



FAKE REVIEW DETECTION ON YELP DATASET

¹ KAVYASHREE JAIN B D, ² SHILPA H L

¹Student, ²Professor
Department of MCA,
PES College of Engineering, Mandya, Karnataka, India

Abstract : Online reviews are often the primary factor in a customer's decision and are a valuable source of information that can be used to determine public opinion on the services. Reviews provide feedback to the companies about their services for any kind of improvement. The huge impact of reviews on customer's decision making motivates wrongdoers to create fake reviews to deliberately promote or demote a services. This is known as Review Spam, where spammers manipulate and poison reviews for their profit.

IndexTerms – Review Spam.

I. INTRODUCTION

With the growth of online information today, people tend to see reviews first for the places they want to visit, such as restaurants, hotels, or other businesses they need or before they go and buy some product. Yelp is a corporate directory and crowded review forum that is usually used by people to post so mere view about their opinion for business. Statistics show that by the end of 2018, there have been more than 177 million reviews on the Yelp website. It is benefiting both consumers and businesses. For a business owner, they get free advertising from people who give a useful and positive review of their business. Unfortunately, the problem arises when a small portion of irresponsible business owners try to boost up their market by hiring people to create some fake reviews about their business on Yelp website.

II. LITERATURE SURVEY

Zaitul Iradah Mahid, Selvakumar Manickam, Shankar Karuppayah in 2018 had published the paper on the "Fake News on Social Media: Brief Review on Detection Techniques" The current review of the state-of-the-art fake news detection mechanisms in social media is described in this article. Container-based approaches, social context based and hybrid based categories of fake news detection mechanisms. Conclude the paper by emphasising problems and challenges in the fake mechanism of news detection. [1].

Ata-Ur-Rehman, Nazir M. Danish , Sarfraz M. Tanzeel , Nasir Usama, Aslam Muhammad , Martinez-Enriquez A. M in 2019 had published the paper on the "Intelligent Interface for Fake Product Review Monitoring and Removal" In this paper proposed FARMS, an intelligent interface that includes a Uniform Resource Locator (URL) for the Amazon, Flipkarts and Daraz products and analyses the reviews and gives the customer the original evaluation. In this age, it is very important to detect fake product assessments. As in each ecommerce website, two different kinds of purchases are verified and non-verified. The n-gram approach is used for multi-lingual detection of fake reviews. There is evidence of the best approach for detecting fake reviews in the text categorization with the SVM classifier. [2].

Xian Wu, Yuxiao Dong, Jun Tao, Chao Huang, Nitesh V. Chawla in 2017 had published the paper on the "Reliable Fake Review Detection via Modeling Temporal and Behavioral Patterns" In this work they present a generative probabilistic model – Reliable Fake Review (RFRD) – that provides reliable online system review detections. The first is to explore user and object characteristics and their temporal patterns in the RFRD framework. Second, it enables users to learn the credibility of their latent review and the sparse distributions of objects. The RFRD model also uses its inter-dependencies to detect spam reviews, both with explicit and implicit factors. [3].

Huaxun Deng, Linfeng Zhao, Ning Luo, Yuan Liu, Guibing Guo, Xingwei Wang, Zhenhua Tan, Shuang Wang and Fucai Zhou in 2017 had published the paper on the "Semi-supervised Learning based Fake Review Detection" In this paper, they Propose a new PU-Learning-based algorithm which includes a number of labelled data and a large number of unlabeled data for the classification of review features. The features are taken from two aspects of the proposed model, including metadata characteristics and content review features. For dimensionality reduction, they use autoencoder and K-means to classify the data. [4]

Nidhi A. Patel, Prof. Rakesh Patel in 2018 had published the paper on the "A Survey on Fake Review Detection using Machine Learning Techniques" In this paper, they discussed different fake reviews based on unmonitored, monitored and semi-monitored

detection methodologies. They use different classifying algorithms such as logistic regression, k-nearest neighbour, random forests, naive bays and a support vector machine. [5].

Anna V. Sandifer, Casey Wilson, Aspen Olmsted in 2017 had published the paper on the “Detection of fake online hotel reviews” In this paper, they were data collected, part of the speech was extracted and the fake online reviews were identified by three different classification models. They found that the Multinomial Naïve Bayes classification model was used for their dataset to achieve the highest precision. [6].

Yuejun Li, Xiao Feng, Shuwu Zhang, Yuejun Li in 2016 had published the paper on the “Detecting Fake Reviews Utilizing Semantic and Emotion Model” This paper proposes three types of new features, including density, semantime and emotion and model and algorithm for each feature. Experiments indicate that in the false examination task the proposed model, algorithm and features are efficient. (Naïve Bayes algorithm, decision tree)[7].

Jitendra kumar Rout, Amiya kumar Dash, Niranjan kumar Ray in 2018 had published the paper on the “A Framework for Fake review Detection: Issues and detection” In this Proposed framework can be used to deal with both labelled and unlabeled data but for dealing with huge amount of data Big data analysis techniques need to be integrated with it. From the experiments for the framework , they have obtained an accuracy of 90.19% for supervised learning and 83.70% for unsupervised learning. (Algorithm knn, logistic regression and Random forest)[8].

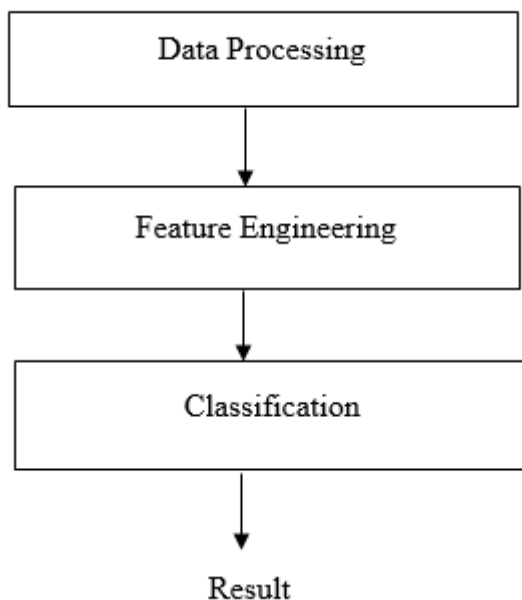
Somayeh Shojaei, Azreen Azman, Masrah Murad, Nurfadhliana Sharef and Nasir Sulaiman in 2015 had published the paper on the “A Framework for Fake Review Annotation” In this paper, They tried to investigate the problem of the review data and proposed a Spam Review Spam Review Machine Learning technique by providing the reviews and the reviewers with more helpful details and meta-data. [9].

Rodrigo Barbado, Oscar Araque, Carlos A. Iglesias in 2019 had published the paper on the “A Framework for Fake Review Detection in Online Consumer Electronics Retailers” This provides a framework for the detection of fake reviews evaluated in the domain of consumer electronics. The contributions are four times that: Construction of a dataset in four different cities to classify false reviews in the consumer electronics domain based on scraping techniques. Definition of a fake review detection feature framework. Development of the fake classification review method based on the framework proposed. Assessment and analysis of the results for each of the studied cities. We achieved an F-score of 82% in the classification task and according to the Friedman test, the Ada Boost classification is proven best by statistical means. [10].

III. PROPOSED MODEL

In order to solve the main problem that websites face thanks to spamming of thoughts, this project proposes to detect such spamming fake reviews by classifying them into fake and true. Tactically, the review is classified using the Naïve Bayes, SVM, and Logistic Recovery algorithms based on customer feedback.

SYSTEM ARCHITECTURE



3.1 System Architecture

DATA PROCESSING

According to the dataset, the ratio of filtered reviews and non-filtered reviews is approximately 1:6, which is extremely unbalanced for the classification. Therefore, we attempt to apply two methods to cope with this problem. First is over-sampling, increases the load of the minority class by replicating the minority class data. During this case, which is to feature more copies of filtered reviews, so we copy the filtered reviews three-time, therefore, the ratio decreasing to approximately 1:3. The second method is under sampling method; the fundamental idea during this method is to get rid of some on filtered reviews from the training data. After we remove the non-reviews reviews, the ratio was decreasing to approximately 1:3. The result show over sampling method gives more good result than under- sampling method. It's reasonable because over sampling method keeps all the data within the training dataset. While in under-sampling method, we lost much information.

➤ Sampling Techniques

Sampling is a popular way of dealing with the problem of distribution in class disequilibriums. Minority and majority class distribution is modified for training. Training. The same number of cases are observed in both the majority and minority classes. Two methods exist to overcome the distribution of imbalance classes: under-sampling and oversampling.

➤ Under-Sampling

In the sampling method, we choose and remove the samples from the main class and thus reduce the number of examples from the transformed data in the majority class.

➤ Over-Sampling

In the Over Sample Method, we select the minority class random example with a replacement and add multiple copies to the training data, so it is possible to select a single instance many times.

FEATURE ENGINEERING

Before doing feature engineering, we do some statistical analysis on the dataset. We found that filtered review tends to give more extreme ratings such as 1 or 5 and also mostly filtered review is shorter review than non-filtered review, even this is not too obvious, but we can use this as additional features.

Besides the basic features we have in the data set such as useful, funny cool and star rating, we tried to extract some other complex features in order to give more characterization for the machine learning classification in training process. We analyzed the business background behind the fake reviews and extracted the possible features which might indicate the signs of suspicious or malicious reviews.

CLASSIFICATION

The assessment according to the words used in the text, rating given to the review, is positive or negative. False reviews are more emotionally positive or negative than real reviews. It is because false reviews affect people's opinions, and it is more important to communicate opinions than to describe the facts clearly. The classification objective is to precisely predict the target class in the data for each case. The weight of each data in the review file is assigned and it is classified in each class - Fake and Genuine.

RESEARCH METHODOLOGY

Algorithms Used:

LOGISTIC REGRESSION

There are fundamentally three kinds of classifiers, namely generative, probabilistic, and discriminative classifier. Logistic regression is a discriminative classifier and belongs to the groups of regression methods. Logistic regression concerns describing the relationship between explanatory variables and discrete predictor, which is achieved by estimating probabilities using the underlying logistic function. Logistic regression assumes the explanatory variables are independent of each other, which can be an advantage or disadvantage depending on the dataset.

While in contrast, our next classifier, the Naive Bayes assumes dependency exists between variables. Logistic regression is ideally suited for binary classification problems meaning the predicted variable or the classes cannot be more than two. As our dataset only has two classes, so, this technique is suitable for our experiment.

Logistic regression will not perform correctly in a sketchy dataset, so, we must first identify all the essential independent variables and remove variable that unrelated and variable that are very similar to each other before we apply this classification technique.

NAIVE BAYSE

Naive Bayes classifier is a probabilistic classifier. It is based on the Bayesian theorem and operates on conditional probability. The Naive Bayes calculates a number of probabilities by means of value combinations in a given data set. Commonly used in text classification, for example, document classification and spam filtering, the Naive Bayes classifier has fast decisions making process compared to other classifiers, and the Naive Bayes often works well on even a small amount of training data. In this experiment, we used the extent of the Naive Bayes, which is Gaussian Naive Bayes; this particular type of algorithm is specifically used when the variables or the features have continued values.

SUPPORT VECTOR MACHINE

Support vector machine is originally developed to solve binary classification problems (although multiclass extensions have been proposed) and have been used successfully across a wide range of machine learning applications. A linear or not linear function is used to classify the supporting vector machine method. The supporting vector machine method is based on estimating the most suitable data separation function. Imagine we have two separate classes in our classification model, the Support vector machine method objective is to find a particular linear boundary to separate the classes.

In the process, there is a possibility that we have more than one linear boundary that can separate the classes, the support vector machine methods then will choose a line that maximizes the margin between classes is determined and therefore the highest error tolerance. The vector points that the margin line touches are known as support vectors, and that is where the name support vector machines come from.

Upon identification of training data and the margin line, test information is classified on the basis of its margin location. Recently, many classification algorithms have been proposed, but support vector machine is still one of the most widely and most popular used classifiers. The accuracy of this algorithm is 84.166%.

In this project we built a SVM Classification model to classify whether the review is fake or real

Step 1

Import Libraries.

The first step is to import Numpy, Pandas libraries.

Step 2

Then, we assign variable X and Y, and import data sets and store them to the variable dataset.

Step 3

Divide the dataset into the set and set of tests.

Divide your data into the workout and test set. In that 25% of the data stored for the test set and 75% will be used for training.

Step 4:

Feature scaling.

This feature scaling step is an added step that can increase the programme speed as the X values are decreased to a smaller range. This reduces the X train and the X test to a limited range between -2 and +2.

Step 5

Training on the training set of the SVM classification model.

After the test has been completed, we can import the SVM Classification Class and fit the training to our model. A variable classifier is attached to the SVC class. For training the model, the classifier.fit () function is used.

Step 6

The results of the test

The classifier.predict() function is used in this step to predict the value of the test set and store the value in the y pred variable.

Step 7

Matrix and accuracy of confusion.

This shows the accuracy of the trained model and describes the matrix of confusion.

IV. CONCLUSION

This project has reviewed three popular machine learning classification methods for finding fake Yelp reviews. Reviews rates such as useful, cool and funny only acquired by non- filtered review mean soon after the reviews get filtered by Yelp, the review will be hidden so it cannot be rated by others.

REFERENCES

- [1] Mahid, S. Manickam and S. Karuppayah, "Fake News on Social Media: Brief Review on Detection Techniques," *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, 2018, pp. 1-5, doi: 10.1109/ICACCAF.2018.8776689.
- [2] Ata-Ur-Rehman *et al.*, "Intelligent Interface for Fake Product Review Monitoring and Removal," *2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 2019, pp. 1-6, doi:10.1109/ICEEE.2019.8884529.
- [3] X. Wu, Y. Dong, J. Tao, C. Huang and N. V. Chawla, "Reliable fake review detection via modeling temporal and behavioral patterns," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 494-499, doi: 10.1109/BigData.2017.8257963.
- [4] H. Deng *et al.*, "Semi-Supervised Learning Based Fake Review Detection," *2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, 2017, pp. 1278-1280, doi: 10.1109/ISPA/IUCC.2017.00195.
- [5] N. A. Patel and R. Patel, "A Survey on Fake Review Detection using Machine Learning Techniques," *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, 2018, pp. 1-6, doi: 10.1109/CCAA.2018.8777594.
- [6] A. V. Sandifer, C. Wilson and A. Olmsted, "Detection of fake online hotel reviews," *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*, 2017, pp. 501-502, doi: 10.23919/ICITST.2017.8356460.
- [7] Y. Li, X. Feng and S. Zhang, "Detecting Fake Reviews Utilizing Semantic and Emotion Model," *2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*, 2016, pp. 317-320, doi: 10.1109/ICISCE.2016.77.
- [8] J. K. Rout, A. K. Dash and N. K. Ray, "A Framework for Fake Review Detection: Issues and Challenges," *2018 International Conference on Information Technology (ICIT)*, 2018, pp. 7-10, doi: 10.1109/ICIT.2018.00014.
- [10] Somayeh Shojaee, Azreen Azman, Masrah Murad, Nurfadhline Sharef and Nasir Sulaiman" A Framework for Fake Review Annotation" *2015 IEEE International Conference on Ubiquitous Computing and Communications*.