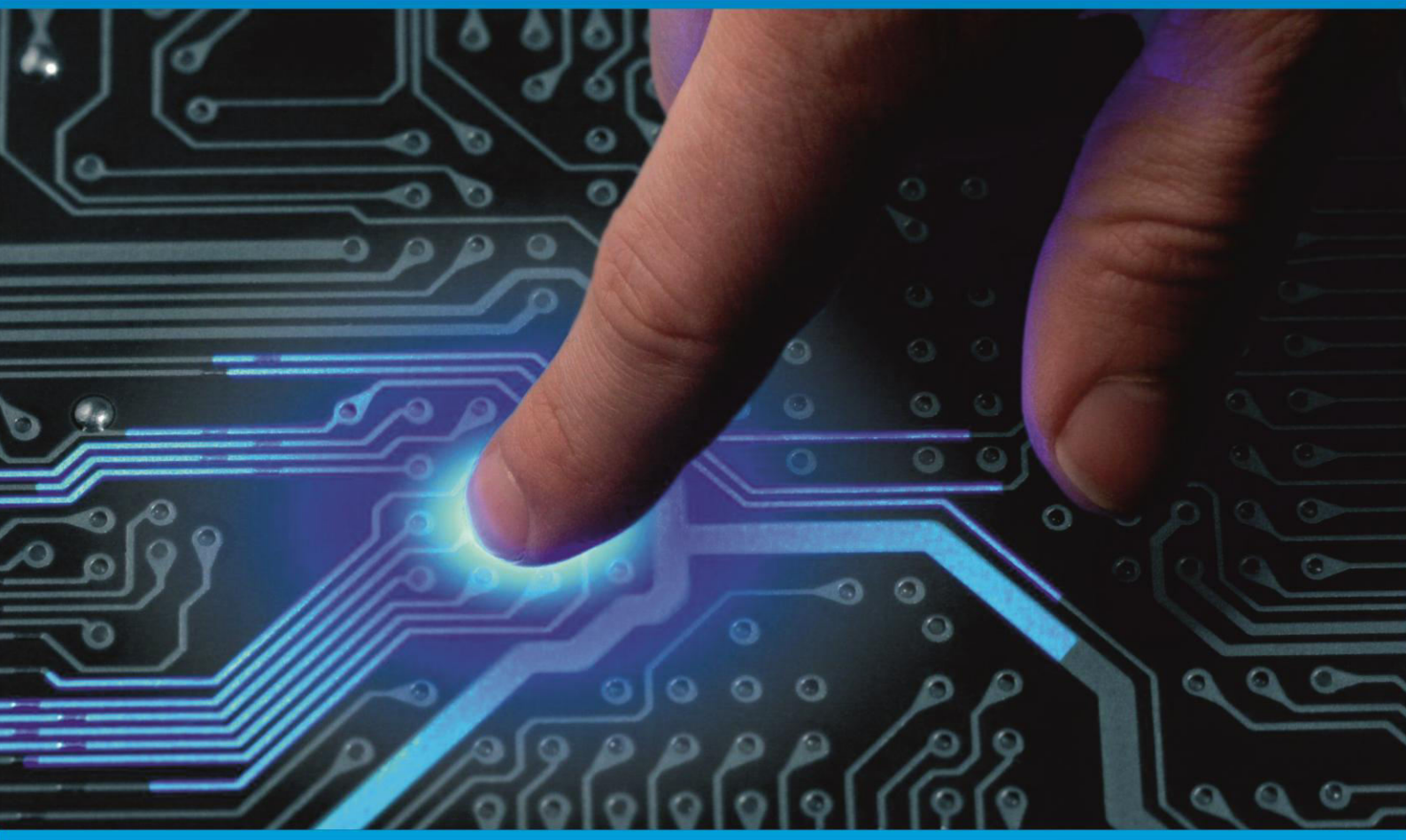
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

Email Spam Filtering Using Naïve Baye's Classifier Method

Blessymol Sunny, Ms. Mrinmoyee,

P.G. Student, Department of Computer Science, St. Joseph's College Bangalore, India

Assistant Professor, Department of Computer Science, St. Joseph's College Bangalore, India

ABSTRACT: E-mail is a useful networking tool because it saves time and money. Companies and individuals submit ads for different goods, harmful news and material, and false proposals, among other things, through email. Unsolicited Commercial E-mail (UCE), also known as email spam, is one of the most serious facing today's Internet users. In recent, machine learning methods have been effective in detecting and filtering spam emails. In this paper I review one of the most popular machine learning method Naive Baye's Classifier to detect spam.

KEYWORDS: Spam Email, Spam Filter, Content Based Filtering, Naïve Baye's Classifier.

I.INTRODUCTION

The internet has become an indispensable part of daily life, and e-mail has evolved into a powerful tool for communication. Many people are making it a habit to read their emails on a daily basis. This is a cost-effective, reliable, and fast mode of communication. Email creates the desired effect in both professional and personal relationships. With the expansion of the Internet and e-mail, spam has increased significantly in recent years. Spam can come from any place on the planet. There is Internet service open. Despite the advancement of anti-spam programs and technology, spam continues to be an issue. The number of spam messages continues to rise at an alarming rate. Companies and individuals submit ads for different goods, harmful news and material, and false proposals, among other things, via email. Spam emails irritate email recipients and waste their valuable time. These emails cause major problems for non-serious and non-technology savvy users because they lead them astray. Spam, or unwanted commercial bulk emails, has become a major issue on the internet in recent years. The spammer is the person who sends out the spam messages. They collect emails from various website addresses, chat rooms and viruses. Spam is a term that is used to describe prevents the customer from making the best use of his or her time and storage space as well as network bandwidth. The massive amount of spam emails flowing through computer networks has a negative impact on email servers memory space, communication bandwidth, CPU capacity, and user experience.. Many users have suffered untold financial losses as a result of internet scams and other deceptive activities by spammers who send emails posing as from reputable businesses in order to force individuals to reveal confidential personal information. Passwords, Bank Verification Numbers (BVNs), and credit card numbers are only a few examples.

. The various types of spam filtering techniques that have been widely used to combat the issue of email spam.

- Content based Filtering Technique
- Case based Spam Filtering method
- Heuristic or Rule based Spam Filtering Technique
- Previous Likeness based Spam Filtering Technique
- Adaptive Spam Filtering Technique

In this paper focused on content based filtering technique to find spam email. Because content based filtering is commonly used to construct automated filtering rules and classify emails using machine learning techniques like Nave Bayesian Classification, Support Vector Machines, K-Nearest Neighbour, and Neural Networks. This approach examines terms, their frequency, and distributions in the content of emails and is commonly used. Then use the created rules to filter the spam emails that come in. In e-mail filtering, two general methods are used: Information engineering and Machine learning. A collection of rules must be defined in the information engineering method, according to which emails are classified as spam or ham. A collection of such rules should be generated by the filter's user or by another authority (for example, the software company that offers a specific rule-based spam filtering tool). Since the rules must

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

be continuously revised and preserved, which is a waste of time and inconvenient for most users, this approach yields no promising results. Machine learning is more effective than information engineering because it does not require any rules to be defined. Instead, a collection of training samples consisting of pre-classified e-mail messages is used. The classification rules are then from these e-mail messages using a special algorithm. The use of machine learning in e-mail filtering, there are several algorithms used. I find the best algorithm to find email spam filtering is Naive Baye's Classifier.

II.BACKGROUND

The architecture of email spam filtering and the stages of email processing were discussed in this section.

2.1. Email spam filtering architecture

Spam filtering aims to keep the number of unsolicited emails to a bare minimum. Email filtering is the method of rearranging emails in compliance with a set of rules. Incoming mail filters are commonly used to handle incoming mail, filter unwanted emails, and detect and prevent spam. Delete any emails that contain malicious code, such as a virus, Trojan, or other malware. Some specific protocols, such as SMTP, have an effect on how email works. Mutt, Elm, Eudora, Microsoft Outlook, Pine, Mozilla Thunderbird, IBM Notes, K mail, and Balsa are some of the most common Mail User Agent (MUAs). They are email clients that help you send and receive emails. The ability to read and write emails by the user Spam filters can be used at any time. Both clients and servers have strategic locations. Many Internet Service providers (ISPs) use spam filters at any layer of the network, including in front of email servers and at mail relays with firewalls. The firewall is a network protection mechanism that controls and handles network traffic based on predetermined security rules. The email server is used to send and receive emails. Filters prevent unsolicited or suspicious emails from reaching the computer system, posing a threat to network security. A consumer may also have a personalized spam filter at the email level, which will block spam emails based on a set of criteria.

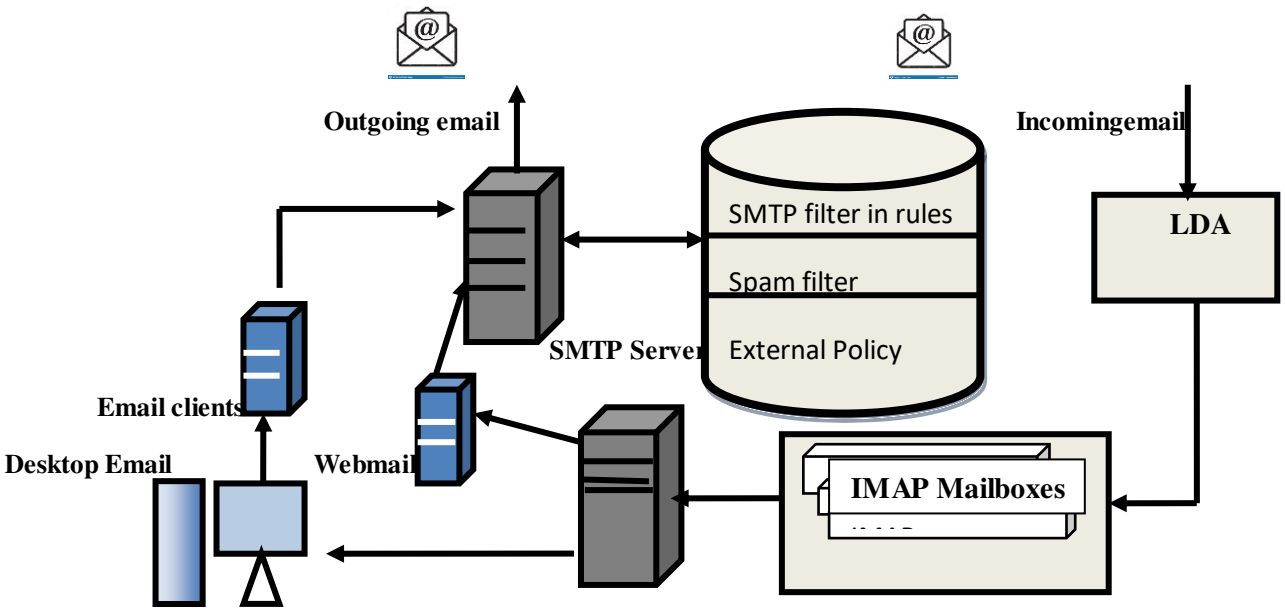


Fig: Architecture of Email spam filtering

2.2. Email Spam Filtering Process

The header and the body are the two most important parts of an email message. The header is the section of the email that contains general information about the content. The subject, sender, and recipient are all included. The email's body is its beating heart. It can include data that hasn't been pre-defined. Web pages, audio, video, analogue data, photographs, files, and HTML mark-up are all examples. The email header includes fields such as the sender's address, the recipient's address, and the timestamp, which display when the message was sent through intermediary servers to the MTAs (Message Transport Agents), which function as an office for the purpose of keeping track of e-mails. The header line normally begins with the word "From," and it is updated as it passes from one server to another via an

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

intermediary server. Headers allow the user to see the content of the page route the email takes, and the time it takes server to process it. Before the classifier can use the available data for filtering, it must first go through some processing. The stages that must be followed when mining data from an email messages can be divided into the following categories:

- **Pre-processing**: When an incoming message is received, this is the first step that is performed.
- **Tokenization:** This is a tool for extracting words from the body of an email. It also breaks down a message into its constituent bits. It divides the email into a set of representative symbols known as tokens. These representative symbols are extracted from the email body, header and topic, according to Subramanian, Jalap, and Taka. Gazelle and Campinas believed that replacing information with distinctive identifying symbols, all of the email's characteristics and words would be lost, with the exception of the context.
- **Feature Selection**: The feature selection stage comes after the pre-processing stage. Feature selection is a form of spatial coverage reduction that effectively exemplifies fascinating email message fragments as a compact feature vector. When the message size is large and a simplified feature representation is needed to make the task of text or image matching easier, this technique is useful. In light of computational complexity and time, recognizing spam e-mails with the fewest possible features is critical. Characteristics Stemming, noise reduction, and stop word removal are all steps in the selection process.

III.PROPOSED METHOD

According to these algorithms I feel the best one is Naive Baye's Classifier algorithm, because this algorithm is extremely fast and can accurately predict the test dataset's class. It can be used to solve multi-class prediction problems because it works well for them. If the assumption of feature independence holds true, the Naive Baye's classifier outperforms other models with less training data. In addition to numerical variables, the Naive Baye's algorithm works remarkably well with categorical input variables.

3.1. Naive Baye's Classifier Method: The naive baye's classifier is a simple but powerful classifier that has been used in a number of data processing applications, including natural language processing, information retrieval, and so on. The naive bayes classifier technique is based on the Bayesian theorem and is particularly useful when the inputs have a high dimensionality. The Bayesian classification method exemplifies both supervised learning and statistical and classification techniques. It acts as a fundamental probabilistic model and let us size ambiguity about the model in an ethical way by influencing the probabilities of the results. It's used to solve problems that are both empirical and predictive. The classification involves functional learning algorithms, as well as the ability to combine prior knowledge and experimental results. Bayesian classification provides a valuable perspective for comprehending and analyzing a number of learning algorithms. It is resilient to noise in input data and computes exact likelihoods for postulation. The influence of a variable value on a given class is assumed to be independent of the values of other variables by Naive Baye's Classifiers. The Naive Bayes inducer calculates the conditional probabilities of the classes given the instance and chooses the one with the highest posterior. In a supervised learning system, naive bayes classifier can be trained very efficiently depending on the precise nature of the probability model.

Autonomous characteristic model is a better expression for the likelihood model. Baye's Theorem: Prob (B given a) = Prob (A and B) / Prob (A). Word probabilities are the key law in this technique for classifying spam emails. If such terms appear often in spam but not in harm, this incoming e-mail is most likely spam. In mail filtering software, the Naive Baye's classifier technique has become very common. To function properly, the Bayesian filter must be trained. Any word in its database has a chance of appearing in spam or ham email. The filter will classify the email into one of two categories if the total number of words probabilities exceeds a certain threshold. Only two categories are needed here: spam and ham. Almost all statistic-based spam filters employ Bayesian probability calculations to aggregate individual token statistics into an average score upon which a filtering decision is based. The most significant statistic for a token T is its Spamminess (spam rating), which is determined as follows:

S[T] = C Spam (T) / C Spam (T) +C Ham (T).

The number of spam or ham messages containing token T is expressed by C Spam (T) and C Ham (T), respectively. To determine the probability of a message M with tokens T1,....., TN, combine the spamminess of the individual tokens to determine the overall message spamminess. Calculating the product of individual token spamminess and contrasting it to the product of individual token hamminess is an easy way to create classifications.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

$$H[M] = \prod_{I=1}^{N} (1 - S[T])$$

If the total spamminess product S [M] is greater than the hamminess product H [M], the message is called spam.

3.2. The Naive Baye's Classification algorithm for Email Spam

Step 1: Training

Parse each email for its individual tokens.

For each token, generate a probability W

S[W] = C Spam (W) / (C Ham (W)) Values of the Spamminess are saved in the database.

Step 2: Filtering

Scanning for the next token Ti for each message M while (M not end)

Check the database for spam S (Ti)

Calculate the total number of message probabilities

S [M] and H [M] are two names for the same guy.

Calculate the overall message filtering indicator by using the following formula:

I [M] = f (S[M]) , H[M])

F is a filter dependent function, such as

I [M] = 1+S[M]-H[M]/2

If I[M]>threshold

Msg is marked as spam

Else

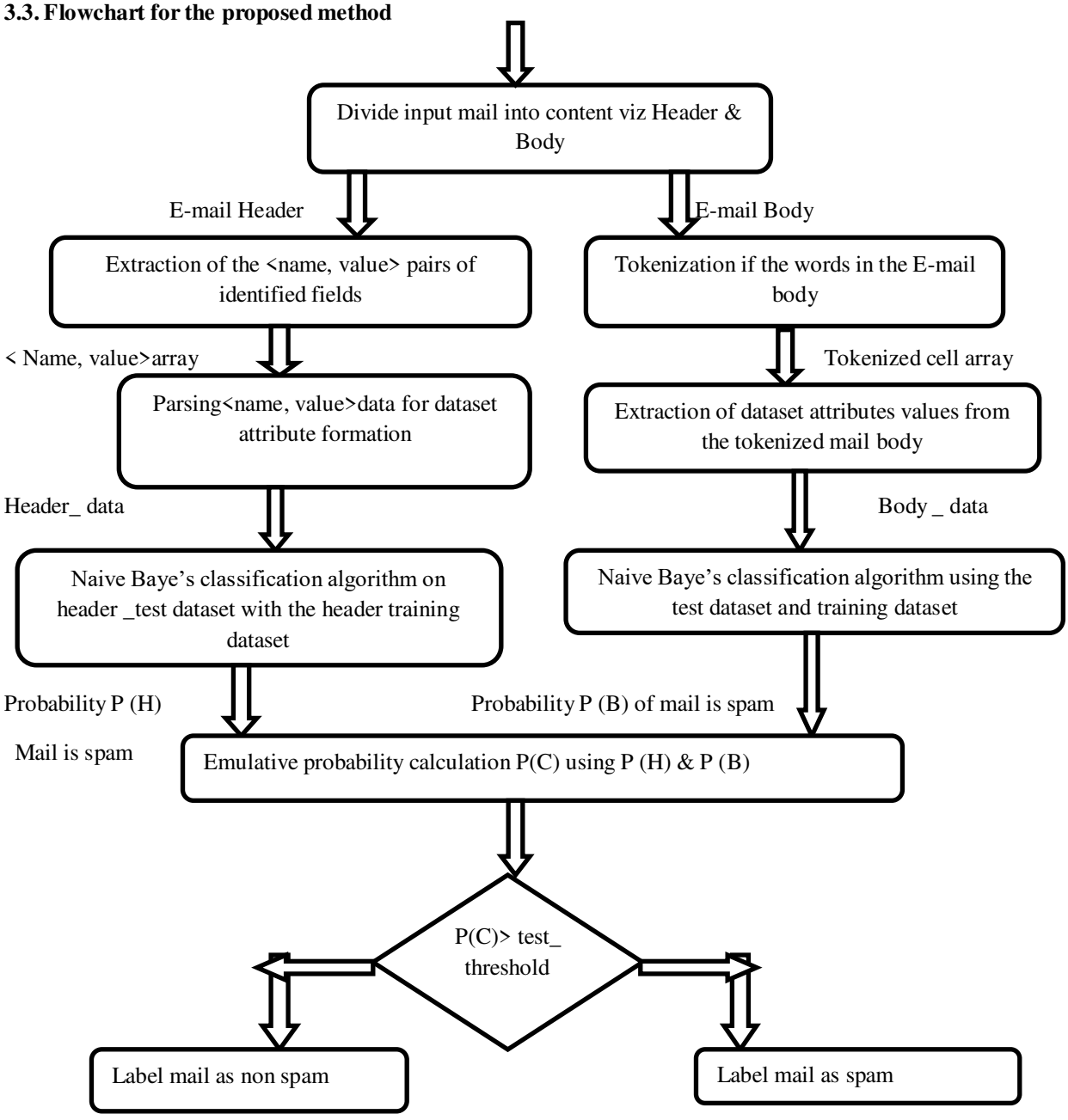
Msg is marked as non-spam

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

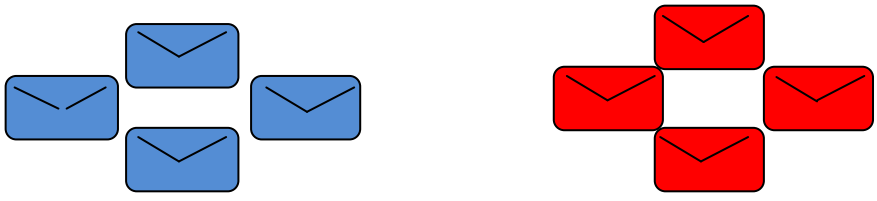
| DOI: 10.15680/IJIRCCE.2021.0905124 |

3.3. Flowchart for the proposed method



IV. EXPERIMENT AND RESULT

Imagine we received normal messages from friends and family and also we received spam (unwanted messages that are usually scams or unsolicited advertisements), we want to filter out the spam messages.



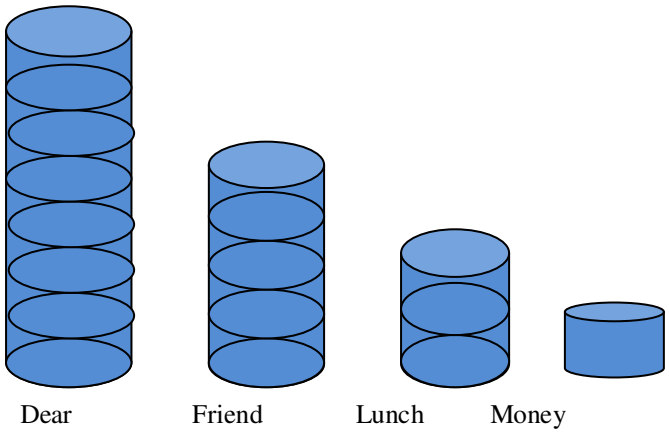
Normal messages spam messages

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

So the first thing we do is make a histogram of all the words that occur in the normal messages from friends and family. We can use the histogram to calculate the probabilities of seeing each word, given that it was in a normal message.



For example, the probability we see the word 'Dear' given .It's a normal message. It is 8 (the total number of times 'Dear' occurred in the normal message) divided by 17 (total number of words in the entire normal message) and that given us 0.47.

P (Dear/Normal)=8/17=0.47

Likewise the probability that we see the word 'Friend 'given .We saw it in a normal message. It is 5, (the total number of times 'Friend' occurred in normal messages) divided by 17 (the total number of word in all the normal messages) and that gives 0.29.

P (Friend/Normal)=5/17=0.29

Likewise we calculate the probability of seeing the remaining words Lunch and Money.

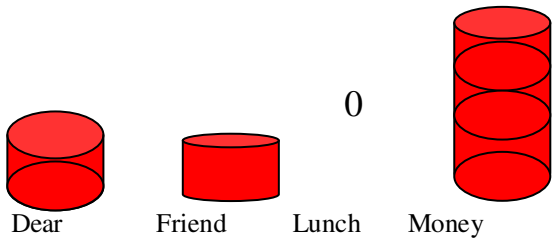
P (Lunch/Normal)=3/17=0.18

P (Money/Normal)=1/17=0.06

And we have to make a histogram for spam emails and also find probabilities

P (Dear/Spam)=2/7=0.29

P (Friend/Spam)=1/7=0.14



P (Lunch/Spam) =0/7=0

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

P (Money/Spam)=4/7=0.57. These histograms are taking up a lot of space. Let's get rid of them, but keep the probabilities. It gives terminology alert. Because we have calculated the probabilities of discrete individual words and not the probability of something continuous like weight or height, these probabilities are also called Likelihoods.

Now imagine we got a new messages, that said 'Dear Friend' we want to decide, if is a normal message or spam. We start with an initial guess about the probability that any message regardless of what it says is a normal message. The guess can be any probability that any message regardless of what it says is a normal message. The guess can be any probability that we want, but a common guess is estimated from the training data.

For example, since 8 of the 12 messages are normal messages, our initial guess will be 0.67

P (N) =8/8+4=0.67

P (N)- The initial guess that we observe a normal message is called a prior probability. Now we multiply that initial guess by the probability that the word 'Dear' occurs in the normal message and the probability that the word 'Friend' occurs in a normal message.

Now we just plug in the values that we worked out earlier and do the math.

P (N)*P (Dear/N)*P (Friend/N)=0.67*0.47*0.29=0.09

We can think of 0.09 as the score that 'Dear Friend' gets if it is a normal message. However, technically it is proportional to the probability that the message is normal given that it says 'Dear Friend'.

0.67*0.47*0.29=0.09 proportional P (N)/Dear Friend

Likewise we have to start with an initial guess about the probability that any message, regardless of what it says is spam. And guess can be any probability that we want, but a common guess is estimated from the training data. Since of 4 of the 12 messages is spam. Our initial guess will be 0.33.

P (S) = 4/4+8=0.33

Now we multiply that initial guess by the probability that the word 'Dear' occurs in spam and the probability that the word 'Friend' occur in spam.

P (S)*P (Dear/S)*P (Friend/S)

0.33*0.29*0.14=0.01

We can think of 0.01 as the score that 'Dear Friend' gets if it is spam. However technically, it is proportional to the probability that the message is spam given that it says Dear Friend.

0.33*0.29*0.14=0.01 proportional P (S/Dear Friend)

The score we got for Normal message, 0.09 is greater than the score we got for spam 0.01. So we can decide that 'Dear Friend' is a Normal message.

V. CONCLUSION

E-mail spam filtering is a critical problem in network security and machine learning techniques. There are many techniques to solve email spam filtering. In this paper we review some of the most popular machine learning methods and of their applicability to the problem of spam e-mail classification. In that, the best method is Naive Baye's Classification, because according to other technique naive baye's needs little training time and speedy assessment to detect and filter email spam. Naive Baye's simply apply Baye's theorem on the context classification of each email with a strong assumption that the words included in the email are independent of each other. Naive Baye's is desirable of email spam filtering because of it's simplicity, ease of implementation and quick convergence compared to conditional models. It needs fewer training data, it is very scalable, no bottleneck is created by increase in the number of predictors and discrete unit of information, It used to solve both classification problems involving two or more

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905124 |

classes. And also used to make forecasting that is subject to or involving probability variation. They can effectively manage continuous and discrete data.

REFERENCES

1). Christina, V., Karpagavalli, S., & Suganya, G. (2010). Email spam filtering using supervised machine learning techniques. International Journal on Computer Science and Engineering (IJCSE), 2(09), 3126-3129.
2). Rusland, N. F., Wahid, N., Kasim, S., & Hafit, H. (2017, August). Analysis of Naïve Bayes algorithm for email spam filtering across multiple datasets. In IOP conference series: materials science and engineering (Vol. 226, No. 1, p. 012091). IOP Publishing.
3). Bhuiyan, H., Ashiquzzaman, A., Juthi, T. I., Biswas, S., & Ara, J. (2018). A survey of existing e-mail spam filtering methods considering machine learning techniques. Global Journal of Computer Science and Technology.
4). Navaney, P., Dubey, G., & Rana, A. (2018, January). SMS spam filtering using supervised machine learning algorithms. In 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 43-48). IEEE.
5). Tretyakov, K. (2004, May). Machine learning techniques in spam filtering. In Data Mining Problem-oriented Seminar, MTAT (Vol. 3, No. 177, pp. 60-79). Citeseer.
6). https://www.youtube.com/watch?v=O2L2Uv9pdDA
7)https://www.youtube.com/watch?v=jS1CKhALUBQ
8)). https://www.youtube.com/watch?v=RixQygYyDKI

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details