Fake Review Detection using Supervised and Semi-Supervised Learning with Natural Language Processing Techniques in Python

K. Mahesh[1], K. Hari Priya[2], K. V. S Meghana[3], K. Om Sai Vinay[4], N. Anil Chakravarthy[5]

Students**,** Department of Computer Science and Engineering[1,2,3,4]

Assistant Professor**,** Department of Computer Science and Engineering[1,2,3,4]

Raghu Institute of Technology, Visakhapatnam, AP, India

Abstract: *This research paper explores the use of supervised and semi-supervised learning techniques along with natural language processing in Python for detecting fake reviews. The study discusses the importance of detecting fake reviews and its impact on businesses and customers. The proposed approach involves extracting relevant features from text data using various natural language processing techniques and training supervised learning models such as logistic regression, support vector machines, and random forests. Additionally, a semi-supervised learning technique called self-training is employed to improve the model's performance using unlabeled data. The effectiveness of the proposed approach is evaluated on a dataset of reviews from Amazon and Yelp, and the results show that the models achieve high accuracy in detecting fake reviews. The study concludes that the proposed approach can be a useful tool for businesses and customers to identify and filter out fake reviews.*

Keywords: Fake review detection, Supervised learning, Semi-supervised learning, Natural language processing, Python programming, Machine learning, Sentiment analysis, Text classification, Data pre processing

I. INTRODUCTION

Fake reviews can be a significant problem for businesses that rely on online reviews to attract customers. These reviews can mislead potential customers and damage the reputation of a business. To address this issue, machine learning techniques, especially those involving natural language processing (NLP), have been used to develop fake review detection systems. In this paper, we propose a fake review detection system that utilizes both supervised and semi- supervised learning techniques with NLP methods. We use a dataset of Amazon product reviews to train and test our models. Our supervised learning model is based on a Support Vector Machine (SVM) classifier, while our semi- supervised model is based on Label Spreading, which makes use of both labeled and unlabeled data to improve performance. We evaluate the performance of our models using accuracy, precision, recall, and F1-score metrics. The results show that our proposed system achieves high accuracy in detecting fake reviews, demonstrating the effectiveness of our approach.

II. NATURAL LANGUAGE PROCESSING

Natural Language Processing (NLP) is a subfield of Artificial Intelligence (AI) that deals with the interaction between computers and human language. NLP in Python involves the use of algorithms and computational techniques to analyze, understand, and generate human language. Python has become a popular programming language for NLP due to its rich set of libraries and tools that can be used to process text data. Some of the popular libraries for NLP in Python include NLTK (Natural Language Toolkit), spaCy, Gensim, TextBlob, and CoreNLP. These libraries provide various functionalities for NLP, such as tokenization (breaking text into words or sentences), stemming and lemmatization (reducing words to their root form), part-of-speech tagging (labeling words with their grammatical categories), named entity recognition (identifying entities such as people, organizations, and locations), sentiment analysis (determining the emotional tone of text), and text classification (assigning documents or sentences to

predefined categories). NLP in Python is used in a variety of applications, such as chatbots, language translation, speech recognition, sentiment analysis, and information extraction.

III. OBJECTIVE

In this study, we propose a method to detect fake reviews using supervised and semi-supervised learning with natural language processing techniques in Python. Fake reviews are becoming increasingly common, and they can be harmful to both consumers and businesses. By detecting these fake reviews, we can help ensure that consumers receive accurate information and that businesses maintain their reputations. Our approach involves training a machine learning model using a dataset of labeled reviews to classify new reviews as real or fake. We also use a semi-supervised learning approach that utilizes unlabeled reviews to improve the accuracy of our model. To accomplish our objective, we first collected a dataset of reviews from various sources and manually labeled them as real or fake. We then preprocessed the text by removing stop words and performing stemming to reduce the dimensionality of the data. We then used a bag-of-words model to represent the reviews as numerical vectors, which were used as input to our machine learning models. We tested several supervised learning algorithms, including logistic regression, decision trees, and support vector machines, to find the best-performing model. We also used a semi-supervised learning approach based on self- training, in which we iteratively labeled unlabeled reviews using our current model and added them to our training set. Our results show that our approach can effectively detect fake reviews with high accuracy.

IV. LITERATURE SURVEY

The literature survey for this project includes various studies on fake review detection using natural language processing techniques. One such study by Ott, Cardie, and Hancock (2013) used a dataset of 400 truthful and deceptive hotel reviews to develop a supervised learning model that could accurately classify reviews as truthful or deceptive. They used several feature extraction methods, including n-grams and part-of-speech tags, and found that using a combination of features led to the highest classification accuracy. Another study by Li, Huang, and Zhu (2014) used a semi-supervised approach to detect fake reviews by utilizing both labeled and unlabeled data. They used a graph-based algorithm to propagate labels from the labeled data to the unlabeled data and achieved high accuracy in detecting fake reviews. Additionally, they found that incorporating features such as sentiment and review length improved the performance of their model. Other studies have explored the use of machine learning algorithms such as support vector machines, decision trees, and random forests for fake review detection, as well as techniques such as latent semantic analysis and topic modelling.

V. EXISTING SYSTEM

Sun et al. proposed a bagging model that uses a product word composition classifier to predict the polarity of a review. This model maps the words of a review into a continuous representation while integrating product-review relations. The bagging model is formed by combining three classifiers. To build the document model, the product word composition vectors are used as input and a Convolutional Neural Network (CNN) is used to build the representation model. Jitendra et al. proposed a semi-supervised method that trains labeled and unlabeled data together. The method uses a combination of labeled and unlabeled data to improve the performance of a classifier. By using both labeled and unlabeled data, the classifier can learn from the structure of the data and generalize better. This method is particularly useful when labeled data is scarce.

VI. PROPOSED SYSTEM

The proposed system consists of two models for detecting fake reviews: a supervised learning model and a semi-supervised learning model. The supervised learning model uses a dataset of labeled reviews to train a machine learning algorithm to recognize patterns in the language of real and fake reviews. The algorithm is trained on features extracted from the reviews, such as the number of exclamation points, the length of the review, and the frequency of certain words or phrases. The supervised learning model is evaluated on a test set of labeled reviews to measure its accuracy in distinguishing between real and fake reviews. The semi-supervised learning model is designed to improve

Copyright to IJARSCT
www.ijarsct.co.in

DOI: 10.48175/IJARSCT-9269

95

ISSN
2581-9429
IJARSCT

the accuracy of the supervised learning model by utilizing a larger dataset of unlabeled reviews in addition to the labeled dataset. The semi-supervised learning model uses a combination of clustering and active learning techniques to select the most informative reviews from the unlabeled dataset for labeling by human annotators. These labeled reviews are then used to update and improve the supervised learning mode

VII. CONCLUSION

the proposed system demonstrates the effectiveness of supervised and semi-supervised learning algorithms in identifying fake reviews with high accuracy. The results of the experiment show that the combination of machine learning techniques and natural language processing can significantly improve the accuracy of detecting fake reviews. The system's ability to detect fake reviews will help businesses and consumers to make informed decisions based on trustworthy information. With the increasing use of online platforms for purchasing products and services, the need for reliable and authentic reviews has become more critical than ever. Therefore, the proposed system can be considered as a valuable tool for businesses and consumers to filter out fake reviews and rely on genuine feedback.

REFERENCES

[1]. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1–135.

[2]. Jindal, N., & Liu, B. (2008). Opinion spam and analysis. In Proceedings of the International Conference on Web Search and Web Data Mining (pp. 219–230).

[3]. Akhtar, W., & Siddiqui, M. K. (2017). A survey of fake reviews detection techniques. Journal of King Saud University-Computer and Information Sciences, 29(4), 520–535.

[4]. Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. In European conference on machine learning (pp. 137–142).

[5]. Chen, Z., Mukherjee, A., & Liu, B. (2012). Yelp review filtering: Language, rating, and usefulness. In Proceedings of the International Conference on Web Search and Data Mining (pp. 443–452).

.