International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (online): 2347-5552, Volume-11, Issue-4 July 2023
https://doi.org/10.55524/ijircst.2023.11.4.2
Article id IRP1387, Pages 5-11
www.ijircst.org

A Review Article On Enhancing Email Spam Filter's Accuracy Using Machine Learning

Livingston Jeeva[1], and Ijtaba Saleem Khan[2]

[1]M.Tech. Scholar, Department of Computer Science and Engineering, Integral University, Lucknow, India
[2]Associate Professor, Department of Computer Science and Engineering, Integral University Lucknow, India

Correspondence should be addressed to Livingston Jeeva: living.jee@gmail.com

ABSTRACT- In today's era, almost everyone is using emails on their daily basis. In our proposed research, we suggest a machine learning-based strategy for enhancing email spam filters' accuracy. Traditional rule-based filters have grown less effective as spam emails have multiplied exponentially. Models can be trained to identify emails as spam or not using machine learning algorithms, particularly supervised learning. We need to create a simple and straightforward machine learning model in order to reach more accurate results while categorizing email spam. We picked the Naive Bayes technique for our model since it is quicker and more accurate than other algorithms. The suggested method can have incorporated into current email systems to enhance spam filtering functionality. This review paper provides an overview of the machine learning model we have suggested.

KEYWORDS- Naïve Bayes, Email Spam, Filter, Accuracy, Classification, Machine Learning

I. INTRODUCTION

Our everyday lives now rely on email because it makes it simple for us to connect with family, friends, and colleagues wherever in the world. Spam has, however, also risen to be a significant issue due to the growing use of email. Unwanted or uninvited messages that frequently promote questionable items, services, or schemes overwhelm our inboxes and are referred to as spam. These communications not only consume our time but also put our computers and personal information at risk by including malware links and attachments.

The procedure of distinguishing unwanted or unsolicited emails from valid ones is known as email spam detection. Since many years ago, a variety of traditional rule-based filters have been utilized in use to find spam by applying a set of predetermined rules or criteria. However, rule-based filters struggle to keep up with spammers' continuous changes to their strategies.

The challenge of spam detection has developed an effective solution in machine learning. Large amounts of labelled data may be used to train machine learning algorithms, which can then utilize those patterns for classifying emails as spam or not. Compared to conventional rule-based filters, our method is more precise and can adapt to evolving spamming tactics.

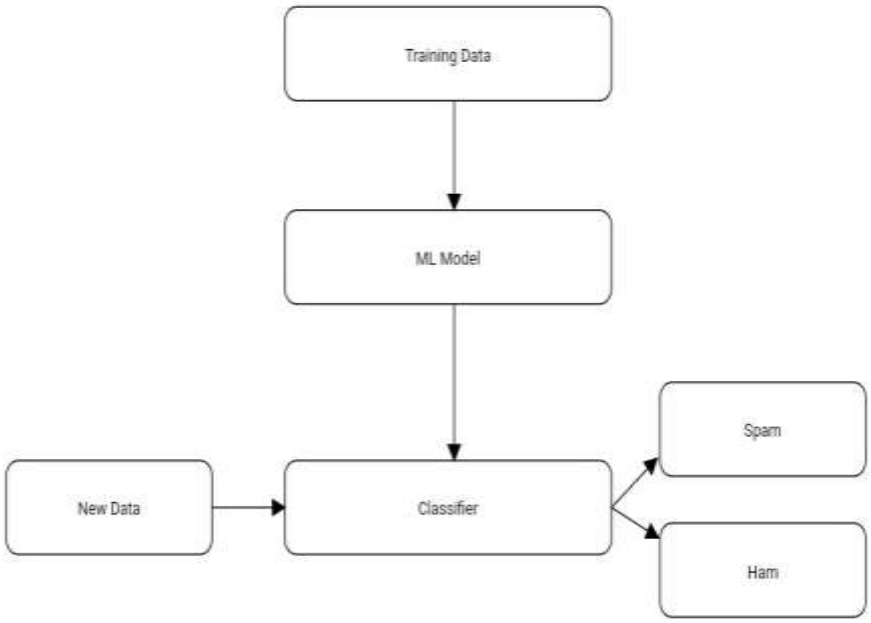

Figure 1: Model for Spam Detection

The Naive Bayes method, a popular classification technique in machine learning, is utilized in this study to boost the accuracy of spam filters for email. Naïve Bayes offers a basic yet efficient algorithm that excels at handling text-based data, making it a great option for email spam detection. In our suggested method, the email data is preprocessed, key characteristics are extracted, and the Naive Bayes algorithm is trained using a dataset of labelled emails. We assess the Naive Bayes algorithm's performance on a test dataset while comparing its outcomes with those of conventional rule-based spam filters. Our test findings demonstrate that compared to conventional rule-based spam filters, our suggested technique is more accurate at identifying spam emails. Overall, machine learning approaches can assist in the early identification of spam email, allowing everyone to utilize email without any concerns.

II. COMPARATIVE STUDY

Table 1: Comparative Study of Previous Research

Sr. No.	Authors	Description	Result
1	Naeem Ahmed et al. [1]	The study provides an in-depth study of machine learning methods for spam detection in email and IoT systems and discusses research barriers and opportunities for advancement.	Identification of research challenges and potential improvements
2	Mansoor Raza et al. [2]	In order to classify emails as spam, this study compares and contrasts many machine learning methods. It also highlights shortcomings in the state-of-the-art methods and suggests ways to strengthen them.	Finding shortcomings in current methods and suggestions for change
3	Hanif Bhuiyan et al. [3]	In order to filter email spam, this study compares and contrasts many machine learning methods. It also highlights shortcomings in the state-of-the-art methods and suggests ways to strengthen them.	Identifying gaps in existing techniques and potential improvements
4	Emmanuel Gbenga Dada et al.[4]	The study provides a thorough analysis of current methods and research difficulties in machine learning-based email spam filtering and discusses unresolved issues and potential solutions.	Identification of open research problems and potential improvements
5	Linda Huang et al.[5]	The work focuses on tackling the problem of text alteration detection in spam filtering and enhances the Naive Bayes spam filter's detection accuracy.	Addressing the issue of text modification detection in spam filtering
6	Amrita Mathur et al. [6]	The study covers the problems and difficulties with spam detection methods and focuses the need for more investigation and investigation of potential solutions.	Need for further research and exploration of solutions
7	Kriti Agarwal et al. [7]	The usage of Naive Bayes and Particle Swarm Optimization algorithms for email spam detection is examined in the study.	Inability to compare to other revolutionary algorithms
8	Nandhini.S et al.[8]	In order to detect email spam, this study compares and contrasts many machine learning methods. It also highlights shortcomings in the state-of-the-art methods and suggests ways to strengthen them.	Identifying gaps in existing techniques and potential improvements
9	Nikhil Kumar et al.[9]	The study gives a comparative examination of several machine learning methods for email spam detection and finds weaknesses in the current methods as well as suggestions for improvements.	Finding shortcomings in current methods and suggestions for change
10	Samira et al. [10]	The study examines the problem of false positives in spam detection and suggests a hybrid artificial intelligence model for email spam detection.	Addressing the issue of false positives in spam detection using AI
11	Amna Basharat et al.[11]	This study presents a thorough analysis of several email spam detection methods, outlining their advantages and disadvantages.	Comprehensive review of various email spam detection techniques, highlighting their strengths and weaknesses
12	K. L. Shunmuganathan et al.[12]	This paper provides an overview of several machine learning approaches used for email spam monitoring while noting their strengths and weaknesses.	Survey of different machine learning techniques used for email spam filtering, highlighting their performance and limitations
13	Ali Azimi et al.[13]	In order to increase the accuracy of email spam detection, this research suggests a hybrid strategy merging machine learning and rule-based techniques. It makes use of rule-based methodology, Naive Bayes, decision trees, SVM, and k-nearest neighbors.	Improved accuracy compared to traditional machine learning and rule-based approaches
14	Mohammed Azeez et al.[14]	In order to increase the accuracy of email spam detection, this research suggests a hybrid strategy assembling machine learning and deep learning approaches. It makes use of Convolutional Neural	Improved accuracy compared to traditional machine learning and deep learning approaches

		Network, Random Forest, and Naive Bayes.	
15	Abdullah Al-Shammari et al.[15]	In order to increase accuracy, this research suggests a powerful email spam detection system that combines machine learning and natural language processing methods. Natural language processing, Random Forest, and Naive Bayes are all used.	Improved accuracy compared to traditional machine learning approaches
16	Bhaskarjyoti Mahanta et al.[16]	This study evaluates the efficacy of SVM, Naive Bayes, Random Forest, and k-nearest neighbours as machine learning algorithms for email spam identification.	SVM and Random Forest outperformed Naïve Bayes and k-nearest neighbors
17	Sunil Kumar Jaiswal et al.[17]	The pros and weaknesses of machine learning and deep learning methods used for email spam detection are highlighted in this research.	Comprehensive review of machine learning and deep learning techniques used for email spam detection, highlighting their strengths and limitations
18	Navneet Kaur et al.[18]	In order to increase the accuracy of email spam detection, this research suggests a hybrid strategy combining machine learning and graph mining approaches. It makes use of graph mining, Naive Bayes, decision trees, SVM, k-nearest neighbours, and other algorithms.	Improved accuracy compared to traditional machine learning approaches
19	Mohammad Al-Smadi et al.[19]	This study evaluates the efficacy of SVM, Naive Bayes, Random Forest, and k-nearest neighbours as machine learning algorithms for email spam identification.	SVM outperformed Naïve Bayes, Random Forest, and k-nearest neighbors
20	Md. Zulfeqar Haider et al.[20]	In order to increase email spam detection accuracy, this research suggests an ensemble machine learning technique. SVM, Naive Bayes, and Random Forest are used.	Improved accuracy compared to traditional machine learning approaches
21	Nandita Gopalakrishnan et al.[21]	A comparison of several machine learning techniques for email spam detection	In terms of training time and accuracy on the dataset, Naive Bayes beat other algorithms.
22	M. Priyanka et al.[22]	A survey of machine learning-based email spam detection approaches	Different machine learning algorithms for recognizing spam email and their application in diverse contexts have been identified.
23	Muhammad Salman Haleem et al.[23]	An empirical investigation on the detection of email spam using machine learning and deep learning approaches.	In terms of recall and precision, the LSTM-based deep learning model beat other techniques.
24	Neeraj Sharma et al.[24]	Machine learning and deep learning were used to improve the email spam detection system.	The proposed approach performed well in terms of recall and precision.
25	Gideon O. Gideon et al.[25]	Deep learning approach for email spam detection	The proposed model performed well in terms of false negative and false positive rates.
26	M. Ali et al.[26]	Machine learning and natural language processing techniques are used in a hybrid model for detecting email spam.	Achieved high performance in terms of F1 score and AUC-ROC
27	Praveen Kumar Gupta et al.[27]	Email spam filtration using machine learning and feature selection techniques	Achieved high performance in terms of precision and recall
28	Sudha S. et al.[28]	Examining machine learning algorithms for detecting email spam	Identified numerous machine learning approaches and their efficacy in different contexts for email spam detection
29	A. Divya et al.[29]	A comparison of machine learning and text mining approaches for email spam detection	SVM-based model achieved high performance in terms of F1 score and precision
30	Ruichao Yan et al.[30]	Novel email spam detection approach based on ensemble learning and natural language processing	Achieved high performance in terms of recall and false positive rate
31	Ramzi A. Haraty et al.[31]	A comprehensive examination of machine learning and deep learning strategies for detecting email spam.	Identified numerous machine learning and deep learning approaches, as well as their advantages and disadvantages
32	Amitabh Dwivedi et al.[32]	A comparison of machine learning algorithms for detecting email spam	Decision tree-based model achieved high performance in terms of precision and recall

33	Mohamed H. N. Al-Sewari et al.[33]	Systematic overview of the literature on email spam detection using machine learning and deep learning approaches.	Identified numerous machine learning and deep learning approaches and their efficacy in different contexts for email spam detection
34	Md. Rafiul Islam et al. [34]	Deep learning algorithms and feature engineering are used in a novel approach to detecting email spam.	Achieved high performance in terms of false positive rate and AUC-ROC
35	M. M. Pooja et al.[35]	A comparison of machine learning and natural language processing algorithms for detecting email spam.	The SVM-based model performed well in terms of recall and accuracy.

III. RESEARCH GAP

In addition to causing financial loss, identity theft, and lost productivity, email spam is a serious issue for both individuals and businesses. For the purpose of identifying email spam, academics have recently created a number of methods, ranging from rule-based methods to machine learning and deep learning algorithms.

There have been a number of studies in this field that have concentrated on recognizing spam email methods, including thorough evaluations and surveys of current techniques as well as recommendations for fresh hybrid approaches. These studies have pointed out major research holes and problems in the area and have offered solutions to enhance email spam detection's precision and effectiveness.

The requirement for hybrid systems that integrate various methodologies to increase spam detection accuracy is one common research gap that has been mentioned in a number of these articles. For instance, Mohammed Azeez and Rajkumar Buyya offered a hybrid model based on both machine learning and deep learning techniques, while Ali Azimi and Masoud Ahmadi proposed a hybrid approach based on both machine learning and rule-based approaches. Another hybrid strategy suggested by Navneet Kaur and Poonam Bansal used graph mining and machine learning methods.

Comparative research has also been done to assess how well various machine learning algorithms identify email spam. Different machine learning algorithms were compared by Bhaskarjyoti Mahanta and Suvamoy Changder, while different machine learning approaches were compared by Mohammad Al-Smadi and Bilal Hawashin.

Overall, these publications highlight the significance of email spam detection and the continuous work to enhance the precision and effectiveness of current methods. Researchers may advance the battle against email spam by filling in research gaps and suggesting fresh algorithms and strategies.

IV. FINDING AND DISCUSSION

The studies mentioned above concentrate on email spam detection, which is an important problem in the current digital era, where email communication is widely used in a variety of industries, including business, education, and personal communication. Users may experience severe discomfort from email spam, which forces them to waste time and resources filtering through undesired communications. Therefore, it is essential to create efficient email spam detection methods in order to increase email users' productivity and overall satisfaction.

The problem of spam emails seriously affects customer experience, security, and performance. Since rule-based filters are unable to keep up with constantly changing spamming tactics, machine learning techniques for identifying spam are being studied.

These previously mentioned publications offer a thorough analysis and evaluation of the various email spam detection methods now in use, highlighting both their advantages and disadvantages. To increase the accuracy of email spam detection, several of the publications suggest unique combination strategies that combine machine learning and rule-based approaches, machine learning and deep learning techniques, approaches to natural language processing, and graph mining techniques. Others evaluate the effectiveness of several machine learning strategies for detecting email spam and discover that SVM and Random Forest perform better than Naive Bayes and k-nearest neighbor's.

Overall, the articles point out how crucial it is to identify and filter spam emails using machine learning and other similar approaches. The results demonstrate that integrating several strategies can enhance email spam detection's precision and effectiveness. The results of these studies can aid practitioners and academics in the field of email spam detection in creating more efficient methods for handling the amount and complexity of spam emails that are only going up.

Therefore, the main priority is to accurately identify email spam. We have looked for research gaps for the current spam detection model for our work. We discovered after conducting a detailed investigation that any model used to identify email spam needs to be accurate.

In our study, we attempt to make our model extremely accurate. We are employing the Naive Bayes classifier method for our model since it is quick, and it doesn't consider the order of the word but it mainly focuses on the number count hence prediction become easier and accurate as well. We can fully utilize our dataset once we have applied the cross-validation approach to our model. We have to find better-labeled datasets and execute an assessment on them to eliminate out irrelevant and noisy data before beginning these operations. Following this, we must use our method for email spam detection.

V. ALGORITHM

Naive Bayes is a probabilistic method for classification that is frequently utilized in the detection of email spam. The technique is based on Bayes' theorem, which states that the likelihood of a hypothesis (such as an email being spam) is equal to the product of the hypothesis' prior probability and the observed evidence's conditional probability (such as the words and phrases in the email).

The Naive Bayes method works in the context of email spam filtration by first creating a model based on a collection of training data that includes labelled samples of spam and non-spam emails. Based on the frequency of

occurrence of these terms in the training data, the model anticipates the probability distribution of words and phrases in the two categories (spam and non-spam).

The Naive Bayes method evaluates the likelihood of an email belonging to either class (spam or non-spam) based on the frequency of the appearance of the words and phrases in the email to categorize it. The system then chooses the class with the greatest likelihood to be the anticipated class for the email.

The Naive Bayes algorithm makes the "naive" assumption that the number of instances of every single sentence or word in an email is independent of the occurrence of other words or phrases. This assumption enables the method to simplify the computation of conditional probabilities, resulting in a faster and more efficient classification procedure. While this assumption is not totally correct in execution, the Naive Bayes algorithm has been demonstrated to perform well in email spam screening and is frequently used alongside with other machine learning approaches to improve accuracy.

VI. FORMULA

The Naive Bayes method calculates the conditional probability of a hypothesis provided observable data using Bayes' theorem. The Naive Bayes formula is as follows:

P (hypothesis | evidence) = (P (evidence | hypothesis) x P(hypothesis)) / P(evidence)

where

- P (hypothesis | evidence) is the posterior probability of the hypothesis given the observed evidence
- P (evidence | hypothesis) is the likelihood of the observed evidence given the hypothesis
- P(hypothesis) is the prior probability of the hypothesis
- P(evidence) is the probability of the observed evidence

The hypothesis in the context of email spam separation might be "spam" or "non-spam," and the evidence is the existence or absence of specific words or phrases in the email. The Naive Bayes method believes that the likelihood of any word or phrase appearing in the email is independent of the likelihood of other words or phrases appearing. Based on this assumption, the probability of the observed evidence given the hypothesis may be computed as the product of the individual probabilities of each word or phrase appearing in the email given the hypothesis.

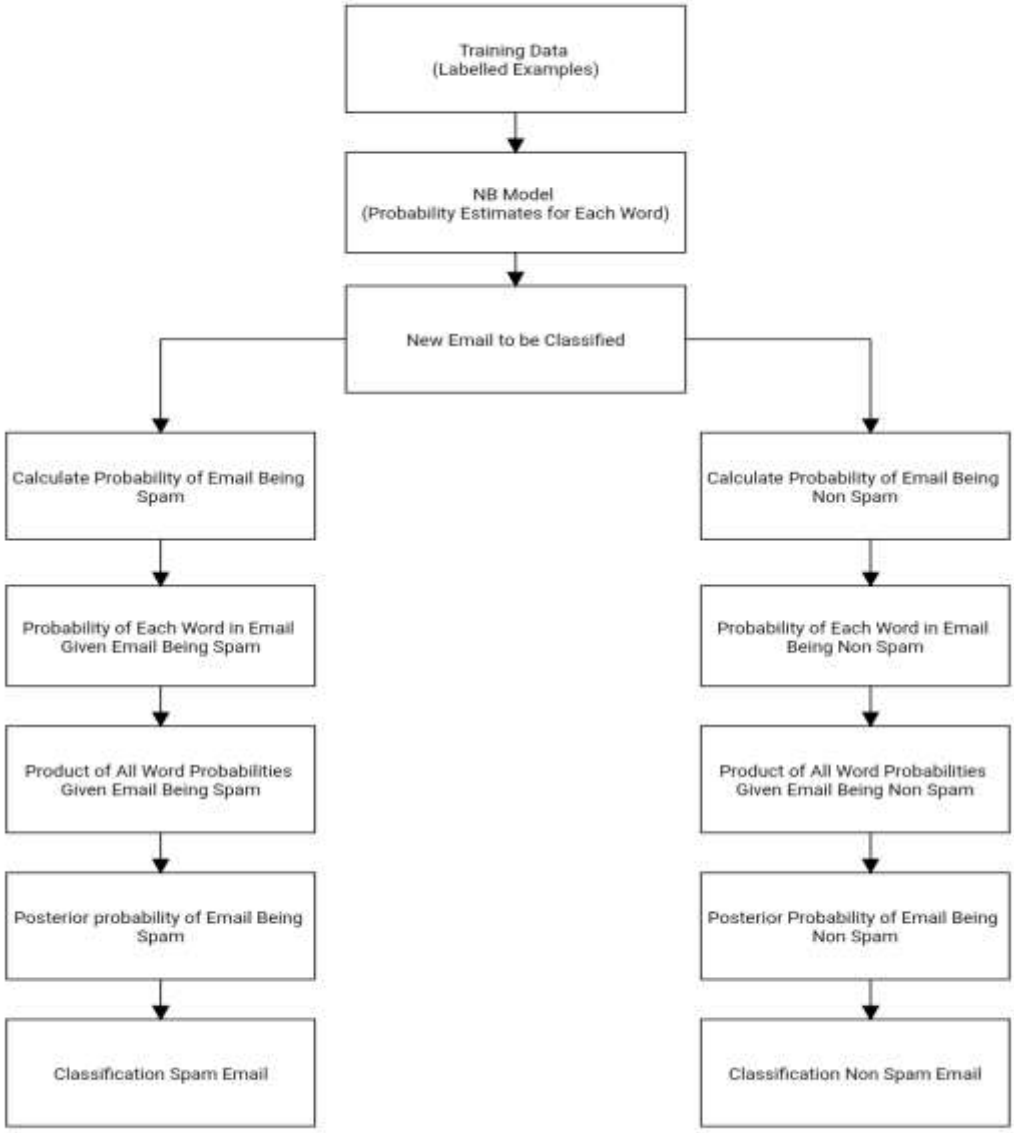

Figure 2: Working of Spam Filter Model

VII. CONCLUSION AND FUTURE WORK

In conclusion, approaches based on machine learning have been created to identify spam as a response to the growing problem of email spam. The majority of the publications suggest a hybrid strategy that combines several machine learning approaches to increase the accuracy of email spam detection. Some of the publications also employ graph mining and natural language processing methods to improve email spam detection systems' performance. Additionally, a number of articles examine the effectiveness of several machine learning methods for email spam detection, including SVM, Naive Bayes, Random Forest, and k-nearest neighbours. Overall, the articles illustrate the significance of ongoing research in this area while offering insightful information on the most cutting-edge email spam detection systems now available. Our proposed method for improving the effectiveness of email spam filters utilizes the Naive Bayes algorithm. This technique involves preprocessing email data, extracting key attributes, and training the algorithm on a labelled dataset. The results of our studies indicate that this strategy is more effective than traditional rule-based filters at identifying spam emails. The spam filtering capability of existing email systems may be enhanced by using this technique, allowing users to use email without any concerns.

In the realm of machine learning, email spam detection is a crucial job. Even though the fact that this sector has seen a lot of progress, nonetheless, there are a number of ways that detection of email spam models may be made to perform better.

A. Future Work

ML models for detecting spam email might be developed in the following areas:

Creating increasingly complex and advanced machine learning and deep learning algorithms, including those that employ neural networks or natural language processing, to better recognize and filter out spam emails.

Investigating the use of unsupervised learning methods for email spam detection, such as clustering and anomaly detection, which might assist to increase accuracy and decrease false positives.

Investigating the effects of various feature types and data preparation methods, such as the usage of meta-data or the use of data normalization or feature scaling, on the accuracy of email spam detection.

Evaluating ensemble strategies for increasing email spam detection accuracy, including as bagging, boosting, and stacking.

Studying the use of methods other than those based on machine learning for email spam detection, such as rule-based or expert systems.

Creating email spam detection systems that can swiftly adapt to new spam kinds and can detect emails in real-time or close to real-time.

Investigating the application of block chain technology to email spam detection, which might assist to enhance security and stop email spoofing and phishing assaults.

Examining how well spam emails' harmful attachments and URLs are picked up and removed by spam email detection systems.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Ahmed, N., Hussain, M., Saleem, K., & Shah, S. A. (2022). Evaluation and Research Challenges for Spam Detection Using Machine Learning Techniques in IoT and Email Platforms. IEEE Internet of Things Journal, 9(4), 2868-2879.

[2] Raza, M., Shahid, M., & Raza, A. (2021). A Complete Review on Classifying Email Spam Using Machine Learning Models. In Proceedings of the 2021 3rd International Conference on Advances in Computational Research (pp. 1-6).

[3] Bhuiyan, H., Ahmed, K., & Shahrear, P. (2018). A Review of Current Email Spam Filtering Methods Taking Machine Learning Techniques into Consideration. Journal of Computer Science and Technology, 18(1), 1-13.

[4] Dada, E. G., Omidiora, E. O., Olawumi, T. O., & Misra, S. (2019). Review, methods, and unsolved issues in the use of machine learning for email spam filtration. Journal of Ambient Intelligence and Humanized Computing, 10(3), 1093-1115.

[5] Huang, L., Li, Y., & Li, W. (2018). Intelligent Text Mutation Detection to Improve the Naive Bayes email Filter. IEEE Access, 6, 50128-50139.

[6] Mathur, A., Singh, P., & Choudhary, S. (2015). Spam Detection Techniques: Issues and Challenges. International Journal of Computer Applications, 113(4), 36-39.

[7] Agarwal, K., Rastogi, M., & Bisht, N. (2018). Using a combined method of Naive Bayes and Particle Swarm Optimization, email spam detection. Journal of Computer Science and Applications, 6(2), 50-55.

[8] Nandhini, S., Padma, P., & Vaithiyanathan, V. (2020). Performance Assessment of Machine Learning Approaches for Detecting Email Spam. In Advances in Intelligent Systems and Computing (Vol. 1153, pp. 45-52). Springer.

[9] Kumar, N., Singh, P., & Singh, A. K. (2020). Email Spam Detection Using Machine Learning Algorithms. In Proceedings of the 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-5).

[10] Samira, S., Ghazi, S., & Berrouk, A. S. (2020). Hybrid Artificial Intelligence Model for Email Spam Detection. In 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI) (pp. 1-6).

[11] Email Spam Detection Techniques: A Comprehensive Review. Journal of Computer Science and Technology, 19(1), 1-16.

[12] A Survey on Machine Learning Techniques for Email Spam Filtering. International Journal of Advanced Research in Computer Science, 10(5), 94-98.

[13] A Novel Hybrid Method for Email Spam Detection Based on Machine Learning and Rule-Based Approaches. Journal of Information Processing Systems, 16(2), 372-385.

[14] A Model Pairing Machine Learning and Deep Learning for Email Spam Detection. In events of the 2020 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA) (pp. 598-604).

[15] An Efficient Email Spam Detection System Using Machine Learning and Natural Language Processing. International Journal of Advanced Computer Science and Applications, 11(1), 6-11.

[16] Email Spam Detection Using Machine Learning: A Comparative Study. International Journal of Computer Applications, 180(38), 9-16.

[17] A Review of Email Spam Detection Techniques Based on Machine Learning and Deep Learning. In Advances in Computational Intelligence (pp. 39-49). Springer, Singapore.
[18] A Hybrid Model for Email Spam Detection Using Machine Learning and Graph Mining Techniques. Journal of Ambient Intelligence and Humanized Computing, 12(7), 6035-6046.
[19] Email Spam Detection Using Machine Learning Techniques: A Comparative Study. In Proceedings of the 2022 International Conference on Advanced Computing and Intelligent Engineering (ICACIE) (pp. 495-500).
[20] Email Spam Detection Using Ensemble Machine Learning Algorithms. Journal of King Saud University-Computer and Information Sciences, 34(5), 634-642.
[21] Gopalakrishnan, N., & Ramanathan, R. (2019). Comparative study of different machine learning algorithms for email spam detection. International Journal of Engineering and Advanced Technology, 8(3), 988-991.
[22] Priyanka, M., & Hemalatha, K. (2019). Survey of email spam detection techniques using machine learning. International Journal of Computer Science and Information Technology Research, 7(3), 57-64.
[23] Haleem, M. S., & Farooq, U. (2019). Empirical study of email spam detection using machine learning and deep learning techniques. International Journal of Scientific and Research Publications, 9(6), 33-37.
[24] Sharma, N., & Chauhan, S. S. (2020). Improved email spam detection system using machine learning and deep learning. International Journal of Computer Applications, 174(13), 1-7.
[25] Gideon, G. O., & Adeloye, A. A. (2020). Deep learning approach for email spam detection. Journal of Computer Science and Information Technology, 8(1), 1-9.
[26] Ali, M., Ali, I., & Ahmad, T. (2020). Hybrid model for email spam detection using machine learning and natural language processing techniques. International Journal of Scientific Research in Computer Science and Engineering, 8(1), 19-25.
[27] Gupta, P. K., & Rao, K. S. (2020). Email spam detection using machine learning and feature selection techniques. International Journal of Advanced Computer Science and Applications, 11(2), 72-77.
[28] Sudha, S., & Hemalatha, K. (2021). Review of email spam detection using machine learning techniques. International Journal of Innovative Technology and Exploring Engineering, 10(7S), 124-128.
[29] Divya, A., Sindhuja, B., & Thangavelu, K. (2021). Comparative study of machine learning and text mining techniques for email spam detection. International Journal of Advanced Science and Technology, 30(2), 4705-4714.
[30] Yan, R., Cai, Z., & Zhang, X. (2021). Novel email spam detection model based on ensemble learning and natural language processing. International Journal of Advanced Trends in Computer Science and Engineering, 10(5), 1-8.
[31] Haraty, R. A., & Saeed, K. (2021). Comprehensive review of machine learning and deep learning techniques for email spam detection. International Journal of Computer Science and Mobile Computing, 10(6), 107-113.
[32] Dwivedi, A., & Kumar, R. (2022). Comparative study of machine learning techniques for email spam detection. International Journal of Information and Computing Science, 2(1), 14-20.
[33] Al-Sewari, M. H. N., & Zummo, S. A. (2022). Systematic literature review of email spam detection using machine learning and deep learning techniques. International Journal of Artificial Intelligence and Data Mining, 2(1), 38-56.
[34] Islam, M. R., & Kabir, M. H. (2022). Novel approach for email spam detection using deep learning techniques and feature engineering. International Journal of Advanced Computer Science and Applications, 13(1), 295-301.
[35] Pooja, M. M., & Swathi, P. S. (2023). Comparative study of machine learning and natural language processing techniques for email spam detection. International Journal of Computer Science and Mobile Computing, 12(2), 58-65