International Journal of Scientific Research and Engineering Development-– Volume 4 Issue 5, 2021

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Fake Review Detection System Using Machine Learning: A Survey

Mayur Shah*, Anmol Valecha**, Fiza Damri***, Naveen Vaswani****

*(Department of Computer Engineering, Watumull Institute of Electronics Engineering and Computer Technology, Ulhasnagar Email: mayurshah048@gmail.com)

**(Department of Computer Engineering, Watumull Institute of Electronics Engineering and Computer Technology, Ulhasnagar, Email: anmolvalecha14@gmail.com)

***(Department of Computer Engineering, Watumull Institute of Electronics Engineering and Computer Technology, Ulhasnagar, Email: damrifiza@gmail.com)

****(Assistant Professor, Department of Computer Engineering, Watumull Institute of Electronics Engineering and Computer Technology, Ulhasnagar, Email: vaswani.naveen@gmail.com)

----------**********----------

Abstract:

In e-commerce, user reviews play an important role in determining the revenue of an organization or understanding the quality of any product. Online users highly rely on the reviews that are being posted. These online reviews may mislead the user about the products quality or about the business. We did a survey of many papers and tried to compare different algorithms that we can use to find the fake reviews. Some of the algorithms that we can use to detect these fake revies are Decision Tree, Convolution Neural Network (CNN), KNN, Multi-layer perceptron, Naye Bayes, etc. Some of these algorithms are described below.

Keywords - **Fake reviews, detect fake reviews, machine learning, deep learning, Decision Tree, CNN, LSTM, Naïve Bayes, Random Forest, Word2Vec (Doc2Vec).**

----------**********----------

I. INTRODUCTION

There are several factors that contribute to the success or failure of a business, and customer satisfaction is always listed as one of the most important ones. Customer satisfactions and good reputation is a matter of life and death for many businesses. It is important to take customer's reviews into account seriously and address their concerns concisely. As we all know that online shopping has become very popular these days. Today everyone is buying products and services online which leads to millions of reviews produced by the consumer about different products and services. These reviews are mainly influencing the users in taking decision about which product to be purchased and which one not to be purchased from a wide range of e-commerce web sites. These online reviews may also have some specified reasons to be generated based on different situations.

Often, in an effort to reinforce and enhance their businesses, online retailers and repair providers may ask their customers to provide feedback about their experience with the products or services they have bought, and whether or not they were satisfied or not. Customers may feel free to review a product or service provided by the respective company based on the exceptionally good or bad experience with it. To obtain higher profits, unscrupulous merchants usually hire professional writers to write fake positive reviews for their products, so as to increase the popularity of products to attract potential consumers, and at the same time write fake negative reviews for competitors to suppress them.

These behaviours not only seriously mislead potential consumers, but also are not conducive to the stable development of e-commerce to avoid misguide for the consumer we are proposing a system which can detect the fake reviews from

which the costumer will able to prefer honest review by the real user and will not be misguided.

II. LITERATURE SURVEY

In [1], The Author shows us that with the continuous evolve of E-Commerce systems, online reviews are mainly considered as a main part of building or destroying a good reputation. The end user's decision is based on these reviews. Due to fake reviews the virtual reputation is build up and to detect these reviews is one of the most important matter of research. This paper takes the key features of the reviews and the behaviour of the reviewers and perform analysis of that data. It uses various Machine Learning techniques to extract the features of reviews and applies various engineering to extract various behaviours of the reviewers. They used the dataset from the yelp and compared the performances of several classifiers like KNN, Naïve Bayes SVM, etc. The results revealed that the KNN (k=7) gave the best results and an accuracy of 82.40%. The result shows that the f-score has increased by 3.80% when reviewers behavioural feature is considered.

In [2], author tells us that Sentiment Analysis has become one of the main sectors in taking decision of the e-commerce products. Along with its main roles it also has a lot of limitations like injecting fake reviews which are of no use. As the spammer writes more fake reviews, the company or the owner of the product is more concerned about how to detect and remove the fake and bluffing reviews. As we know all the reviews in the list are not genuine, so they tried to extract the key features of the span reviews using Natural Language Processing (NLP). They tried various Machine Learning Algorithms that can be used in detecting fake reviews and compared each other's results to find the best and more accurate method among all. They studied various papers and tried to get all the results under one section as there is no one perfect method of detecting the fake reviews.

In [3], the author tries to show the damage caused to the small businesses through the spread of misleading information or the fake reviews. These revies may promote or damage certain businesses. In the last few years because of this reason various approaches have been proposed so that the credibility of the user generated content can be assessed. The analysis of the main review and the reviewer-centric features are proposed to detect the fake reviews by using supervised machine learning approaches rather than the unsupervised approaches which are based on graphical methods.

In [4], they have used the Sentiment Analysis (SA) and Latent sematic Analysis (LSA) with spam detection algorithm of Netspam. LSA is used to reduce the similar type of comments and increase the spam detection accuracy. They crawled the data from amazon and created their own dataset. Pre-processed the data and performed the SA and LSA on the data to get the results of the comments/reviews. They extracted the features like Calculated the overall weight on the features and then classifying the reviews whether they are spam or not.

In [5], the Ensemble Learning is the method in which we combine one or more weak classifiers in order to get the more accuracy by combining 2 or more classifiers. In this paper they have considered 5 classifiers they are Decision Tree, Randon Forest, Support Vector Machine, Extreme Gradient Boosting Trees, Multilayer Perception. They have created their own data set and labelled it as they wanted and also collected some genuine reviews from the restaurants. They applied Random Forest and XGBT with Decision Tree or bagging, and adaboost ensembles along with SVM and MLP's as week classifiers with optimized hyperparameters. Their stand-alone accuracy was 68.2% and using Ensemble learning gave them the accuracy of 77.3%. Their only used Doc2Vec to convert words into vector matrix and tried their model on only 1 dataset.

In [6], the machine learning algorithms helps in detecting the fake reviews. Once the user uploads the new review the Machine Learning algorithm helps in detecting the fake review easily. There are various algorithms that detects the fake reviews. But the accuracy of the algorithm is based on the training of that model. The researchers have found that Supervised Learning have greater accuracy than unsupervised learning. The Naïve Bayes, Neural Network and the SVM are the ones that give best results than other algorithms. The Naïve Byes gives the accuracy of 96.08% while the Neural Network and the SVM gave 90.90% accuracy. To train the algorithm we take the available dataset, clean the dataset by removing the stopping words or stemming. Test the dataset and training the dataset. At last, the classifier is trained with the dataset and the results are displayed.

In [7], the algorithm takes the data form the dataset and cleans the data. The sentimental analysis of the data is performed. Convert its words into the aspects and assign it the sentiment value. The similarity between the two reviews is found by using the Latent Semantic Analysis. After LSA the Term Frequency – Inverse Document Frequency (TF-IDF) is applied. This method directly woks with the numbers so it takes the text and convert it into the matrix by calculating the logarithmic value of the division of the total numbers of the reviews. Finally, this labelled data is passed on to the classifier to find out which review is fake or which is real one. According to the paper SVM gives the best result by the accuracy of the 84.88%. Furthermore, they can test the model by doing some combination of two or more classifiers and getting more accurate results.

III. ALGORITHMS

Below listed are some of the algorithms that we can use for detecting fake reviews from different websites.

A. Decision Tree:

Decision tress consist of root, nodes, branches and leaf nodes. Every internal node exists a test on an attribute which results into the branches from them. And every leaf node holds a class label. Root node always remains in the top of the decision tree. Given an input of attributes together with its classes, a decision tree generates a sequence of rules that can be used for classifying the data.

A decision tree will be built using the whole dataset taking into consideration all features. Basic algorithm for decision trees

i. start with whole training set.
ii. select attribute or feature satisfying criteria that results into the “best” partition.
iii. create child nodes based on partition.
iv. Repeat process on each child using child data until a stopping criterion is reached. [12]

B. Multi-layer Perception:

In the Multi-layer Perceptron, there is one input layer and one output layer, in between layers are called as the hidden layers. As MLP is sensitive to the feature scaling, so we have to use built-in StandardScalar for standardization. We can define the size of the hidden layers. Suppose we use 3 layers, each consisting of 170 neurons. 5-fold cross validation can be used along with the unigram, bigram and trigram. Then the processed and scaled data is used to train the model by fitting the data into the model. We have to use the predict() method to get the predictions from the model. Finally, we have to evaluate all the predictions results. We can do this by using skikit-learn built-in performance matrix.

C. Long Short-Term Memory (LSTM):

Recurrent Neural Networks suffer from short-term memory because of vanishing gradient problem. So RNNs may leave out relevant information from earlier if a paragraph of text is processed to do predictions. LSTM is a specialized Recurrent Neural Network that is created to mitigate the short-term memory problem of RNN [14]. LSTMs function just like RNNs, but they are capable of learning long-term dependencies using mechanisms called gates. These gates are different

tensor operations that can learn what information to add or remove to the hidden state. The convolution layer inside the hidden layer uses a kernel or filter same as CNN but the activation function used is different which is known as Adam. The initial learning rate is 0.001. The initial weight value is 6. The output of the hidden layer is passed to a feed-forward MLP that uses Softmax activation function to generate the prediction.

D. Convolution Neural Network (CNN):

In simpler way the process of CNN can be explained by a single review process. At first, the review text is pre-processed by performing the NLP technique. The input text is represented as a matrix. Each row of the matrix is a vector that represents a word and the vector that are known as word embedding which is converted using Word2Vec with the dimensions that user require. A convolution is used by sliding the input into the filter or kernel. At every location the matrix multiplication is performed and then the result is summed up onto a feature map. Next, to continuously reduce the dimensionality of the number of parameters and computation in the network we have used Max Pooling as pooling layer that reduces the training time and controls overfitting.

E. Traditional Machine Learning:

K Nearest Neighbours (KNN), Naive Bayes (NB) and Support Vector Machine (SVM) classifiers are used as traditional machine learning techniques for spam detection for the datasets. Based on these traditional classifiers we can say that naïve Bayes (NB) is best with the highest accuracy and score. According to the experiments of the researchers the SVM and Naïve Bayes classifiers are best for detecting fake news. These two are better than other classifiers on the basis of accuracy they provide. A classifier with more accuracy is considered as a better classifier. The major thing is the accuracy that is provided by any classifier. Classifier with more accuracy will help in detecting more fake news.

TABLE I
COMPARISION BEWTWEEN DIFFERENT ALGORITHMS FOR DETECTING FAKE REVIEWS

Reference	Algorithm	Feature	Dataset	Result	Comments
[1]	KNN	Review's key features and reviewer's behavioral features	Yelp Dataset of restaurants with and without Features.	The highest f1-score value in Tri-gram is achieved in KNN with f1-score value of 86.20%	They have not considered the behavioral feature of the reviewers into consideration that can be put into the Future Work.
[2]	Random Forest	Sentimental Analysis	Yelp Dataset	Random Forest gave 90% accuracy after comparing with other ML algorithms.	Random Forest is best in achieving effective and accurate results than compared with other primitive Machine Learning algorithms.
[3]	Supervised Algorithm	Textual Features, Review-Centric Features,	Yelp Dataset	Performed data selection, data cleaning and data transformation on the given dataset and then propose the result after	Supervised classifiers are in general more effective, and usually employ review and reviewer-centric features. Unsupervised solutions are in general

				performing ML algorithm.	less effective, but have the advantage that they do not need labelled datasets for training.
[4]	Sentiment Analysis, Latent sematic Analysis, with Netspam	User Linguistic Behavior, User Behavior, and Review Behavior	Crawled data from Amazon	The decision tree algorithm gave a precision of 65.4%, Recall of 89.7% and Accuracy of 92.06 %	There was an existing system using Naïve Bayes Algorithm and they proposed a system using Decision Tree Algorithm and showed comparison between them.
[5]	Random Forest and XGBT with Decision Tree or The Bagging/adaboost ensembles with their respective SVMs or MLPs	Identifying whether the comment is positive or negative.	Created their own dataset named "Restaurants Database"	Bagging/adaboost ensembles with their respective SVMs or MLPs gave more accuracy as compared with others i.e., of 77.31%	They trained the model on only 1 dataset and not on any other. Need to use other datasets. Also, they can use more technique to convert word to vector matrix other than Doc2Vec.
[7]	Support Vector Machine (SVM)	Extract words and assign them the sentimental value for each.		The got the best result using SVM machine learning algorithm i.e., 84.88%.	They can do more research by combining two or more machine learning techniques and working towards more accuracy.
[10]	CNN	Sentiment analysis, Numeric rating.	Google Play store Dataset	The CNN outperformed here from all the other algorithms and gave us the fake reviews from the google play store app.	Making use of different algorithms together is still an option to get more accuracy in the results.
[12]	Linear SVC, Random Forest, Naïve Bayes, SVM	Doing sentimental Analysis on the data extracted from the URL.	Extracts the data from the URL of the product and find all the reviews.	The ensemble of these 4 algorithms gave the best results in finding the fake reviews.	The admin has to remove the fake removes manually. Also, we can work towards to get the fake reviewer's id or name.
[14]	Long Short-Term Memory (LSTM)	The used n-grams, TF-IDF and word Embedding to convert words into vector matrix with similarity between the words.	Ott Dataset, yelp dataset	Used Deep learning techniques to find the fake reviews i.e., Word2Vec and LSTM gave best accuracy on both the datasets.	Variation of the CNN and RNN along with the CNN-RNN model can be tried and tested for finding the fake reviews along with the spammer reviewers' detection.

IV. CONCLUSION

In this paper, we have performed comprehensive survey which shows how Machine Learning Techniques can enhance fake review detection. The excessive use of internet has also increased the challenges and problems of fake online reviews. As there are many methods for detecting spams, the selection of best technology and algorithm depends on value of data analysed, the time taken to process the data set into useful information and also the percentage of accuracy of the result achieved. In the above survey, the comparative results state that the Naïve Bayes algorithm gives the maximum accuracy. We can conclude that supervised learning provides better result than the unsupervised learning.

V. FUTURE WORK

There is a lot of scope improvement in the research of this topic. In the future we can try to extend this topic with some more different algorithms and ideas proposed by new authors and try to extract the best solution for the online opinions and spammer identification. Future works can also include various behavioural features such as the features that depend on the frequency of reviews, the time span in between the reviews by a particular reviewer. We can use many other different data sets which would help in training the model precisely. The number of classifiers can be increased for the betterment of accuracy rate. Accessing the classifiers on the real-world application is preferable as it would help in developing algorithms that would work efficiently in the real world.

REFERENCES

[1] Ahmed M. Elmogy, Usman Tariq, Atef Ibrahim, Ammar Mohammed, "Fake Reviews Detection using Supervised Machine Learning". International Journal of Advanced Computer Science and Applications. Vol. 12, No. 1, 2021.

[2] Pilaka anusha, kaki leela prasad, "survey on fake online reviews using Machine learning algorithms", Journal of critical reviews ISSN- 2394-5125 vol 7, issue 18, 2020.

[3] Aishwarya M. Kashid, Ankita K. Lalwani, Samiksha S. Gaikwad, Rajal A. Patil, R. G. Sonkamble, S. S. More, "Fake Review Detection System Using Machine Learning", International Journal of Research in Engineering, Science and Management Volume-1, Issue-12, December-2018.

[4] Deepika Vachane, G.D. Upadhye, "Online Products Fake Reviews Detection System Using Machine Learning", Turkish Journal of Computer and Mathematics Education Vol.12 No.1S (2021).

[5] Luis Gutierrez-Espinoza, Faranak Abri, Akbar Siami Namin, Keith S. Jones, and David R. W. Sears, "Fake Reviews Detection through Ensemble Learning", IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)14 Jun 2020.

[6] Alim Al Ayub Ahmed, Ayman Aljarbouh, Praveen Kumar Donepudi, Myung Suh Choi, "Detecting Fake News using Machine Learning: A Systematic Literature Review".

[7] D. Viji, Nikhil Asawa, Tanay burreja, "Fake Reviews of Customer Detection Using Machine Learning Models", International Journal of Advanced Science and Technology. Vol. 29, No. 6, (2020)

[8] Jingdong Wang, Haitao Kan, Fanqi Meng, Qizi Mu, Genhua Shi, And Xixi Xiao, "Fake Review Detection Based on Multiple Feature Fusion and Rolling Collaborative Training", IEEE Access. October 16. Volume 8, 2020.

[9] Rami Mohawesh, Shuxiang Xu, Son N. Tran, Robert Ollington, Matthew Springer, Yaser Jararweh, And Sumbal Maqsood, "Fake Reviews Detection: A Survey", IEEE Access, May 6. Volume 9, 2021.

[10] Saima Sadiq, Muhammad Umer, Saleem Ullah, Seyedali Mirjalili, Vaibhav Rupapara and Michele NAPPI, Discrepancy "Detection between Actual User Reviews and Numeric Ratings of Google App Store using Deep Learning", Research Gate, April 2021.

[11] Pankaj Chaudhary, Abhimanyu Tyagi, Santosh Mishra, "Fake Review Detection through Supervised Classification", International Journal of Creative Research thoughts, April 2018.

[12] Ms. Rajshri P. Kashti, Dr. Prakash S. Prasad, "Enhancing NLP Techniques for Fake Review Detection", International Research Journal of Engineering and Technology (IRJET), Volume -06, Issue: 02, Feb 2019.

[13] Mansi Pakhale, Swati Pandey, Gauri Patil, "Spam Review Analysis and Detection System", International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 04, Apr 2021.

[14] G. M. Shahariar, Swapnil Biswas, Faiza Omar, Faisal Muhammad Shah, Samiha Binte Hassan, "Spam Review Detection Using Deep Learning", ResearchGate IEEE, October 2019.

