



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

# Fake Review Detection Using Data Mining

Mrs.H.A.Shinde<sup>1</sup>, Purva Sarode<sup>2</sup>, Prasanna Bhujbal<sup>2</sup>, Sanskruti Netke<sup>2</sup>, Shruti Kamble<sup>2</sup>

Lecturer, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>1</sup>

Students, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>2</sup>

**ABSTRACT :** Online reviews play a very important role for decision-making in today's e-commerce. Large parts of the population, i.e. customers read product or store reviews before deciding what to buy or where to buy and whether to buy or not. Because writing fake / fraudulent reviews comes with monetary gain, online review websites there has been a huge increase in tricky opinion spam. Basically, an untruthful review is a fake review or fraudulent review or opinion spam. Positive reviews of a target object can attract more customers and increase sales; negative reviews of a target object can result in lower demand and lower sales. Fake review detection has attracted considerable attention in recent years. Most review sites, however, still do not filter fake reviews publicly. Yelp is an exception that over the past few years has filtered reviews. Yelp's algorithm, however, is a business secret. In this work, by analyzing their filtered reviews, we try to find out what Yelp could do. The results will be useful in their filtering effort for other review hosting sites. Filtering has two main approaches: supervised and unmonitored learning. There are also about two types in terms of the characteristics used: linguistic characteristics and behavioral characteristics. Through supervised learning approach we have tried to make a model which can identify the fake review with almost 70 percent accuracy.

**KEYWORDS:** Text detection, Inpainting, Morphological operations, Connected component labelling.

## I. INTRODUCTION

As the Internet continues to grow in size and importance, the quantity and impact of online reviews is increasing continuously. Reviews can influence people across a wide range of industries, but they are particularly important in e-commerce, where comments and reviews on products and services are often the most convenient, if not the only, way for a buyer to decide whether to buy them. Online reviews can be generated for a variety of reasons. Online retailers and service providers may often ask their customers to provide feedback on their experience with the products or services they have purchased in order to improve and enhance their businesses. Customers may also feel inclined to review a product or service if they had an exceptionally good or bad experience with it. While online reviews can be helpful, blind trust of these reviews is dangerous for both the seller and buyer. Many look at online reviews before placing any online order; however, the reviews may be poisoned or faked for profit or gain, thus any decision based on online reviews must be made cautiously. Furthermore, business owners might give incentives to whoever writes good reviews about their merchandise or might pay Someone write bad reviews of the products or services of their competitor. These fake reviews are considered spam review and due to the importance of reviews can have a great impact on the online marketplace. Someone to write bad reviews about their competitor's products or services. As review spam is a pervasive and harmful issue, it is an important but challenging issue to develop methods to help businesses and consumers distinguish true reviews from fake ones.

## II. RELATED WORK

We have tried multiple algorithms and used grid search to find best parameters to get the best model with highest accuracy.

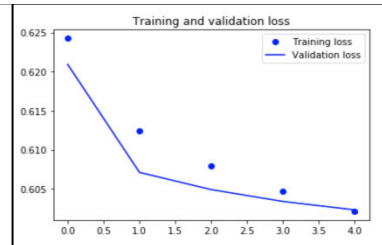
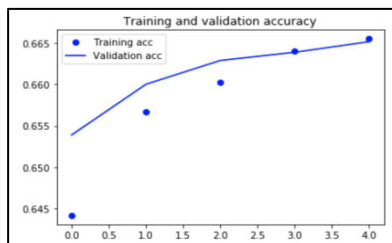
- 1) **Logistic Regression** - We have used logistic regression on the all the features combined selected through feature selection. It didn't give good accuracy so we moved on with other models for better results
- 2) **Multinomial Naive Bayes** : This is a generative model, in which label is generated using a Bernoulli distribution. For multinomial naive Bayes we initially trained the model on 'Review Content' and it gave almost 67 percent accuracy. To improve the accuracy we used GridSearchCV with following parameters to select *parameters = {'vect\_ngram\_range': [(1, 1), (1, 2)], 'tfidf\_use\_idf': (True, False), 'clf\_alpha': (1e-2, 1e-3)}*. We got the best parameter values as *gram\_range': (1, 2), Tfidf\_use\_idf: True, clf\_alpha : .0001*. After using the above parameters and cross validation accuracy improved to almost 74 percent. We used 3-fold cross-validation in GridSearchCV. We also changed parameters *njobs = -1* , so that it can run all jobs in parallel for both fit and predict by using all the available cores to make the execution faster.
- 3) **Neural Network with LSTM**: This model had an embedding layer, an LSTM layer, and the LSTM output was fed

into a fully connected layer that was hidden. The output size of this layer is 1, which means it will always output 1 or 0, 1 for Spam reviews and 0 for real reviews. 10,000 most common words were passed as features, 64 as the second, and input\_length of 200 as the length of each sequences to the embedding layer. In this first argument came from the second argument from the Embedding layer. We kept 20% chance of dropout. Training model with a batch size of 128, 5 epochs and validation split=0.2, which means that the neural network would learn from 80% of the data and test itself on the remaining 20% of the data.

```

Train on 440530 samples, validate on 110133 samples
Epoch 1/5
440530/440530 [=====] - 681s 2ms/step - loss: 0.6242 - acc: 0.6442 - val_loss: 0.6209 - val_acc: 0.6539
Epoch 2/5
440530/440530 [=====] - 669s 2ms/step - loss: 0.6124 - acc: 0.6566 - val_loss: 0.6071 - val_acc: 0.6600
Epoch 3/5
440530/440530 [=====] - 665s 2ms/step - loss: 0.6080 - acc: 0.6603 - val_loss: 0.6049 - val_acc: 0.6629
Epoch 4/5
440530/440530 [=====] - 662s 2ms/step - loss: 0.6047 - acc: 0.6640 - val_loss: 0.6034 - val_acc: 0.6639
Epoch 5/5
440530/440530 [=====] - 660s 1ms/step - loss: 0.6021 - acc: 0.6655 - val_loss: 0.6023 - val_acc: 0.6651
CPU times: user 1h 16min 23s, sys: 5min 51s, total: 1h 22min 15s
Wall time: 55min 39s

```



Figures 1 and 2. Implementation of Neural Network

the first three epochs, the training accuracy was less than validation accuracy which means model was underfitting on the dataset, but in the last 2 epochs, the model generalize well and gave the same training and validation accuracy. This model gave the 66.26 % accuracy on the test set.

### III. METHODOLOGY

The most challenging part of this project is to extract predictive features from reviews and the corresponding reviewer information. For this particular task we basically worked on the following review centric features -

- 1) **Structural features:** review length, average word length, Standard deviation of the review length.
- 2) **Semantic features:** We used pretrained model *Textblob* to find out the polarity and sentiment for the review, and also found out the positive and negative sentiment bearing words in the review.
- 3) **n-gram features:** We have extracted unigram, Bigram, Trigram features from review content.
  - We have majority normal reviews, so we use oversampling to handle biased dataset.
  - Use 80-20 train, test split
  - Tokenize using Spacy and Keras to split review into words
  - Remove punctuations, convert to lowercase.
  - Remove stop words using Spacy word-labels ("POS", "stop word", etc.)
  - Vectorizer to convert word to index, TF-IDF to scale using inverse document frequency.
  - Plot word frequency graph and word cloud.

#### Feature Engineering for review spam detection

Using the reviews, we add review length feature. Additionally we used a pretrained model from TextBlob to detect sentiment polarity [-1,+1] and sentiment subjectivity [0,1] for each review. Along with this we use other features like rating, usefulCount to train models on features other than review itself.

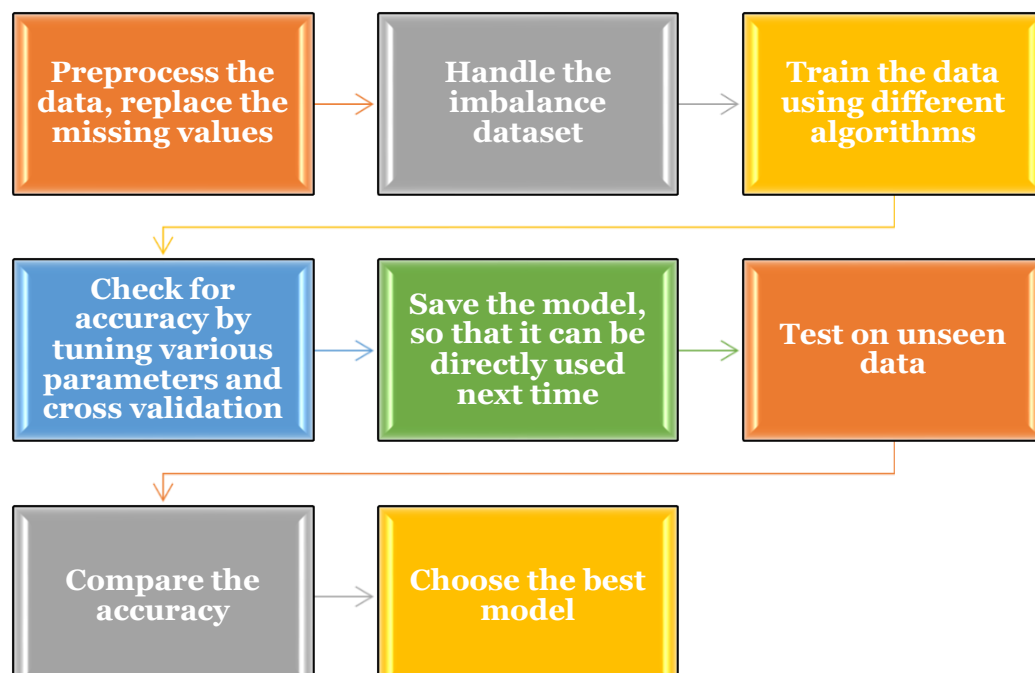


Fig 3. Data Flowchart

#### IV. EXPERIMENTAL RESULTS

##### Data mining based system to check review

Text:

##### Data mining based system to check review

Text:

Given review is Fake

probability score: 0.5897656314026887





### Data mining based system to check review

I especially like the frames they carry -  
- they are occasion appropriate and look  
great. If you're looking for a small,  
memorable gift, this is pretty much your  
one stop shop -- you can get the gift,  
greeting card, and a gift bag all in one  
shop!  
I live in Clifton -- the capitol of  
Cincinnati's Indian restaurant scene.

Text:

Submit

Given review is Not fake

probability score: 0.817836962297468

## V. CONCLUSION

Contrary to our expectations, adding reviewer-centered features does not significantly boost model performance. We believe this is because most reviewers have only one review record, making most of our reviewer-centered features poorly defined (e.g., maximum number of reviews in a day, standard deviation of ratings). Sometimes using the simple approach works best and that happened with this project, using Grid Search cross validation with Multinomial Naive Bayes Classifier and using the best parameter, gave us the accuracy of almost 77 percent which was the highest amongst all the other classifiers.

## REFERENCES

1. Arjun Mukherjee, Vivek Venkataraman, Bing Liu, and Natalie Glance. What Yelp Fake Review Filter Might Be Doing. Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013), July 8-10, 2013, Boston, USA.
2. M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, H. A. Najada, "Survey of review spam detection using machine learning techniques", Journal of Big Data, vol. 2, no. 23, Dec 2015.
3. [https://scikit-learn.org/stable/supervised\\_learning.html#supervised-learning](https://scikit-learn.org/stable/supervised_learning.html#supervised-learning)
4. <https://github.com/zzhang83/Neural-Network-3-projects/blob/master/Yelp%20Review%20Classification.ipynb>
5. <https://github.com/lukas/ml-class/blob/master/keras-cnn/cnn.py>
6. <https://towardsdatascience.com/sentiment-analysis-with-text-mining-13dd2b33de27>



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



**www.ijirset.com**

Scan to save the contact details