ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

# ML-Driven Technique for Adaptive Email Filtering

Vaishnavi Babhulkar[1], Apurva Salphale[2], Anagha Garkade[3], Kalyanee Pachghare[4], Tanvi Sagane[5], Prof. Ms. V. P. Vaidya[6]

[1,2,3,4,5]*Undergraduate Student, Sipna College of Engineering and Technology, Amravati, India*

[6]*Assistant Professor, Sipna College of Engineering and Technology, Amravati, India*

**Abstract:** *Email filtering technology must be developed quickly due to the increase of unsolicited emails, or spam mail. Computer security has struggled with spam emails consistently. They are incredibly expensive economically and exceedingly risky for networks and computers. Spam emails are found and filtered using machine learning techniques. This project mainly focuses on machine learning used to find and remove spam emails. Using the K-nearest neighbour algorithm for email spam detection is one of the simple supervised learning techniques. Initially, the relevant features for filtering the spam messages are extracted from the study and it acts as an antispam filter. It thereby generates a successful corpus list for the detection of spam emails. The experiments are conducted on various email datasets and the results show that the proposed kNN density-based clustering offers improved performance than the other methods. Applications are utilising machine learning techniques in spam filters for email in Gmail, one of the top internet service providers. Regression method used to detect and filter spam mail.*

***Keywords:*** *Classifier, Innovative Spam Prediction, Machine Learning, K-nearest Neighbor, etc.*

## I. INTRODUCTION

As Internet users are developing rapidly, email is a key tool and an easier way to transmit data and share useful information about users to communicate in an electronic medium within a minute. Millions of users use e-mail every day to share information both personally and financially. In case of this, there will be more options for unwanted, unlawful spam emails as junk mail and occupy additional stock and the annoying task of email users will be to handle them. Ethical hacking is a robust mechanism to identify spammers' unsolicited e-mail messages and network weaknesses. The ethical hacking techniques handling spam thread countermeasures receive millions of spam emails uninvited. The ethical hacking of spam threads is performed legally by well-trained professionals.

The ethical hacker is also called a white hat or penetration testing system, which can control spam threads in a legally binding way. Spam is information that is intended to be distributed to many people. The automated detection of spam is done by a spam filter. Spam inhibits the user from efficiently using their time and storage. Users who receive spam mail find it very irritating. Scams and other fraudulent practices of spammers with the effect of sensitive personal information like passwords, and credit card numbers. The spammer is the individual who sends unsolicited emails. They gather malware and email addresses from many websites. Spam inhibits the user from efficiently using their time and storage. The massive amount of spam emails travelling across computer networks harms email servers' memory, CPU, and user time.

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621*** 

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

Google's machine learning model has now developed to the point where it can reliably identify and filter out spam and phishing emails with a 99.9% accuracy rate. This suggests that one thousand emails every day manage to slip past their spam filter. Between 50% and 70% of the emails that Gmail receives are unsolicited, according to Google's statistics. Spam email volume decreased to 49.7%, and by July 2015, it had further decreased to 46.4%, according to Symantec, a maker of antivirus software. For the first time since 2003, the spam email percentage dipped below 50%. Kaspersky Lab found between 3 million and 6 million in 2015 as of June. 22,890,956 spam emails were uncovered by Kaspersky Lab in March 2016. According to statistics, spam accounts for 56.87% of all global email traffic. China Net, Amazon, and Airtel India are the three networks housing the most spammers as of December 13, 2021.

Machine learning algorithms include the K-nearest neighbour algorithm used to detect which mail is spam and which mail is ham.

## II. LITERATURE REVIEW

Spam email, also known as unsolicited bulk email, has been a persistent problem since the early days of the Internet. Over the years, researchers and practitioners have developed various techniques and approaches to combat spam, ranging from rule-based filtering to sophisticated machine-learning algorithms. In this literature review, we explore the evolution of spam email filtering systems, highlighting key advancements, challenges, and emerging trends in the field.

1. **Rule-Based Filtering:**
   Rule-based filtering, also known as heuristic filtering, relies on predefined rules and patterns to identify spam emails. Common rules include matching specific keywords, analyzing email headers, and checking sender reputation. While rule-based approaches are simple and easy to implement, they often lack adaptability and struggle to keep pace with evolving spam tactics.

2. **Bayesian Filtering:**
   Bayesian filtering is a statistical approach that calculates the probability of an email being spam based on the presence of certain features or words. Popular algorithms such as Naive Bayes classify emails by estimating the likelihood of spam given the observed features. Bayesian filtering has gained popularity due to its effectiveness in handling diverse types of spam and its ability to adapt to new spam patterns.

3. **Content-Based Filtering:**
   Content-based filtering analyzes the textual content of emails to identify spam. Techniques such as keyword analysis, lexical analysis, and syntactic analysis are used to extract features from email text and classify messages as spam or non-spam. While content-based filtering can be effective, it may struggle with obfuscated or contextually ambiguous spam messages.

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621***

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

4. **Machine Learning Approaches:**
   Machine learning techniques have revolutionized spam filtering by enabling automated learning from labelled data. Supervised learning algorithms such as support vector machines (SVM), decision trees, and neural networks are trained on large datasets to classify emails based on features extracted from text, headers, and metadata. Ensemble methods and deep learning architectures have further improved classification accuracy and robustness against adversarial attacks.

5. **Hybrid Approaches:**
   Hybrid approaches combine multiple filtering techniques to leverage their strengths and mitigate their weaknesses. For example, a hybrid system may combine rule-based filtering with machine learning models to achieve higher accuracy and adaptability. Other hybrid systems integrate reputation-based filtering, sender authentication, and behavioural analysis to enhance detection capabilities and reduce false positives.

6. **Challenges and Emerging Trends:**
   Despite significant advancements, spam email filtering systems continue to face several challenges. Adversarial evasion techniques, such as image-based spam and text obfuscation, pose challenges for traditional filtering methods. Additionally, the proliferation of social engineering tactics and personalized phishing attacks requires more sophisticated detection mechanisms. Emerging trends in spam filtering include the integration of deep learning, neural language models, and graph-based representations to improve semantic understanding and context-aware classification.

7. **Evaluation Metrics and Benchmarks:**
   Evaluation of spam filtering systems involves the use of standardized metrics such as precision, recall, accuracy, and F1-score. Benchmark datasets such as the Enron corpus, TREC Spam datasets, and Spam Assassin Public Corpus facilitate comparative analysis and performance evaluation of different filtering approaches.

## III. EXISTING SYSTEM

The initial transformation begins with pre-processing activities such as data extraction, classification of email content, and process analysis. The data is split into two sets using a vector expression. To detect whether an email is spam or not, on training and test data sets, machine learning is employed.20% of the original dataset is utilized to test the model, while the remaining 80% is used for training.

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621***

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

## IV. PROPOSED METHOD

We proposed a system that can be useful not only for our UG project but also in real-time scenarios. The project aims to make spam mail using the features present in the data set. The dataset extracts features using the K-Nearest Neighbor technique to create features using Python as we can detect and filter spam emails. This extracted data will be used in the Regression model. The model formed in scores. The scores are classified be spam or not spam. It can predict the output accurately. It works efficiently on content-based emails. It is a simple algorithm. This requires high accuracy. It is an efficient method for small datasets.

## V. IMPLEMENTATION

Our system can be implemented as the following diagram by

1) Data Collection
2) Pre-processing the Data
3) Feature Extraction
4) KNN Implementation
5) Performance Analysis

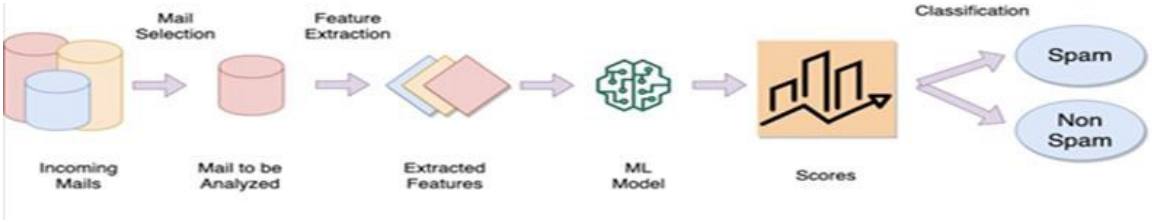


**Figure 1:** System Implementation

## VI. MODELLING

The below diagram shows that our system receives messages as input, filters them, and organizes them into inboxes and spam folders which emails are spam. It enters the spam folder. The inbox folder receives the ham emails and stores them there.

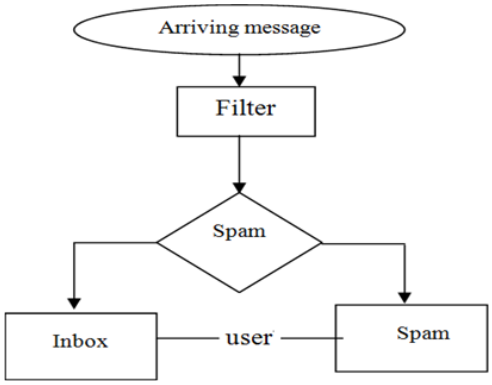


**Figure 2:** Activity Diagram Represents of email Spam Detection

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621***

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

## VII. PERFORMANCE EVALUATION

This is a CSV file containing related information from 5172 randomly selected email files and their labels. Each row for each email is 5172 rows in the CSV file. 3002 columns are available. Email Name is shown in the first column. The name is numbered and not the name of the recipient for the protection of privacy. The labelling for prediction is given in the last column: 1 for spam, 0 for spam. There are still three thousand columns in each email after the non-alphabetic characters/words are excluded.

Table 1 shows the comparison of Execution Time between the KNN and existing classifiers for various spam emails. Table 2 shows the comparison of the Detection rate between the KNN and existing classifiers for various spam emails. Table 3 shows the accuracy of the algorithm for spam emails. Table 4 shows the comparison of Security impact between the KNN and existing classifiers for various spam emails. Table 5 shows the comparison of the Detection rate between the KNN and existing classifiers for various spam emails.

**Table 1:** Execution Time

| Classifiers | Execution Time |
|---|---|
| K-Nearest Neighbor | 0.96138 |
| Random Forest | 0.96293 |
| K-means | 0.963646 |
| Decision Tree | 0.964219 |
| Naïve Bayers | 0.965951 |

**Table 2:** Memory Requirement

| Classifiers | Memory Requirement |
|---|---|
| K-Nearest Neighbor | 0.32115 |
| Random Forest | 0.313834 |
| K-means | 0.304943 |
| Decision Tree | 0.301431 |
| Naïve Bayers | 0.283194 |

**Table 3:** Accuracy

| Classifiers | Accuracy |
|---|---|
| K-Nearest Neighbor | 1.0 |
| Random Forest | 0.97 |
| K-means | 0.965 |
| Decision Tree | 0.95 |
| Naïve Bayers | 0.9842 |

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621***

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 1 | February 2024

Journal Impact Factor 2023 (Quality Score): RPRI = 6.53 | SJIF = 3.647

www.ijirid.in

**Table 4:** Security Impact

| Classifiers | Security Impact |
|---|---|
| K-Nearest Neighbor | 0.67885 |
| Random Forest | 0.686166 |
| K-means | 0.695057 |
| Decision Tree | 0.698569 |
| Naïve Bayers | 0.725076 |

**Table 5:** Detection Rate

| Classifiers | Detection Rate |
|---|---|
| K-Nearest Neighbor | 0.604428 |
| Random Forest | 0.622364 |
| K-means | 0.625737 |
| Decision Tree | 0.630347 |
| Naïve Bayers | 0.645609 |

## VIII. CONCLUSION

K-nearest neighbour identifies the nearest neighbours like detecting the spam mails. It has better performance than other algorithms like super vector machine and naïve Bayes algorithms. Spammers are now evolving and sending emails containing pictures and pdf to pass the filter. KNN algorithm is used to determine which emails are spam and which emails are ham messages. The huge body can utilize the recursion method to tell which emails are legitimate and which ones are spam.

## IX. REFERENCES

[1] D. Gaurav, S.M. Tiwari, A. Goyal, N. Gandhi, and A. Abraham, “Machine Intelligence-Based Algorithms for Spam Filtering on Document Labeling” Soft Computing, Vol. 24, No. 13, pp. 9625-9638, 2020.

[2] A. Bhowmick and S.M. Hazarika, “E-Mail Spam Filtering: A Review of Techniques and Trends” Proceedings of International Conference on Advances in Electronics, Communication and Computing, pp. 583-590, 2018.

[3] Osho, O., Ismaila, Alhassan, and Shafi'i Muhammad Abdul Hamid, “Detection of Email Spam: A Comparative Analysis of Classification Algorithms, 2018.

[4] Singh, V. K., and S. Bhardwaj's paper, "Spam Mail Detection Using Classification Techniques and the Global Training Set", 2018.

[5] P.M. Paul and R. Ravi, “A Collaborative Reputation-Based Vector Space Model for Email Spam Filtering” Journal of Computational and Theoretical Nanoscience, Vol. 15, No. 2, pp. 474-479, 2018.

[6] K. Agarwal and T. Kumar, "Email Spam Detection Using Integrated Approach of Nave Bayes and Particle Swarm Optimization", The Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018. Pages 685–690.

*Content from this work may be used under the term of the Creative Commons Attribution-Non-commercial (CC BY-NC) 4.0 licence. This license allows refusers to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator. Any further distribution of this work must maintain attribution to the creators. © copyright at IJIRID.* ***DOI: 10.5281/zenodo.10822621*** 