International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

Fake Reviews Detection Using Supervised Machine Learning

C. Rekha[1], Mr. G. Lakshmikanth[2]
M. Tech Scholar[1], Associate Professor[2]
[1,2]Department of Computer Science, Sri Ramachandra Engineering and Technology, Chennai, Tamil Nadu, India

Article Info
Volume 9, Issue 5
Page Number : 124-133

Publication Issue
September-October-2022

Article History
Accepted : 08 Sep 2022
Published : 20 Sep 2022

ABSTRACT

With the ongoing evolution of E-commerce platforms, online evaluations are increasingly seen as a critical aspect in establishing and maintaining a positive reputation. Furthermore, they play an important role in end-user decision making. A positive evaluation for a target object typically draws more customers and results in a significant rise in sales. Deceptive or phoney evaluations are now intentionally generated in order to build a virtual reputation and attract potential clients. Identifying bogus reviews is thus an active and ongoing research topic. Detecting phoney reviews is dependent not only on the primary elements of the reviews, but also on the reviewers' behaviour. This research suggests using machine learning to detect bogus reviews. In addition to the review features extraction approach, this research utilises different features engineering techniques to extract distinct reviewer behaviours. The study examines the performance of machine learning classifiers KNN, Naive Bayes (NB), and Logistic Regression using a genuine Yelp dataset of restaurant reviews. In terms of accuracy, the results show that Logistic Regression surpasses all other classifiers. The results reveal that the algorithm is more capable of distinguishing between genuine and false reviews.

Keywords: Machine learning, fake, reviews, Logistic Regression

I. INTRODUCTION

Reviews have become the primary source of information for clients looking to make a judgement regarding services or items. For example, when customers decide to book a hotel, they read reviews about previous customers' experiences with the hotel's services. They select whether or not to reserve a room based on the feedback from the reviews. If they get great feedback from the reviews, they will most likely book the room. As a result, past reviews have become incredibly reputable sources of information for the majority of individuals in various online services. Since reviews are regarded real forms of exchanging input about positive or negative services, any attempt to manipulate such evaluations by providing misleading or inauthentic content is deemed dishonest behaviour, and such reviews are labelled as phoney. This example makes us wonder what if not all of the published reviews are truthful or genuine. What if any of these testimonials are false? As a result, detecting fraudulent reviews has become, and continues to be, an important and necessary study topic.

The advent of social media has blurred the barrier between genuine content and advertising, resulting in an increase in false endorsements across the industry. Fake online reviews and other false endorsements are frequently used to promote things on the internet. Consequently, the FTC is now using its Penalty Offense Authority to remind advertisers of the law and deter them from breaking it. By mailing a Notice of Penalty Offenses to over 700 businesses, the FTC is informing them that they might face hefty civil penalties—up to $43,792 per violation—if they use endorsements in ways that contradict previous FTC administrative actions.

“Fake reviews and other forms of deceptive endorsements cheat consumers and undercut honest businesses,” said Samuel Levine, Director of the FTC’s Bureau of Consumer Protection. “Advertisers will pay a price if they engage in these deceptive practices.”

The Notice of Penalty Offenses allows the agency to seek civil penalties against a company that engages in activity that it is aware has been declared illegal in a previous FTC administrative ruling, other than a consent order.

The Notice filed to the corporations lists a variety of behaviours that the FTC has previously judged to be unfair or deceptive in administrative cases. These include, but are not limited to, falsely claiming a third-party endorsement; misrepresenting whether an endorser is an actual, current, or recent user; using an endorsement to make deceptive performance claims; failing to disclose an unexpected material connection with an endorser; and falsely claiming that endorsers' experience represents consumers' typical or ordinary experience.

Companies receiving the notice represent an array of significant corporations, leading advertisers, leading merchants, leading consumer goods manufacturers, and important advertising agencies. A full list of the businesses receiving the Notice from the FTC is available on the FTC’s website. A recipient’s presence on this list does not in any way suggest that it has engaged in deceptive or unfair conduct.

In addition to the Notice, the FTC has created multiple resources for business to ensure that they are following the law when using endorsements to advertise their products and services, which can be found on the FTC’s website.

To that end, this work employs a number of machine learning classifiers to detect phoney reviews based on the text of the reviews as well as various derived reviewer attributes. We test the classifiers on a genuine corpus of reviews collected from open source websites. In addition to using natural language processing on the corpus to extract and feed review features to classifiers, the research uses multiple features engineering techniques on the corpus to extract diverse reviewer behaviours. The research compares the influence of extracted reviewer features when they are taken into account by classifiers. The research examines the results of the extracted features in the absence and presence of the extracted features in two distinct language models, TF-IDF. The results show that the created characteristics improve the detection of false reviews.

The Internet's rapid expansion has altered many of our daily routines. Ecommerce is one of the fastest growing sectors. Generally, e-commerce sites allow customers to write feedback about their services. These reviews might be utilised as a source of information. Companies, for example, can use it to make design decisions for their products or services, whereas potential customers can use it to determine whether or not to buy or use a product. Unfortunately, the relevance of the review is misunderstood by certain parties that strive to fabricate fake reviews, either to boost the popularity or to discredit the product. The goal of this project is to detect phoney product reviews by using a review's text and rating properties.

The rapid growth of the Internet influenced many of our Machine learning approaches can make a significant contribution to detecting bogus web content reviews. In general, web mining approaches use a variety of machine learning algorithms to identify and extract important information. Content mining is one of the web mining duties. Opinion mining is a traditional example of content mining in which a classifier is trained to analyse the features of the reviews as well as the sentiments to determine the sentiment of the text (positive or negative). Detecting phoney reviews is typically dependent not only on the type of reviews but also on specific features that are unrelated to the content. Text and natural language processing NLP are commonly used in the development of review features. Fake reviews, on the other hand, may necessitate the development of additional information related to the reviewer themselves, such as review time/date or writing styles. Thus, the successful detection of phoney reviews is dependent on the building of relevant reviewer features extraction.

Detecting phoney reviews is typically dependent not only on the type of reviews but also on specific features that are unrelated to the content. Text and natural language processing NLP are commonly used in the development of review features. Fake reviews, on the other hand, may necessitate the development of additional information related to the reviewer themselves, such as review time/date or writing styles. Thus, the successful identification of bogus reviews is dependent on the development of relevant features extraction.

II. RELATED WORKS

R. Barbado, O. Araque, and C. A. Iglesias: The importance of online evaluations on businesses has expanded considerably over the last year [1], and they are now critical to determining business performance in a wide range of industries, from restaurants to hotels to e-commerce. Unfortunately, some people utilise unscrupulous methods to promote their internet reputation, such as generating fictitious reviews of their businesses or competitors. Previous research has focused on the detection of fraudulent reviews in a variety of sectors, including product or business reviews in restaurants and hotels. Despite its economic importance, the consumer electronics business domain has yet to be adequately researched. This paper evaluates a feature framework for detecting false reviews in the consumer electronics market. Four contributions are made: I Using scraping techniques, construct a dataset for classifying fake reviews in the consumer electronics domain in four different cities; (ii) define a feature framework for fake review detection; (iii) develop a fake review classification method based on the proposed framework; and (iv) evaluate and analyse the results for each of the cities under consideration. We achieved an 82 percent F-Score on the classification challenge, and the Ada Boost classifier was statistically demonstrated to be the best by the Friedman test.

Tadelis: Online markets have become widespread, with billions of people frequenting sites like eBay, Taobao, Uber, and Airbnb [2]. The success of these marketplaces is linked not just to the simplicity with which buyers and sellers may discover each other, but also to the trust that these marketplaces assist to create through reputation and feedback systems. I begin by quickly discussing the fundamental concepts underlying the importance of reputation in fostering trust and trade, and then provide an overview of how feedback and reputation systems function in online marketplaces. The literature on the implications of reputation and feedback systems on online marketplaces is then described, and some of the difficulties with bias in feedback and reputation systems as they currently exist are highlighted. I explore how to solve these issues in order to enhance the practical design of online marketplaces, as well as possible future research prospects.

M. J. H. Mughal Web data mining has evolved into a simple and crucial platform for obtaining meaningful information. Users choose to upload and download data from the World Wide Web [3]. Discovering informative knowledge and patterns is becoming more difficult and time consuming as the amount of material on the internet grows. It is difficult to extract knowledgeable and user-requested information from unstructured and inconsistent data on the internet. To retrieve meaningful information from the web, there are numerous mining techniques in use (hyperlinks, contents, web usage logs). Web data mining is a data mining specialty focused on the internet. There are three types of web data Web structure mining, web content mining, and web usage mining are all types of mining. All of these strategies, tools, methodologies, and algorithms are used to discover knowledge from massive amounts of data on the internet.

C. C. Aggarwal Users can now express their opinions regarding entities [4, individuals, events, and subjects] in a variety of formal and informal settings, thanks to the recent proliferation of social media. Reviews, forums, social media posts, blogs, and discussion boards are some examples of such environments. The computational analytics connected with such text are defined as the problem of opinion mining and sentiment analysis.

A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance Online reviews have evolved into an important decision-making resource [5]. However, its use brings with it a curse: misleading opinion spam. In recent years, there has been a lot of interest in bogus review identification. Most review sites, however, still do not publicly filter bogus reviews. Yelp is an exception, having begun moderating reviews in recent years. The algorithm used by Yelp, on the other hand, is a trade secret. We try to figure out what Yelp is up to by studying its filtered reviews in this paper. Other review hosting sites will benefit from the results in their filtering efforts. Filtering can be divided into two types: supervised and unsupervised learning. There are two sorts of features used: linguistic features and behavioural features. We shall utilise a supervised strategy in this work because we can train using Yelp's filtered reviews. Existing supervised learning techniques all rely on phoney reviews rather than reviews regulated by a commercial website. Recently, it was demonstrated that supervised learning employing linguistic n-gram features performs remarkably well (with greater than 90% accuracy) in spotting crowdsourced fraudulent Amazon Mechanical Turk reviews (AMT). We put these current research methodologies to the test and evaluated their efficacy using real-world Yelp data. Surprisingly, the behavioural aspects outperform the language features. A unique information theoretic methodology is proposed to determine the precise psycholinguistic difference between AMT assessments and Yelp reviews in order to examine (crowdsourced vs. commercial fake reviews). We discover something intriguing. This research and testing results lead us to believe that Yelp's filtering is acceptable, and that its filtering method appears to be associated with anomalous spamming activities.

III. METHODS AND MATERIAL

Proposed system:

In this paper, we propose this application that can be considered a useful system since it helps to reduce the limitations obtained from traditional and other existing methods. The objective of this study to develop fast and reliable method which detects and estimates anaemia accurately. To design this system is we used a powerful algorithm in a based Python environment with Django frame work.

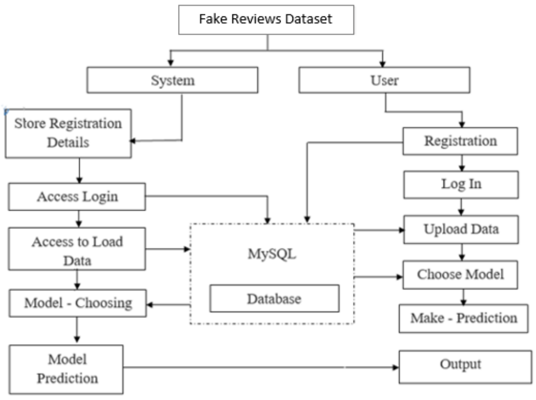


Figure 1 : Fake reviews dataset block diagram

IV. Implementation

1. Naive Bayes:

A Naive Bayes classifier is a probabilistic machine learning model that's used for classification task. The crux of the classifier is based on the Bayes theorem.

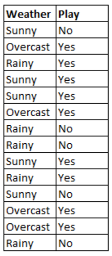
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Using Bayes theorem, we can find the probability of **A** happening, given that **B** has occurred. Here, **B** is the evidence and **A** is the hypothesis. The assumption made here is that the predictors/features are independent. That is presence of one particular feature does not affect the other. Hence it is called naive.

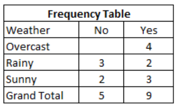
Let's look at an example to better comprehend it. I've attached a weather training data set and the matching goal variable 'Play' (suggesting possibilities of playing). Now we must decide whether or not to play based on the weather. Follow the steps below to complete it.

Step 1: Make a frequency table out of the data collection.

Step 2: Create a Likelihood table by calculating probabilities such as Overcast probability = 0.29 and Playing probability = 0.64.

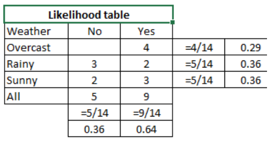


Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No



Frequency Table

Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9



Likelihood table

Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

Step 3: Apply the Naive Bayesian equation to each class to determine the posterior probability. The outcome of prediction is the class with the highest posterior probability.

Problem: If the weather is sunny, the players will play. Is this correct?

We can solve it using the posterior probability method outlined above.

P(Yes | Sunny) = P(Yes | Sunny) * P(Yes) / P(Yes) (Sunny)

P (Sunny |Yes) = 3/9 = 0.33, P (Sunny |Yes) = 5/14 = 0.36, and P (Yes) = 9/14 = 0.64.

P (Yes | Sunny) = 0.33 * 0.64 / 0.36 = 0.60, indicating a higher likelihood.

A similar strategy is used by Naive Bayes to forecast the probability of distinct classes based on various attributes.

This approach is commonly used in text classification and multi-class situations.

• Predicting the class of test data set is simple and quick. It also excels at multi-class prediction.

• When the assumption of independence is met, a Naive Bayes classifier outperforms other models such as logistic regression and requires less training data. • It performs well with categorical input variables as opposed to numerical inputs (s). The normal distribution is assumed for numerical variables (bell curve, which is a strong assumption).

Applications of Naive Bayes Algorithms:

Real-time Prediction: Naive Bayes is a quick-learning classifier. It may thus be used to make real-time forecasts.

This algorithm is very widely recognised for its multi class prediction feature. In this case, we can forecast the likelihood of many target variable classes.

Text classification/ Spam Filtering/ Sentiment Analysis: Naive Bayes classifiers have a greater success rate than other algorithms in text classification (owing to better results in multi class issues and independence rule). As a result, it is commonly utilised in Spam filtering (determining spam e-mail) and Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments)

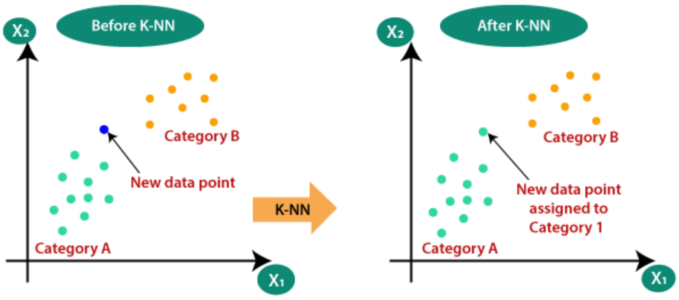
Recommendation System: Naive Bayes Classifier and Collaborative Filtering together builds a Recommendation System that uses Machine learning and data mining techniques are used to filter unknown information and forecast whether or not a user will want a certain resource.

KNN:

- K-Nearest Neighbour is a simple Machine Learning algorithm that relies on the Supervised Learning technique.
- The K-NN method assumes similarity between the new case/data and existing cases and assigns the new case to the category that is most similar to the existing categories.
- The K-NN method maintains all available data and uses similarity to classify a new data point. This means that when fresh data arrives, it may be quickly sorted into a well-suited category using the K- NN algorithm.
- The K-NN algorithm can be used for both regression and classification, but it is primarily employed for classification.
- K-NN is a non-parametric algorithm, which implies that it does not make any assumptions about the underlying data.
- It is also known as a lazy learner algorithm since it does not instantly learn from the training set; instead, it saves the information and performs a classification operation on it.
- **Example**: Assume we have an image of a critter that resembles a cat or a dog and we want to know whether it is a cat or a dog. So, because it works on a similarity measure, we may utilise the KNN method for this identification. Our KNN model will look for similarities between the new data set and the photographs of cats and dogs and place it in either category depending on the most similar attributes.

What is the purpose of the K-NN Algorithm?

Assume there are two categories, A and B, and we get a new data point x1, and we want to know which of these categories this data point belongs to. We require a K-NN algorithm to solve this type of problem. We can quickly determine the category or class of a dataset using K-NN. Consider the illustration below:

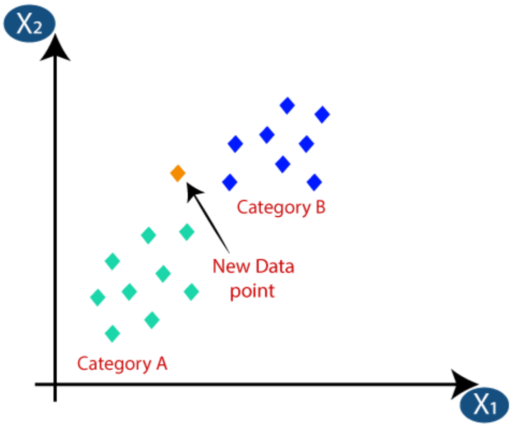


What is the mechanism of K-NN?

The following algorithm explains how K-NN works:

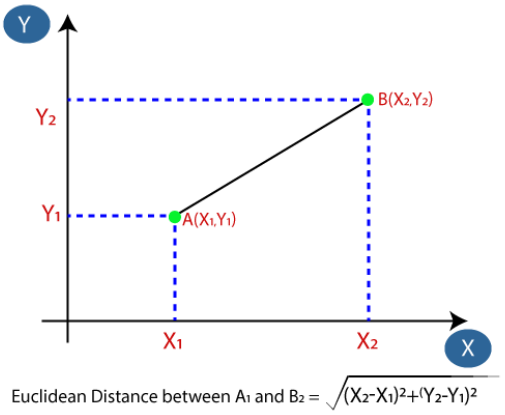
Step 1: Determine the number K of neighbours.
Step 2: Determine the Euclidean distance between the K neighbours.
Step 3: Determine the K nearest neighbours based on the Euclidean distance.
Step 4: Count the number of data points in each category among these k neighbours.
Step 5: Assign the new data points to the category with the highest number of neighbours.
Step 6: Our model is completed.

Assume we have a new data point that has to be assigned to the appropriate category. Consider the following illustration:

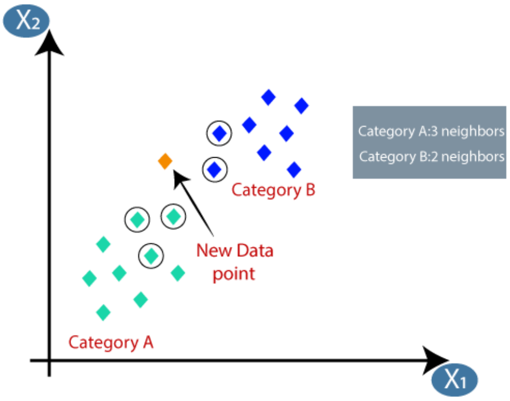


To begin, we will select the number of neighbours, so k=5.

The Euclidean distance between the data points will then be computed. The Euclidean distance is the distance between two locations studied in geometry. It can be computed as follows:



We obtained the nearest neighbours by computing the Euclidean distance, which yielded three nearest neighbours in category A and two nearest neighbours in category B. Consider the following illustration:



As we can see, the three closest neighbours are all from category A, thus this new data point must be from that as well.

In the K-NN Algorithm, how should the value of K be chosen?

Here are some considerations to make while determining the value of K in the K-NN algorithm:

- Because there is no specific way to discover the ideal value for "K," we must experiment with different values to find the best one. The most popular K value is 5.
- A very low K value, such as K=1 or K=2, might be noisy and cause outlier effects in the model.
- High K values are preferable, although they may cause problems.

KNN Algorithm Advantages:

- It is easy to deploy,
- It is resistant to noisy training data.
- If the training data is large, it may be more effective.

It is always necessary to determine the value of K, which can be difficult at times.

• Because the distance between the data points for all training samples is determined, the computation cost is significant.

Logistic Regression:

In the early twentieth century, the biological sciences adopted logistic regression. It was afterwards employed in a variety of social scientific applications. When the dependent variable (target) is categorical, logistic regression is utilised.

For example,

To predict whether an email is spam (1) or (0)

Whether the tumor is malignant (1) or not (0)

Consider a scenario where we need to classify whether an email is spam or not. If we use linear regression for this problem, there is a need for setting up a threshold based on which classification can be done. Say if the actual class is malignant, predicted continuous value 0.4 and the threshold value is 0.5, the data point will be classified as not malignant which can lead to serious consequence in real time.

From this example, it can be inferred that linear regression is not suitable for classification problem. Linear regression is unbounded, and this brings logistic regression into picture. Their value strictly ranges from 0 to 1.

Purpose and examples of logistic regression:

Logistic regression is one of the most commonly used machine learning algorithms for binary classification problems, which are problems with two class values, including predictions such as “this or that,” “yes or no” and “A or B.”

The purpose of logistic regression is to estimate the probabilities of events, including determining a relationship between features and the probabilities of particular outcomes.

One example of this is predicting if a student will pass or fail an exam when the number of hours spent studying is provided as a feature and the variables for the response has two values: pass and fail.

Organizations can use insights from logistic regression outputs to enhance their business strategies so they can achieve their business goals, including reducing expenses or losses and increasing ROI in marketing campaigns, for example.

An e-commerce company that mails expensive promotional offers to customers would like to know whether a particular customer is likely to respond to the offers or not. For example, they’ll want to know whether that consumer will be a “responder” or a “non responder.” In marketing, this is called propensity to respond modeling.

Likewise, a credit card company develops a model to decide whether to issue a credit card to a customer or not will try to predict whether the customer is going to default or not on the credit card based on such characteristics as annual income, monthly credit card payments and number of defaults. In banking parlance, this is known as default propensity modeling.

Uses of logistic regression:

Logistic regression has become particularly popular in online advertising, enabling marketers to predict the likelihood of specific website users who will click on particular advertisements as a yes or no percentage.

- Logistic regression can also be used in:
- Healthcare to identify risk factors for diseases and plan preventive measures.
- Weather forecasting apps to predict snowfall and weather conditions.
- Voting apps to determine if voters will vote for a particular candidate.

Insurance to predict the chances that a policy holder will die before the term of the policy expires based on certain criteria, such as gender, age and physical examination.

Banking to predict the chances that a loan applicant will default on a loan or not, based on annual income, past defaults and past debts.

Logistic regression vs. linear regression:

The main difference between logistic regression and linear regression is that logistic regression provides a constant output, while linear regression provides a continuous output.

The outcome, such as a dependent variable, has a finite number of possible values in logistic regression. However, the output of linear regression is continuous, which means it can take any of an unlimited number of values.

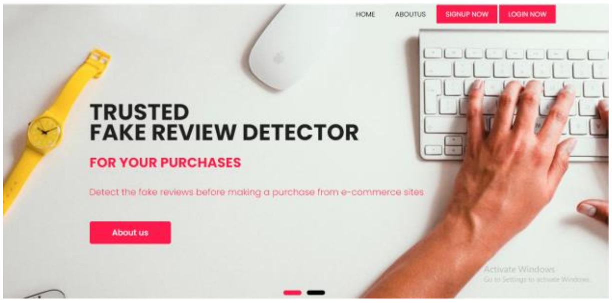
When the answer variable is categorical, such as yes/no, true/false, or pass/fail, logistic regression is utilised. When the response variable is continuous, such as time, height, or weight, linear regression is used.

For example, given data on the time a student spent studying and that student's exam scores, logistic regression and linear regression can predict different things.

With logistic regression predictions, only specific values or categories are allowed. Therefore, logistic regression can predict whether the student passed or failed. Since linear regression predictions are continuous, such as numbers in a range, it can predict the student's test score on a scale of 0 -100.

V. RESULTS AND DISCUSSION

The following images will visually depict the process of our project.

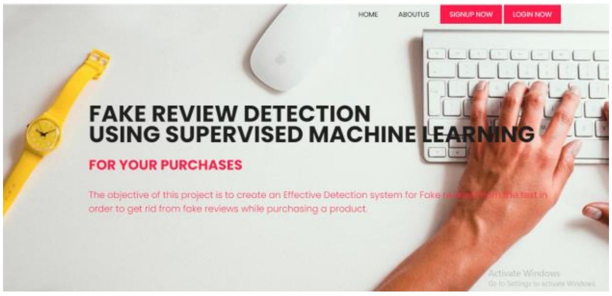


TRUSTED
FAKE REVIEW DETECTOR

FOR YOUR PURCHASES

Detect the fake reviews before making a purchase from e-commerce sites

Home page: In this home page we can see the logo designing of our website and here we are detecting the fake reviews from the review entered by the user.

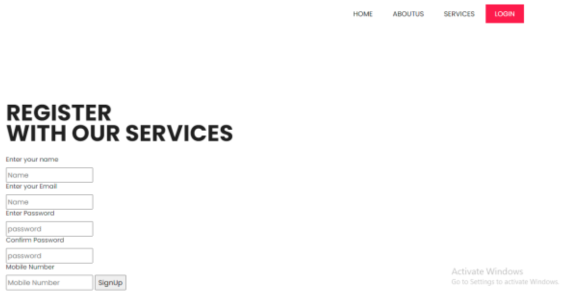


FAKE REVIEW DETECTION
USING SUPERVISED MACHINE LEARNING

FOR YOUR PURCHASES

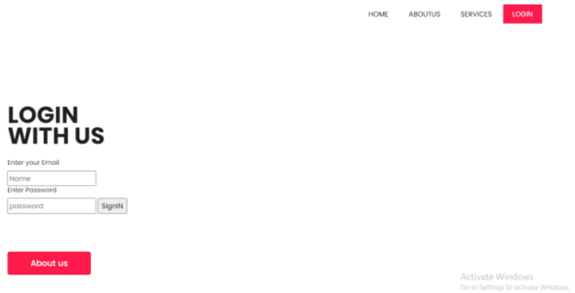
The objective of this project is to create an Effective Detection system for Fake [illegible] the text in order to get rid from fake reviews while purchasing a product.

About page: This is about page, here the application describes what main objective of this project is.



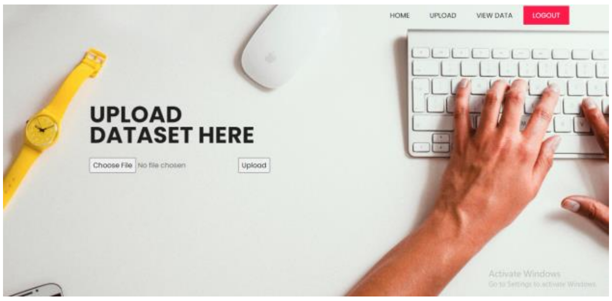
REGISTER
WITH OUR SERVICES

Registration page: Registration page in which user need to register to start.



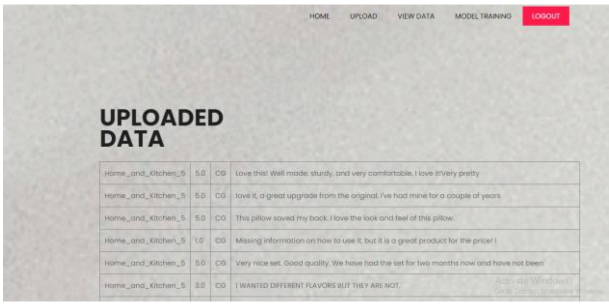
LOGIN
WITH US

Login page: In this login page, user need to enter valid credentials in order to enter.



UPLOAD
DATASET HERE

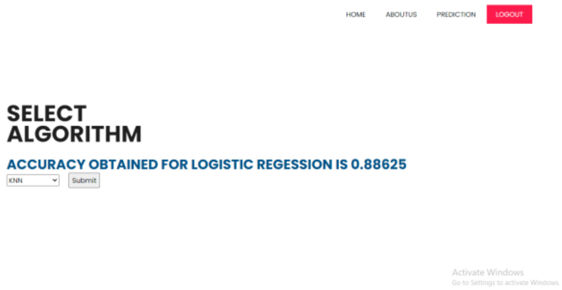
Upload page: In this upload Page in order to upload the dataset.



UPLOADED
DATA

Home_and_Kitchen_5	5.0	CG	Love this! Well made, sturdy, and very comfortable. I love it!Very pretty
Home_and_Kitchen_5	5.0	CG	love it, a great upgrade from the original. I've had mine for a couple of years
Home_and_Kitchen_5	5.0	CG	This pillow saved my back. I love the look and feel of this pillow.
Home_and_Kitchen_5	1.0	CG	Missing information on how to use it, but it is a great product for the price! I
Home_and_Kitchen_5	5.0	CG	Very nice set. Good quality. We have had the set for two months now and have not been
Home_and_Kitchen_5	3.0	CG	I WANTED DIFFERENT FLAVORS BUT THEY ARE NOT.

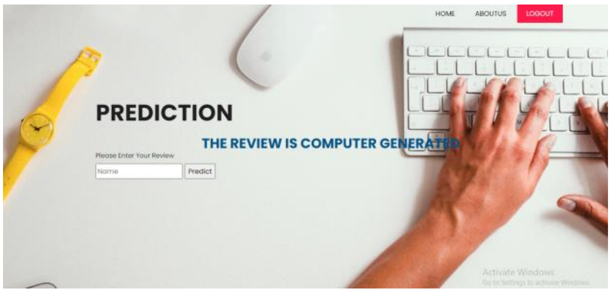
View data page: In this user views the data which he was uploaded to the system.



SELECT ALGORITHM

ACCURACY OBTAINED FOR LOGISTIC REGESSION IS 0.88625

Model training page: In this model training page, training of your model takes place and display the model's accuracy



PREDICTION

THE REVIEW IS COMPUTER GENERATED

Please Enter Your Review

Prediction page: In this prediction page, user need to enter the required fields in order to get the response from the data whether the review is computer generated or original.

VI. CONCLUSION

We have successfully developed a system to detect fake reviews in this application. This is created in a user-friendly environment with Python programming and Django framework. The system is likely to gather data from the user in order to determine whether the review is fake or not.

VII. REFERENCES

[1]. R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," Information Processing & Management, vol. 56, no. 4, pp. 1234 – 1244, 2019.

[2]. S. Tadelis, "The economics of reputation and feedback systems in e-commerce marketplaces," IEEE Internet Computing, vol. 20, no. 1, pp. 12–19, 2016.

[3]. M. J. H. Mughal, "Data mining: Web data mining techniques, tools and algorithms: An overview," Information Retrieval, vol. 9, no. 6, 2018.

[4]. C. C. Aggarwal, "Opinion mining and sentiment analysis," in Machine Learning for Text. Springer, 2018, pp. 413–434.

[5]. A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What yelp fake review filter might be doing?" in Seventh international AAAI conference on weblogs and social media, 2013.

[6]. N. Jindal and B. Liu, "Review spam detection," in WWW '07, Proceedings of the 16th International Conference on the World Wide Web.

[7]. Elmurngi, E., and Gherbi, A., Detecting Fake Reviews Using Sentiment Analysis and Machine Learning Techniques. DATA ANALYTICS/IARIA, 2017.

[8]. V. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews and blog posts," in Advance Computing Conference (IACC), 2013, pp. 893–898.

[9]. A. Molla, Y. Biadgie, and K.-A. Sohn, "Detecting Negative Deceptive Opinion from Tweets." in International Conference on Mobile and Wireless Technology. Singapore: Springer, 2017.

[10]. S. Shojaee et al., "Detecting deceptive reviews using lexical and syntactic features." 2013.

[11]. Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study," Information Sciences, vol. 385, pp. 213– 224, 2017.

[12]. H. Li et al., "Spotting fake reviews via collective positive-unlabeled learning." 2014.

Cite this article as :

C. Rekha, Mr. G. Lakshmikanth, "Fake Reviews Detection Using Supervised Machine Learning", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 5, pp. 124-133, September-October 2022.
Journal URL : https://ijsrst.com/IJSRST229534