iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43202

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Fake Reviews Detection Using Supervised Machine Learning

Nikhil Chandra Sai Ram[1], Gowtham Vakati[2], Jagadesh Varma Nadimpalli[3], Yash Sah[4], Sai Karthik Datla[5]

Abstract: ***With the continuous evolve of E-commerce systems, online reviews are mainly considered as a crucial factor for building and maintaining a good reputation. Moreover, they have an effective role in the decision making process for end users. Usually, a positive review for a target object attracts more customers and lead to high increase in sales. Nowadays, deceptive or fake reviews are deliberately written to build virtual reputation and attracting potential customers. Thus, identifying fake reviews is a vivid and ongoing research area. Identifying fake reviews depends not only on the key features of the reviews but also on the behaviours of the reviewers. This paper proposes a machine learning approach to identify fake reviews. In addition to the features extraction process of the reviews, this paper applies several features engineering to extract various behaviours of the reviewers. The paper compares the performance of several experiments done on a real Yelp dataset of restaurants reviews, we compare the performance of machine learning classifiers; KNN, Naive Bayes (NB), Logistic Regression. The results reveal that Logistic Regression outperforms the rest of classifiers in terms of accuracy achieving best. The results show that the system has better ability to detect a review as fake or original.***

Keywords: ***Machine learning, fake, reviews, Logistic Regression.***

I. INTRODUCTION

Nowadays, when customers want to draw a decision about services or products, reviews become the main source of their information. For example, when customers take the initiation to book a hotel, they read the reviews on the opinions of other customers on the hotel services. Depending on the feedback of the reviews, they decide to book room or not. If they came to a positive feedback from the reviews, they probably proceed to book the room. Thus, historical reviews became very credible sources of information to most people in several online services. Since, reviews are considered forms of sharing authentic feedback about positive or negative services, any attempt to manipulate those reviews by writing misleading or inauthentic content is considered as deceptive action and such reviews are labelled as fake Such case leads us to think what if not all the written reviews are honest or credible. What if some of these reviews are fake. Thus, detecting fake review has become and still in the state of active and required research area. The rise of social media has blurred the line between authentic content and advertising, leading to an explosion in deceptive endorsements across the marketplace. Fake online reviews and other deceptive endorsements often tout products throughout the online world. Consequently, the FTC is now using its Penalty Offense Authority to remind advertisers of the law and deter them from breaking it. By sending a Notice of Penalty Offenses to more than 700 companies, the agency is placing them on notice they could incur significant civil penalties—up to $43,792 per violation—if they use endorsements in ways that run counter to prior FTC administrative cases.

"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 conduct that it knows has been found unlawful in a previous FTC administrative order, other than a consent order.

The Notice sent to the companies outlines a number of practices that the FTC determined to be unfair or deceptive in prior administrative cases. These include, but are not limited to: falsely claiming an endorsement by a third party; 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 misrepresenting that the experience of endorsers represents consumers' typical or ordinary experience.

Companies receiving the notice represent an array of large companies, top advertisers, leading retailers, top consumer product companies, and major 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.

©IJRASET: All Rights are Reserved

To this end, this paper applies several machine learning classifiers to identify fake reviews based on the content of the reviews as well as several extracted features from the reviewers. We apply the classifiers on real corpus of reviews taken from open source sites. Besides the normal natural language processing on the corpus to extract and feed the features of the reviews to the classifiers, the paper also applies several features engineering on the corpus to extract various behaviours of the reviewers. The paper compares the impact of extracted features of the reviewers if they are taken into consideration within the classifiers. The papers compares the results in the absence and the presence of the extracted features in two different language models namely TF-IDF. The results indicates that the engineered features increase the performance of fake reviews detection process.

The rapid growth of the Internet influenced many of our daily activities. One of the very rapid growth area is ecommerce. Generally e-commerce provide facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services, while potential customers can use it to decide either to buy or to use a product. Unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. This research aims to detect fake reviews for a product by using the text and rating property from a review. The rapid growth of the Internet influenced many of our daily activities. One of the very rapid growth area is Ecommerce. Generally e-commerce provide facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services, while potential customers can use it to decide either to buy or to use a product. Unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. This research aims to detect fake reviews for a product by using the text and rating property from a review. The rapid growth of the Internet influenced many of our daily activities. One of the very rapid growth area is ecommerce. Generally e-commerce provide facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services, while potential customers can use it to decide either to buy or to use a product. Unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. This research aims to detect fake reviews for a product by using the text and rating property from a review. Machine learning techniques can provide a big contribution to detect fake reviews of web contents. Generally, web mining techniques find and extract useful information using several machine learning algorithms. One of the web mining tasks is content mining. A traditional example of content mining is opinion mining which is concerned of finding the sentiment of text (positive or negative) by machine learning where a classifier is trained to analyse the features of the reviews together with the sentiments. Usually, fake reviews detection depends not only on the category of reviews but also on certain features that are not directly connected to the content. Building features of reviews normally involves text and natural language processing NLP. However, fake reviews may require building other features linked to the reviewer himself like for example review time/date or his writing styles. Thus the successful fake reviews detection lies on the construction of meaningful features extraction of the reviewers.

Usually, fake reviews detection depends not only on the category of reviews but also on certain features that are not directly connected to the content. Building features of reviews normally involves text and natural language processing NLP. However, fake reviews may require building other features linked to the reviewer himself like for example review time/date or his writing styles. Thus the successful fake reviews detection lies on the construction of meaningful features extraction of the reviewers.

II. RELATED WORKS

R. Barbado, O. Araque, and C. A. Iglesias: The impact of online reviews on businesses has grown significantly during last year's [1], being crucial to determine business success in a wide array of sectors, ranging from restaurants, hotels to e-commerce. Unfortunately, some users use unethical means to improve their online reputation by writing fake reviews of their businesses or competitors. Previous research has addressed fake review detection in a number of domains, such as product or business reviews in restaurants and hotels. However, in spite of its economic interest, the domain of consumer electronics businesses has not yet been thoroughly studied. This article proposes a feature framework for detecting fake reviews that has been evaluated in the consumer electronics domain. The contributions are fourfold: (i) Construction of a dataset for classifying fake reviews in the consumer electronics domain in four different cities based on scraping techniques; (ii) definition of a feature framework for fake review detection; (iii) development of a fake review classification method based on the proposed framework and (iv) evaluation and analysis of the results for each of the cities under study. We have reached an 82% F-Score on the classification task and the Ada Boost classifier has been proven to be the best one by statistical means according to the Friedman test.

Tadelis: Online marketplaces have become ubiquitous, as sites such as eBay, Taobao, Uber, and Airbnb are frequented by billions of users [2]. The success of these marketplaces is attributed to not only the ease in which buyers can find sellers, but also the trust that these marketplaces help facilitate through reputation and feedback systems. I begin by briefly describing the basic ideas surrounding the role of reputation in facilitating trust and trade, and offer an overview of how feedback and reputation systems work in online marketplaces. I then describe the literature that explores the effects of reputation and feedback systems on online marketplaces and highlight some of the problems of bias in feedback and reputation systems as they appear today. I discuss ways to address these problems to improve the practical design of online marketplaces and suggest some directions for future research.

M. J. H. Mughal Web data mining became an easy and important platform for retrieval of useful information. Users prefer World Wide Web more to upload and download data [3]. As increasing growth of data over the internet, it is getting difficult and time consuming for discovering informative knowledge and patterns. Digging knowledgeable and user queried information from unstructured and inconsistent data over the web is not an easy task to perform. Different mining techniques are used to fetch relevant information from web (hyperlinks, contents, web usage logs). Web data mining is a sub discipline of data mining which mainly deals with web. Web data mining is divided into three different types: web structure, web content and web usage mining. All these types use different techniques, tools, approaches, algorithms for discover information from huge bulks of data over the web.

C. C. Aggarwal The recent proliferation of social media has enabled users to post views about entities [4], individuals, events, and topics in a variety of formal and informal settings. Examples of such settings include reviews, forums, social media posts, blogs, and discussion boards. The problem of opinion mining and sentiment analysis is defined as the computational analytics associated with such text. A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance Online reviews have become a valuable resource for decision making [5]. However, its usefulness brings forth a curse – deceptive opinion spam. In recent years, fake review detection has attracted significant attention. However, most review sites still do not publicly filter fake reviews. Yelp is an exception which has been filtering reviews over the past few years. However, Yelp's algorithm is trade secret. In this work, we attempt to find out what Yelp might be doing by analyzing its filtered reviews. The results will be useful to other review hosting sites in their filtering effort. There are two main approaches to filtering: supervised and unsupervised learning. In terms of features used, there are also roughly two types: linguistic features and behavioral features. In this work, we will take a supervised approach as we can make use of Yelp's filtered reviews for training. Existing approaches based on supervised learning are all based on pseudo fake reviews rather than fake reviews filtered by a commercial Web site. Recently, supervised learning using linguistic n-gram features has been shown to perform extremely well (attaining around 90% accuracy) in detecting crowdsourced fake reviews generated using Amazon Mechanical Turk (AMT). We put these existing research methods to the test and evaluate performance on the real-life Yelp data. To our surprise, the behavioral features perform very well, but the linguistic features are not as effective. To investigate, a novel information theoretic analysis is proposed to uncover the precise psycholinguistic difference between AMT reviews and Yelp reviews (crowdsourced vs. commercial fake reviews). We find something quite interesting. This analysis and experimental results allow us to postulate that Yelp's filtering is reasonable and its filtering algorithm seems to be correlated with abnormal spamming behaviors.

III. METHODOLOGY

A. 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.

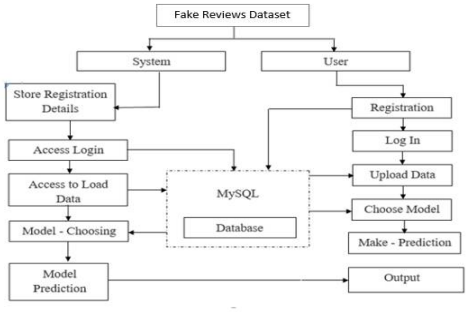


Figure 1: Fake reviews dataset block diagram

IV. IMPLEMENTATION:

A. 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 understand it using an example. Below I have a training data set of weather and corresponding target variable 'Play' (suggesting possibilities of playing). Now, we need to classify whether players will play or not based on weather condition. Let's follow the below steps to perform it.

Step 1: Convert the data set into a frequency table

Step 2: Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of playing is 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: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Problem**:** Players will play if weather is sunny. Is this statement is correct?

We can solve it using above discussed method of posterior probability.

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

Here we have P (Sunny |Yes) = 3/9 = 0.33, P (Sunny) = 5/14 = 0.36, P (Yes) = 9/14 = 0.64

Now, P (Yes | Sunny) = 0.33 * 0.64 / 0.36 = 0.60, which has higher probability.

Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification and with problems having multiple classes.

1) It is easy and fast to predict class of test data set. It also perform well in multi class prediction
2) When assumption of independence holds, a Naive Bayes classifier performs better compare to other models like logistic regression and you need less training data.
3) It perform well in case of categorical input variables compared to numerical variable(s). For numerical variable, normal distribution is assumed (bell curve, which is a strong assumption).

B. Applications of Naive Bayes Algorithms

1) *Real time Prediction:* Naive Bayes is an eager learning classifier and it is sure fast. Thus, it could be used for making predictions in real time.
2) *Multi class Prediction:* This algorithm is also well known for multi class prediction feature. Here we can predict the probability of multiple classes of target variable.
3) *Text classification/ Spam Filtering/ Sentiment Analysis:* Naive Bayes classifiers mostly used in text classification (due to better result in multi class problems and independence rule) have higher success rate as compared to other algorithms. As a result, it is widely used in Spam filtering (identify spam e-mail) and Sentiment Analysis (in social media analysis, to identify positive and negative customer sentiments)
4) *Recommendation System:* Naive Bayes Classifier and Collaborative Filtering together builds a Recommendation System that uses machine learning and data mining techniques to filter unseen information and predict whether a user would like a given resource or not

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

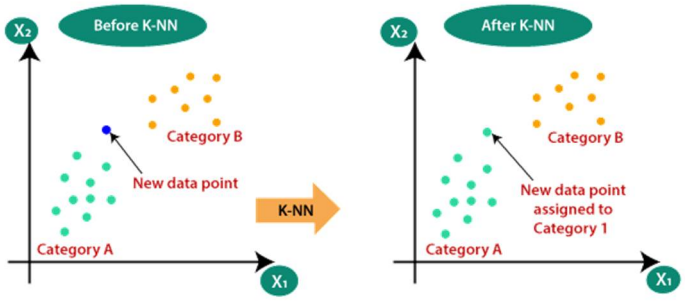
C. KNN

- K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.
- K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.
- K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.
- K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.
- K-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data.
- It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.
- KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.
- Example: Suppose, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog. So for this identification, we can use the KNN algorithm, as it works on a similarity measure. Our KNN model will find the similar features of the new data set to the cats and dogs images and based on the most similar features it will put it in either cat or dog category.



1) Why do we need a K-NN Algorithm?

Suppose there are two categories, i.e., Category A and Category B, and we have a new data point x1, so this data point will lie in which of these categories. To solve this type of problem, we need a K-NN algorithm. With the help of K-NN, we can easily identify the category or class of a particular dataset. Consider the below diagram:



2) How does K-NN work?

The K-NN working can be explained on the basis of the below algorithm:

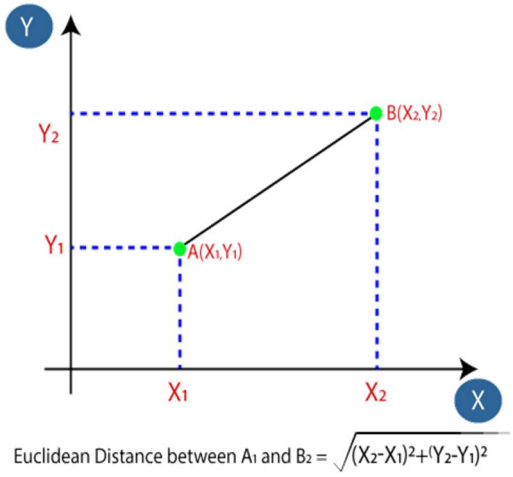
- Step-1: Select the number K of the neighbors
- Step-2: Calculate the Euclidean distance of K number of neighbors
- Step-3: Take the K nearest neighbors as per the calculated Euclidean distance.
- Step-4: Among these k neighbors, count the number of the data points in each category.
- Step-5: Assign the new data points to that category for which the number of the neighbor is maximum.
- Step-6: Our model is ready.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

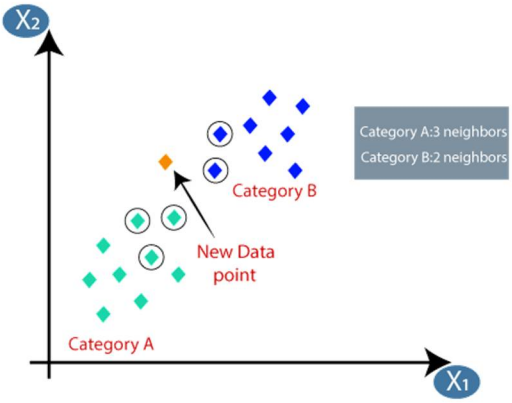
Suppose we have a new data point and we need to put it in the required category. Consider the below image:



Firstly, we will choose the number of neighbors, so we will choose the k=5.
Next, we will calculate the **Euclidean distance** between the data points. The Euclidean distance is the distance between two points, which we have already studied in geometry. It can be calculated as:



By calculating the Euclidean distance we got the nearest neighbors, as three nearest neighbors in category A and two nearest neighbors in category B. Consider the below image:



As we can see the 3 nearest neighbors are from category A, hence this new data point must belong to category A.

3) How to select the value of K in the K-NN Algorithm?

Below are some points to remember while selecting the value of K in the K-NN algorithm:

- There is no particular way to determine the best value for "K", so we need to try some values to find the best out of them. The most preferred value for K is 5.
- A very low value for K such as K=1 or K=2, can be noisy and lead to the effects of outliers in the model.
- Large values for K are good, but it may find some difficulties.

4) Advantages of KNN Algorithm

- It is simple to implement.
- It is robust to the noisy training data
- It can be more effective if the training data is large.

5) Disadvantages of KNN Algorithm

- Always needs to determine the value of K which may be complex some time.
- The computation cost is high because of calculating the distance between the data points for all the training samples.

D. Logistic Regression

Logistic Regression was used in the biological sciences in early twentieth century. It was then used in many social science applications. Logistic Regression is used when the dependent variable (target) is categorical.

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.

E. 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.

F. 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.

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

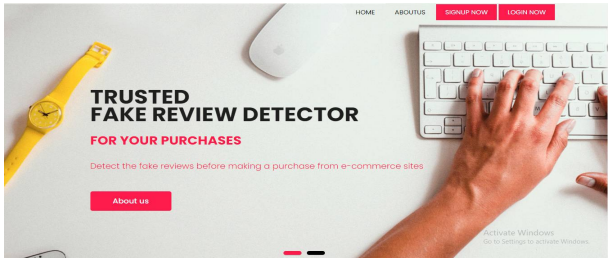
Insurance to predict the chances that a policy holder will die before the term of the policy expires based on certain criteria, such as gender, age and physical examination.
Banking to predict the chances that a loan applicant will default on a loan or not, based on annual income, past defaults and past debts.

G. Logistic Regression vs. Linear Regression

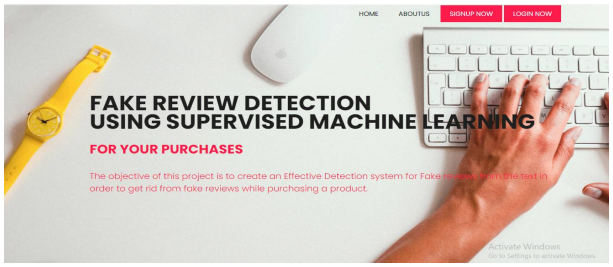
The main difference between logistic regression and linear regression is that logistic regression provides a constant output, while linear regression provides a continuous output.
In logistic regression, the outcome, such as a dependent variable, only has a limited number of possible values. However, in linear regression, the outcome is continuous, which means that it can have any one of an infinite number of possible values.
Logistic regression is used when the response variable is categorical, such as yes/no, true/false and pass/fail. Linear regression is used when the response variable is continuous, such as number of hours, height and weight.
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:

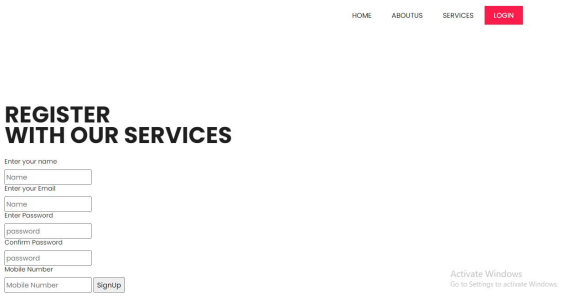
The following images will visually depict the process of our project.



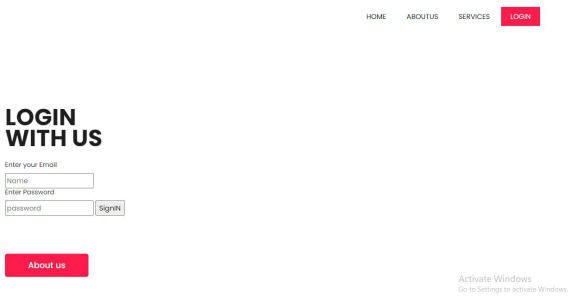
1) *Home Page:* In this home page we can see the logo designing of our website and here we are detecting the fake reviews from the review entered by the user.



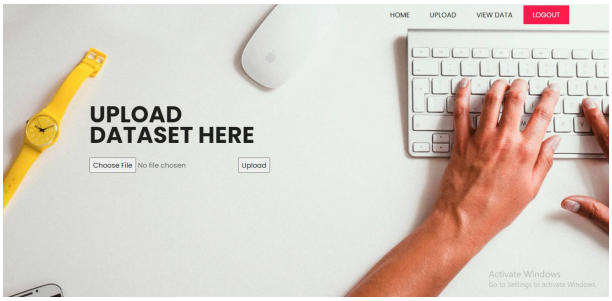
2) *About Page:* This is about page, here the application describes what main objective of this project is.



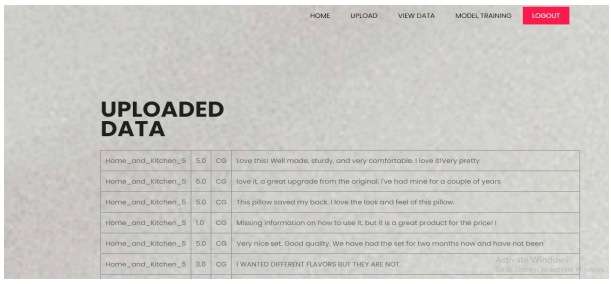
3) *Registration Page:* Registration page in which user need to register to start.



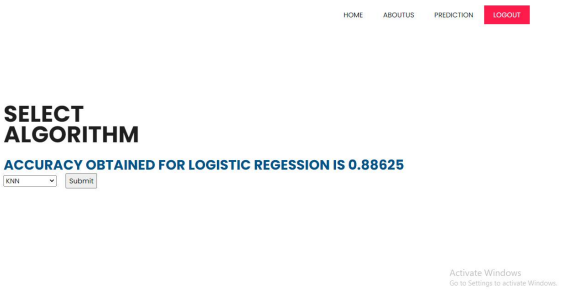
4) *Login Page:* In this login page, user need to enter valid credentials in order to enter.



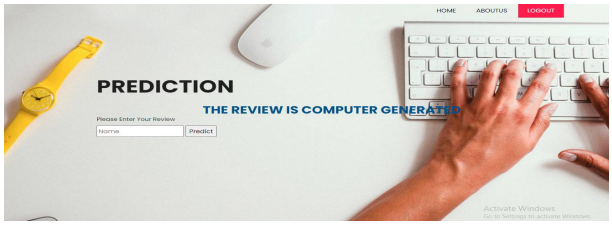
5) *Upload Page:* In this upload Page in order to upload the dataset.



6) *View Data Page:* In this user views the data which he was uploaded to the system.



7) *Model Training Page:* In this model training page, training of your model takes place and display the model's accuracy



8) *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.

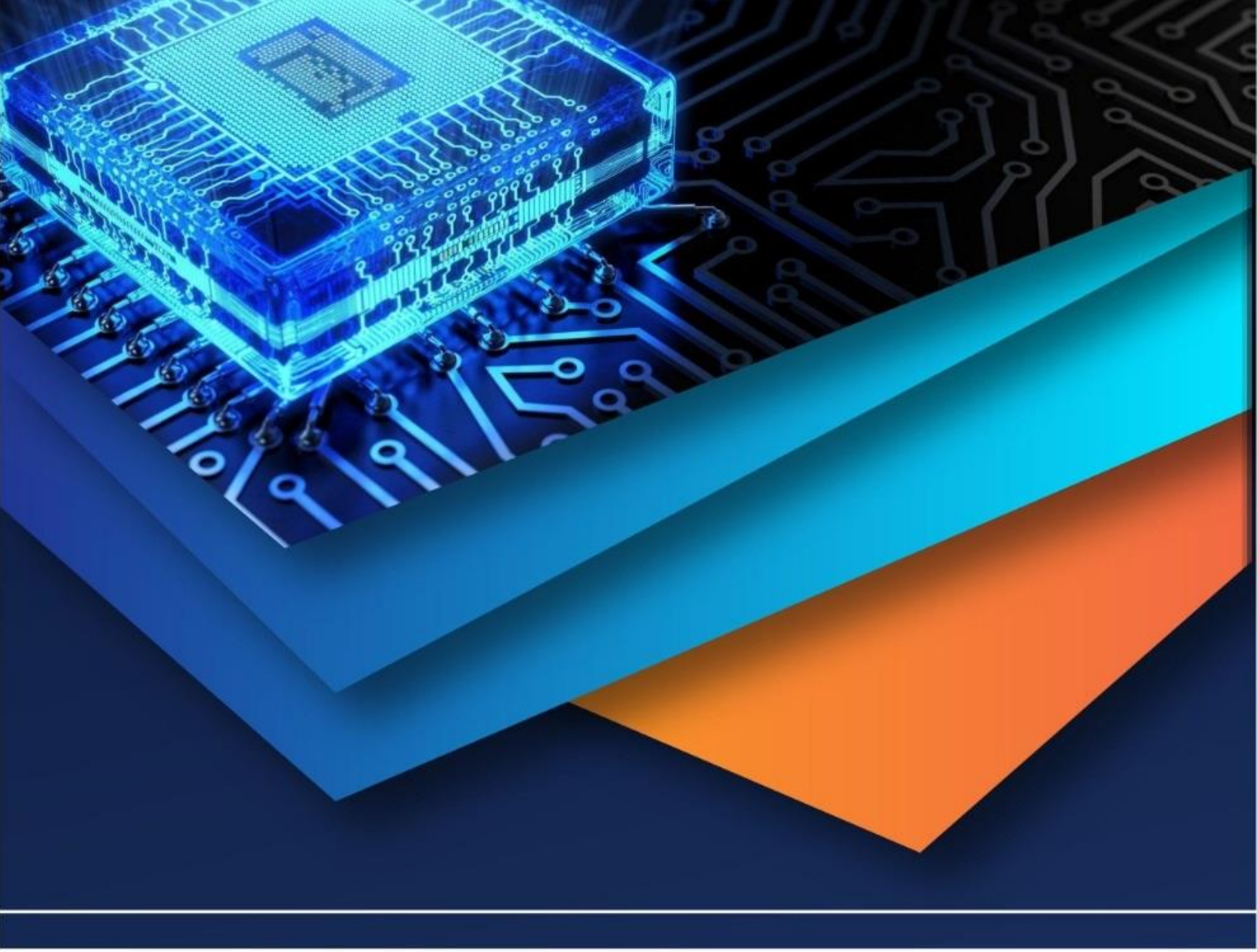
VI. CONCLUSION

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.

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 Proceedings of the 16th International Conference on World Wide Web, ser. WWW '07, 2007.
[7] E. Elmurngi and A. Gherbi, Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques. IARIA/DATA ANALYTICS, 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.
[11] H. Li et al., "Spotting fake reviews via collective positive-unlabeled learning." 2014.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)