JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A SURVEY OF EMAIL SPAM DETECTION USING ARTIFICIAL INTELLIGENCE

Dr. P.Selvarani [1],A. Kanagalashmi[2], M. Keerthika[3], M. Kathambari[4], Mancy Arokiya Mary[5].
Associate Professor [1] , UG Scholar [2,3,4,5]
Department of Computer Science and Engineering
Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College

Abstract

Email spam has grown significantly in recent years along with the rapid growth of internet users. People are utilizing them in unlawful and unethical ways. Fraud, phishing, and conducts. Spam is waste of time to the user since they have to sort the unwanted junk mail and it consumes storage space and communication bandwidth. They can easily send their shady message to so many email addresses in a single stroke. In our Proposed work using machine learning techniques. This paper will discuss machine learning algorithms and apply all of them to our data. The task of classifying an incoming email as either spam or not spam can be described as an AI-based binary classification issue. Accuracy, precision, recall, TPR – True Positive Rate, TNR – True Negative Rate, FPR – False Positive rate, FNR - False Negative Rate are a few metrics that can be used to assess the effectiveness of a spam detection system.

Keywords: Email Spam Detection, Machine Learning Algorithms, Security, TPR, FPR, TNR, FNR.

I INTRODUCTION

Spam detection is a technique [1] [9] to filter all the mail we receive and find out which mail is spam and which is genuine. The processing of email to organize it in accordance with predetermined criteria is known as email filtering. The word could refer to a human intelligence's intervention, although it most frequently describes the automatic message processing at an Email server, sometimes involving anti-spam measures. Both incoming and outgoing emails are subject to filtering. Since the user receives emails that have passed through different spam checks, the security risk can be decreased thanks to the advantages of email spam filters. Additionally, these email spam filters eliminate emails that include viruses, malware,[12] and other harmful content and safeguard user security. Since the last ten years, spam emails have become more and more common. Spam has grown to be a significant online problem. Spam is wasteful. Although automatic email service [8] filtering may be the best way to stop spam, modern spammers may quickly get around all of these apps. Prior to a few years ago, the majority of spam that came from particular email addresses could be manually stopped.
For spam detection, [14] a machine learning approach will be utilized major techniques used to filter junk mail include "text analysis, white and blacklists of domain names, and community-based techniques.

II RELATED WORKS

In our work collecting the datas from kaggle dataset for email spoofing [2] in comma separated value {CSV} format. which includes 5000 spam and ham emails. Additionally, we are gathering this data forms [3] targeted literature review of artificial intelligence (AI) in four segments according to the structure of emails that can be

used for intelligent analysis. which also includes (MTA) that offers details such as email and the IP address of each sender and receiver of where the email originated, (SMTP) envelope that contains mail exchanges, from, to, date, subject (which is shown in the majority of email clients), email content, and attachment.

III PROBLEM STATEMENT

Since the user must go through the undesired mail and it uses up storage space and communication bandwidth, spam is a time waster for the user. The issue with email is that spammers may easily defraud their way to a large payday even though they only expect a tiny number of recipients to interact or respond to their message because they can send their dubious message to numerous email accounts at once. Because of this, spam is still a significant issue in the contemporary digital economy. Additionally, identifying email spam is a binary classification issue. The rationale behind this is straightforward: by identifying undesirable and unauthorised emails, we can stop spam from infiltrating users' inboxes and consequently enhance user experience.

IV EXISTING SYSTEM

Before a spammer's email address, IP address, or domain gets banned, thousands of clustering spam emails [13] may make it to Inboxes. Because spam filtering [15] is automated, there is a chance for errors known as "false positives." Spammers may trick Bayesian filters by employing big [1]blocks of legitimate content.

Table 1 represents some of the Existing Techniques and Methods, and Table 2 represents Approaches that are used in Email Spam Detection is given below.

Table 1 : Existing Techniques and Methods used in Email Spam Detection.

Existing Techniques	Existing Methods
• Syntactial • ELM - Extreme Learning Machine • SLFN - Single Hidden Feed Forward Neural Networks • SVM - Support Vector Machine • BRR- Burst Review Ratio	• Texical • Amalgam • LIWC(core logic of Linguistic Inquiry and Word Count) and POS(Point of Sale) • Stylometric • N-grams model • RAVP - Ratio of Amazon Verified Purchase • RD - Rating Deviation • RCS - Review Content Similarity • RB - Reviewer Burstiness

Table 2 : Existing Approaches used in Email Spam Detection

SN	Year	Author Name	Algorithms	Data Set	Performance Metrics	Evaluation Parameters
1	2020	N.Kumar et.al	• **SVM** – Support Vector Machine [7] • **KNN** - K-Neareast Neighbour [5]	**5573 Emails Dataset**	Max accuracy achieved is **98%**	• Accuracy

			• **RF** – Random Forest			
2	2019	S.Suryawanshi et.al	• **SVM** - Support Vector Machine • **KNN** - K-Neareast Neighbour	**5674 Labelled Dataset**	Max accuracy achieved is **97.5%** using SVM	• Accuracy, • Precision, • Recall, • F-measure
3	2018	K.Iyyengar et.al	• **INB** - Integrated Navie Bayes [6] • **PSO** - Particle Swarm Optimization [10]	**Spam Base Emails Dataset**	Max Accuracy achieved is **95.5%** using INB	• Accuracy, • Precision, • Recall, • F-measure
4	2018	P.Sharma et.al	• Hybrid bagging	**1000 Spam Base Emails**	HB approach has obtained max accuracy **98.3%**	• Accuracy, • Precision, • Recall, • F-measure
5	2016	S.K.Tuteja et.al	• K- means Clustering	**200 Spam-Based Emails**	They obtained **98.42%** accuracy.	• Accuracy, • ROC-AUC
6	2016	Roman Urdu tweets	• DMNB - Discriminative Multinomial Naive Bayes	**1463 Roman Urdu Tweets**	Obtained Max accuracy **95.42%** using DMNB.	• Accuracy, • ROC-AUC
7	2015 April	M.Mohamad & A.Selmat	• SVM - Support vector machine [7] • NB - Naive Bayes	**English emails**	Max accuracy achieved is **86.40%**	• Accuracy, • Recall

V PROPOSED SYSTEM:

Figure 1 Represents Proposed Email Spam Detection Architecture. It Consists Data Gathering , Data Preprocessing, Feature Extraction , Training the Model Data Using AIML Technique, Testing the data and Finally Detecting the Email as SPAM & HAM.

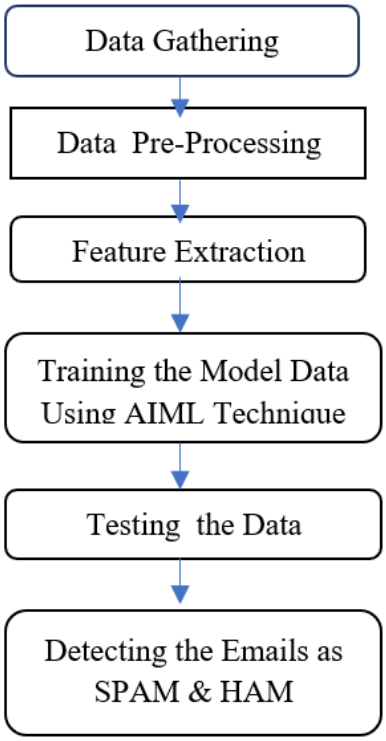


Figure 1 – Proposed Email SPAM Architectur

DATA GATHERING:

In order to find patterns and traits that are typical of spam communications, data collection in email spam detection refers to the process of obtaining and analyzing data connected to email messages. The sender's email address, the message's content, the subject line, and any links or attachments may all be included in this data.

Machine learning algorithms are frequently trained on enormous datasets of email messages, both spam and non-spam, to understand the characteristics of each type of message in order to successfully detect spam emails. Since the data collection process enables the algorithms to more accurately differentiate between spam and legitimate emails and adapt to new spamming strategies, it is crucial for the creation of accurate and successful spam detection systems.

Table 3: Email SPAM Detection Data Set Details

LABEL	TEXT	LABEL NUMBER
Ham - 71%	4993	It is spam
Spam - 29%	Unique values	It is 1,else it us 0
100 %	Binary Values	0 or 1

Table 3 represents Email SPAM Detection Data Set. It referes column labeled as “type” having possible values spam and ham… is used to classify the emails. The 2nd column is contains text of emailed date content in that are it was delivered up to 80% of the emails will be used to train and models and 20% of the email will be used to test models contain 4993 unique values is there.

DATA PRE - PROCESSING:

The alteration and analysis of the data acquired during the data collection phase are referred to as "data processing" in the context of email spam detection. In most cases, to accurately categorise emails as spam or valid, important features are first extracted from the data using algorithms and procedures, and then these features are used to train machine learning models.
Among the common methods for detecting email spam used in data processing are:

1**. Preprocessing**: This entails activities like cleaning and normalising the data, eliminating stop words, and stemming or lemmatizing words to decrease the dimensionality of the data and enhance the precision of the models.
2. **Identifying and extracting pertinent elements** from the data, like the sender's email address, the subject line,and the message's content. Then, machine learning models can be trained using these attributes.

3. **Model training**: In this step, machine learning models like decision trees, random forests, support vector machines [7] , or neural networks [4] are trained using the retrieved features. Usually, a sizable dataset of emails that have been classified as spam or valid is used to train the models.

4. **Model evaluation**: In this step, labelled emails from a different dataset that wasn't utilised during training are used to test the trained models' accuracy. This makes sure that the models are correctly identifying emails as real or spam.

Overall, data processing is an important phase in email spam detection since it enables the creation of precise

FEATURE EXTRACTION:

The process of finding and extracting pertinent data from email messages that can be utilised to differentiate spam from valid emails is known as feature extraction in the context of email spam detection. Typical characteristics that are frequently applied in email spam identification include:

1. **Sender details**: These contain the sender's name, domain, and email address. While genuine emails are more likely to come from well-known and reliable sources, spam emails frequently originate from dubious or unidentified email addresses or domains.
2. **Email subject line**: Whether an email is spam or not can often be determined by its subject line. Spam emails frequently include eye-catching or deceptive subject lines to get the recipient to open the message.
3. **Text of the message**: The email's content is one of the most crucial components included in spam detection. Certain words or phrases that are frequently linked with spam, such "urgent," "free," or "click here," are frequently used in spam emails.
4. **Attachments and links:** Spam emails frequently include files that might infect a recipient's machine with malware or links to questionable websites. An email may be spam if it contains questionable attachments or links, which can be easily determined.

The timestamp of the email, the email client used to send the message, and the sender's IP address are examples of metadata. Using this data, it may be possible to spot trends or irregularities that frequently appear in spam emails.In conclusion, feature extraction is an essential stage in email spam detection.

VI TRAINING THE MODEL DATA USING AIML TECHNIQUE:

The following steps are commonly involved in training a model for email spam detection using artificial intelligence and machine learning (AIML):

1. **Data gathering and preparation**: A large and varied dataset of both spam and legitimate emails is required in order to train a model. The dataset has to be labelled so that the model can distinguish between authentic and spam emails. Once the dataset has been gathered, it might need to be cleaned up and preprocessed to get rid of extraneous or redundant data.

2. **Feature extraction:** After the data has been gathered and prepared, features that will be used to train the model must be extracted from each email. This could include information like the sender's email address, the message's content, and the subject line, body of the message, and any files or links that were attached.

3**. Model choice**: Different machine learning methods and models, such as decision trees, random forests, support vector machines, or neural networks, can be used to identify email spam. The individual project needs and the dataset's features will determine which model is used.

4**. Model training**: Using the labelled dataset of characteristics culled from the emails, the selected model is subsequently trained. The objective is to accurately train the model to distinguish new, unseen emails as spam or legitimate.

5. **Model evaluation**: After the model has been trained, it needs to be assessed to see how well it performs and how accurate it is. To do this, test the model on.

VII TESTING THE DATA:

In a spam detection project, testing the data entails assessing how well the trained model performs on a different dataset of emails that weren't utilised during the training stage. This is done to assess the model's generalizability to fresh, untested data as well as to gauge its degree of accuracy when applied to actual data.

The following steps are often included in the testing phase:

1. **Data preparation**: Preprocessing and feature extraction for the testing dataset should be done in the same manner as for the training dataset.

2. **Model prediction**: The trained model is then applied to forecast the emails in the testing dataset's class labels (legitimate or spam).

3. **Performance assessment:** To assess the model's performance, the predicted and actual class labels are comparedto the actual class labels in order to assess the performance and correctness of the model. Precision, recall, and F1-score are often used performance measures in spam detection applications. These metrics give a broad indication of how effectively the algorithm can differentiate between real emails and spam.

4. **Fine-tuning**: The model may need to be adjusted or modified based on the performance evaluation's findings in order to increase accuracy and performance. This can entail changing the machine learning algorithm's settings or applying a completely other model.

5. **Implementation:** The model can be implemented in a real-world scenario to categorize fresh emails as spam or legitimate after it has been tested and improved.

Overall, any spam detection effort [11] must include testing as a crucial component.

VIII DETECTING THE EMAILS AS SPAM AND HAM :

Machine learning algorithms are frequently used to categorize Figure 2 represents emails based on a collection of attributes or characteristics that are frequently linked with spam or genuine emails in order to detect spam or ham (legitimate) emails. The following steps are often included in the process:

1. **Data collection and preprocessing**: To eliminate redundant or irrelevant data, a sizable dataset of emails is collected and processed. Based on the content and context of the emails, they are either classified as spam or ham.
2. **Feature extraction:** From the emails that are pertinent to the categorization task, features are extracted. These could include metadata, the sender's identity, the message's content, the subject line, and any attachments.
3. **Model selection and training:** A machine learning model is chosen based on the project's unique specifications and the properties of the data.

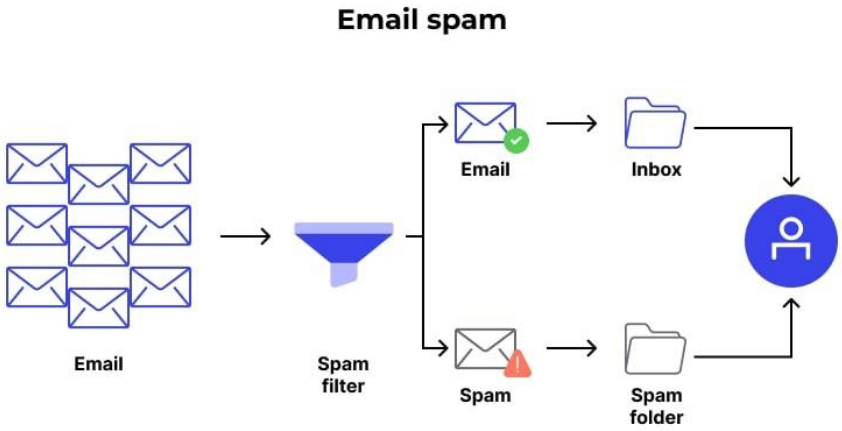


Figure 2 : Email SPAM Detection

IX EVALUATION PARAMETER

True Negative Rate= True Negative / True Negative +False Positive

False Negative Rate = False Negative / False Negative + True Positive

False Positive Rate = False Positive / False Positive + True Negative.

Precision Formula: True Positive / True Positive + False Positive

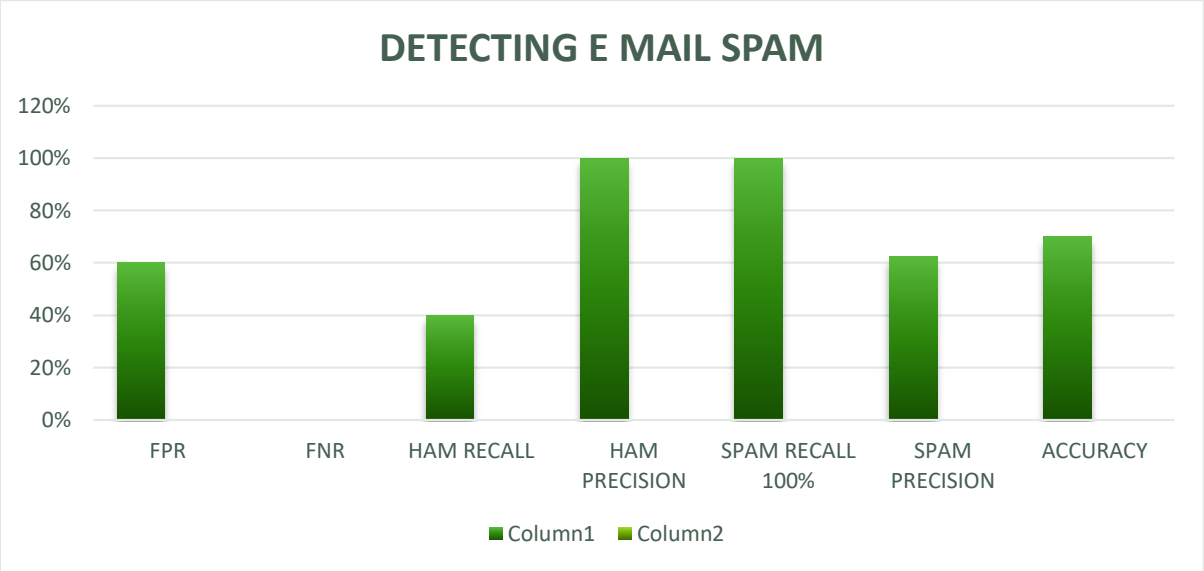
Recall Formula: True Positive / True Positive + False Negative

F- Measure Formula = Precision.Recall / Precision – Recall

Accuracy Formula = True Positive Rate + True Negative Rate / True Positive Rate + True Negative Rate False Positive Rate + False Negative Rate.

Table 4: FORMULA FOR DETECTING THE EMAIL SPAM / HAM

Measure	Formula
FPR	nH=S/nH=H+nH=S = 60%
FNR	nS=H/nS=S+nS=H = 0%
Ham Recall	Nh=H/nH=H+nH=S = 40%
Ham Precision	nH=H/nH=H+nS=H = 100%
Spam Recall	nS=S/nS=S+nS=H = 100%
Spam Precision	nS=S/nS=S+nH=S = 62.5%
Accuracy %	nH=H+nS=S/nH=H+nS=S+nS=H+nH=S = 70%



X EXPERIMENTAL RESULT:

FEATURE	SPAM EMAILS	HAM EMAILS
Sender information	Often from unknown or suspicious addresses	Often from known or trusted sources.
Subject Line	Often contains misleading or clickbait phrases	Often accueately reflects the content of the email.
Message Content	Often contains keywords or phrases associated with spam	Often contains relevent and useful information.
Attachments and Links	Often contain suspicious attachments or links	Often contain legitimate attachments or links
Metadata	Often sent from multiple IP addresses or locations	Often sent from a consistent IP address or location
Frequency Messages	Often sent in large volumes ormat irregular intervals	Often sent at consistent intervals of frequencies

The particular characteristics used to categorise emails may differ depending on the spam detection algorithm or technique being utilised; this comparison table is merely an example. It's also crucial to keep in mind that not all spam emails will display every one of the traits mentioned above, and some valid emails might display some of the traits typically connected with spam. To effectively classify emails as spam or ham, it is crucial to employ a combination of characteristics and machine learning techniques.

XI CONCLUSION:

For the purpose of removing undesired and potentially hazardous emails from our inboxes, email spam detection utilising AI and machine learning has emerged as a crucial tool. We can accurately categorise emails as spam or ham and take the necessary action by analysing the content and attributes of emails using machine learning algorithms. We may combine indicators like the False Positive Rate (FPR), False Negative Rate (FNR), Ham Recall, Ham Precision, and Accuracy to gauge how well a spam detection system performs. These metrics enable us to evaluate the efficacy and accuracy of the system's classification of emails as spam or ham. When compared to FNR, FPR gauges the proportion of valid emails that are mistakenly labelled as spam.

REFERENCES:

[1] A. Zamir, H. U. Khan, W. Mehmood, T. Iqbal, and A. U. Akram, "A feature-centric spam email detection model using diverse supervised machine learning algorithms," e Electronic Library, vol. 38, no. 3, 2020.

[2] TheStar. Company Cheated of RM 4.5 Mil Due to Email Spoofing. Accessed: Jul. 30, 2019. [Online]. Available: https://www.thestar.com. my/news/nation/2017/06/11/kedah-based-company-cheated-due-toemail-spoofing .

[3] S. Suryawanshi, A. Goswami, and P. Patil, "Email spam detection: an empirical comparative study of different ml and ensemble classifiers," in Proceedings of the 2019 IEEE 9th International Conference on Advanced Computing (IACC), pp. 69–74, IEEE, Tiruchirappalli, India, Dec 2019.

[4] S. Smadi, N. Aslam, and L. Zhang, ''Detection of online phishing email using dynamic evolving neural network based on reinforcement learning,'' Decis. Support Syst., vol. 107, pp. 88–102, Mar. 2018.

[5] C.-Y. Chiu, A. Prayoonwong, and Y.-C. Liao, ''Learning to index for nearest neighbor search,'' 2018, arXiv:1807.02962. [Online]. Available: https://arxiv.org/abs/1807.02962.

[6] A. Iyengar, G. Kalpana, S. Kalyankumar, and S. GunaNandhini, "Integrated spam detection for multilingual emails," in Proceedings of the 2017 International Conference on Information Communication and Embedded Systems (ICICES), pp. 1–4, IEEE, Chennai, India, February 2017.

[7] T. Kumaresan and C. Palanisamy, "E-mail spam classification using s-cuckoo search and support vector machine," International Journal of Bio-Inspired Computation, vol. 9, no. 3, pp. 142–156, 2017.

[8]] H. M. Al-Mashhadi and M. H. Alabiech, ''A survey of Email service: Attacks, security methods and protocols,'' Int. J. Comput. Appl., vol. 162, no. 11, pp. 31–40, 2017.

[9] P. Parveen and P. Halse, "Spam mail detection using classification," International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, pp. 347– 349, 2016.

[10] H. Kaur and A. Sharma, "Improved email spam classification method using integrated particle swarm optimization and decision tree," in Proceedings of the 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), pp. 516–521, IEEE, Dehradun, India, Oct 2016.

[11] D. Sipahi, G. Dalkılıç, and M. H. Özcanhan, ''Detecting spam through their sender policy framework records,'' Secur. Commun. Netw., vol. 8, no. 18, pp. 3555–3563, Dec. 2015.

[12] Y. Li, S. C. Sundaramurthy, A. G. Bardas, X. Ou, D. Caragea, X. Hu, and J. Jang, ''Experimental study of fuzzy hashing in malware clustering analysis,'' in Proc. 8th Workshop Cyber Secur. Experimentation Test (CSET), Berkeley, CA, USA: USENIX Association, 2015, p. 8.

[13] J. Chen, R. Fontugne, A. Kato, and K. Fukuda, ''Clustering spam campaigns with fuzzy hashing,'' in Proc. AINTEC Asian Internet Eng. Conf. (AINTEC), Nov. 2014, p. 66.

[14] M. Prilepok, P. Berek, J. Platos, and V. Snasel, ''Spam detection using data compression and signatures,'' Cybern. Syst. Int. J., vol. 44, nos. 6– 7, pp. 533–549, Mar. 2013.

[15] G. Caruana and M. Li, ''A survey of emerging approaches to spam filtering,'' ACM Comput. Surv., vol. 44, no. 2, p. 9, Feb. 2012.