SPAM DETECTOR:A TOOL TO MONITOR AND DETECT SPAM ATTACKS

A Minor Project Synopsis Submitted to

Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal

Towards the Partial Fulfillment for the Award of

Bachelor of Technology

(Information Technology)

Under Supervision of

Prof. Asif Ali

Submitted By:-

Prayag Gavshinde

Raghav Agrawal

Rishi Somani

Sambhav Jain

JUNE - DECEMBER 2020

PROJECT SYNOPSIS

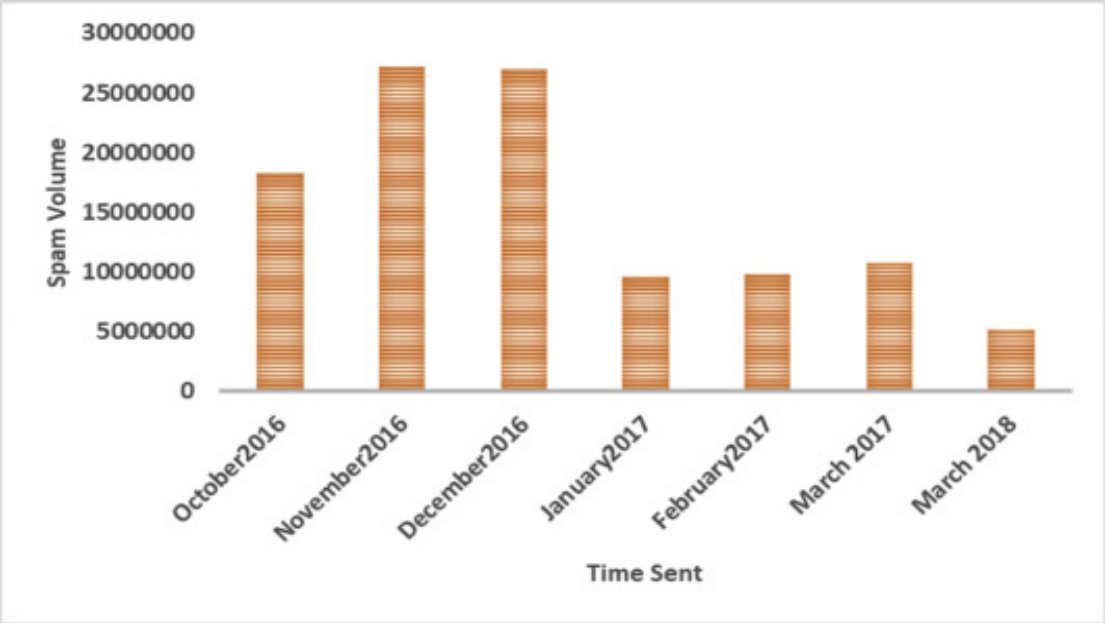
★ Abstract :-

Machine learning is a branch of artificial intelligence concerned with the creation and study of systems that can learn from data. A machine learning system could be trained to distinguish between spam and non-spam (ham) data. We aim to analyze current methods in machine learning to identify the best techniques to use in content-based spam filtering.

★ Introduction :-

Due to the wide popularity of the internet and its communication with no cost, it was recognized as the premium tool for advertising and marketing. With respect to economic constraints, most numbers of people started sending emails to thousands of people across the world in bulk. This made the internet a commercial network with the association of electronic mail as one of the quick resources of communication. The major problem in today's internet world is sending bulk or unsolicited emails to numerous users. This adds an additional advantage of launching other attacks and wasting of resources. E-mail spam comes under the electronic spam which sends bulk of unnecessary or junk mail of duplicate emails to the recipients.

E-mail spam filtering is a very widely discussed and studied topic in the field of pattern classification. E-mails can be filtered as spam or non-spam based on many features such as the frequency of occurrence of a few words in the e-mail, the length of the e-mail, or the domain from which it is being sent. researchers have come up with many techniques to identify an email as spam or non-spam.



★ Objective :-

The goal of our project is to analyze machine learning algorithms and determine their effectiveness as content-based spam filters.

Other objectives are:

- ➢ The objective of this research is to design and implement a tool to detect the spam attacks in a network.
- ➢ To give Knowledge to the user about the fake e-mails and relevant emails
- ➢ To classify that mail spam or not
- ➢ To run the test cases for various functionalities of an application in real time.
- ➢ To evaluate the performance of the tool based on the amount of spam involved in the network, it's detecting capacity.

★ Scope of the Project :-

- ➢ It provides sensitivity to the client and adapts well to future spam techniques.
- ➢ It considers a complete message instead of single words with respect to its organization.
- ➢ It increases security and control.
- ➢ It reduces IT Administration costs.

Problems with spam:

1) **Viruses**: Viruses are the most dangerous threats across the network. With the increase in internet technology, a wide variety of viruses produced to attack the machines. Spam is one of basic sources to launch such types of viruses. Spam viruses in modern technology are more dangerous as it controls the machine itself and then annihilates them. There are so many techniques used by spammers in order to allow users to click or use the links to launch thousands of spam viruses across the network. Due to increase in the intensity of spam, it captures the user's email address and passes numerous messages to the customer list, through which it disturbs the customer trust and destroys the system

2) **Server Problems:** Most of the time servers are being targeted by the spammers. Due to the increase in the intensity and volume of the spam, the company or any system has to use huge resources to maintain the server.

Due to this frequency of spam, the performance also gets affected.

3) **Hacking and Phishing:** As the computers and technologies are transforming and becoming more and more advanced and secure. So, the spammers face more difficulty to capture confidential details.
In this case, they tend to use various methods to break through the security of different IT departments. Spammers make use of hacking methods like entering into the trusted employee system without the user's awareness. Then, spammers perform different activities and keep a record of the confidential data or hold vital information either for the cost or for self-happiness. Another way is to trap the employees of the companies to enter the passwords or any valuable information into the spammer's website, so that it keeps track of the password to reveal important credentials

Review existing methods:-

- **List-based filters** - In this attempt to stop spam by categorizing senders as spammers or trusted users, and blocking or allowing their messages accordingly.
- **Content-based filters** - content-based filters evaluate words or phrases found in each individual message to determine whether an email is spam or legitimate.
- **Word-based filters** - *A word-based spam filter is the simplest type of content-based filter. Generally speaking,*

word-based filters simply block any email that contains certain terms.

- **Bayesian Filters -** Bayesian filters, considered the most advanced form of content-based filtering, employ the laws of mathematical probability to determine which messages are legitimate and which are spam.

★ Existing Systems

Many models came into picture in order to reduce the amount of spam's transfer around the internet. Despite the awareness of spam emails, productive efforts have not been developed for the network administrator to monitor the status of clients in the network. Each model has its own advantages and disadvantages. Following are few existing systems and their drawbacks.

1)Spam signature generation based framework:-

The aim of this model is to analyze the behavior of spam in the network through its characteristics and properties. This behavioral analysis identifies the spam loopholes in the network thus helping the future strategies for the prevention of spam spread across the network. This mechanism is developed based on the framework AutoRE which identifies the spam attackers based on the signatures from the exchange of messages.

DISADVANTAGE:

- ➢ Even though this framework is accessible for the URL groups embedded in email, it lacks in action taking for IT security group
- ➢ It does not provide any clarity about how well it is performed in real time for the spam campaigns.
- ➢ The results that are obtained do not provide any aggregate view of the large groups of emails.

2) Characterizing botnets from Email Spam Records:

This framework presents techniques, which use traces of spam email to map botnets in groups. This is done by viewing several bots involved in the same spam campaign. This has been used against a sample of spam email from Hotmail web mail system and has successfully detected multiple botnets. This technique uses a large set of spam emails as input, which are destined at Hotmail in a regular period.

DISADVANTAGE:

- ➢ This framework deals better with the estimation of sizes and activities based on the results generated through common characteristics
- ➢ Most importantly, as in the previous framework it lacks to provide online detection and monitoring of the networks.

3) Botsniffer:-

Botsniffer is defined as “a prototype system which is used to capture the spatial temporal-correlation in the network traffic and utilize statistical algorithms to detect botnets with theoretical bounds on the false-positive and false-negative rates.” Botnets here are the network of zombies and are recognized as a very

serious threat these days. Command & Control (C&C) channel is one of the main characteristics of botnets when compared to other malware. They use certain protocols like, for example: IRC, HTTP.

DISADVANTAGE:-

- ➢ It requires observing multiple rounds of response crowds. If there are only a few response crowds the accuracy of the algorithm may suffer.
- ➢ Sometimes not all bots respond within the similar time window when there is a relatively loose C & C.
- ➢ It does not provide aggregate global characteristics of spam botnets involved in spamming.

4)Botminer:

Botnet is a compromised machine under the influence of malware code. These compromised machines work under the influence of boot master and utilizes all the resources to counter denial of service (DOS), spam attacks, phishing, and identity theft. Command and control channel (C&C) is used by the boot master to issue the commands 19 to the bots and to coordinate between different computers.

DISADVANTAGE:-

- ➢ It is easy to manipulate the communication patterns between botnet members.

➢ If the bot master tries to generate the new commands in any unspecified manner, it tends to promote previous patterns.

★ Project Description

In this project, We aim at implementing various filtering techniques that are widely used for text classification,text-mining and content-based filters. Some of the algorithms and models are listed below which we aim to implement with the help of a rich-set of python libraries, classes and objects.

- **➢ support vector machine**
- **➢ Naive Bayes Classifier**
- **➢ Decision Trees**
- **➢ Logistic Regression**

Support vector machine:-

Support Vector Machines (SVM) has proved over the years to be one of the most powerful and efficient state-of-the-art classification techniques for solving the email spam problem. They are supervised learning models that analyze data and identify patterns. SVM algorithms are very potent for the identification of patterns and classifying them into a specific class or group. SVM is a good classifier due to its sparse data format and satisfactory recall and precision value. SVM has high classification accuracy.

Naive Bayes Classifier:-

The Naive Bayesian classifier takes its roots in the famous Bayes Theorem. Bayes Theorem essentially describes how much we should adjust the probability that our hypothesis (H) will occur, given some new evidence (e).

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

where A and B are events and $P(B) \neq 0$.

- $P(A \mid B)$ is a conditional probability: the likelihood of event A occurring given that B is true.
- $P(B \mid A)$ is also a conditional probability: the likelihood of event B occurring given that A is true.
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independently of each other; this is known as the marginal probability.

The naive bayes classifier (NB) is a simple but effective classifier which has been used in numerous applications of information processing including, Natural language processing(NLP), information retrieval, etc. The Naive Bayes Classifier technique is based on Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Naïve Bayes classifiers assume that the effect of a variable value on a given class is independent of the values of other variables. The Naive-Bayes inducer computes conditional probabilities of the classes given the instance and picks the class with the highest posterior. Depending on the

precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting.

Decision Trees:-

Decision Tree generates the output as a binary tree-like structure called a decision tree, in which each branch node represents a choice between a number of alternatives, and each leaf node represents a classification or decision. A Decision Tree model contains rules to predict the target variable. This algorithm scales well, even where there are varying numbers of training examples and considerable numbers of attributes in large databases

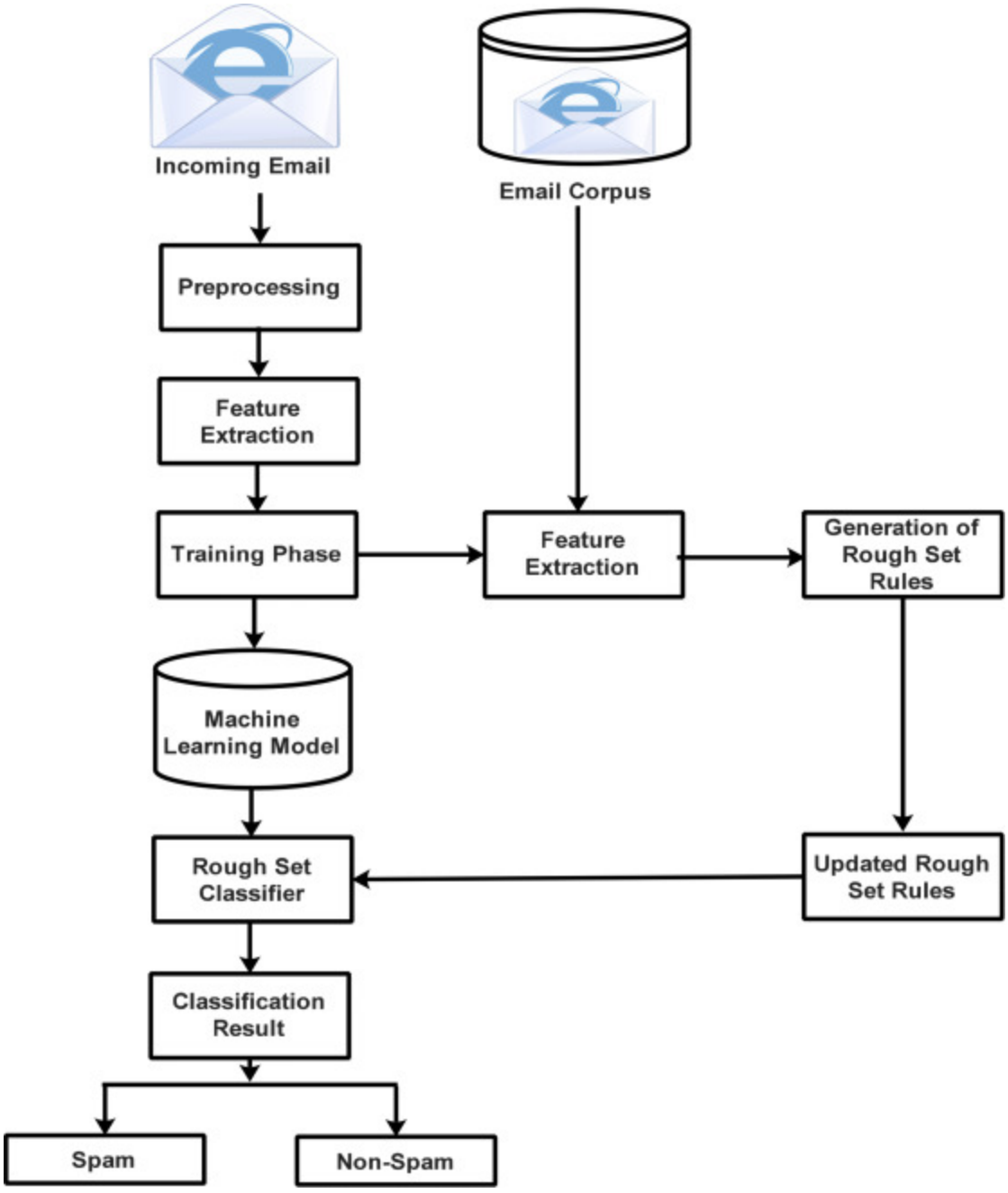
Logistic Regression:-

In statistics, the logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick. This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc. Each object being detected in the image would be assigned a probability between 0 and 1, with a sum of one.

Equation of logistic Regression

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

Below shows the email filtering process workflow of rough set approach for spam detection.

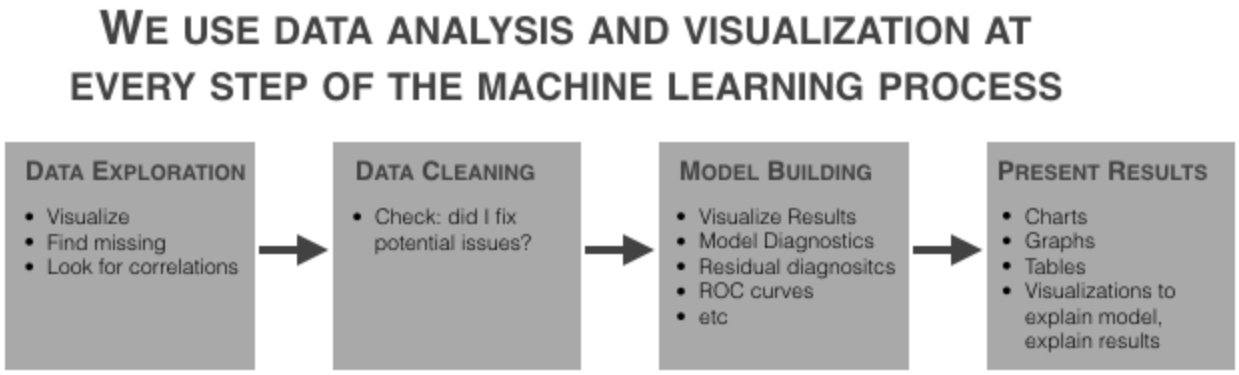


★ Implementation Of the Project

The Project consist of 4 major parts-

- ❖ Data gathering and preprocessing
- ❖ Writing and implementing different Machine Learning Algorithms with python as mentioned.
- ❖ Evaluation And analysis of Results.
- ❖ Model deployment for ready to use.

The first phase consists of collecting possible spam-mail datasets to work with. There are simple text datasets of email. We choose kaggle as our first priority for datasets to have a real-time data set and collect the spambase data sets from it. The criterion for choosing the particular dataset was that dataset should be flexible and consist utmost each field and instances with which we are able to classify the model and able to deploy in a production environment.



REQUIREMENT:-

Software- Python IDE, Jupyter-Notebook, cloud Server(Heroku)

Hardware-

- ❖ Modern Operating System:
  - ❑ Windows 7 or 10
  - ❑ Mac OS X 10.11 or higher, 64-bit
  - ❑ Linux: RHEL 6/7, 64-bit (almost all libraries also work in Ubuntu)
- ❖ 4 GB RAM
- ❖ x86 64-bit CPU (Intel / AMD architecture)

★ Conclusion:-

- We are able to classify the emails as spam or non-spam. Given a set of words, we used feature selection to obtain words which allow us to distinguish between spam and ham emails. We also compared the accuracy of various classifiers in predicting the class attribute.
- In future work, this tool can be extended to image spam detection as this one is completely based on the content spam filtering. It can be further enhanced by incorporating the sending message service feature to personal contact numbers if the spam exceeds the assumed threshold value. And finally, apart from spam attacks several other attacks can also be focused along with the protective measures.

★ Bibliography:-

https://www.techsoupcanada.ca/en/learning_center/10_sfm_explained
https://www.sciencedirect.com/science/article/pii/S2405844018353404#fd11

https://www.researchgate.net/publication/50235326_Email_Spam_Filtering_using_Supervised_Machine_Learning_Techniques