iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** V **Month of publication:** May 2019

DOI: https://doi.org/10.22214/ijraset.2019.5279

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: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Detecting Spam in Twitter and Email using Machine Learning Approach

Sneha Linganur[1], Krishnagouda Patil[2], Vishwa Gangur[3], Bhavani Patil[4]
[1, 2, 3, 4] *Student, Department of CSE, DSCE, Bengaluru, India*

Abstract: ***Online social Networks have turned out to be increasingly famous in the global. Individuals utilize these social platforms (e.g., Facebook, Twitter, etc.) for storing and sharing their personal information, activities, views, documents, posts and so on. At the same time, Social spam turned into an incessant issue nowadays which is incredibly hazardous to both individual clients and companies. Therefore, it is an imperious need to develop more accurate and powerful spam recognition models. This Paper essentially centers around machine Learning technique for distinguishing spam in Twitter and Email so as to make it increasingly secure. We choose twitter and email as our study targets. We apply Naive Bayes algorithm for discovery of spam using techniques such as preprocessing, word extraction by taking emails and real-time tweets as input to our project. The outcomes show an efficacious classification of spam and non-spam. Finally, we verify the perceptibility of spam tweets and mails as well as accounts through evaluation. Our proposed method could accomplish better outcomes contrasted with other existing supervised machine learning methods.***

Keywords: Naive Bayes, Spam, Spam Classification, Spam Detection, Non-Spam, Online Social Network.

I. INTRODUCTION

From a previous couple of years, Online Social Networks such as Twitter and Facebook have become progressively popular platforms which are an essential part of our everyday lives. Users spend more time on well-known social platforms use those for sharing ideas, personal thoughts and information, discuss events, make friends worldwide. Due to fast developing, these stages draw a millions of people along with spammers because of the vast measure of information present on these sites. Twitter is appraised as the most well-known social network in the midst of youngsters [1]. Spammers steal personal information, spreading malware and misinformation, malicious links for the purpose of financial gain [2]. These spam attacks not only affect individual users but also damage the whole internet services worldwide.

In recent days, Online Social Networks are known as popular communication channels still, the email regnant as the most popular internet communication form in our daily life as revealed by survey results. One of the major problems in Email communication is spam emails. spam emails, also known as unwanted or group of spontaneous messages sent in bulk by email. In some instances, they may even be deleterious. for e.g. directing to fake advertising sites, extending viruses and hoard network bandwidth, storage space, and computational power.

In addition to wasting bandwidth and time, spam email also costs money to users to get out.

Thus issues related to spam are increasingly growing day by day, this has led to the imperious need to create more accurate and efficient spam detection model which needs to correctly distinguish as spam and ham. There is no spam filter that works 100%. therefore, more sophisticated and precise classifier model need to be developed to dispense with the issue of spam. To detect this, machine learning algorithms present various classifier models [3] for example Naive Bayes, Random Forest, Support Vector Machine etc. are used. One of machine learning classifier model is Naive Bayesian Classifier is additionally utilized in [4] to isolate the spam and ham emails. In this paper, for efficient and accurate purpose more sophisticated classifier model Naive Bayes classifier is used to eliminate the problem of spam.

II. LITERATURE SURVEY

Tianda Yang describes how to identify and classify the spam emails utilizing the naive Bayes classifier algorithm by using 3 different steps: they are pre-processing, testing. and training Based on the training data set. And detailed process and technique is described in the paper [5].

Rohit Kumar Solanki describes the identification of spam emails by 2 filtering techniques based on the local-global classifier. The 2 techniques used are (1) Training (2) Learning message. Training model it has 2 parts, (1) Pre-processing (2) Tokenization .and it will help to build the model to identify the spam, predict the probability and classify them, and about nearly 93% accuracy is obtained for correctly classified data. the rest discussed in the paper [6].

Himank Gupta gave a model depends on multiple machine learning algorithms approach which deals with different issues related to probability, accuracy, duplicate tweets, time lag, etc. and also showed the accuracy to handle thousands of tweets in short time. used multiple algorithms such as SVM, random forest etc. and the detailed operation is discussed in paper [7].

W.A. Awaad and S.M. ELseuoifi describe the different machine learning techniques to filter spam emails with efficient level. Details of all the methods are explained and comparison of their analogy of their execution is done on the Spam Assassin spam corpus is presented. In terms of accuracy we can see naive Bayes algorithm has a satisfactory result. [8].

Haiyi Zhang, Di Li describes Email Spam Detection by Classifying the text/message using Naive Bayes algorithms and this paper focused on training, testing and classifying the Data. we selected this paper because classification of emails is well explained here. And also investigated strategies utilized for text pre-processing and probability computation in text classification [9].

Ms. Ashwini Athawale,

Mrs. Deepali M. Gohil describes how to identify spam in twitter using content, user based feature to analyze spam behavior of user. and in this paper anomaly detection is being done by mistreatment trained data-set and Naive Bayes rule for the classification. They showed that accuracy of the result by comparing SVM and Naive Bayes algorithm and Naive Bayes is more efficient than SVM [10].

III. PROPOSED METHODOLOGY

Spam Classification has a noteworthy issue in the present computerized area. Various spam classification methods are utilized to tackle this issue. Utilizing this spam recognition approach, we can recognize the spam and non-spam tweets and emails. On this wise, we are utilizing the Naive Bayesian Classifier for spam classification and also we likewise use preprocessing and Tokenization and clustering techniques as part of classification process.

A. Preprocessing

Preprocessing is a datamining technique which converts the raw data into justifiable format or clean data before feeding it to the algorithm. Real word data (tweets and emails) may be inconsistent, incomplete and may also contain errors or outliers. Data preprocessing is a method which fix such issues. By transforming into structured data made it useful and predictable for analysis.

B. Tokenization

It is the process of separating/breaking an arrangement of strings or dividing the text into pieces, for instance, words, catchphrases, expressions, and other meaningful elements called tokens. Tokens can be Phrases, singular words, or also entire sentences. The series of tokens becomes the input for further processing such as text mining or parsing.

C. Naive Bayes

Naive Bayes Performs on the basis of Bayes theorem which uses hypothesis of strong anatomy. This algorithm strongly depends on the probability models. It works at its fullest in a supervised learning environs. The significant benefit of this classification algorithm is that, it only requires a minimum volume of training data to determine the parameters important for classification and the classifier can be trained incrementally.

Bayes' Theorem identifies the likelihood of an incident happening given the likelihood of another incident that has previously happened. The mathematical equation for Bayes' theorem is given as:

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

D. Clustering

Clustering is a technique of unsupervised learning and is a typical method for statistical data analysis used in numerous fields. Clustering contains the grouping of data points. data points having homogenous features or properties are ought to have in the same group. Whereas data points are having heterogeneous features or properties ought to have in different groups.

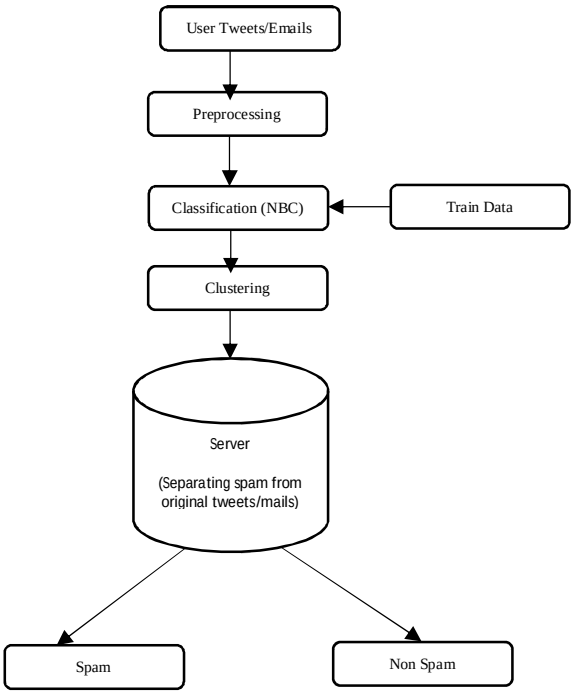


Fig 1: The Proposed Methodology

1) *Step 1:* Emails and Real-time tweets are given as input.
2) *Step 2:* After cleaning up the data, Tweet terms and email terms are extracted.
3) *Step 3:* The Naive Bayes algorithm compares the extracted words with train data and classifies as spam and non-spam tweets/ Emails.
4) *Step 4:* The classified similar type of spam and non-spam tweets/Emails are grouped and stored separately.
5) *Step 5:* The server is a database where actually it separates spam tweets/Emails from original tweets and Emails.
6) *Step 6:* Finally, Classifies and display as Spam and Non-spam tweets/Emails.

IV. RESULTS

The Naive Bayes Classifier model tested on the standard inputs. The inputs are taken from Twitter and Email. The Naive Bayes classifier initially computes the likelihood probabilities of all feature available in the training set after that uses the earlier likelihood (either spam or non-spam) to determine the label of a document. When the model is trained, most representative and least representative words for spam alongside their calculated probabilities shown by the classifier. The Outcome of a model is shown in Fig 2.

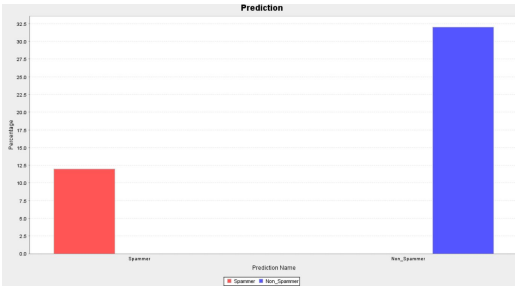


Fig 2: Percentage graph of Spam and Non-spam

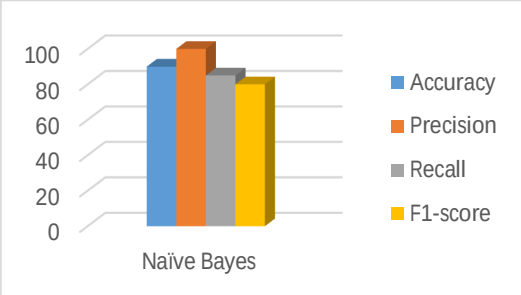


Fig 3: Evaluation graph of Naive Bayes

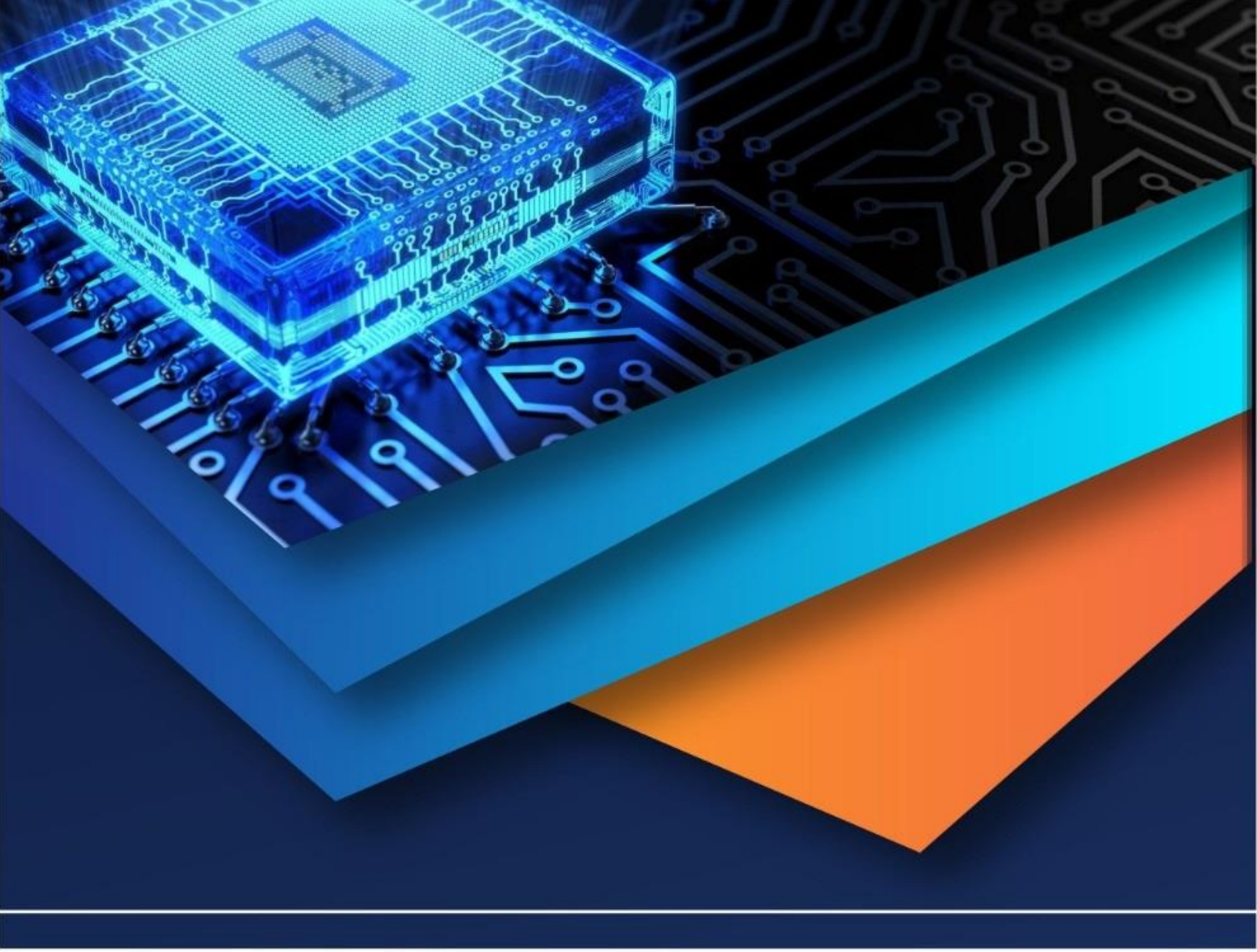
The accuracy of the classifier is also produced by the model. The outcomes acquired from Naive Bayes classifier are organized in Fig 3.

V. CONCLUSION

In this paper, we introduce a perspective method to distinguish between spam and ham contents in Twitter and Email. For this purpose, we are utilizing Naive Bayesian Classifier which is a basic and effective method for spam classification. In this project, we are developing a spam classification framework for classifying the spam and non-spam tweets and emails. As Twitter API is accessible to all clients, we used Real-time tweets and manual emails as input(dataset) to our project. The results demonstrate the effective classification of spam and non-spam tweets and emails. After evaluation, we discover that when we use Naive Bayesian Classifier has more definite and the error rate is exceptionally low. Thus we can say that Naive Bayesian Classifier delivers a superior outcome than other supervised machine learning methods. We conclude that the Naive Bayes algorithm is the ultimate classification algorithm in machine learning as it has an accuracy of 80.11% and F1-score 0.6005[11].

REFERENCES

[1] Greig, "Twitter Overtakes Facebook as the Most Popular Social Network for Teens,According to Study, DailyMail, accessed on Aug. 1, 2015 ," http://www.dailymail.co.uk/news/article-2475591/ Twitter-overtakes-Facebook-popular-socialnetwork-teens-according-study. html, 2015.
[2] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in In Collaboration, Electronic messaging, AntiAbuse and Spam Conference CEAS, 2010.\
[3] Sharma K. and Jatana N. (2014) "Bayesian Spam Classification: Time Efficient Radix Encoded Fragmented Database Approach" IEEE 2014 pp. 939-942.
[4] Sharma A. and Anchal (2014), "SMS Spam Detection Using Neural Network Classifier",ISSN: 2277 128X Volume 4, Issue 6, June 2014, pp. 240-24.
[5] Tianda Yang, Kamal AI Nasr and Ying Qian "Spam Filtering using Association Rules and NaIve Bayes Classifier" IEEE International Conference on Progress in Informatics and Computing,2015.
[6] Rohit Kumar Solanki, Karun Verma, Ravinder Kumar "Spam filtering using hybrid local-global Naive Bayes classifier" IEEE International Conference on Advances in Computing, Communications and Informatics,2015.
[7] Himank Gupta, Mohd. Saalim Jamal, Sreekanth Madisetty and Maunendra Sankar Desarkar "A Framework for Real-Time Spam Detection in Twitter" IEEE 10th International Conference on Communication Systems & Networks, 2018.
[8] W.A. Awad and S.M. ELseuofi "machine learning methods for spam e-mail classification" International Journal of Computer Science & Information Technology Vol 3, No 1, Feb 2011.
[9] Haiyi Zhang, Di Li "Naïve Bayes Text Classifier" IEEE International Conference on Granular Computing,2007.
[10] Ms. Ashwini Athawale , Mrs. Deepali M. Gohil "Spam Detection on Collection of Twitter Data Using Naive Bayes Algorithm" International Journal of Innovative Research in Science, Engineering and Technology, Vol. 7, Issue 6, June 2018
[11] https://www.analyticsindiamag.com/7-types-classification-algorithms/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY