Role of Machine Learning in Fake Review Detection

P. Manish Kumar[1], S. Shri Harrsha[1], K. Abhiram [1], Dr. M. Kavitha [1], M Kalyani[2]

[1] Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.

manishpemmasani123@gmail.com, shriharrsha1@gmail.com,

kondaabhideeksha12@gmail.com , mkavita@kluniversity.in

[2]Department of Mathematics, PACE institute of technology and sciences, Ongole, AP, India.
modepallikalyani123@gmail.com

Abstract—

In today's culture the growing technology is promoting a lot of products and events in a very positive way. Technology usage in current generation has taken a new step in reaching great heights. But when a technology brings in so much positiveness it also has its own negative usage and one among them is the fake reviews. Fake reviews are weakening the actual worth of the product. To be more specific, the reviews can be divided into two categories: legitimate fake reviews and reviews written intentionally to decapitate the product or brand value. On the other hand, the machine learning algorithms are extensively used. The incorporation of machine learning techniques into the classification of the reviews is considered as an excellent combination. In this work, various datasets from different industries such as airline industry, movie industry and food industry are considered and fake reviews are classified using various algorithms including K-Nearest Neighbors, Naive Bayes, Random Forest, Decision tree, Support Vector Machine, Logistic Regression from Machine learning. There are reviews which can be decoded using the sentiment analysis from Natural Language Programming. Sentiment analysis is used to find the emotion in a text. The accuracy parameter result is analyzed for all the implemented models. The results demonstrate support vector machine technique giving high accuracy compared to other machine learning classification techniques.

Keywords—Fake reviews, Machine learning, Natural language processing, Sentiment analysis

I. INTRODUCTION

Any user who comes online to checkout a product first reaches to the review section to know what the other users experience on the intended product. But what if those reviews are polluted with fake reviews? How can a user decide what to buy and what not to buy? Hence there is demand for fake review detection mechanisms.

The fake review system segregates the reviews which are really genuine, and which are not. There are reviews which are truthful to the product and some of them mislead the buyer. Different datasets are taken from different industries. When a garbage data is taken and used in the algorithms then the result might not be appropriate. A raw data will be in different forms. While considering different datasets, the raw data should be converted into a single form. Further, the pre-processing techniques are applied by considering the raw data, which is then converted to fit in to algorithms. Data quality assessment, Data cleaning, Data transformation, Data reduction are few steps in "Data Pre-processing". Any sentence which is said contains its own emotion and it can be understandable. But when an emotion of a written sentence is to be known then there is the real task. For that, the Natural language programming techniques are considered to know more about the emotion of a sentence or a word in particular. Stop words are frequently employed in text mining and natural language processing (NLP) to filter out terms that are overused and provide very little valuable information. "a", "the", "is", "are" and etc. are some of the stop words. Tokenization is also one of the techniques used and it is the process of tokenizing or splitting a string, text into a list of tokens. A sentence is a token in a paragraph. Classification techniques such as Naive Bayes [1] and Decision Tree SVM are available [11] similarly, Linear SVP is also used. In some of the papers authors used Multidimensional Feature Engineering for better results [9].

Many small scale industries completely reside their business on word of mouth which are also called as reviews. Industries like Movie industry ,Amazon shopping [5] gets most of its revenue from the positive word of mouth. And these fake reviews are misleading the common audience or the user by not letting them to give a try.

As the Algorithms are such that the most liked reviews are shown first. The misleading fake reviews are shown on the top and the genuine ones are pushed to the bottom This work mainly aims to clear the bogus in the reviews which are destroying not only the product owner but

also the user experience and present the positive reviews in any platform. A model is proposed with the help of various machine learning techniques [2, 12]. Cclassifiers are applied to classify the reviews into Fake and Genuine reviews. Any user who goes online and checks the products ask for a genuine review. This proposed model will clear the issues for the user.

II. LITERATURE WORK

The author utilised TF-IDF to efficiently distinguish false and true hotel reviews using a dataset of gold standard hotel ratings. Author discussed three Naive Bayes, logistic regression, and support vector machines in his work. They have obtained a validation set using a multinomial Naive Bayes classifier [1].

The author took movie review dataset and used various machine learning algorithms like K*, Naive bayes, SVM, KNN algorithms. After testing all the algorithms SVM surpasses as the best accuracy among the other classification algorithms [2].

The author has done the literature survey of various papers to know which algorithm is giving accurate value and he has gone through techniques like Naive baye's from Machine learning and LSTM, Bidirectional LSTM, GRNN. And he finally got the highest accuracy in Naive bayes from machine learning algorithms and LSM gave the highest accuracy from couple of techniques. He got 98.9% accuracy in deep learning that in bidirectional- LSTM for filtering words. He also used maximum entropy, KNN,K-star algorithms and checked various publications and concluded the above accuracy from naive bayes which he got as the highest amongst them [3,13].

The author has used LIAR dataset and used pre-processing techniques to know the sentiment analysis and then have used various algorithms in machine learning. There are learning techniques like RNN, CNN, LSTM, GRU, Logistic regression and SVM. Among which CNN has done extremely well showing its best accuracy of 0.270 and the other test accuracies are as follows SVM(0.255), Logistic regression(0.247), Bi-LSTM(0.233), GRU(0.217), LSTM(0.2166) [4].

The author utilised the Amazon Review Data (2018) dataset to analyse up to 10M reviews from Amazon.com in an effort to identify different sorts of opinion spam. The fact that they automatically labelled totally copied and nearly replicated reviews as false reviews undermines the legitimacy of the results even if they attained a respectable performance. To create false product reviews, they utilised two language models. ULMFiT- Universal language Model Finetuning and GPT-2. The author asserts that of the four prediction sources, the fake Roberta model performed the best Compared to the other ML model, the OpenAI model fared much worse [5].

WEKA Tool (Waikato Environment for Knowledge Analysis) used in data mining jobs, it is a tool for gathering machine learning algorithms. categorization, regression, clustering, association rules, and visualisation are all methods for processing data. The author has used NB, DT-J48, LR and SVM algorithms to analyse Amazon reviews datasets [6].

The test results for the algorithms utilised by the author, who used the Restaurant Dataset, are as follows: The Decision Tree has the best training accuracy, followed by the XGBT, SVMs, Random Forest, and MLP utilising Doc2Vec document embedding. After hyperparameter adjustment, stand-alone classifiers may obtain up to 68.2% accuracy in the case of MLP. The adaboost ensemble of MLPs allows ensemble learning-based classifiers to reach up to 77.3% accuracy [7].

The author uses the amazon dataset. The author in the paper studies the Fake review system by using convolutional neural network model. For integrating the product related features with the product owned person. The author used bagging model to reduce overfitting and high variance. By using the Bi grams and Tri grams the author got the result [8].

The author used manual annotated dataset. The author in the paper studies the Fake review system by using multidimensional feature engineering. To recognise fake reviews, six feature criteria are created. Relativity of review items and content is determined by -

(1) Analyse through reviews for product characteristics.

(2) Create word vectors from product reviews depending on the features of the item.

(3) Use the x2 statistical approach to choose the correlation product characteristics. [9]

The author combined actual and seemingly fraudulent reviews. Following Author study, behavioural and contextual aspects are crucial for spotting phoney reviews. Their study made use of the crucial reviewer behaviour trait known as "reviewer deviation." NNC, LTC, and BM25 term weighting systems had all been tested. As per authors observation BM25 beat other term weighting schemes [10].

The author had taken a tourism hotel review dataset. Author used Support Vector Machine model for fake review detection and for the second analytical component a spelling checker software tool was developed according to their usage. They used python for programming the software[11].

III. PROPOSED WORK

The literature works reveals that fake review detection is an important research issue because it has great impact in user's daily life. And also literature work demonstrate that machine learning techniques are playing vital role in fake review detection. This research study has applied multiple machine learning techniques to perform fake review detection. The results are verified with multiple data sets. This work supports to demonstrate the role of machine learning techniques in fake review detection.

A. Architecture

Figure 1 represents the working model of proposed fake review detection using machine learning techniques.
The first step is to take a dataset and perform the data mining techniques which are cleaning, clustering, classification. A\and in the next step we need to proceed with the NLP techniques which are removal of tokens and tokenization. And then the sentiment analysis is done. And then the training and testing is performed using six machine learning

algorithms which are discussed below. And we get the final accuracies, precession and recall which are shown in result discussion.
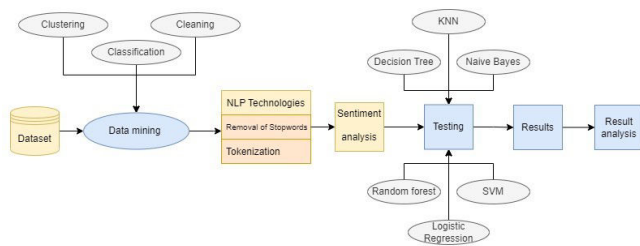

Fig.1. Working model of proposed work

Fig.2. Sample Data Set

B. Discussion on Working Algorithms

KNN

K nearest neighbors' algorithm is used to find the nearest node. And is helpful in finding the group of similar instances. KNN is also a pattern recognition algorithm. And also used in statistical estimation.

SVM

SVM algorithm will work to find out the best separable hyper plane by separating the training data into many classes.it is a discriminated classifier and it is a supervised learning algorithm.

Decision Tree

Decision Tree is one of the most used machine learning algorithms which is used to make the right decision by splitting the data into nodes. It is there to construct a tree on best possible features. It takes entropy, information gain etc. into consideration and relies most on it.

Random Forest

Random forest is a Machine learning algorithm which is there to handle the overfitting problems of decision tree. It constructs a bag of trees from various datasets, and it generates a small number of features in constructing the trees in the forest i.e., the bag of trees.

Naive Bayes

Naive bayes is completely based on bayes theorem which generally relies on the probability of the occurrences. Naive bayes is frequently used in many occasions like text classification and fake reviews detection. And also mainly used in recommendation systems.

Logistic Regression

Logistic Regression is also a supervised learning algorithm which is used in machine learning. It finds the hyperplane which classifies the training data.

C. Dataset Description

To perform any algorithm, the first thing we need is the dataset. We have taken the dataset called "fake reviews dataset" from OSF website. 20k false reviews and 20k actual product reviews make up the dataset of created phoney reviews. OR = Original reviews (human created and authentic),CG = Computer-generated fake reviews. The data set contains the attributes like Category , Rating , text ,Label. Figure 2 represents the Dataset which we have used for the implementation.

IV. IMPLEMENTATION

For implementation, python programming language is used on anaconda platform. Anaconda platform supports vast libraries under python. The collected dataset is loaded into platform and all the proposed models are implemented on it. The accuracy parameter observed to see the performance of the implemented model. Accuracy is one of the parameters to assess the model performance or to find which model is best in comparative with other models in machine learning.

The execution is carried out as per below steps:

--> The dataset of various reviews is taken from OSF and it contains 20,000 CG reviews and 20,000 OG reviews.

-->Pre-processing of the data is done using various methods like Data Quality assessment and data cleaning.

-->Text summarization is done by removal of stop words using NLTK in python

-->To get the meaning of the word to decide whether the word is positive or negative Lemmatization is done.

-->For any algorithm in machine learning we need to have Training and testing data, so we created training and testing data. While implementation 80% of the dataset is considered for training part and 20% of the dataset is used for testing part.

-->Training and testing with the six-machine learning Algorithm we listed earlier on the pre-processed data is done

-->And finally, we get the final accuracies, precession, recall, f1 score, support of all the algorithms, and we get a chance to choose the best among them.

Accuracy of all proposed algorithm is arrived to show the best model among them more than 86%. The Random Forests Classifier and Multinomial Naive Bayes algorithm predicted a precision level of approximately 84%. However, the Decision Tree Classifier performed fake reviews prediction up to an accuracy of just over 73%. The worst performing algorithm was the K Nearest Neighbors algorithm which could only perform the predictions up to an accuracy level of nearly 58%. The evaluation has also got the values of Precession, recall, F1 score and support. Precession is about predicting the specific model. Recall picks the number of positive predictions. F1 score is the evaluation of precession and recall getting the prediction value. Support shows how many times the class has occurred in the dataset.

V. RESULTS DISCUSSION

The proposed model's performance in terms of precision, recall, f1-score and support is given in table-1 and the performance in terms of accuracy is shown in table-2. The SVM Classifier performed the most accurate predictions regarding the fake nature of reviews having a predictive accuracy of just over 88%, closely followed by Logistic Regression which had a prediction accuracy of 86% and the least among all is KNN algorithm which Is 57.66%. which is displayed in table 2.

TABLE I
MODELS PERFORMANCE ANALYSIS OF MODELS PRECISION, RECALL, F1-SCORE AND SUPPORT

Algorithm	Category	Precision	Recall	F1-Score	Support
logistic regression	OR	0.86	0.87	0.86	7119
	CG	0.87	0.86	0.86	7032
Random Forest	OR	0.88	0.79	0.83	7119
	CG	0.80	0.89	0.85	7032
Naïve Bayes Classifier	OR	0.89	0.80	0.84	7119
	CG	0.81	0.90	0.85	7032
Support Vector Machine	OR	0.88	0.89	0.89	7119
	CG	0.89	0.88	0.88	7032
Decision Tree	OR	0.74	0.72	0.73	7119
	CG	0.72	0.75	0.74	7032
K-Nearest Neighbors	OR	0.88	0.18	0.29	7119
	CG	0.54	0.97	0.69	7032

TABLE 2: PERFORMANCE ANALYSIS OF WORKING MODELS IN FAKE REVIEW DETECTION

S.No	Algorithm	Accuracy
1	logistic regression	86.33
2	Random Forest	83.92
3	Naïve Bayes Classifier	84.65
4	Support Vector Machine	88.48
5	Decision Tree	73.27
6	K-Nearest Neighbors	57.30

Fig.3 represents the accuracy values of all the implemented models in graphical manner. The graphical representation helps to understand or to analyze the performance of the models easily. Where the X-axis shows the various algorithms, and the Y-axis shows the accuracies.

The bar graph in figure 3 shows the various accuracies of six algorithms we have listed above. And from that we observe that SVM is highest and KNN is least. By the usage of this current proposed system the sorting of the reviews becomes much easier and make a way for genuine customers.
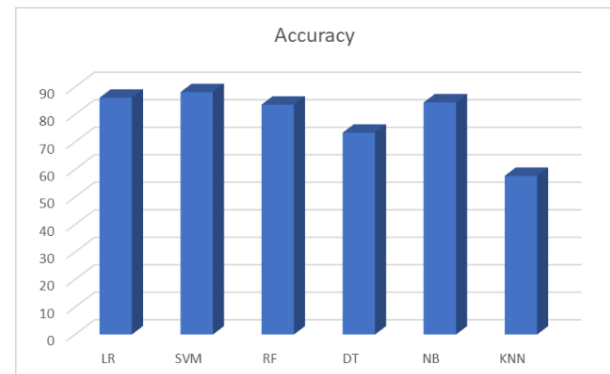

Fig. 3. Accuracy Analysis of Working Models in Fake review detection.

VI. CONCLUSION

As the fake reviews are declining the purchase of the products by the customer. This system of removing the fake reviews and junk from the reviews by the usage of the effective methods from machine learning gives an edge for the product owners. We have got the accuracies of the six techniques we have used from machine learning and the best accuracy is from SVM classifier.

VII. FUTURE WORK

We need to develop more efficient model which has much more accuracy and identify the Fake reviews more accurately. And also, we will focus on removing those fake reviews so that the customer can truly rely on the reviews to but his/her desired product. So, we want to introduce deep learning technique techniques to the existing introduced system to get more accurate results.

REFERENCES

[1] R. Hassan and M. R. Islam, "A Supervised Machine Learning Approach to Detect Fake Online Reviews,"
2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020, pp. 1-6, doi: 10.1109/ICCIT51783.2020.9392727.
[2] L. Gutierrez-Espinoza, F. Abri, A. Siami Namin, K. S. Jones and D. R. W. Sears, "Ensemble Learning for Detecting Fake Reviews," 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 2020, pp. 1320-1325, doi: 10.1109/COMPSAC48688.2020.00-73.
[3] J. C. Rodrigues, J. T. Rodrigues, V. L. K. Gonsalves, A. U. Naik, P. Shetgaonkar and S. Aswale,"Machine & Deep Learning Techniques for Detection of Fake Reviews: A Survey," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-8, doi: 10.1109/ic-ETITE47903.2020.063.
[4] Girgis, Sherry, Eslam Amer, and Mahmoud Gadallah. "Deep learning algorithms for detecting fake news in online text." 2018 13th international conference on computer engineering and systems (ICCES). IEEE, 2018. [5] Salminen, Joni, et al. "Creating and detecting fake reviews of online products." Journal of Retailing and Consumer Services 64 (2022): 102771.

[6] Elmurngi, Elshrif & Gherbi, Abdelouahed. (2018). Unfair reviews detection on Amazon reviews using sentiment analysis with supervised learning techniques.
JCS. 14. 714-726. 10.3844/jcssp.2018.714.726.
[7]L. Gutierrez-Espinoza, F. Abri, A. Siami Namin, K. S. Jones and D. R. W. Sears, "Ensemble Learning for Detecting Fake Reviews," 2020 IEEE 44th Annual Computers,Software, and Applications Conference (COMPSAC), 2020, pp. 1320-1325, doi: 10.1109/COMPSAC48688.2020.00-73.
[8]Sun, Chengai, Qiaolin Du, and Gang Tian. "Exploiting product related review features for fake review detection." Mathematical Problems in Engineering 2016 (2016).
[9]Wang, Ge, et al. "Fake Review Identification Methods Based on Multidimensional Feature Engineering." Mobile Information Systems 2022 (2022).
[10] Kumar, Jay. "Fake Review Detection Using Behavioral and Contextual Features."
arXiv preprint arXiv:2003.00807 (2020).
[11]Möhring, Michael, et al. "HOTFRED: A Flexible Hotel Fake Review Detection System."
Information and Communication Technologies in Tourism 2021. Springer, Cham, 2021. 308
[12]. Kumar, C. N., Keerthana, D., Kavitha, M., & Kalyani, M. (2022, June). Customer Loan Eligibility Prediction using Machine Learning Algorithms in Banking Sector. In 2022 7th International Conference on Communication and Electronics Systems (ICCES) (pp. 1007-1012). IEEE.
[13] Kavitha, M., Srinivas, P. V. V. S., Kalyampudi, P. L., & Srinivasulu, S. (2021, September). Machine Learning Techniques for Anomaly Detection in Smart Healthcare. In 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 1350-1356). IEEE.

ABOUT AUTHOR

Pemmasani Manish Kumar, studying B. Tech in Department of Computer Science and Engineering from Koneru Lakshmaiah Education Foundation, his research interests are Artificial Intelligence, Cloud Computing, Internet of Things. He is certified at Amazon web Services in Solution Architect Associate, Oracle Foundation Associate. Aviatrix Multi-Cloud Networking Associate.

Shri Harrsha Samala, studying B. Tech in Department of Computer Science and Engineering from Koneru Lakshmaiah Education Foundation, his research interests are Artificial Intelligence, Cloud Computing, Internet of Things. He is certified at Amazon web Services in Solution Architect Associate, Aviatrix Multi-Cloud Networking Associate, Wipro Talent Next.

Konda Abhiram, studying B. Tech in Department of Computer Science and Engineering from Koneru Lakshmaiah Education Foundation, his research interests are Artificial Intelligence, Cloud Computing. He is certified at Amazon web Services in Solution Architect Associate.

Dr M Kavitha, working as Associate Professor in the Department of CSE, K L University. She has 11 years of teaching and research experience and has published more than 25 articles in reputable journals and conferences. Participated in more than 50 global challenges on various coding platforms like codechef, hacker rank, hacker earth, leet code, techgig, code forces and top coder. Her interested research areas are Internet of Things, Artificial Intelligence and Machine learning techniques.

M Kalyani working as Assistant Professor in the department of mathematics at Pace institute of technology and sciences. She has 3 years of teaching experience and published a good number of research articles through Scopus indexed conferences and journals. Her research interests are mathematical and computational models.