms, deep learning models, and text processing tasks. Commonly used libraries include TensorFlow, PyTorch, Scikit-learn, NLTK, and Gensim for their extensive functionalities and community support.

MODULE DESCRIPTION

1. Collection of Raw Data

The foundation of any machine learning project lies in the quality and quantity of the dataset. In this step:

Identify Data Sources: Determine the sources from which the raw data will be collected. Websites like Amazon, Kindle, TripAdvisor, and others can provide a diverse range of reviews.

Data Retrieval: Extract reviews from chosen platforms either through web scraping techniques or by accessing publicly available datasets.

Data Quality Assurance: Perform data quality checks to ensure the authenticity and relevance of reviews. Remove duplicates, spam, or irrelevant reviews to maintain dataset integrity.

Dataset Balance: Strive to achieve a balanced dataset with an equal representation of fake and genuine reviews to prevent bias in model training.

2. Data Pre-processing

Data preprocessing serves as a crucial preparatory step in the fake review detection process, wherein raw textual data undergoes various transformations to enhance its suitability for subsequent analysis. The suggested method incorporates a series of meticulous internal processes, including stopword removal, stemming, normalization, and tokenization, each contributing significantly to the refinement and optimization of the dataset. Here is a detailed breakdown of each preprocessing step:

Normalization:

Normalization aims to standardize the textual data by removing symbols, numerals, and converting all characters to a consistent case format (either uppercase or lowercase). This process involves several standardized steps:

Eliminating all numerical digits from the text.

- Removing symbols such as punctuation marks (!, ?, #, $, etc.), special characters, and non-alphanumeric characters.
- Converting each word to a uniform case format, typically lowercase, to ensure consistency and eliminate case-related discrepancies.

Tokenization:

Each review phrase is divided into discrete tokens, allowing for efficient manipulation and analysis of the text at a granular level. Tokenization facilitates subsequent tasks such as feature extraction, sentiment analysis, and machine learning model training by breaking down the text into manageable units.

Stemming:

The primary objective of stemming is to normalize variations of words to their common linguistic root, thereby reducing the vocabulary size and computational complexity of subsequent analysis. The suggested method utilizes the Porter stemming algorithm, a widely adopted technique in natural language processing (NLP), to perform efficient stemming operations. By consolidating related word forms into their base stems, stemming enhances the coherence and consistency of the text corpus, facilitating more accurate analysis.

Stopword Removal:

Stopword removal entails filtering out these non-informative words from the textual data to focus on content-bearing terms and improve the relevance of analysis results. Common stopwords such as articles (e.g., "the," "a," "an"), prepositions (e.g., "to," "for," "in"), and conjunctions (e.g., "and," "but," "or") are removed to eliminate noise and enhance the

discriminative power of the dataset. This phase is particularly crucial in sentiment analysis, where the presence of stopwords can distort sentiment polarity and affect the accuracy of classification models.

3. Exploratory Data Analysis (EDA)

Before diving into model development, it's essential to gain insights into the dataset's characteristics and distributions. EDA involves:

Statistical Analysis: Compute descriptive statistics such as mean, median, standard deviation, etc., to understand the central tendency and variability of review attributes.

Visualization: Generate visualizations like histograms, bar plots, word clouds, etc., to visualize the distribution of fake and genuine reviews, word frequencies, sentiment distributions, and other relevant metrics.

Feature Engineering: Identify potential features or attributes that could aid in distinguishing between fake and genuine reviews, such as sentiment scores, review length, readability metrics, etc.

4. Model Selection

Sentiment Analysis

Sentiment analysis, also known as opinion mining, is a research subject focused on analyzing people's attitudes, views, feelings, and assessments of various entities such as goods, services, organizations, and events. It aims to determine the relative polarity of text, categorizing it as constructive, destructive, or neutral. Users often express their opinions through product reviews on social networking and shopping websites, making these platforms ideal for sentiment analysis. The technique requires a training set to effectively analyze reviews, particularly in e-commerce settings where sentiment analysis is commonly applied. Sentiment analysis aids in classifying free-form language as positive, negative, or neutral, enabling the summarization of customer judgments and the understanding of product strengths and weaknesses from other customers' perspectives.

5. Feature Extraction

Feature extraction is a pivotal aspect of pattern recognition or machine learning systems, aiming to distill data into its essential components to enhance the system's performance. It involves eliminating extraneous characteristics from the data, thus providing more relevant information to the model. This section explores various feature extraction techniques, focusing on Term Frequency-Reverse Document Frequency (TF-RDF) and Term Frequency-Inverse Document Frequency (TF-IDF).

6. Model Training and Evaluation

Once the dataset is pre- it's time to train and evaluate the selected model:

Training: Divide the dataset into sets for testing, validation, and training. Train the model using the training data and fine-tune hyperparameters using the validation set to prevent overfitting.

Evaluation: Evaluate the trained model's performance on the test set using appropriate evaluation metrics. Assess metrics like accuracy, precision, recall, F1-score, etc., to gauge the model's effectiveness in fake review detection.

Cross-validation: Perform cross-validation techniques (e.g., k-fold cross-validation) to ensure robustness and generalization of the model across different data splits.

The proposed approach for fake review detection employs a combination of machine learning algorithms, including Stochastic Gradient Descent (SGD), and XGBoost, to effectively identify deceptive reviews. The dataset is divided into training and testing sets with an 80-20 split. An analysis of the efficacy of fake review filters in bolstering user confidence reveals potential drawbacks, such as increased costs leading to heightened skepticism. Feature extraction and sentiment analysis are conducted on the dataset, incorporating features like spam (1) and not spam (0), along with features devoid of stop words. Leveraging multiple machine learning algorithms simultaneously enhances the system's accuracy and robustness in detecting fake reviews.

RESULT

A series of tests were conducted to evaluate the performance of the trained and tested model. A number of evaluation criteria were used to evaluate the model's performance, including F1 score, accuracy, precision, and recall. The proposed system was particularly scrutinized for its ability to identify false reviews using the dataset, with findings duly justified. In testing our models, we experimented with both oversampling and under sampling techniques to address class imbalances.

The sentiment analysis of fake reviews serves as a valuable tool for consumers in selecting the best services and goods while gaining insights from public opinions. Table 1 presents studies related to detecting fake reviews in kindle, outlining various approaches utilized by researchers worldwide.

These approaches, including the algorithms employed and accuracy conclusions, contribute to the understanding of kindle perspectives and the efficacy of fake review detection methods.

Overall, the results obtained from these tests and analyses provide valuable insights into the performance of the proposed system and its effectiveness in identifying fake reviews in online platforms. The combination of sentiment analysis and machine learning techniques offers promising opportunities for improving consumer decision-making and ensuring the authenticity of online reviews.

Class	Precision	Recall	F1-Score	Support
Fake	0.922553	0.960906	0.941339	8876
Real	0.959205	0.919324	0.938841	8875
macro avg	0.940879	0.940115	0.94009	17751
weighted avg	0.940878	0.940116	0.94009	17751

Accuracy: 0.9401160498000113
Rand Index: 0.7747958835990076

Table 1. Performance metrics

CONCLUSION

This paper has presented a comprehensive analysis of fake review detection utilizing three distinct machine learning approaches, yielding notably similar results across the methodologies employed. The current work demonstrates a robust capability to accurately identify false reviews within the Kindle dataset, showcasing effectiveness in handling both balanced and unbalanced datasets.

Our methodology has showcased significant efficacy, particularly in addressing imbalanced datasets with binary classes. In comparison to prior studies, the proposed model in this research accurately detects fake reviews within the Kindle dataset, highlighting advancements in fake review detection methodologies.

REFERENCES

[1] Julia Kiseleva, Gianmarco De Francisci Morales, and Alejandro Bellogín. "Fake review detection: a tutorial." In International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1443-1446. 2021.

[2] Qianqian Han, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. "Towards fine-grained fake reviews detection via adversarial training." In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 2325-2328. 2021.

[3] Fatemeh Niksirat, Hamed Zamani, and W. Bruce Croft. "Mitigating the Impact of Cross-Domain Sentiment Analysis on Fake Review Detection." In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 2046-2049. 2021.

[4] Jinoh Oh, Byungsoo Jeon, Sang-Wook Kim, and Changki Lee. "Multimodal fake review detection with fine-grained attention mechanism." Information Processing & Management 58, no. 1 (2021): 102316.

[5] Yi Yang, Liang Wu, Mingyang Zhang, and Jianjun Ma. "Fine-grained fake review detection: a multi-task learning approach." Information Processing & Management 58, no. 1 (2021): 102432.