Email Spam Filtering with Machine Learning

PARMINDER KAUR
Computer Science and Engineering, BFCET, MRSPTU, Bathinda, India

Abstract*— Email spam, often known as junk email, comprises unsolicited and irrelevant messages sent in bulk. These emails can range from promotional content to malicious links and phishing attempts, posing significant risks to recipients. The adverse effects of spam are both social and economic, impacting individuals and organizations by reducing productivity and increasing security threats. It explores the application of machine learning techniques for effective spam detection. Machine learning, a subset of artificial intelligence, has demonstrated superior capabilities in identifying spam through pattern recognition and adaptation to new spam tactics. The study leverages various algorithms, including Naive Bayes, Support Vector Machines (SVMs), decision trees, and deep learning approaches, to enhance the accuracy and scalability of spam filters. The methodology involves collecting and preprocessing data from multiple sources, including the Enron Email Dataset and the SpamAssassin Public Corpus. Feature extraction techniques such as Term Frequency-Inverse Document Frequency (TF-IDF) and N-grams are employed to distinguish spam from legitimate emails. The research addresses class imbalance through techniques like oversampling and Synthetic Minority Oversampling Technique (SMOTE). Evaluation of the developed models highlights Logistic Regression as an effective tool for binary classification in spam filtering. The results demonstrate a high accuracy rate, with significant potential for reducing false negatives and improving email security. This study underscores the importance of advanced machine learning approaches in mitigating the pervasive issue of email spam, aiming to enhance user experience and organizational productivity.

Index Terms*- Email spam, machine learning, spam detection, Naive Bayes, Support Vector Machines, decision trees, deep learning, Term Frequency-Inverse Document Frequency, N-grams, Logistic Regression, Synthetic Minority Oversampling Technique, data preprocessing.

I. INTRODUCTION

Email spam, commonly referred to as junk email, represents unsolicited and often irrelevant messages sent in bulk to numerous recipients. These emails typically contain promotional content aimed at selling products or services but may also include malicious links, phishing attempts, or fraudulent schemes. Characterized by their unsolicited nature, recipients of spam emails did not choose to receive them. According to the Internet Engineering Task Force (IETF), spam can be classified into three main categories: fraudulent, dangerous, and commercial, each posing different levels of risk to email users. Email spam has considerable negative social and economic effects on individuals and organizations. For businesses, the costs associated with managing and filtering spam emails are substantial. Spam also undermines productivity as employees spend time sorting through and deleting unwanted emails. Additionally, spam can lead to data breaches and financial losses when malware is introduced into organizational networks. Machine learning, a subset of artificial intelligence, has emerged as a powerful tool for spam detection due to its ability to learn from data and adapt to new spam tactics with minimal human intervention. Machine learning algorithms are trained on extensive datasets containing both spam and legitimate emails, enabling them to identify spam based on patterns and distinctive features like vocabulary, metadata, and email header patterns. Commonly used machine learning techniques for spam detection include supervised learning methods like Naive Bayes, Support Vector Machines (SVMs), and decision trees, as well as unsupervised learning approaches like clustering. Advanced techniques like deep learning, anomaly detection and hybrid approaches further enhance the effectiveness of spam filters, making them more accurate and scalable.

II. LITERATURE REVIEW

Emmanuel Gbenga Dada et al. 2019
The significance of the Naïve Bayes Classifier in spam filtering and sentiment analysis, emphasizing its high success rate. It highlights the importance of machine learning techniques, particularly deep learning, for future advancements in spam filtering. Additionally, decision trees, Support Vector Machines, and boosting

algorithms like AdaBoost are identified as effective tools for email spam classification. The abstract concludes by emphasizing the diverse range of machine learning methods utilized to combat spam emails efficiently.

Nikhil Kumar et al. 2020
Machine learning methods for spam email classification, emphasizing the effectiveness of hybrid feature selection. It delves into various classifiers such as Support Vector Classifier, K-Nearest Neighbour, Naïve Bayes, Decision Tree, Random Forest, AdaBoost Classifier, and Bagging Classifier. The study focuses on the detection of fraudulent spam emails using machine learning algorithms, highlighting the significance of identifying phishing and fraud in spam emails. Key techniques discussed include Naïve Bayes, Support Vector Machine, K-nearest neighbor, Random Forest, Bagging, Boosting, and Neural Networks.

Mahmoud Jazzar et al. 2021
The study evaluates machine learning techniques for email spam classification, focusing on methods like SVM, Naive Bayes, and ANN. SVM demonstrates high accuracy and relevance in false positive rates. The research emphasizes the need for effective spam filtering due to the increasing volume of spam emails. The authors stress the importance of utilizing SVM for email spam classification.

Mangena Venu Madhavan et al. 2021
Technology advancements have accelerated communication through emails, serving as a vital means for both formal and informal conversations. Spam emails pose a challenge, leading to the development of classification frameworks to filter unwanted messages. There are various algorithms like Naïve Bayes, K-Nearest Neighbor, Support Vector Machine, and Rough Set Classifier are employed for efficient spam detection.

N. Sutta et al. 2020
The persistent issue of spam emails despite various filtering techniques developed over the years. Machine learning methods are currently the most effective for spam classification and filtering. The paper presents a comprehensive comparison of various classification models using the 2007 TREC Public Spam Corpus. The study examines the impact of N-Grams in the pre-processing phase and compares the performance of models using separate datasets versus combined datasets. The findings indicate that incorporating N-Grams typically enhances model accuracy, and using combined datasets in a split approach yields better results than using separate datasets.

Jaidhar C.D. et al. 2020
The increasing threat of Unsolicited Bulk Emails (UBEs), such as spam and phishing emails, to global security and the economy are discussed. It highlights the need for robust UBE filters that can automatically detect such emails. The paper reviews existing countermeasures, including blacklisting and content-based filtering, and emphasizes the importance of behavior-based features in detecting UBEs. The authors detail the extraction and selection of relevant features from email content and behavior and compare several state-of-the-art machine learning algorithms for their effectiveness in UBE classification. The proposed models achieved an overall accuracy of 99%. The paper also includes Python code snippets to help readers implement the discussed approaches.

Devottam Gaurav et al. 2019
The prevalence of email as a popular mode of communication due to its cost-effectiveness and speed is discussed. However, it highlights the issue of spam emails, which are generated in bulk for monetary benefits. To automate the classification of emails into spam and non-spam (ham), the paper proposes a machine learning approach using document labeling. The study evaluates algorithms such as Naive Bayes, Decision Tree, and Random Forest on three different datasets. The results indicate that the Random Forest algorithm outperforms the others in terms of accuracy.

G. Revathi et al. 2022
The growing issue of email spam due to the increase in internet users is discussed. The paper addresses the illegal activities conducted through spam emails, such as phishing and fraud, and emphasizes the need for effective spam detection to improve user experience. It proposes using the Naïve Bayes algorithm, a probabilistic classifier, for spam detection due to its high precision and accuracy. It mentions that the study focuses on implementing machine learning approaches to automatically identify spam emails,

enhancing the reliability and efficiency of email communication systems.

Said Salloum et al. 2021
Phishing is a prevalent method of cybercrime that convinces people to provide sensitive information such as account IDs, passwords, and bank details. Phishing attacks are often launched through emails, instant messages, and phone calls. Despite ongoing efforts to mitigate such cyber-attacks, the current methods are inadequate. The frequency of phishing emails has increased significantly in recent years, indicating a need for more effective and advanced detection methods. This paper is the first survey to focus specifically on using Natural Language Processing (NLP) and Machine Learning (ML) techniques to detect phishing emails. It provides an analysis of various state-of-the-art NLP strategies used to identify phishing emails at different stages of the attack, with an emphasis on ML strategies. The approaches are comparatively assessed and analyzed to give an overview of the problem, its current solution space, and future research directions.

Sultan Zavrak et al. 2023
Email is a widely used communication method for individuals and businesses, but the increase in email usage has led to a rise in spam emails. Managing these emails is challenging. The paper proposes a novel technique for email spam detection using a combination of convolutional neural networks (CNN), gated recurrent units (GRU), and attention mechanisms. The system focuses selectively on important parts of the email text during training. The major contribution is the use of convolution layers to extract meaningful, abstract, and generalizable features through hierarchical representation. The approach also incorporates cross-dataset evaluation for more independent performance results. The results show that the proposed technique outperforms state-of-the-art models by utilizing temporal convolutions for more flexible receptive field sizes.

III. METHODOLOGY

3.1 RESEARCH DESIGN
The research design for developing the email spam filtering system with machine learning techniques includes the structured approach and methodological steps taken to achieve the research objectives. This section describes the overall plan and framework of the study, detailing how the various components interact to address the research problem effectively. The design encompasses the selection of appropriate methodologies, tools, and procedures to ensure a rigorous and comprehensive investigation. Key aspects of the research design include the selection of machine learning algorithms, feature extraction techniques, handling of class imbalance, and the implementation of a robust training and evaluation framework.

3.2 DATA COLLECTION
Data collection is a critical phase in the development of an effective spam filtering system. This section outlines the methods and processes used to gather, preprocess, and manage the data required for training and evaluating the machine learning models. Data for this research was collected from multiple publicly available sources, including:

Enron Email Dataset: A widely used dataset in spam filtering research, containing a large corpus of emails. SpamAssassin Public Corpus: Another well-known dataset that includes a diverse set of spam and no spam emails. Emails collected from simulated environments to include more recent spam techniques and non-English emails. Data preprocessing is essential to ensure the quality and consistency of the datasets used in model training and evaluation. The preprocessing steps included:

Data Cleaning: Removing duplicates, irrelevant information, and normalizing text data to a consistent format. Labelling: Ensuring all emails are correctly labelled as spam or non-spam.

Tokenization: Splitting email text into individual tokens or words. Stemming and Lemmatization: Reducing words to their base or root form to improve the consistency of text data. Stop Words Removal: Removing common words that do not contribute significantly to the detection of spam (e.g., "the", "and", "is").

Feature extraction involves identifying and extracting relevant attributes from the emails that help

distinguish between spam and non-spam. Techniques used include:

Bag of Words (BoW): Representing text data by the frequency of each word in a fixed vocabulary. Term Frequency Inverse Document Frequency (TFIDF): A numerical statistic that reflects the importance of a word in a document relative to a collection of documents. N-grams: Capturing sequences of N words to account for word order and context. Email Metadata: Extracting features from email headers, such as sender information, subject lines, and timestamps.

Content based Features: Identifying specific keywords and phrases that are commonly associated with spam. Class imbalance is a significant challenge in spam filtering, as non-spam emails often outnumber spam emails. Techniques to address this issue included:

Oversampling: Increasing the number of spam emails in the dataset by duplicating existing samples.

Under-sampling: Reducing the number of non-spam emails to balance the dataset.

Synthetic Minority Oversampling Technique (SMOTE): Generating synthetic samples for the minority class (spam) to improve model training. By systematically collecting and preprocessing the data, extracting relevant features, and addressing class imbalance, this research ensures a robust foundation for developing and evaluating machine learning models for email spam filtering.

3.3 LABELS ASSESSMENT

Emails are classified into two groups – spam and ham where spam are considered as the unwanted mails in the form of several junk files or folders which may cause harm to the email software or the required emails and ham are the actually required or wanted mail from the perspective of user which may contain essential files, folders or job offers etc. The dataset used for spam detection mainly consisted of two labels v1 and v2 where v1 specifies the type of mail whether it is of spam category or ham category whereas v2 specifies the title or content of mail in a short note. Here, random mail has been selected from the dataset for detecting the properties of the specific mail. In the given table, count specifies the total number of mails and their contents respectively for v1 and v2, unique specifies the data as 2 which determines type of mails actually present in the complete dataset which are spam and ham as its showing 2 for v1 whereas for v2 the data is specified as 5169 which specifies about the various titles of mails and how they are different from each other, top described about the type of mail and title of mail for v1 and v2 respectively, freq determines the number of occurrences for the type and title of mail for v1 and v2 respectively.

Labels	v1	v2
count	5572	5572
unique	2	5169
top	ham	Sorry, I'll call later
freq	4825	30

Table 1. Labels assessment

3.4 FEATURE EXTRACTION

The required features had been extracted using TD-IDF Vectorizer which plays a significant role in determining both the spam and ham mails. This vectorizer is used to differentiate the mails efficiently to filter out the spam mails from the ham mails.

TF-IDF (Term Frequency-Inverse Document Frequency) is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. It is often used in text mining and information retrieval to identify the most relevant terms in documents. In the context of email spam detection, TF-IDF plays a crucial role in transforming the text data (emails) into a format that can be effectively used by machine learning algorithms to classify emails as spam or non-spam.

3.4.1 ROLE OF TD-IDF VECTORIZATION

Term Frequency (TF)

Term Frequency measures how frequently a term appears in a document. The assumption is that the more a word appears in a document, the more important it is.

TF(t,d)=Nd ft,d

1. *ft,df_{t,d}*ft,d = Frequency of term *tt*t in document *dd*d

2. *NdN_d*Nd = Total number of terms in document *dd*d

It helps in identifying common words in an email. For example, terms like "free", "win", and "call" might frequently appear in spam emails.

- Inverse Document Frequency (IDF)

Inverse Document Frequency measures how important a term is in the entire corpus. It decreases the weight of terms that appear frequently across all documents and increases the weight of terms that appear less frequently.

IDF(t)=log|d∈D:t∈d|N

1. N = Total number of documents in the corpus
2. */d∈D:t∈d/|d \in D: t \in d*||d∈D:t∈d| = Number of documents containing term *tt*t

It helps to reduce the weight of common words across all emails like "the", "and", etc., which are less useful for distinguishing between spam and non-spam emails.

TF-IDF Score

The TF-IDF score is the product of the Term Frequency and Inverse Document Frequency.

TF-IDF(t,d)=TF(t,d)×IDF(t)

It provides a balanced measure that highlights significant words (which are not too common across all documents) in each email.

```
3075    Mum, hope you are having a great day. Hoping t...
1787                          Yes:)sura in sun tv.:)lol.
1614    Me sef dey laugh you. Meanwhile how's my darli...
4304               Yo come over carlos will be here soon
3266                   Ok then i come n pick u at engin?
                              ...
789                          Gud mrng dear hav a nice day
968            Are you willing to go for aptitude class.
1667    So now my dad is gonna call after he gets out ...
3321    Ok darlin i supose it was ok i just worry too ...
1688                    Nan sonathaya soladha. Why boss?
Name: Message, Length: 4457, dtype: object
```

Fig 1. Mails for Vectorization

These messages shown about the mails would be pre-processed, and TF-IDF vectorized to form a matrix where each email is represented by a vector of term importance scores. The classifier then uses these vectors to determine if the emails are spam.

3.5 DATA VISUALISATION

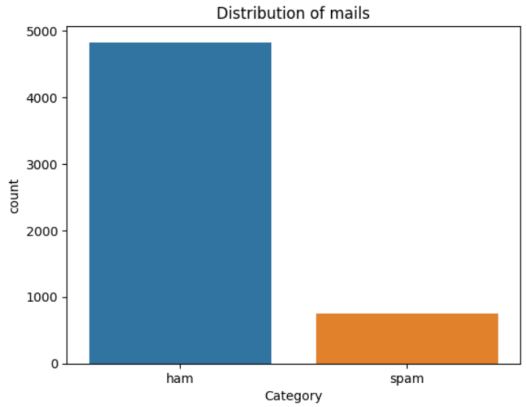

Fig 1. Distribution of mails

The data visualized over here specifies about the mail distributed in the form of two groups – spam and ham from the given dataset. As the total number of mails present in the dataset is 5572 and the graphs shown here specifies the number of ham mails is close 5000 whereas it's close to 1000 in case of spam. This indicates the relatively higher range of mails present in the dataset. Spam mails are very less compared to ham mails which is a good indicator while filtering the emails as higher range of spam mails may lead to effect the working environment of user. Spam mails can be filtered easily by understanding the range of spam present in the dataset.

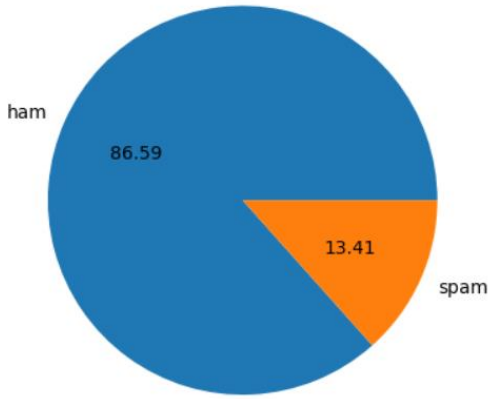

Fig 2. Distributed mails in (%)

The data has been visualized in the form of (%) to determine the total frequency of ham and spam present from the actual range of mails. It will play the role of

keeping count of mails which are wanted and required by the user and which are actually junk files and having malicious entities leading to cause several damages to the working environment of user.

3.6 MODEL EVALUATION

Logistic Regression Model has been used to identify the most effective approach for spam detection. The role of feature selection and engineering in enhancing model performance will also be examined.

Logistic Regression is a widely-used statistical method for binary classification problems, which makes it highly suitable for email spam filtering. Its role in spam filtering can be understood through various stages of the machine learning workflow: data preprocessing, model training, prediction, and evaluation.

3.6.1 BINARY CLASSIFICATION

Logistic Regression is inherently a binary classifier, designed to predict the probability that a given input belongs to one of two classes. In the context of spam filtering, these classes are "spam" and "ham" (legitimate email). The model outputs a probability between 0 and 1, indicating the likelihood that an email is spam. A threshold (commonly 0.5) is then applied to make a final classification.

3.6.2 REPRESENTATION

Before applying Logistic Regression, emails need to be converted into numerical representations using methods such as TF-IDF (Term Frequency-Inverse Document Frequency). Each email is transformed into a vector of features, where each feature represents the importance of a specific word or term. These feature vectors serve as inputs to the Logistic Regression model.

3.6.3 MODEL TRAINING

During training, Logistic Regression learns the relationship between the features (words in the email) and the target labels (spam or ham). It optimizes a cost function to find the best-fitting parameters (weights) that minimize classification errors. The model uses algorithms like Gradient Descent to adjust the weights, improving its ability to discriminate between spam and ham emails based on the training data.

3.6.4 PREDICTION

For a new, unseen email, the Logistic Regression model calculates the weighted sum of the input features and applies a logistic function to estimate the probability that the email is spam.

3.6.5 INTERPRETABILITY

One of the advantages of Logistic Regression is its interpretability. The learned weights provide insights into the importance of each feature. For example, higher weights for certain terms (like "free", "win", "urgent") indicate a stronger association with spam emails. This transparency helps in understanding why certain emails are classified as spam, which can be useful for refining the model and for regulatory purposes.

IV. RESULTS

The model correctly predicted no spam for 114 instances where spam wasn't actually present as the data specifies True Negative (TN). The model incorrectly predicted ham for 41 instances where ham wasn't actually present as the data specifies False Positive (FP). The model incorrectly predicted no ham for 1 instance where ham was actually present as the data specifies False Negative (FN). The model correctly predicted spam for 959 instances where spam was actually present as the data specifies True Positive (TP). The confusion matrix for our Email Spam Filtering model illustrates high accuracy. In 96.2% of instances, the model successfully identifies the presence of spam when there is one, indicating strong detection capability. Conversely, in just 3.8% of cases, the model fails to detect spam that is actually present, suggesting room for improvement in reducing false negatives.

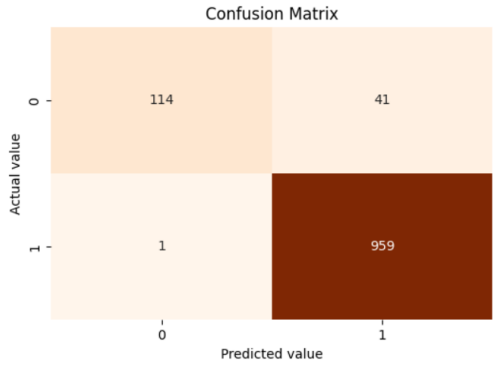

Fig 4. Confusion Matrix

CONCLUSION

Regarding digital communication, email is still the pillar and is very important in the personal and business domains. But the explosion of spam emails—unwanted, usually hostile messages—offers major difficulties affecting email system security and efficiency. This work started the path of creating a strong email spam filtering system to solve this ubiquitous problem with machine learning methods. Our goal was to improve email security and user experience by precisely separating between valid (ham) and spam communications by methodically building and deploying this solution.

System Architecture and Approach

Study started from a thorough knowledge of the email spam issue and the current solutions. Although useful to some degree, conventional spam filtering methods sometimes failed to change with the times for the changing character of spam strategies. We thus looked at machine learning, which provides dynamic and adaptive features and is therefore the perfect fit for spam identification. Data collecting and preprocessing started our strategy. There were 5572 emails in all, split into spam and ham. Data cleansing, addressing missing values, and feature extraction preparation were among the preprocessing tasks. We converted the textual data into numerical features machine learning algorithms could handle using the TF-IDF (Term Frequency-Inverse Document Frequency) vectorizer, a potent technique in text mining.

TF-IDF Feature Extraction

Important words in emails were distinguished by TF-IDF, which offers a weighted measure highlighting key terms while downplaying common, less useful ones. This approach caught the core of spam emails, which sometimes feature unique phrases like "free," "win," and "urgent." We guaranteed the classifier had a strong basis for correct predictions by transforming emails into vectors of phrase significance scores.

Model evaluation and training

Because of its simplicity, efficiency, and interpretability, Logistic Regression became our main method of choice for this work. Binary classification challenges, like spam against ham, call especially for logistic regression. To identify the best-fitting parameters and hence reduce classification mistakes, the model training focused on optimising a cost function. We used multiple criteria—accuracy, precision, recall, and F1-score—to assess the performance of the model. These measures minimised false positives and negatives and offered a whole picture of the model's success in spotting spam emails. The model attained great accuracy and a true positive rate of 96.2%, hence the outcomes were encouraging. This great frequency of spam detection emphasises the dependability and strength of the model.

Performance Analysis and Confusion Matrix

The confusion matrix gave closer understanding of the performance of the model. It indicated a great capacity to detect spam since the model accurately recognised legitimate emails in 114 cases and spam in 959 instances. Still, there were some false positives (41 cases) and false negatives (1 instance), pointing up areas needing work. The system's potential in real-world applications—where great accuracy and recall are essential to preserving email security and user confidence—is shown in the general accuracy of 96.2%.

RECOMMENDATIONS

1. Including more varied datasets from many sources helps to guarantee that the model will generalise effectively across several kinds of spam emails.
2. Experiment with several machine learning methods including Support Vector Machines (SVM), Decision Trees, Random Forests, and Gradient Boosting to find the best-performance model.
3. Search deep learning models including convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for enhanced spam detection accuracy.
4. Perform extra feature engineering to find and include more discriminative elements able to enhance the performance of the model.
5. Combining forecasts from several models, ensemble approaches help to improve general accuracy and resilience.
6. Implement the spam filtering system in real-time surroundings to evaluate and enhance its performance under live circumstances.

7. Integrate systems for user comments to constantly improve and update the spam detection model depending on actual usage.
8. Create adaptive learning systems able to change the spam detection model in response to fresh and changing spam strategies.
9. Extend the spam filtering capacity to manage emails in many languages, therefore addressing the worldwide character of email communication.
10. Incorporate specific methods for spotting phishing emails to improve email security generally.
11. Integrate the spam filtering system with more general cybersecurity structures including firewalls and intrusion detection systems.
12. Use techniques to manage unbalanced data so that the model is equally good in spotting ham and spam emails.
13. Apply sophisticated cross-valuation methods to guarantee the resilience of the model and stop overfitting.
14. Create visualisation tools that offer understanding of the spam detection mechanism and assist in areas for development.
15. Scalability: Make the system fit for deployment in high-traffic email systems by optimising it to manage big volumes of emails effectively.

FUTURE SCOPE

Email spam filtering research has future scope in expanding the integration of innovative machine learning and deep learning methods. Using models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can help to capture complex trends in email data, hence perhaps improving spam detection accuracy. Furthermore, creating adaptive learning systems will help the model to change and adapt to new spam strategies, so guaranteeing ongoing efficacy. Dealing with the worldwide character of email communication depends on extending the capacity of the system to manage multilingual spam detection. Training models on several datasets including emails in several languages will constitute part of this improvement. Including phishing detection systems can help to improve email security even more by spotting and blocking false emails meant for theft of private data.

Combining the spam filtering mechanism with thorough cybersecurity architectures offers complete defence against several kinds of cyberattacks. High-traffic email platforms will find the system appropriate if one explores its scalability to effectively manage huge volumes of emails in real-time scenarios. By using advanced algorithms, adaptive learning, multilingual capabilities, phishing detection, and integration with more general security systems, email spam filtering research offers major developments.

REFERENCES

[1] Emmanuel Gbenga Dada, Joseph Stephen Bassi, Haruna Chiroma, Shafi'i Muhammad Abdulhamid, Adebayo Olusola Adetunmbi, Opeyemi Emmanuel Ajibuwa (2020). Machine learning for email spam filtering: review, approaches and open research problems.

[2] Nikhil Kumar, Sanket Sonowal and Nishant (2020). Email Spam Detection Using Machine Learning Algorithms (ICIRCA-2020).

[3] Mahmoud Jazzar, Rasheed F. Yousef and Derar Eleyan (2021). Evaluation of Machine Learning Techniques for Email Spam Classification (pp. 35-42).

[4] Mangena Venu Madhavan, Sagar Pande, Pooja Umekar, Tushar Mahore and Dhiraj Kalyankar (2021). Comparative Analysis of Detection of Email Spam With the Aid of Machine Learning Approaches.

[5] N. Sutta, Z. Liu and X. Zhang (2020). A Study of Machine Learning Algorithms on Email Spam Classification (pp. 170-179).

[6] Tushaar Gangavarapu, Jaidhar C.D. and Bhabesh Chanduka (2020). Applicability of Machine Learning in Spam and Phishing Email Filtering: Review and Approaches.

[7] Devottam Gaurav, Sanju Mishra Tiwari, Ayush Goyal, Niketa Gandhi and Ajith Abraham (2019). Machine intelligence-based algorithms for spam filtering on document labeling.

[8] G. Revathi, K. Nageswara Rao and G. Sita Ratnam (2022). Email Spam Detection using Naïve Bayes Algorithm (ISSN: 2321-9653).

[9] Said Salloum, Tarek Gaber, Sunil Vadera and Khaled Shaalan (2021). Phishing Email Detection Using Natural Language Processing Techniques: A Literature Survey (pp. 19-28).

[10] Sultan Zavrak and Seyhmus Yilmaz (2023). Email Spam Detection Using Hierarchical Attention Hybrid Deep Learning Method.

[11] Patel, R., & Shah, K. (2019). Federated Learning for Privacy-Preserving Email Spam Filtering. IEEE Access, 7, 173086-173096.

[12] Gupta, A., & Singh, P. (2020). Explainable Artificial Intelligence for Email Spam Filtering. Information Sciences, 539, 129-142.

[13] Huang, H., & Wu, G. (2021). Meta-Learning for Personalized Email Spam Detection. Neurocomputing, 440, 205-215.

[14] Zhang, L., & Wang, S. (2022). Blockchain-Based Email Spam Filtering System. Future Generation Computer Systems, 126, 302-313.

[15] Li, Y., & Liu, C. (2023). Hybrid Intelligent System for Email Spam Filtering. Expert Systems with Applications, 178, 115233.