International Journal of Innovative Research in Engineering & Management (IJIREM)
ISSN: 2350-0557, Volume-9, Issue-2, April 2022
https://doi.org/10.55524/ijirem.2022.9.2.103
Article ID IJIRP-2302, Pages 648-652
www.ijirem.org

An Evaluation on the efficiency of E-Mail Spam Detection Using Naive Bayes Classifier

Gaddam Chakradhar Reddy[1], Ramanadham Rohith Kumar[2], Pikkili Siva Kasi[3], Navuluri Sarath Chandra[4], R. Pavan Kumar[5], and P. Prabakaran[6]

[1,2,3,4,5,6] Department of Computer Science and Engineering, PACE Institute of Technology & Sciences, Vallur, Ongole, Andhra Pradesh, India

Correspondence should be addressed to Gaddam Chakradhar Reddy; 19kq5a0509@pace.ac.in

ABSTRACT- Nowadays, electronic mail is ubiquitous, being used everywhere from the business sector to the classroom. Emails can be broken down into two distinct categories: ham and spam. Email spam, also known as junk email or unwanted email, is a form of email that can be used to harm any user by wasting his or her time, draining system resources, and stealing sensitive data. Every day, the proportion of spam emails increases dramatically. Today's email and IoT service providers face a large and formidable task in detecting and filtering spam. One of the most prominent and widely-known approaches of detecting and avoiding spam is email filtration. It's also one of the most discussed tactics out there. Several machine learning and deep learning techniques, including Naive Bayes, decision trees, neural networks, and random forests, have been employed to reach this objective. After completing a survey of the available machine learning approaches, this article groups them into the most acceptable categories for usage in spam screening on email and IOT platforms. Accuracy, precision, memory requirements, and other metrics are used to thoroughly assess the methodologies. In the final section, we examine both the overall takeaways and directions future studies could go in.

KEYWORDS- E-Mail Spam Detection, Naive Bayes Classifier, Spam Filtering, Spam Filtering

I. INTRODUCTION

In current age of information technology, information sharing is rapid and easy owing to the internet. Users in any part of the world can use any of the available platforms to share and receive information. Sending an email is the quickest, cheapest, and most convenient way to communicate with someone on the other side of the world. When compared to other channels, its reach is unparalleled. However, due to the simplicity of emails, they are vulnerable to a wide range of attacks, the most common of which is spam [1]. Emails that have nothing to do with the interests of the recipients are annoying and annoying because they waste the time and effort of the receivers. Furthermore, these emails may contain malicious content within attachments or URLs, which can compromise the host system's security [2]. Spam [3] occurs when an offender sends a message or email to many people that they know will not want to read it. Spam can be distributed by any medium that allows for the exchange of data, not just email. This means that ensuring the security of the email system must be a top priority. Viruses, worms, and Trojan horses are all forms of malware that could be concealed in spam emails. In most cases, this is how attackers would get victims to visit malicious websites or use compromised accounts. The multiple-file-extension-carrying files and densely-packed links to malicious and advertising websites in spam emails are just the beginning of the potential damage [4, 5]. Numerous email providers enable their users to set up rules based on keywords that filter incoming messages. Scammers target users' inboxes because it's easy for them to gain access without doing any work, and because few users bother to customize their emails.

The Internet of Things (IoT) has grown in prominence and permeated modern culture over the past couple decades. The Internet of Things is becoming increasingly important in creating "smart" neighborhoods. Many popular social media sites and useful mobile apps are based on IoT infrastructure. The explosion in the number of connected devices is directly contributing to a rise in malware. Several methods for detecting spam were presented by the researchers as possible answers to the issues of identifying and blocking spam and offenders. Existing spam detection systems are generally divided into two broad categories: behavioral pattern-based approaches and conceptual pattern-based approaches. Each of these approaches has advantages and disadvantages. There has been a huge increase in the volume of spam transmitted by email in tandem with the growth of the Internet and worldwide communication [6]. Spam can be sent from any computer, anywhere in the world, and the sender's identity will remain hidden thanks to the anonymizing power of the Internet. There is a lot of anti-spam software and approaches, yet the spam rate is still very high. Malicious emails that link to other malicious websites that can harm the victim's data are the most serious form of spam. In addition to reducing response times, spam emails can overload servers and force them to shut down before their time is up. In order to properly identify spam emails and avoid the growing number of difficulties created by email spam, each

organization should do a thorough study of the anti-spam solutions at their disposal. Emails that have been received can be evaluated using a variety of well-known processes, including as whitelist/blacklist [7], mail header analysis, keyword verification, and similar approaches.

Experts in the field of social networking have concluded that roughly 40% of all profiles on social networking sites are used solely for the purpose of sending spam [8]. Popular social networking tools are being used by marketers to send camouflaged links in the text to pornographic or other merchandise sites with the intent of selling something from fraudulent accounts. Malicious emails targeting the same group of recipients often have similar layouts and subject lines. More research into these high points can help improve the detection of such communications. Through the use of AI, we can distinguish between spam and non-spam email [9]. The initial step in putting this strategy into action is to execute feature extraction on the communication's headings, subjects, and bodies. Having extracted this data based on their features, we can now divide them into two distinct categories: spam and ham. Spam detection in today's online contexts typically makes use of learning-based classifiers [10]. The identification procedure in learning-based classification is based on the hypothesis that spam emails have a distinct set of features that distinguish them from legitimate emails [11]. The inclusion of numerous variables complicates the process of spam detection in learning-based models. Spam's subjective nature, the spread of ideas, language barriers, processing load, and lag time in text messages are just a few examples.

II. COMPARISON WITH PREVIOUS SURVEYS

Messages that are either untrue or unwanted and are sent in big quantities through any account or an autonomous system are examples of email spam. Spam emails can be sent from any account. Over the course of the past ten years, the rate at which emails that are regarded as unsolicited commercial correspondence, also known as spam, have continued to proliferate at an alarmingly high rate. Spambots, which are computer programs that cruise the internet in search of email addresses, are typically responsible for compiling the lists of email addresses that are targeted by fraudulent email messages. Several different applications of machine learning, each of which plays an important part, have been of tremendous assistance in the identification of fraudulent emails. For the purpose of developing novel spam detection and filtration models, researchers are drawing from its vast repertoire of models and methods [13]. Kaur and Verma [14] give a report on the identification of email spam with that strategy. They do so by employing a supervised method and feature selection in their research. They discuss the steps involved in the knowledge discovery process for spam detection algorithms. In addition to this, they devise a great deal of tactics and tools that are proposed for the detection of spam. Another topic that is covered in this study is the selection of features in accordance with the N-Gram. An N-Gram [15, 16] is a predictive-based technique that can be used to forecast the chance of the next word occurrence in a sentence or text corpus after identifying N minus 1 words in the sentence or text corpus. This may be done after recognizing N words in a sentence or text corpus. N-Gram uses a variety of techniques that are based on probability in order to guess what the following word will be. In this study, a number of techniques for identifying spam email are contrasted and compared. These techniques include machine learning (multilayer perceptron neural network, support vector machine, and Naive Bayes), as well as non-machine learning (Signatures, Blacklist and Whitelist, and analyzing message headers).

A report on sophisticated scam email detection is presented by Saleh and colleagues [7]. They discuss the many security risks that are linked with emails, most notably spam emails, as well as the scope of spam analysis and several techniques, some of which are based on machine learning and others of which are not, for detecting and filtering spam. They arrived at the conclusion that supervised learning [8] strategies are used extensively all around the world in order to recognize spam in email. They say that the extensive application of supervised learning is due to the precision and consistency of supervised methods, which are the grounds for the broad use of supervised learning. In addition to this, they discussed multialgorithm frameworks, and after doing so, they arrived at the conclusion that multialgorithm frameworks are superior to a single algorithm in terms of effectiveness. They arrived at the conclusion that practically all of the research that makes use of the content of emails for the identification of spam, and more precisely fraudulent emails, depends on word-based categorization or clustering algorithms. This was the main finding that led them to this conclusion.

A discussion of the methods for learning-based email spam filtration can be found in Blanzieri and Bryl's [8] article. They presented an overview of learning-based techniques to spam filtration and explored the challenges that are created by spam in this article. They discuss the various qualities of unsolicited electronic mail, also known as "spam." The purpose of this study was to investigate the effects that phishing emails have had on a number of different email accounts. This research also examines a wide range of ethical and professional issues that are connected to spam. Both the conventional antispam approach and the learning-based filtration system have progressed to the stage of development known as "mature." The majority of the widely used filters are based on a range of categorization strategies that are applied to various characteristics of email messages. These strategies can be implemented in a variety of different ways. The outcomes of this study indicate that among the many various kinds of learning algorithms that are used for spam filtration, the Naive Bayes classifier holds a special and distinct position in the industry. It accomplishes its goals in a lightning-fast and uncomplicated manner, and the results it generates are extremely accurate.

Bhuiyan et al. [10] give an overview of the many different strategies that are being employed to filter out junk email. They do this in order to present a summary of a number of different spam filtration approaches as well as a summary of the precision on various aspects of various suggested systems. This is accomplished by analyzing a huge number of processes. They address the

fact that all of the strategies that are now available for screening trash emails are efficient to some degree. While some individuals have been successful in accomplishing their objectives, others are exploring additional methods through which they can raise the level of precision in their performances. In spite of the fact that they are all successful, there are still certain issues with the methods of spam filtration, which is the primary concern of researchers. They are focusing on building a spam filtration mechanism for the next generation that is capable of comprehending a huge number of different types of audiovisual data and removing junk emails from the inbox. They arrived at the conclusion that the SVM algorithm and the Naive Bayes method are used in the filtering of the vast majority of spam emails. It is feasible to train these models using a number of datasets, such as "ECML" and the UCI dataset [10], which can be done in order to validate the efficacy of the spam filtration models that have been developed.

Ferrag et al. [3] presented a review of the deep learning approaches used in intrusion detection systems and spam detection databases. They discussed the various detection strategies that are based on the different deep learning-based detection models and assessed the effectiveness of a variety of deep learning-based detection models. Before doing so, they analyzed 35 well-known internet datasets and classified each of those datasets according to one of seven categories. These categories include datasets that are based on Internet traffic, network traffic, Interanet traffic, electrical network traffic, virtual private network traffic, Android app traffic, and Internet of Things traffic. Additionally, these classifications contain datasets that are based on devices that are linked to the internet. When it comes to the identification of spam and incursions, they came to the realization that deep learning models have the ability to beat more standard machine learning and vocabulary models.

In their paper [2,] Vyas et al. offer an overview of supervised machine learning approaches that can be used to screen junk email. They came to the conclusion that out of all the other approaches that were investigated, the Naive Bayes method gives speedier results and better precision than any other method, with the exception of the SVM and ID3 methods. This was the conclusion that they arrived to. Although SVM and ID3 produce results that are more accurate than those produced by Naive Bayes, the construction of a system using these two methods takes a considerable amount more time. Timeliness and accuracy are two competing priorities that require careful management in order to avoid conflict. They arrived at the conclusion that the conditions, as well as the required level of precision and amount of time, play a significant role in determining the learning algorithm that should be utilized. They argue that in the future, in order to establish a more thorough structure for filtering junk email, all components of the email should be taken into consideration, and this should occur in the future.

III. SPAM MESSAGES

Due to the fact that everyone has a different perspective on it, the description of email spam is murky. At this time, everyone's attention is being drawn to the problem of spam in email. Email spam typically consists of one-off, spur-of-the-moment communications that are distributed en masse by people you are not familiar with. The word "spam" originates from a comedy performed by Monty Python [3], in which an item of tinned beef manufactured by Hormel is given numerous tiresome highlights. Although the word "spam" was said to have been used for the first time in 1978 to refer to unsolicited email, its usage became much more widespread in the middle of the 1990s as the Internet became more widely used outside of academic and research communities [4]. One prominent example is the "development expense trick," in which a customer is sent an email containing a purported offer that will result in the awarding of a reward. In this day and age, the dodger or spammer demonstrates a story in which the unfortunate victim requires direct financial assistance so that the con artist can gain a significantly larger sum of money, which they would then split. When the unfortunate victim finally pays off the debt, the con artist will either make a profit or simply stop responding to his victim's messages.

The Standard Spam Filtering Method

Standard spam filtering refers to a filtering system that employs a collection of rules and works with those rules in order to function as a classification. This type of filtering system is known as a "filtering system." The conventional approach to the elimination of spam is depicted in Figure 1. In the first stage, content filters are put into place, and in order to identify malware, these filters make use of various forms of artificial intelligence. The second stage involves the implementation of the email header filter, which takes information from the email's heading and displays it. After that, backlist filters are applied to the emails in order to eliminate any scam emails that may have originated from the backlist file. Following this step, rule-based filters are put into place, which identify the originator by making use of the subject line and the specifications that the user has specified. In the end, allotment and assignment filters are utilized by putting into place a method that enables the account proprietor to send the email themselves.

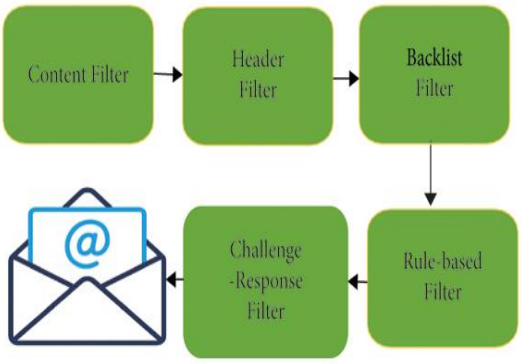

Figure 1: Enterprise Level Spam Filtering

The identification of spam emails at the enterprise level makes use of a methodology that calls for the installation of a number of different filters frameworks on the server. These frameworks interact with the mail transmission agent in order to classify the gathered emails as either spam or ham, depending on the content of the message.

The user of this system makes use of it in a dependable and effective manner on a network that filters the emails using an enterprise filtering approach. This customer of this system uses the network to access the system. The currently available techniques for identifying junk email follow the principle of sorting the emails in chronological order. This principle requires the specification of a scoring system, which ultimately results in the production of a number for each entry. A certain number or rating is given to each kind of communication, regardless of whether it is spam or ham. It is required to adjust all obligations on a continuous basis by adopting a list-based technique to automatically reject communications. This is necessary since con artists use a range of strategies in their scams.

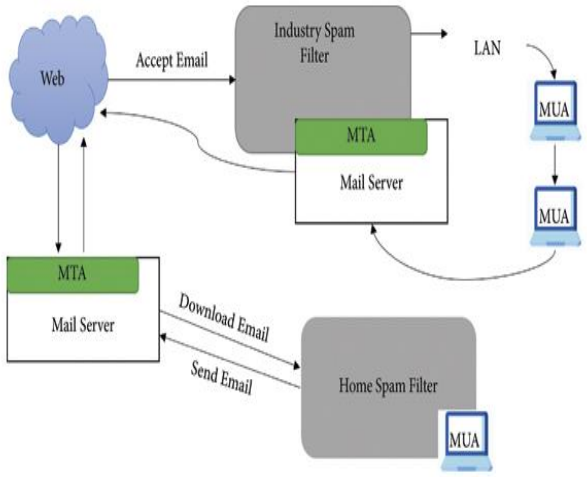

Figure 2: Case-Based Spam Filtering

IV. FILTERING METHODS

Large amounts of data can be processed more easily thanks to machine learning. It can take more time and money to train its models for a high level of performance, even though it normally gives quicker and more accurate results to detect harmful content.

Handling enormous amounts of data can be even more effective when machine learning is combined with AI and cognitive computing [7].

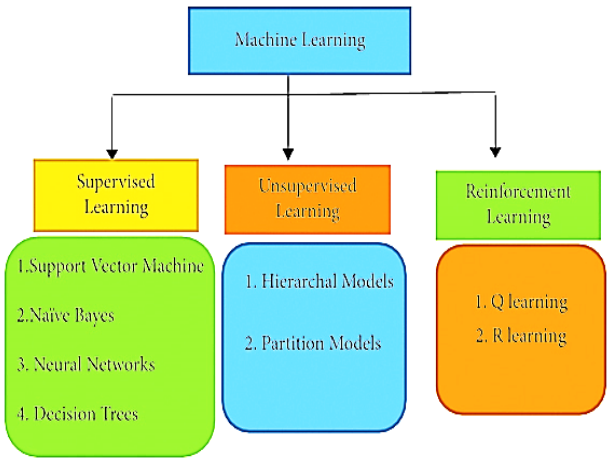

Figure 3: Naïve Bayes Classifier (NB)

The Bayes theorem serves as the foundation for the Naive Bayes classification. It makes the assumption that the indicators are unrelated to one another, which indicates that knowing the value of one attribute has an effect on the value of any other attribute. Classifiers based on the Naive Bayes algorithm are simple to develop because they do not involve any repetitive processes. Furthermore, they are capable of working with very large databases in an effective and accurate manner. Present research on email spam filtration and conduct the analysis using a machine learning technique called Naive Bayes. Naive Bayes is known to have often outperformed other categorization methods in different situations, despite its simplicity. They used two different datasets that were analyzed based on the importance of accuracy, F-measure, precision, and memory. As is common knowledge, the categorization method known as Naive Bayes makes use of probability, and tallying the frequency and combination of values in a collection constitutes the probability. This study filters emails in three stages: first, it preprocesses the messages, then it selects the features to be used, and finally, it uses the Naive Bayes classifier to put those features into action. During the preprocessing phase, the entirety of the email content is purged of any and all conjunction words, articles, and stop words. After that, they created two databases with the help of the WEKA tool, which they referred to as the spam data dataset and the spam base dataset. Utilizing two datasets resulted in an accuracy of 89.59% on average, whereas the trash data achieved 91.13% accuracy. The precision of the spam basis collection was determined to be 82.54%. The precision of the findings for spam data came in at an average of 83%, while the precision of the results for spam base was 88%. They asserted that the Naive Bayes classifier performs better on spam base data when compared with spam data when it comes to machine learning-based spam identification strategies for Internet of Things devices. They employed five different machine learning algorithms and examined the outcomes with a variety of performance measures. During the training process for the suggested models, a large number of input characteristics were utilized. Each model arrives at a spam number by applying the input characteristics to a calculation. This value is a representation of how trustworthy an Internet of Things gadget is based on a number of different variables. Using the REFIT smart house dataset, the validity of the recommended strategy is demonstrated. They assert that their suggested system can identify spam more effectively than the systems that are presently being used to detect spam. The results of their study could be applied in "smart homes" and other settings that make use of sophisticated technology.

Detection of malware in emails using a variety of machine learning techniques Their article investigates ML techniques and how to apply them to datasets in various ways. From among the many ML algorithms, the optimum algorithm for email spam detection is determined to be the one that possesses the greatest possible precision and accuracy. They came to the conclusion that the Multinomial Naive Bayes algorithm yields the best results; however, it does have some restrictions because of its dependence on class-conditional independence, which causes the computer to

incorrectly categorize some inputs. Following Multinomial Naive Bayes in terms of producing the most accurate and trustworthy findings, Ensemble models come in second. Only junk that is contained within the bodies of emails can be identified by the suggested method in this research.

methodology based on semisupervised machine learning for the identification of abuse in social IoT platforms. They used a structure that is built on an ensemble and comprises of four different classifications. The architecture is built on the utilization of probabilistic data structures (PDS) such as Locality Sensitive Hashing (LSH) for similarity search and Quotient Filter (QF) to examine the database of URLs, spam users, and databases of spam keywords. The suggested model seeks to optimize, and as a result, it makes decisions using a flexible weighted voting strategy that is based on the outcome of each classification. The computational efforts required to sample the data in accordance with each classification are kept to a minimum by using the hybrid sampling technique. According to the findings of this research, the model that was suggested can be utilized for the identification of spam on massive databases. The suggested model's usefulness was assessed by contrasting its performance with that of conventional data models and the conventional assessment measures, which included accuracy, recall, and F-score.

V. CHALLENGES OF SPAM DETECTION

- The ever-increasing data accessible on the Internet, along with its many novel features, presents a major challenge for spam identification algorithms.
- It can be challenging for spam algorithms to evaluate characteristics based on a variety of criteria, including but not limited to time, writing style, grammatical structure, and statistical significance.
- Most models are taught on fair datasets, but developing self-learning models is currently impossible.
- Different spam detection models can be undermined by adversarial machine learning attacks, which is a problem because it lowers the quality of spam detection. It's possible for enemies to use a wide range of attacks against machine learning models while they're being trained and tested. An attacker can introduce contamination into training data, leading to incorrect classifications by a classifier, create adverse instances during testing, and obtain private training data via a learning model, all without being detected. (privacy attack)

VI. CONCLUSION

Over the past two decades, researchers have focused on spam identification and avoidance. This field is heavily researched due to its possible effects on customer behavior and incorrect evaluations. This study reviews machine learning methods for detecting and removing advertising from IoT networks and interactions. Governed, uncontrolled, reinforcement learning, etc. This study compares various methodologies and summarizes the main findings from each group. This research found that most email and IoT malware detection systems use directed machine learning. Annotating a dataset is time-consuming but necessary to train a supervised model. Two supervised learning methods that detect garbage are Naive Bayes and SVMs.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] J. Dean, “Large scale deep learning,” in Proceedings of the Keynote GPU Technical Conference, San Jose, CA, USA, 2015.

[2] E. Blanzieri and A. Bryl, E-mail Spam Filtering with Local SVM Classifiers, University of Trento, Trento, Italy, 2008.

[3] J. K. Kruschke and T. M. Liddell, “Bayesian data analysis for newcomers,” Psychonomic Bulletin & Review, vol. 25, no. 1, pp. 155–177, 2018.

[4] K. S. Adewole, N. B. Anuar, A. Kamsin, K. D. Varathan, and S. A. Razak, “Malicious accounts: dark of the social networks,” Journal of Network and Computer Applications, vol. 79, pp. 41–67, 2017.

[5] D. Lee, M. J. Lee, and B. J. Kim, “Deviation-based spam-filtering method via stochastic approach,” EPL (Europhysics Letters), vol. 121, no. 6, Article ID 68004, 2018.

[6] PrasaduPeddi (2019), Data Pull out and facts unearthing in biological Databases, International Journal of Techno-Engineering, Vol. 11, issue 1, pp: 25-32

[7] S. E. Kille, Mapping Between X. 400 and RFC 822, University College Department of Computer Science, London, UK, 1986.

[8] A. Makkar and N. Kumar, “An efficient deep learning-based scheme for web spam detection in IoT environment,” Future Generation Computer Systems, vol. 108, pp. 467–487, 2020.

[9] W.-F. Hsiao and T.-M. Chang, “An incremental cluster-based approach to spam filtering,” Expert Systems with Applications, vol. 34, no. 3, pp. 1599–1608, 2008.

[10] P. Lison, An Introduction to Machine Learning, Language Technology Group, Edinburgh, UK, 2015.

[11] G.-H. Lai, C.-M. Chen, and C.-S. LaihT. Chen, “A collaborative anti-spam system,” Expert Systems with Applications, vol. 36, no. 3, pp. 6645–6653, 2009.