SPAM REVIEW IDENTIFICATION USING PRE-TRAINED WORD EMBEDDING AND WEIGHTED SWARM SUPPORT VECTOR MACHINES

[1]Nellore Mounika, [2]Tulasi Venkata Naga Sujitha
[1,2]UG Student, [1,2]Department of Computer Science & Engineering, Geethanjali Institute of Science and Technology, Gangavaram, Andhra Pradesh, India

ABSTRACT
Online reviews are important information that customers seek when deciding to buy products or services. Also, organizations benefit from these reviews as essential feedback for their products or services. Such information required reliability, especially during the Covid19 pandemic which showed a massive increase in online reviews due to quarantine and sitting athome. Since all reviews are not trustworthy. to address this issue, the proposed approach utilizes a combination of weighted swarm support vector machines and harries hawks optimization.
The HHO algorithm is used to optimize hyperparameters and feature weighting, while the WSVM is employed for spam review detection. In addition, pre-trained word embedding using bi-directional encoder representation from transformers (BERT) and three-word representation methods (NGram-3, TFIDF, andOne-hot encoding) are applied.

INTRODUCTION
NATURAL LANGUAGE PROCESSING
NLP stands for Natural Language Processing, which is a part of Computer Science, Human language, and Artificial Intelligence**.** It is the technology that is used by machines to understand, analyse, manipulate, and interpret human's languages. It helps developers to organize knowledge for performing tasks such as translation, automatic summarization, Named Entity Recognition (NER), speech recognition, relationship extraction, and topic segmentation.

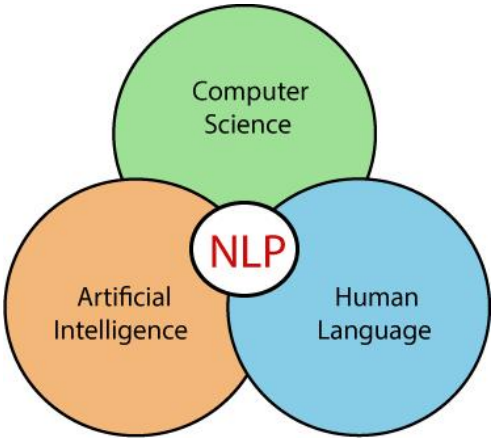

Fig: 1 NLP

Natural Language Processing (NLP) allows machines to break down and interpret human language. It's at the core of tools we use every day – from translation software, chatbots, spam filters, andsearch engines, to grammaqr correction software, voice assistants, and social media monitoring tools.

SPAM REVIEWS
Spam reviews are reviews that are not genuine or authentic. They are usually created with the intention to deceive or manipulate others. Spam reviews can be found on various platforms such as online marketplaces, review websites, or social media platforms. These reviews often contain false information,

exaggerated claims, or biased opinions. They can be harmful because they mislead consumers and can negatively impact businesses or products. It's important to be cautious and critical when reading reviews to ensure that you're getting reliable and trustworthy information.

Fig.2. what is Spam

OBJECTIVE

The main objective of our project is to detect the spam reviews. A dataset containing a large number of reviews, some of which are spam (fake or misleading) and others are legitimate, the task is to develop a model capable of automatically identifying spam reviews. The model should leverage pre-trained word embeddings to capture semantic information from the text and utilize a weighted Support Vector Machine (SVM) classifier to make predictions. The goal is to achieve high accuracy in distinguishing between spam and legitimate reviews while handling class imbalances in the dataset.

LITERATURE SURVEY

In this chapter we review some papers to get knowledge and understanding on the techniques had been proposed. All those techniques have the same aim which is identify the reviews being spam or not. As Archimedes once said, “Man has always learned from the past. After all, you can't learn history in reverse!” it is essential for man to learn from history. Thus, considering all past researches, the most relevant research glimpses have been picked to be explained in detail. The overview shall discuss relevant aspects contributing to our research.

Spam review detection using self attention based CNN andbi-directional LSTM

Dr Bhuneswari, A.Nagaraja Rao and Haraldo Robinson

Opinion reviews are a valuable source of information in e-commerce. Indeed, it benefits users in buying decisions and businesses to enhance their quality. However, various greedy organizations employ spammers to post biased spam reviews to gain an advantage or to degrade the reputation of a competitor. This results in the explosive growth of opinion spamming. Due to its nature and their increasing volume, spam reviews are a fast-growing serious issue on the internet. Until now, researchers have developed many Machine Learning (ML) based methods to identify opinion spam reviews. However, the traditional ML methods cannot effectively detect spam messages due to the limited feature representations and the datamanipulations done by spammers to escape from the detection mechanism. As an alternative to ML-based detection, in this paper, we proposed a Deep Learning (DL) based novel framework called Self Attention-based CNN Bi-LSTM (ACB) model to learn document levelrepresentation for identifying the spam reviews. Our approach computes the weightage of eachword present in the sentence and identifies the spamming clues exists in the document with an attention mechanism. Then finally, sentence vectors are combined using Bi-directional LSTM (Bi-LSTM) as document feature vectors and identify the spam reviews with contextual information. The evaluated experiment results are compared with its variants and the result shows that ACB outperforms other variants in terms of classification accuracy.

HOTFRED: A Flexible Hotel Fake Review Detection System

Barbara Kelle, Rainer Schmidt, Matthias Gutmann and Michael Mohring

The importance to cope with online fake reviews in Tourism becomes more and more evident. In the hotel sector hoteliers as well as guests often struggle with the challenges to separate trueand fake reviews from each other. Therefore, our research introduces HOTFRED - a flexible hotel fake review detection system - as part of an on-going research project. By combining different analytical approaches, the HOTFRED system indicates via an aggregated probabilitywhether a review is true or fake. As the evaluation of the prototypical implementation showed,this approach can support to detect fake reviews. Many different stakeholders in the Tourism sector can profit from this automatic tool. Thus, hoteliers can take measures to safe their reputation, guests can benefit in their decision-making process and research might use the toolas an initial starting point for future research in the area of fake information.

Improving Opinion Spam Detection by Cummulative RealativeFrequency Distribution

Marinella Petrocchi, Francesco Buccafurri, and Michela Fazzolari

Over the last years, online reviews became very important since they can influence the purchase decision of consumers and the reputation of businesses, therefore, the practice of writing fake reviews can have severe consequences on customers and service providers. Various approaches have been proposed for detecting opinion spam in online reviews, especially based on supervised classifiers. In this contribution, we start from a set of effectivefeatures used for classifying opinion spam and we re-engineered them, by considering the Cumulative Relative Frequency Distribution of each feature. By an experimental evaluation carried out on real data from Yelp.com, we show that the use of the distributional features is able to improve the performances of classifiers.

Fact or Factious? Contextualized spam detection

Kennedy S, Sloka K and Walsh N

In recent years, it has been shown that falsification of online reviews can have a substantial, quantifiable effect on the success of the subject. This creates a large enticement for sellers to participate in review deception to boost their own success, or hinder the competition. Most current efforts to detect review deception are based on supervised classifiers trained on syntactic and lexical patterns. However, recent neural approaches to classification have been shown to match or outperform state-of-the-art methods. In this paper, we perform an analytic comparison of these methods, and introduce our own results. By fine-tuning Google's recently published transformer-based architecture, BERT, on the fake review detection task, we demonstrate near state-of-the-art performance, achieving over 90% accuracy on a widely used deception detection dataset.

Spam review detection using spiral cuckoo search clustering

Avinash Pandey and Dharmveer Singh Rajput

Nowadays, online reviews play an important role in customer’s decision. Starting from buyinga shirt from an e-commerce site to dining in a restaurant, online reviews has become a basis ofselection. However, peoples are always in a hustle and bustle since they don’t have time to payattention to the intrinsic details of products and services, thus the dependency on online reviewshave been hiked. Due to reliance on online reviews, some people and organizations pompouslygenerate spam reviews in order to promote or demote the reputation of a person/product/organization. Thus, it is impossible to identify whether a review is a spam or aham by the naked eye and it is also impractical to classify all the reviews manually. Therefore,a spiral cuckoo search based clustering method has been introduced to discover spam reviews. The proposed method uses the strength of cuckoo search and Fermat spiral to resolve the convergence issue of cuckoo search method. The efficiency of the proposed method has been tested on four spam datasets and one Twitter spammer dataset. To validate the efficacy of proposed clustering method it is compared with six metaheuristics clustering methods namely; particle swarm optimization, differential evolution, genetic algorithm, cuckoo search, K-means, and improved cuckoo search. The experimental results and statistical analysis validate that the proposed method outruns the existing methods.

Spam review detection using Ensemble Machine Learning
Ayushi Jain, Prabhat Kumar, and Shewtha mani
The importance of consumer reviews has evolved significantly with increasing inclination towards e-Commerce. Potential consumers exhibit sincere intents in seeking opinions of other consumers. These consumers have had a usage experience of the products they are intending to make a purchase decision on. The underlying businesses also deem it fit to ascertain commonpublic opinions regarding the quality of their products as well as services. However, the consumer reviews have bulked over time to such an extent that it has become a highly challenging task to read all the reviews and detect their genuineness. Hence, it is crucial to manage reviews since spammers can manipulate the reviews to demote or promote wrong product. The paper proposes an algorithm for detecting the fake reviews. Since the proposed work concentrates only on text. So, n-gram (unigram + bigram) features are used. Supervised learning technique is used for reviews filtering. The proposed algorithm considers the combination of multiple learning algorithms for better predictive performance. The obtained results clearly indicate that using only simple features like n-gram, Ensemble can boost efficiency of algorithm at significant level.

2.8 Fake Review Detection in Big Data Using Parallel BBO
Ashish Kumar Tripathi, Kapil Sharma and Manju Bala
Online reviews are increasingly used by the customers while purchasing a product or service. In the last few years, significant growth in the number of fake reviews has been observed due to the increasing competition among the e-commerce sites. Thus, fake review detection is an open and challenging problem. However, the majority of research in this field has focused on sequential algorithms which provide inferior results when scaled on the big datasets. To mitigate this problem, this paper presents a novel parallel bio-geography optimization based method to unfold the fake review detection in the big data environment. The experimental analysis is performed on two standard fake review datasets and compared with K-means and 4state-of-the-art methods in terms of accuracy. The results affirmed that the proposed method has surpassed all the other considered methods on both the dataset. Further, the parallel performance potency of the proposed method is validated by analyzing the speedup performance.

EXISTING SYSTEM
Review Spam Detection Using Word Embeddings and Deep Neural Networks is used in existing model.
The goal is to develop a system that can automatically detect spam in project reviews.
To achieve this, the project utilizes word embeddings, which are vector representationsof words that capture their semantic meaning. These embeddings help the system understand the context and relationships between different words in the reviews.
The project also leverages deep neural networks, which are powerful machine learning models capable of learning complex patterns and relationships in data. The neural network is trained on a labeled dataset of project reviews, where each review is labeledas spam or non-spam.
During the training process, the neural network adjusts its internal parameters to minimize the difference between its predicted labels and the true labels of the training examples. This enables the network to generalize its learning and make accurate predictions on new, unseen reviews.

PROPOSED SYSTEM
In this project we proposed two techniques for detecting, identifying and filteringspam reviews. There are
Pre-trained word embeddings
Weighted swarm support vector machines

Pre-trained word embeddings are pre-computed vector representations of words. They capture semantic and syntactic information about words based on their context. These embeddings are trained on large amounts of text data, allowing them to capture meaningful relationships between words. They are often used in natural language processing tasks like sentiment analysis, machine translation, and text classification.

Weighted swarm support vector machines (WSSVMs) are a variant of support vector machines (SVMs) that incorporate the concept of swarm intelligence. In WSSVMs, the training data is divided into multiple subsets, or "swarms," and each swarm is assigned a weight basedon its importance. These weights are then used to adjust the contribution of each swarm duringthe training process. The goal is to optimize the classification performance by giving more emphasis to the swarms that contain more informative data points. WSSVMs have been shownto improve the accuracy of SVMs in certain scenarios, particularly when dealing with imbalanced dataset.

SYSTEM STUDY

This study is carried out to check the economic impact that the system will have onthe organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

SYSTEM ARCHITECTURE

The System consists of the following steps :-

1. Data collection
2. Data Preprocessing
3. Word Embeddings
4. Feature Extraction
5. Classification

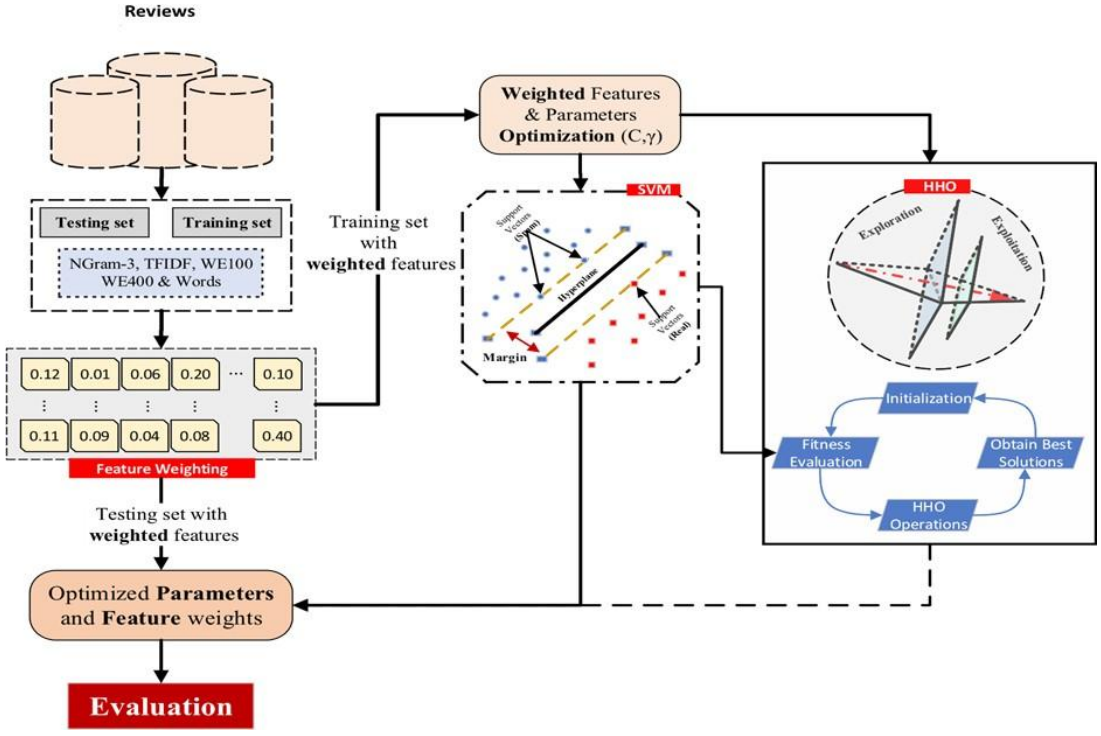

Fig.3. System Architecture

DATA COLLECTION:

Data collection is the process of gathering information or data from various sourcesfor analysis, research, or

decision-making purposes. It involves systematically collecting relevant data points or observations to address specific questions, test hypotheses, or achieve objectives. Data collection can take place through various methods, including surveys, interviews, observations, experiments, and automated data retrieval from digital sources.

DATA PREPROCESSING:

Data preprocessing is a fundamental step in data analysis and machine learning pipelines. It involves cleaning, transforming, and organizing raw data into a format suitable for further analysis or model training. Data preprocessing aims to improve the quality, consistency,and usability of the data by addressing issues such as missing values, noise, outliers, and inconsistencies.

DATA CLEANING: Cleaning the data involves removing any irrelevant or duplicate data, correcting errors, and standardizing formats. This ensures that the data is consistent and accurate.

DATA TRANSFORMATION: Data transformation includes scaling or normalizingthe data to ensure that all variables are on a comparable scale. This is important for certain algorithms that are sensitive to the magnitude of the variables.

Results & Analysis

Evaluation Criteria

In this study, we construct a system for spam reviews detection by using pre-trainedword embeddings and weighted swarm support vector machines. In this method, there areseveral key criteria to consider: a) Accuracy of the model. b) Precision and recall. c) F1 score, which is the harmonic mean of precision and recall d) Calculate the model performson unseen data by splitting the dataset into training and testing sets.

The execution of the process will be explained clearly with the help of the continuous screenshots. The whole process in the execution is taking a dataset which contains the online reviews. and it converts the given text data into vectors by using wordembeddings after that system will classify which is spam review or not by using weightedswarm support vector machine algorithm. Finally display the review spam or not. This whole process is done in five simple steps. Each figure mentioned below are the simultaneous process of screening outputs.

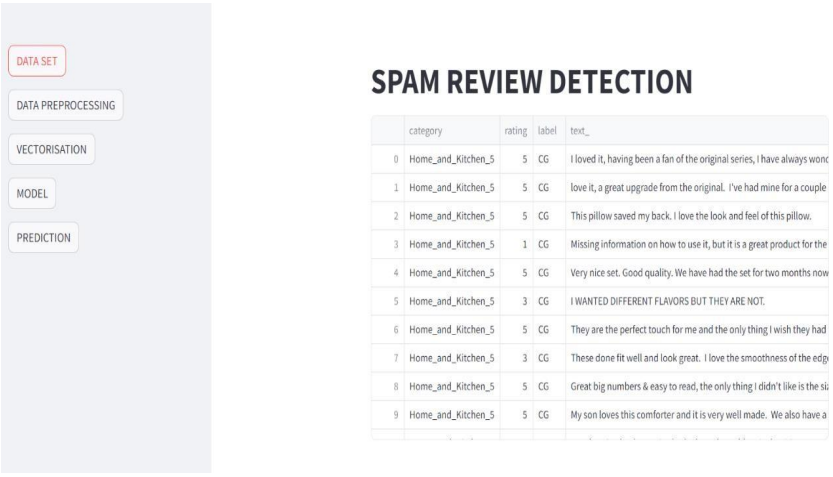

Fig.4. Dataset Containing Reviews

Description : Fig 1, describes the dataset which contains the combination of spam reviewsand original reviews.

This dataset is collected from online. After uploading the dataset the further steps such asdata preprocessing, vectorization and prediction will be started.

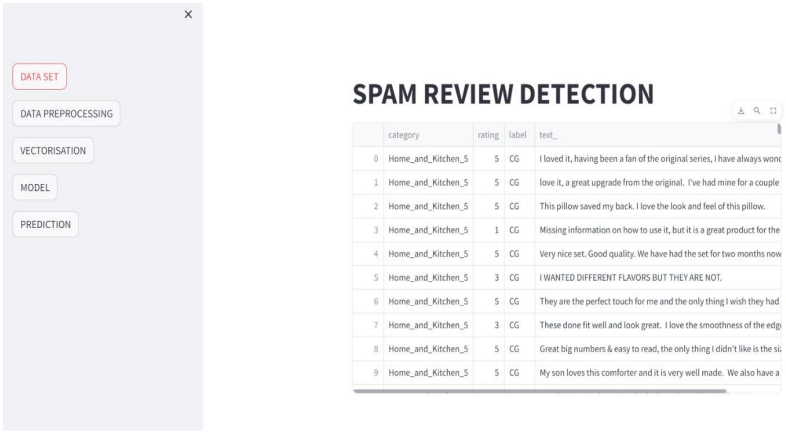

Fig.5. Data Preprocessing Description :

The above figure 5 represents that the preprocessed data.

Data preprocessing aims to improve the quality, consistency, and usability of thedata by addressing issues such as missing values, noise, outliers, and inconsistencies.

Fig.6.Data Vectorisation

Fig.7. Logistic Regression

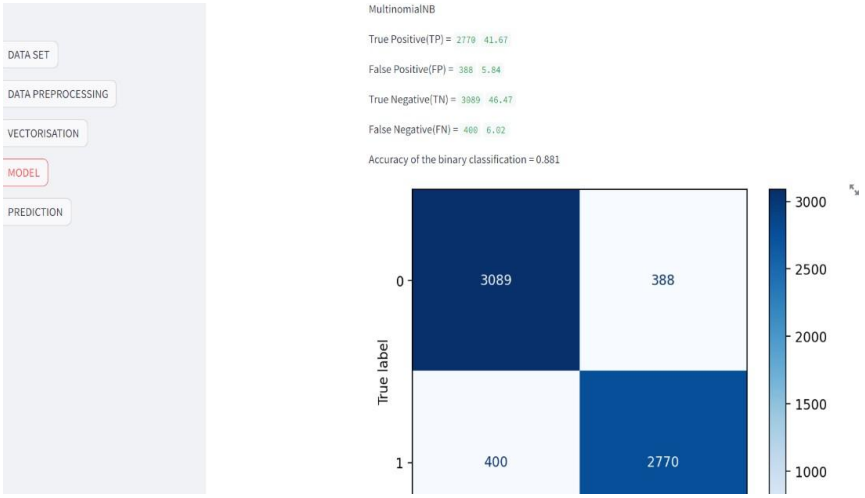

Fig.8. Multinomial NB

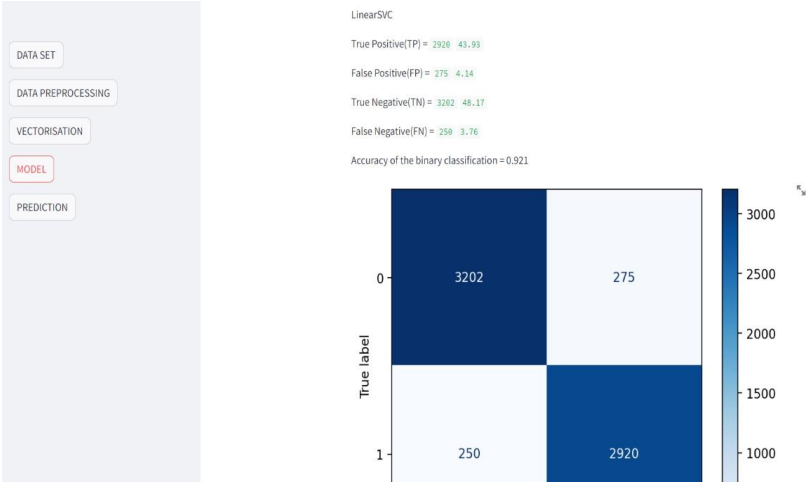

Fig.9. Linear SVC

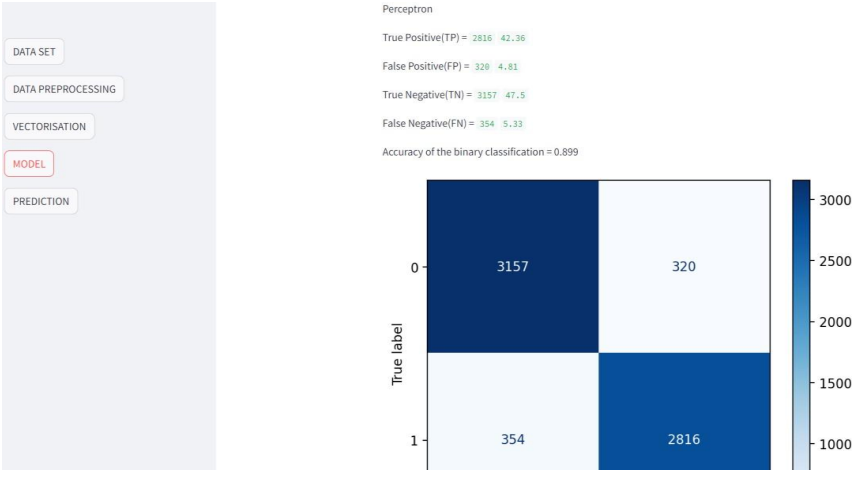

Fig.10. Perceptron

Description : Above figures represents the different types of models prediction like linear regression, multinomial NB, linear SVC and perceptron.

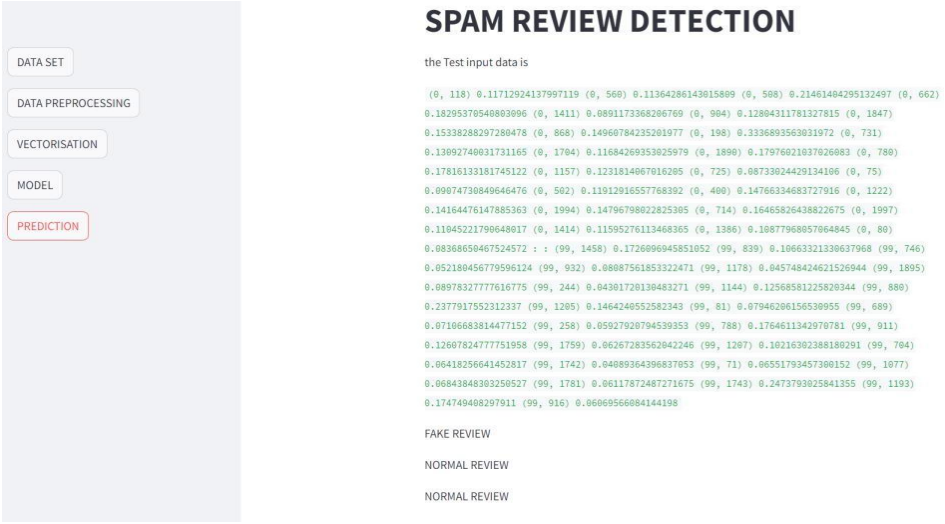

Fig.11. Reviews Prediction

Description : Above figure represents the report displays the given test data and also given reviews is spam or not.

CONCLUSION

In this project we have effectively identifying and filtering out the spam reviews using Pre-trained word embeddings and Weighted swarm support vector machines. By utilizing pre-trained word embeddings, which capture the semantic meaning of words, the algorithm can better understand the context and sentiment of reviews. This helps in distinguishing between genuine and spam reviews. The weighted swarm support vector machines (WSVM) technique an make moreaccurate predictions and effectively classify reviews as spam or not. Overall, the combination ofpre-trained word embeddings and weighted swarm SVM offers a robust and efficient approach to detect spam reviews. It has the potential to improve the quality and reliability of online review platforms by filtering out deceptive or misleading content.

FUTURE ENHANCEMENTS

After many efforts we had successfully identifying and filtering out the spam reviews very quickly and efficiently by using Pre-trained word embeddings and Weighted swarm support vector machines. But,we have to add the some more features to enhance the performance.

Language Expansion: Include support for additional languages to enhance the system's coverage and usability.

User Feedback Integration: Integrate user feedback mechanisms to further improve model accuracy and relevance to specific domains or contexts.

Multimodal Data Analysis: Extend the system to analyze multimodal data, such as text combined with images or audio, for a more comprehensive spam detection approach.

References

[1] M. Mohring, B. Keller, R. Schmidt, M. Gutmann, and S. Dacko, ‘‘HOTFRED: A flexible hotel fake review detection system,’’ in Information and Communication Technologies in Tourism 2021. Berlin, Germany: Springer, 2021, pp. 308–314.

[2] P. Bhuvaneshwari, A. N. Rao, and Y. H. Robinson, ‘‘Spam review detection using self attention based

CNN and bi-directional LSTM,'' Multimedia Tools Appl., vol. 80, pp. 18107–18124, Feb. 2021

[3] A. K. Tripathi, K. Sharma, and M. Bala, ''Fake review detection in big data using parallelBBO,'' Int. J. Inf. Syst. Manag. Sci., vol. 2, no. 2, pp. 288–298, 2019.

[4] S. Mani, S. Kumari, A. Jain, and P. Kumar, ''Spam review detection using ensemble machine learning,'' in Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit. Cham, Switzerland: Springer, 2018, pp. 198–209.

[5] A. C. Pandey and D. S. Rajpoot, ''Spam review detection using spiral cuckoo search clustering method,'' Evol. Intell., vol. 12, no. 2, pp. 147–164, Jun. 2019.

[6] A. Barushka and P. Hajek, ''Review spam detection using word embeddings and deep neural networks,'' in Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov. Cham, Switzerland: Springer, 2019, pp. 340–350.

[7] S. Kennedy, N. Walsh, K. Sloka, J. Foster, and A. McCarren, ''Fact or factitious? Contextualized opinion spam detection,'' 2020, arXiv:2010.15296.

[8] S. P. Rajamohana, K. Umamaheswari, and B. Abirami, ''Adaptive binary flower pollination algorithm for feature selection in review spam detection,'' in Proc. Int. Conf. Innov. Green Energy Healthcare Technol. (IGEHT), Mar. 2017, pp. 1–4.