|| Volume 6 || Issue 5 || May 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

SPAM DETECTION BY USING KNN ALGORITHM TECHNIQUES

Nikita Deshmukh[1], Vrushali Dhumal[2], Nikita Gavasane[3], Prof. Suchita V. Jadhav[4]

Student in Computer Engineering , Zeal College of Engineering and Research Pune,
Savitribai Phule Pune University Maharashtra[1,2,3]
Asst. Professor, Computer Engineering, Zeal College of Engineering and Research Pune,
Savitribai Phule Pune University Maharashtra[4]

Abstract: - Preceding purchasing an item, individuals typically advise themselves by perusing on the web reviews. To make more benefit dealers frequently attempt to fake client experience. As clients are being misdirected along these lines, perceiving and eliminating fake reviews is critical. This paper examines spam recognition techniques, in light of AI, and presents their outline and results. With the nonstop develop of E-trade frameworks, online reviews are essentially considered as a pivotal factor for building and keeping a decent standing. Besides, they have a powerful part in the dynamic interaction for end clients. Generally, a positive audit for an objective article pulls in more clients and lead to high expansion in deals. These days, tricky or fake surveys are intentionally composed to construct virtual standing and drawing in possible clients. Hence, distinguishing fake reviews is a clear and progressing research region. Recognizing fake reviews depends not just on the critical highlights of the surveys yet additionally on the practices of the commentators. This proposed system is designed to deal with distinguish fake reviews. Notwithstanding the highlights extraction interaction of the surveys, this paper applies a few highlights designing to extricate different practices of the commentators.

Keywords: *Online Reviews, Spam, KNN, fake reviewers*

I INTRODUCTION

These days, when clients need to draw an attention about products or services, reviews become the main source of Information. For instance, when clients take the inception to book an inn, they read the reviews on the assessments of Different clients on the inn products. Contingent upon the criticism of the surveys, they choose to book room or not. In the event that they went to a positive criticism from the reviews, they most likely continue to book the room. In this manner, chronicled reviews turned out to be entirely believable wellsprings of data to the vast majority in a few online products. Since, reviews are viewed as types of sharing bona fide input about certain or negative products, any endeavour to control those surveys by composing deceiving or inauthentic substance is considered as misleading activity and such surveys are named as phony. Such case drives us to think consider the possibility that not every one of the composed surveys are straightforward or believable. Imagine a scenario in which a portion of these reviews are phony. Subsequently, identifying fake audit has become and still in the condition of dynamic and required examination territory. AI methods can give a major commitment to recognize fake reviews of web substance. For the most part, web mining strategies find and concentrate valuable data utilizing a few AI calculations. One of the web mining errands is content mining. A conventional illustration of substance mining is assessment mining which is worried of discovering the feeling of text (good or negative) by AI where a classifier is prepared to break down the highlights of the surveys along with the assumptions. Generally, fake reviews identification depends on the class of surveys as well as on specific highlights that are not straightforwardly associated with the substance. Building highlights of reviews typically includes text and regular language preparing NLP. In any case, fake reviews may require building different highlights connected to the analyst himself like for instance survey time/date or his composing styles. Consequently the effective phony reviews location lies on the development of significant highlights extraction of the analysts.

AI is perhaps the main innovative patterns which lies behind numerous basic applications. The fundamental force of AI is assisting machines with consequently taking in and develop themselves from past experience. There are a few kinds of AI calculations; to be specific directed, semi managed and unaided AI. In the astonished methodology, both information and yield information are given and the preparation information should be marked and ordered. In the unaided learning approach, just the information is given with no order or marks and the job of the methodology is to track down the best fit grouping or arrangement of the information. Consequently, in unaided learning, all information are unlabeled and the job of

the methodology is to name them. At long last, in the semi managed approach, some information are named however the most are unlabeled. In this part, we present a rundown of the regulated learning calculations as they are the fundamental focal point of this proposed framework.

II LITERATURE SURVEY

Their [1] study reflect on the classifier which was good for text classification. Also evaluate machine learning algorithm on spam emails detection and outcomes shows naive Bayes algorithm gives effective accuracy and precision using WEKA and our email management system which utilize php-ml library hosted. Also comparative study with SVM and previous existing system in terms of accuracy and were dataset used.

A brief overview of spam detection methods published during the last decade was discussed [2] . It was shown that using different datasets yields extremely different results. Moreover, the lack of a proper gold standard dataset was recognized as a major problem in spam detection. Although linguistic approaches dominate in number of research papers, spammer detection techniques have shown promising results.

An intelligent system [3] for the detection and filtering of spam e-mails was described. Machine learning methods aim to create the best models using the available data and analyze new data most accurately, with the help of the model created using previous data. They studied, spam detection was carried out using machine learning methods. The K-nearest neighbors, support vector machine, and decision trees were used in the classification stage. The classification achieved an accuracy was best achieved in spam detection

Many different solutions exist to categorize incoming messages such as white list, grey list, blacklist, Machine Learning, Rule-based filtering, etc. However, no one definitively. A possible reason is since spammers are high resilient, once a spam filtering method is compromised spammers adapt to it. The aim of their work had the objective of detecting in a more effective way spam email with the Multinomial Naïve Bayes approach, in addition to text sanitation and TF-IDF. Results given by their proposed model gave an accuracy improve than Multinomial Naïve Bayes by its own. They elucidates the different Machine Learning Techniques such as J48 classifier, Adaboost, K-Nearest Neighbor, Naive Bayes, Artificial Neural Network, Support Vector Machine, and Random Forests algorithm for filtering spam emails using different email dataset. However, here the comparison of different spam email classification technique was presented and summarizes the overall scenario regarding accuracy rate of different existing approaches.

III. PROBLEM STATEMENT

Untruthful opinions represent purposefully fake reviews. Reviews on brands only aren't focused on products, but rather on brands or manufacturers. Non-reviews include advertisements or other irrelevant reviews containing no opinions. Although types two and three fail to address specific products, they aren't fraudulent. These types of spam are also easy to spot manually and traditional classification approaches have no problem in detecting them. Untruthful reviews are shown to be much harder task for a machine as well as for a human observer. For those reasons this type of spam is considered and system is developed.

IV . IMPLEMENTATION DETAILS OF MODULE

Machine learning can be used to detect fake or spam reviews in this system. The yelp dataset is considered and used in this system, yelp dataset is collected and we are applying algorithm (KNN: - k-nearest neighbor) and prepare a trained file to compare with further review data.

The proposed approach consists of three basic phases in order to get the best model that will be used for fake reviews detection. These phases are explained in the following:

- Data Preprocessing
- Tokenization
- Lemmatization
- Feature Extraction
- Classification

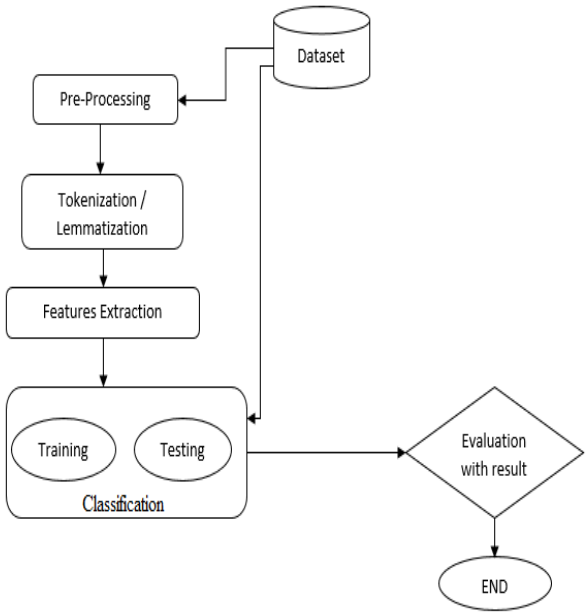

Fig 1: - System Architecture

Algorithm: - KNN is the K Nearest Neighbor classification algorithm. KNN is the simplest and best technique for classification. For classification and regression it's a non-parametric technique. Two stages are involved in KNN algorithm i.e. one is filtering and other is training.

- Stage 1 Training- Trained message are store after training process.
- Stage 2 Filtering- In filtering stage for a given message x, its k nearest neighbors among the messages in the training set are determined. If there are more spam's among these neighbors, then classify given message as spam. Otherwise classify it as ham.
- Steps:-

1. Determine parameter D = number of nearest neighbors.

2. Calculate the distance between the training samples and query-instance.

3. Sort the distance and determine nearest neighbors based on the N-th minimum distance

4. Gather the category X of the nearest neighbors

5. Utilize straightforward larger part of the classification of closest neighbors as the neighbors as the prediction value of the query instance

V. RESULT AND EXPERIMENTAL

To perform test assessment of the models, different execution measurements like Accuracy, F1-Score, Sensitivity, Specificity, Precision, and Recall can be utilized in this work.

On testing the dataset and model in system we have achieved following parameter.

$$\text{Accuracy} = \frac{(TP + FN)}{(TP + TN + FP + FN)},$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)},$$

$$\text{Precision} = \frac{TP}{(TP + FP)},$$

$$\text{Recall} = \frac{TN}{(TN + FN)},$$

$$\text{F1} - \text{Score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}.$$

Accuracy = 80

Precession = 88

Recall Sore = 72 and

F1 score = 79

VI. CONCLUSIONS

As discussed the significance of reviews and what they mean for pretty much all that identified with web data. Clearly surveys assume an essential role in people's decision. Consequently, fake reviews location is a distinctive and continuous examination territory. Machine Learning based fake reviews identification approach is introduced. In the proposed approach, both the behavioural features of the reviewers and features of the reviews are thought of. The Yelp dataset is utilized to assess the proposed approach. Various techniques from KNN are carried out in the proposed approach.

REFERENCES

[1] Drasko Radovanovic, Bozo Krstajic, "***Review Spam Detection using Machine Learning***", 2018 23rd International Scientific-Professional Conference on Information Technology (IT)

[2] Asma Bibi, Rasia Latif, Samina Khalid, Waqas Ahmed,"***Spam Mail Scanning Using Machine Learning Algorithm***", Spam Mail Scanning Using Machine Learning Algorithm, Journal of Computers, January 24, 2020

[3] Mete Yaganoglu1, Erdal Irmak, "***Separation of Incoming E-Mails Through Artificial Intelligence Techniques***", European Journal of Science and Technology, January 2021

[4] Alan Chavez, "***TF-IDF classification based Multinomial Naive Bayes model for spam filtering***"

[5] V. Sri Vinitha and D. Karthika Renuka," ***Performance Analysis of E-Mail Spam Classification using different Machine Learning Techniques***", IEEE , 2019