Spam Detection in Emails Using Machine Learning Techniques: A Review

Stanley Munga Ngigi[1]
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
Email: sngigi [AT] kyu.ac.ke

Josphat Karani[2]
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
Email: jkarani [AT] kyu.ac.ke

Richard Mathenge[3]
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
Email: mathengerichard29 [AT] gmail.com

Nicholus Muriithi[4]
School of Pure and Applied Sciences
Kirinyaga University, Kutus, Kenya
Email: nickmuri123 [AT] gmail.com

Abstract—**Despite the vast amounts of data available within email communication systems, spam remains a persistent issue, posing challenges for both users and organizations. Analyzing this data holds the potential to develop more effective methods for detecting and mitigating spam emails. However, extracting actionable insights from this data and leveraging them to construct robust spam detection systems presents a significant challenge. Traditional approaches to combating spam, such as rule-based filtering and heuristic methods, have become increasingly inadequate due to the evolving tactics of spammers. Machine learning techniques offer a promising solution by enabling the training of predictive models using historical email data. However, the effectiveness of these models is influenced by factors such as class imbalance and the identification of relevant features essential for spam detection. This paper provides a comprehensive review of various machine learning techniques employed in spam detection within email communication systems. By examining the strengths and weaknesses of different approaches, we aim to identify strategies for improving the efficiency and accuracy of spam detection. Additionally, we propose a spam detection framework centered around ensemble learning models trained on balanced datasets using techniques like SMOTE, and featuring only the most relevant features. This approach is intended to enhance detection performance while reducing false positives, thereby offering a more effective solution to the challenge of spam detection in email systems.**

Keywords— Spam, Class Imbalance, SMOTE, Feature Selection, Ensemble Learning

I. INTRODUCTION

Email communication has become an integral part of modern life, serving as a primary mode of interaction for individuals and businesses alike. (Kipkebut et al., 2019) However, alongside the convenience of email communication comes the persistent challenge of spam. Spam, characterized by unsolicited and often fraudulent or malicious emails, poses a threat to the efficiency, security, and user experience of email systems (Ahmed et al., 2022).Spam detection, therefore, plays a crucial role in safeguarding users and organizations from the detrimental effects of spam emails. By identifying and filtering out spam messages, email users can focus on legitimate correspondence, while organizations can mitigate the risks associated with phishing attacks, malware distribution, and other spam-related threats.(Faris et al., 2019)Machine learning techniques have emerged as a promising approach to address the complexities of spam detection. (Kumar et al., 2020)Leveraging the vast amounts of email data available, machine learning models can be trained to distinguish between spam and legitimate emails with high accuracy. (Ahmed et al., 2022)However, achieving effective spam detection requires overcoming various challenges, including class imbalance, feature selection, and the dynamic nature of spamming tactics. This review aims to provide an overview of the machine learning techniques utilized in spam detection within email communication systems. By examining the strengths and limitations of different approaches, this paper seeks to identify strategies for improving the efficiency and accuracy of spam detection. Specifically, the focus will be on supervised learning models, which have shown promise in distinguishing between spam and non-spam emails based on labeled training data.(Krishnamoorthy et al., 2024)The rest of this paper is organized as follows: Section II outlines the methodology employed in conducting this review. In Section III, we delve into the fundamental concepts of spam, spamming tactics, and the challenges posed by spam emails. Section IV discusses classical approaches to spam detection, while Section V provides a comprehensive review of machine learning techniques and their application in spam detection, Section VI literature review. Finally, the paper concludes with a summary of key findings and avenues for future research in Section VII and VIII.

II. METHODOLOGY

The primary objective of this study was to systematically investigate various machine learning techniques employed in the detection of spam emails within email communication systems. We aimed to assess the effectiveness of different approaches in distinguishing between spam and legitimate emails, while also identifying their strengths and weaknesses. To conduct this review, we utilized a structured methodology to gather relevant literature from reputable academic databases. Our search encompassed Google Scholar, IEEE Xplore, Science Direct, Springer Link, and Elsevier. The search criteria were designed to retrieve papers discussing spam detection in email communication systems specifically focusing on the utilization of machine learning techniques. In the process of selecting articles for review, we initially screened papers based on their abstracts. Specifically, we targeted papers that addressed topics such as spam detection, machine learning, and email communication systems. Additionally, we ensured that selected papers discussed the challenges posed by spam, the techniques utilized for spam detection, and the application of machine learning algorithms in this domain. To maintain the quality and relevance of the review, we excluded sources such as book chapters, magazines, and review papers published more than seven years ago (prior to 2017). Furthermore, classical papers reviewing machine learning techniques were omitted to focus on more recent developments in the field. After screening and filtering the literature, we identified a total of 15 papers deemed suitable for inclusion in this review. These papers were subsequently analyzed in detail to extract insights into the various machine learning techniques employed in spam detection and to evaluate their efficacy in addressing the challenges posed by spam emails.

III. FUNDAMENTAL CONCEPTS OF SPAM, SPAMMING TACTICS, AND CHALLENGES

Spam, within the context of email communication, encompasses unsolicited and often unwanted messages sent in bulk to a large number of recipients. (Jáñez-Martino et al., 2023).The proliferation of spam emails poses a significant challenge to both users and organizations, leading to concerns regarding privacy, security, and productivity. Spamming tactics employed by perpetrators of spam emails are diverse and constantly evolving. Common tactics include the dissemination of promotional content, phishing attempts aimed at stealing sensitive information such as login credentials and financial data, and the distribution of malware through malicious attachments or links embedded within emails.(Gupta et al., 2021) Additionally, spammers may employ techniques such as email spoofing, where the sender's address is forged to appear as a legitimate source, further complicating detection and mitigation efforts. Challenges in combating spam emails arise from several factors. One major challenge is the sheer volume of spam messages received daily, which can overwhelm email filters and consume valuable resources (Mallampati & Hegde, 2023). Furthermore, spammers often employ sophisticated techniques to evade detection, such as obfuscating text or using image-based spam that bypasses text-based filters. The dynamic nature of spamming tactics requires constant adaptation and refinement of detection mechanisms to stay ahead of emerging threats.

Moreover, the diverse nature of spam content makes it challenging to develop universal detection algorithms capable of accurately identifying all spam emails while minimizing false positives (legitimate emails incorrectly classified as spam) and false negatives (spam emails incorrectly classified as legitimate). Addressing these challenges requires a multi-faceted approach that leverages advanced machine learning techniques, robust feature engineering, and continuous monitoring and adaptation to evolving spamming tactics. In the following sections, we will explore various machine learning techniques utilized in spam detection within email communication systems, examining their efficacy in addressing the challenges posed by spam emails and enhancing the overall security and usability of email platforms.

IV. CLASSICAL APPROACHES TO SPAM DETECTION

Numerous classical approaches have been employed in the realm of spam detection to predict, identify, and mitigate the onslaught of spam emails.(Akinyelu, 2021) These approaches often rely on human expertise, intuition, and domain knowledge to combat the pervasive nature of spam. One effective method utilized by email service providers is the implementation of spam filters, which employ predefined rules and heuristics to categorize incoming emails as either spam or legitimate. (Mallampati & Hegde, 2023)These filters analyze various attributes of emails, such as sender reputation, email headers, and content characteristics, to make classification decisions. Another classical approach involves the use of blacklists and whitelists, which maintain lists of known spam sources and trusted senders, respectively. (Azeez et al., 2021)Emails originating from blacklisted sources are automatically flagged as spam, while those from whitelisted sources are deemed legitimate. However, these approaches may be limited in their effectiveness as spammers continually change tactics and evade detection. Furthermore, (van Meteren & van Someren, 2022.)Content-based filtering techniques analyze the textual content of emails to identify patterns indicative of spam. This may involve keyword analysis, linguistic analysis, and the detection of common spam phrases or characteristics. While content-based filtering can be effective, it may also lead to false positives if legitimate emails contain similar language or content. Additionally, statistical approaches such as Bayesian filtering utilize probabilistic models to assess the likelihood that an email is spam based on observed features and past occurrences(Chen, n.d.2021). Bayesian filtering calculates the probability of an email being spam given its characteristics, incorporating feedback from users to continuously refine the classification model. While these classical approaches have demonstrated some degree of effectiveness in spam detection, they also have limitations. (Sokhangoee & Rezapour, 2022)They may struggle to adapt to evolving spamming tactics, require manual tuning and maintenance, and can produce false positives or false negatives. In the following sections, we will explore how machine learning techniques can complement these classical approaches and offer more robust solutions to the challenge of spam detection in email communication systems.

V. MACHINE LEARNING TECHNIQUES FOR SPAM DETECTION

Machine learning serves as a powerful tool in the arsenal of methods for spam detection, allowing computers to learn from data without explicit programming.(Baecker et al., 2021) This process involves training algorithms on labeled datasets and creating predictive models based on learned patterns. (Singh et al., 2017)Machine learning encompasses various learning paradigms, including supervised, unsupervised, semi-supervised, and deep learning.(Ge et al., 2017)

1.Supervised Learning: Supervised learning entails the use of labeled datasets, where each email instance is tagged as either spam or legitimate. (Ligthart et al., 2021)Supervised models are trained on extensive labeled data to classify incoming emails accurately.
2.Unsupervised Learning: In unsupervised learning, algorithms analyze unlabeled datasets to group emails based on similarities (clustering) or uncover hidden relationships between them (association). (Liu, 2020)This approach can aid in identifying patterns indicative of spam.
3.Semi-supervised Learning: (Yang et al., 2023)Semi-supervised learning involves training models on datasets that contain both labeled and unlabeled examples. This hybrid approach can leverage the available labeled data while also extracting insights from the unlabeled portion, potentially enhancing the detection of subtle spam patterns.
4.Deep Learning: Deep learning employs artificial neural networks with multiple layers to automatically extract intricate representations of the data.(Sarkar et al., 2018) This technique is particularly adept at feature extraction and can enhance the detection of subtle patterns indicative of spam, making it a promising approach for spam detection in email communication systems.

Given the objective of flagging emails as either spam or legitimate, supervised learning techniques are predominantly employed in spam detection. However, the integration of semi-supervised and deep learning methodologies offers additional avenues for improving the accuracy and robustness of spam detection systems.

VI. LITERATURE REVIEW

While building a spam detection model, the sampling methodology employed, parameter selection, and identification techniques used all have a significant impact on the effectiveness of the model. In this section, several researches relating to spam detection and prediction using machine learning techniques have been reviewed.

(Nthurima et al., 2023) combined two random forest & to detect and prevent phishing attacks, design and develop a supervised classifier which can detect phishing and prevent phishing emails and test the model with existing data. A dataset consisting of both benign and phishing emails will be used to conduct a supervised learning by the model. Expected accuracy is 99.9%, False Negative (FN) and False Positive (FP) rates of 0.1% and below.

According to (Kipkebut et al., 2019), SMS spam causes significant financial losses for mobile users in Kenya. They propose using machine learning, particularly the Naive Bayes algorithm, for client-side SMS spam detection. Their study, conducted using data from mobile service providers and users, achieved an impressive 96.1039% classification accuracy.

(Das et al. 2021), used various techniques to filter spam emails, including non-machine learning methods such as list-based filtering and content-based filtering, as well as machine learning methods. Machine learning techniques utilize lists of email content, email addresses, and IP addresses to detect incoming unknown mails. The results of experiments using machine learning techniques showed an accuracy of 99.72% for the Random Forest algorithm and 78.94% for the Decision Tree algorithm.

(AbdulNabi & Yaseen, 2021) addressed the challenge of detecting unsolicited emails, focusing on phishing and spam emails, which incur significant financial losses annually. They proposed a BERT-based approach for classifying spam emails, comparing it with a baseline DNN model and classic classifiers like k-NN and NB. The BERT model achieved the highest accuracy of 98.67% and an F1 score of 98.66%.

(V.Christina et al., 2010) employed supervised machine learning techniques such as Decision tree classifier, Multilayer Perceptron and Naïve Bayes Classifier to filter the email spam messages. The machine learning techniques are used in learning the features of spam emails and the model is built by training with known spam emails and legitimate emails. The results of the experiment using the supervised machine learning techniques, showed an accuracy of 98.6% for Naïve Bayes, 96.6% for Decision Tree classifier and 99.3% for Multilayer perception.

(Kontsewaya et al., 2021) focused on reducing spam through a classifier implementation. They employ various machine learning algorithms, including Naive Bayes, K-Nearest Neighbors, SVM, Logistic Regression, Decision Tree, and Random Forest, to analyze email text for spam detection. Training on an existing dataset reveals that Logistic Regression and Naive Bayes achieve the highest accuracy levels, up to 99%. These findings suggest promising avenues for enhancing spam detection classifiers through algorithmic combinations or filtering methods.

(Alhogail & Alsabih, 2021) , proposed a phishing email classifier model utilizing deep learning algorithms, specifically a graph convolutional network (GCN), and natural language processing (NLP) to analyze email body text. Their supervised learning approach yielded promising results, with the classifier achieving a high accuracy rate of 98.2%. Additionally, it demonstrated a low false-positive rate of 0.015. These statistics underscore the effectiveness of their model in detecting phishing emails, emphasizing its potential to bolster email security against such threats.

(Nallamothu & Khan, 2023) surveyed different existing email spam filtering system regarding Machine Learning Technique (MLT) such as Naive Bayes, SVM, K-Nearest Neighbor, Bayes Additive Regression, KNN Tree, and rules. However, here they present the classification, evaluation and comparison of different email spam filtering system and summarize the overall scenario regarding accuracy rate of different existing approaches. They achieved 97.4% precision using SVM

(Mallampati, 2018) employed a supervised machine learning techniques to filter the email spam messages. The supervised machine learning techniques used are Decision tree classifier, Multilayer Perceptron and Naïve Bayes. They used Naïve Bayes Classifier for learning the features of spam emails and the model is built by training the mentioned Classifiers with known spam emails and legitimate emails. They came up with a predicted accuracy, Naïve Bayes Classifier 98.6%, Decision tree classifier 96.6% and Multilayer Perceptron 99.3%

(Roy et al., 2013) presented an efficient spam filter technique to spam email based on Naive Bayes Classifier. They collected a statistical data which they used in training the Bayesian Classifier. This Bayesian filtering works by evaluating the probability of different words appearing in legitimate and spam mails and then classifying them based on those probabilities.

(Kaur & Verma, 2019) discussed the process of filtering the mails into spam and ham using various techniques. This technique are Machine Learning Based Technique (Support Vector Machine, Multi-Layer Perceptron, Naïve Bayes Algorithm, Decision Tree Based etc.) and Non-Machine Learning Based Technique (signature based, heuristic scanning, black and whitelist, sandboxing, mail header scanning. They concluded by saying no algorithm guarantees 100% results in spam detection but still there are some algorithms that provide high accuracy for detection of spam emails when used with feature selection technique like MLP neural network but MLP has a limitation of selecting initial infor mation point using a randomized approach which increases the execution and model building time of the MLP algorithm.

(Mohammed et al., 2019) highlighted some current problems and improved on an anti-spam model. They proposed a new agent-based Multi-Natural Language Anti-Spam (MNLAS) model. The Multi-Natural Language Anti-Spam model process in the spam filtering process of an email handles both visual information such as images and texts in English and Arabic languages. The Jade agent platform and Java environments are employed in the implementation of MNLAS model. The MNLAS model was tested on a 200 emails' dataset and the results showed that it was able to detect and filter various kinds of spam emails with high accuracy of about 93.32%

(Svadasu & Adimoolam, 2022) aimed to detect spam in social media using Support Vector Machine (SVM) and compare its accuracy with Artificial Neural Network (ANN). They utilized a dataset of 5489 messages, with 80% used for training and 20% for testing. The classification was done using K-Nearest Neighbors (KNN) with N=10 and compared with SVM (N=10). Results showed ANN achieved 98.2% accuracy, while SVM reached 96.2%. They concluded that ANN slightly outperformed SVM in spam detection accuracy.

(Hasas et al., 2024) proposed an innovative approach integrating Long Short-Term Memory (LSTM), K-Nearest Neighbors (KNN), and Random Forest techniques for dynamic intrusion detection, highlighting the significance of advanced methodologies in fortifying digital security. Through rigorous evaluation on a robust dataset, the LSTM model achieves an accuracy of 99.11%, demonstrating exceptional proficiency in capturing sequential dependencies within network traffic. KNN exhibits resilience with a high accuracy of 99.23%, while the Random Forest model emerges as the standout performer, boasting an accuracy of 99.63% along with exceptional precision, recall, and F1-score metrics.

(Gayathri et al., 2021) used two ensemble learning algorithms named naive bayes and k-nearest neighbors are applied to data. The algorithms have been implemented and tested over a dataset which consists of 5574 records. Ensemble learning methods combined several models trained with a given learning algorithm to improve accuracy. After performing the experiment as result shows mean accuracy of 88.05 % by using naive bayes algorithm and compared k-nearest neighbor algorithm mean accuracy is 58.04% for SMS spam detection.

VII. SUMMARY OF EXISTING E-MAIL SPAM CLASSIFICATION APPROACHES

In recent decades, researchers have endeavored to enhance the security of email communication. A fundamental aspect of this endeavor is spam filtering, aimed at securing the email platform. Despite numerous research efforts in this domain, there remain untapped potentials. Email spam classification continues to be a significant area of research aimed at addressing these gaps. Consequently, a plethora of studies have been conducted on email spam classification utilizing various machine learning techniques to enhance email efficiency for users. Thus, this paper aims to provide a summarized overview of existing machine learning approaches in this field.

TABLE I Summary: supervised learning techniques for spam detection

Technique	**Author**	**Frequency**	**Accuracy**
Naive Bayes (NB)	(Kipkebut et al., 2019)	1	96.1039% Accuracy Achieved
	(Das et al. 2021)	1	78.94% Accuracy Achieved
	(V.Christina et al., 2010)	1	98.6% Accuracy Achieved
	Kontsewaya et al. (2021),	1	99% Accuracy Achieved
	(Mallampati, 2018)	1	98.6% Accuracy Achieved
	(Gayathri et al., 2021)	1	88.05% Accuracy Achieved
Support Vector Machines (SVM)	(Nallamothu & Khan, 2023)	1	97.4% Accuracy Achieved
	(Svadasu & Adimoolam, 2022)	1	96.2% Accuracy Achieved
Decision Trees	(V.Christina et al., 2010)	1	96.6% Accuracy Achieved
	(Mallampati, 2018)	1	96.6% Accuracy Achieved
Random Forest (RF)	(Nthurima et al., 2023)	1	99.9% Accuracy Achieved
	(Das et al. 2021)	1	99.72% Accuracy Achieved
	(Hasas et al., 2024)	1	99.63% Accuracy Achieved
K-Nearest Neighbour (KNN)	(Hasas et al., 2024)	1	99.23% Accuracy Achieved
	(Gayathri et al., 2021)	1	58.04% Accuracy Achieved
Logistic Regression (LR)	Kontsewaya et al. (2021)	1	99% Accuracy Achieved
BERT model	(AbdulNabi & Yaseen, 2021)	1	86.7% Accuracy Achieved
Multilayer perception	(V.Christina et al., 2010)	1	99.3% Accuracy Achieved

	(Mallampati, 2018)		99.3% Accuracy Achieved
Natural language processing (NLP)	(Alhogail and Alsabih 2021),	1	98.2% Accuracy Achieved
Multi-Natural Language Anti-Spam (MNLAS)	(Mohammed et al., 2019)	1	93.32% Accuracy Achieved
ANN	(Svadasu & Adimoolam, 2022)	1	98.2% Accuracy Achieved
Long Short-Term Memory (LSTM)	(Hasas et al., 2024)	1	99.11% Accuracy Achieved

VIII. DISCUSSION

The review highlights a diverse array of machine learning techniques utilized for spam detection, with Naive Bayes (NB), Random Forest (RF), and K-Nearest Neighbour (KNN) being the most frequently mentioned. These techniques appear multiple times across the reviewed literature, indicating their popularity and relevance in spam detection research.

The accuracy scores provided for each technique underscore their effectiveness in spam detection. Most techniques achieve high accuracy rates, with scores ranging from approximately 58% to almost 100%. Particularly, Random Forest (RF) (Nthurima et al., 2023) and K-Nearest Neighbour (KNN) (Svadasu & Adimoolam, 2022) demonstrate exceptionally high accuracy rates, often exceeding 99%, showcasing their robust performance in distinguishing between spam and legitimate emails.

While some techniques consistently achieve high accuracy across different studies, others display more variability. For example, Naive Bayes (NB) (Kipkebut et al., 2019) and Logistic Regression (LR) (Kontsewaya et al., 2021), exhibit a range of accuracy scores across different studies, suggesting potential sensitivity to dataset characteristics or implementation specifics. This variability highlights the importance of carefully considering the context and nuances of each study when interpreting accuracy metrics.

Certain techniques, such as Support Vector Machines (SVM), Decision Trees, and Multilayer Perception (Svadasu & Adimoolam, 2022) , demonstrate consistent performance across studies, with accuracy scores consistently hovering around the mid to high 90s. This consistency underscores the reliability and robustness of these techniques in spam detection tasks.

The review also discusses the inclusion of emerging techniques such as BERT model, Natural Language Processing (NLP), and Long Short-Term Memory (LSTM). These advanced methods represent the exploration of innovative approaches beyond traditional machine learning algorithms. While these techniques show promise, their varying accuracy scores suggest the need for further refinement and investigation to unlock their full potential in spam detection applications.

Overall, the review provides valuable insights into the landscape of machine learning techniques for spam detection, highlighting both the strengths and limitations of existing approaches. By leveraging these insights, researchers and practitioners can make informed decisions when selecting and implementing spam detection techniques tailored to their specific needs and objectives.

IX. CONCLUSION

This survey paper elaborates on various existing spam filtering systems using machine learning techniques by exploring several methods. It concludes by providing an overview of different spam filtering techniques and summarizing the accuracy of various proposed approaches across several parameters. While all the existing methods are effective for email spam filtering, some yield more effective outcomes than others, while some are exploring additional processes to enhance their accuracy rates. Despite their effectiveness, current spam filtering systems still exhibit some shortcomings, which remain a major concern for researchers. Efforts are underway to develop next-generation spam filtering processes capable of effectively handling a large volume of multimedia data and filtering spam emails more prominently.

REFERENCES

[1] AbdulNabi, I., & Yaseen, Q. (2021). Spam Email Detection Using Deep Learning Techniques. *Procedia Computer Science*, *184*, 853–858. https://doi.org/10.1016/j.procs.2021.03.107

[2] Ahmed, N., Amin, R., Aldabbas, H., Koundal, D., Alouffi, B., & Shah, T. (2022). Machine Learning Techniques for Spam Detection in Email and IoT Platforms: Analysis and Research Challenges. *Security and Communication Networks*, *2022*, e1862888. https://doi.org/10.1155/2022/1862888

[3] kinyelu, A. A. (2021). Advances in spam detection for email spam, web spam, social network spam, and review spam: ML-based and nature-inspired-based techniques. *Journal of Computer Security*, *29*(5), 473–529. https://doi.org/10.3233/JCS-210022K. Elissa, "Title of paper if known," unpublished.

[4] zeez, N. A., Misra, S., Margaret, I. A., Fernandez-Sanz, L., & Abdulhamid, S. M. (2021). Adopting automated whitelist approach for detecting phishing attacks. *Computers & Security*, *108*, 102328. https://doi.org/10.1016/j.cose.2021.102328.

[5] Baecker, L., Garcia-Dias, R., Vieira, S., Scarpazza, C., & Mechelli, A. (2021). Machine learning for brain age prediction: Introduction to methods and clinical applications. *eBioMedicine*, *72*, 103600. https://doi.org/10.1016/j.ebiom.2021.103600.

[6] AbdulNabi, I., & Yaseen, Q. (2021). Spam Email Detection Using Deep Learning Techniques. *Procedia Computer Science*, *184*, 853–858. https://doi.org/10.1016/j.procs.2021.03.107

[7] Ahmed, N., Amin, R., Aldabbas, H., Koundal, D., Alouffi, B., & Shah, T. (2022). Machine Learning Techniques for Spam Detection in Email and IoT Platforms: Analysis and Research Challenges. *Security and Communication Networks*, *2022*, e1862888. https://doi.org/10.1155/2022/1862888

[8] Akinyelu, A. A. (2021). Advances in spam detection for email spam, web spam, social network spam, and review spam: ML-based and nature-inspired-based techniques. *Journal of Computer Security*, *29*(5), 473–529. https://doi.org/10.3233/JCS-210022

[9] Alhogail, A., & Alsabih, A. (2021). Applying machine learning and natural language processing to detect phishing email. *Computers & Security*, *110*, 102414. https://doi.org/10.1016/j.cose.2021.102414

[10] Azeez, N. A., Misra, S., Margaret, I. A., Fernandez-Sanz, L., & Abdulhamid,S. M. (2021). Adopting automated whitelist approach for detecting phishing attacks. *Computers & Security*, *108*, 102328. https://doi.org/10.1016/j.cose.2021.102328

[11] Baecker, L., Garcia-Dias, R., Vieira, S., Scarpazza, C., & Mechelli, A. (2021). Machine learning for brain age prediction: Introduction to methods and clinical applications. *eBioMedicine*, *72*, 103600. https://doi.org/10.1016/j.ebiom.2021.103600

Chen, Z. (n.d.). *Bayesian filtering: From Kalman filters to particle filters, and beyond*.

[12] Faris, H., Al-Zoubi, A. M., Heidari, A. A., Aljarah, I., Mafarja, M., Hassonah, M. A., & Fujita, H. (2019). An intelligent system for spam detection and identification of the most relevant features based on evolutionary Random Weight Networks. *Information Fusion*, *48*, 67–83. https://doi.org/10.1016/j.inffus.2018.08.002

[13] Gayathri, A., Aswini, J., & Revathi, A. (2021). Classification Of Spam Detection Using Naive Bayes Algorithm Over K-Nearest Neighbors Algorithm Based On Accuracy. *NVEO - NATURAL VOLATILES & ESSENTIAL OILS Journal | NVEO*, 8516–8530.

[14] Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE Access*, *5*, 20590–20616. https://doi.org/10.1109/ACCESS.2017.2756872

[15] Gupta, S. D., Saha, S., & Das, S. K. (2021). SMS Spam Detection Using Machine Learning. *Journal of Physics: Conference Series*, *1797*(1), 012017. https://doi.org/10.1088/1742-6596/1797/1/012017

[16] Hasas, A., Zarinkhail, M., Hakimi, M., & Quchi, M. M. (2024). Strengthening Digital Security: Dynamic Attack Detection with LSTM, KNN, and Random Forest. *Journal of Computer Science and Technology Studies*, *6*, 49–57. https://doi.org/10.32996/jcsts.2024.6.1.6

[17] Jáñez-Martino, F., Alaiz-Rodríguez, R., González-Castro, V., Fidalgo, E., & Alegre, E. (2023). A review of spam email detection: Analysis of spammer strategies and the dataset shift problem. *Artificial Intelligence Review*, *56*(2), 1145–1173. https://doi.org/10.1007/s10462-022-10195-4

[18] Kaur, H., & Verma, P. (2019). *IJESRT INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY SURVEY ON E-MAIL SPAM DETECTION USING SUPERVISED APPROACH WITH FEATURE SELECTION*. https://doi.org/10.5281/zenodo.496096

[19] Kipkebut, A., Thiga, M., & Okumu, E. (2019). *Machine Learning Sms Spam Detection Model*. http://ir.kabarak.ac.ke/handle/123456789/386

[20] Kontsewaya, Y., Antonov, E., & Artamonov, A. (2021). Evaluating the Effectiveness of Machine Learning Methods for Spam Detection. *Procedia Computer Science*, *190*, 479–486. https://doi.org/10.1016/j.procs.2021.06.056

[21] Krishnamoorthy, P., Sathiyanarayanan, M., & Proença, H. P. (2024). A novel and secured email classification and emotion detection using hybrid deep neural network. *International Journal of Cognitive Computing in Engineering*, *5*, 44–57. https://doi.org/10.1016/j.ijcce.2024.01.002

[22] Kumar, N., Sonowal, S., & Nishant. (2020). Email Spam Detection Using Machine Learning Algorithms. *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 108–113. https://doi.org/10.1109/ICIRCA48905.2020.9183098

[23] Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification. *Applied Soft Computing*, *101*, 107023. https://doi.org/10.1016/j.asoc.2020.107023

[24] Liu, J. (2020). From Statistics to Data Mining: A Brief Review. *2020 International Conference on Computing and Data Science (CDS)*, 343–346. https://doi.org/10.1109/CDS49703.2020.00073

[24] Mallampati, D. (2018). An Efficient Spam Filtering using Supervised Machine Learning Techniques. *International Journal of Scientific Research in Computer Science and Engineering*, *6*(2), 33–37.

Mallampati, D., & Hegde, N. (2023). *A Machine Learning Based Email Spam Classification Framework Model: Related Challenges and Issues*. *9*, 3137–3144.

[26] Mohammed, M. A., Mostafa, S. A., Obaid, O. I., Zeebaree, S. R. M., Ghani, G., Mustapha, A., Fudzee, M. F. M., Jubair, M. A., Hassan, M. H., Ismail, A., Ibrahim, D. A., & AL-Dhief, F. T. (2019). An Anti-Spam Detection Model for Emails of Multi-Natural Language. *Journal of Southwest Jiaotong University*, *54*(3), 6. https://doi.org/10.35741/issn.0258-2724.54.3.6

[27] Nallamothu, P. T., & Khan, M. S. (n.d.). *Machine Learning for SPAM Detection*. *6*(1).

[28] Nthurima, F., Mutua, A., & Stephen Titus, W. (2023). Detecting Phishing Emails Using Random Forest and AdaBoost Classifier Model. *Open Journal for Information Technology*, *6*(2), 123–136. https://doi.org/10.32591/coas.ojit.0602.03123n

[29] Roy, S., Patra, A., Sau, S., Mandal, K., & Kunar, S. (2013). An Efficient Spam Filtering Techniques for Email Account. *American Journal of Engineering Research*.

[30] Sarkar, D., Bali, R., & Sharma, T. (2018). Deep Learning for Computer Vision. In D. Sarkar, R. Bali, & T. Sharma (Eds.), *Practical MachineLearning with Python: A Problem-Solver's Guide to Building Real-World Intelligent Systems* (pp. 499–520). Apress. https://doi.org/10.1007/978-1-4842-3207-1_12

[31] Singh, S. P., Kumar, A., Darbari, H., Singh, L., Rastogi, A., & Jain, S. (2017). Machine translation using deep learning: An overview. *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, 162–167. https://ieeexplore.ieee.org/abstract/document/8003957/

[32] Sokhangoee, Z. F., & Rezapour, A. (2022). A novel approach for spam detection based on association rule mining and genetic algorithm. *Computers & Electrical Engineering*, *97*, 107655. https://doi.org/10.1016/j.compeleceng.2021.107655

[33] Svadasu, G., & Adimoolam, M. (2022). *Spam Detection in Social Media using Artificial Neural Network Algorithm and comparing Accuracy with Support Vector Machine Algorithm* (p. 5). https://doi.org/10.1109/ICBATS54253.2022.9758927

[34] van Meteren, R., & van Someren, M. (n.d.). *Using Content-Based Filtering for Recommendation*.

[35] V.Christina, S.Karpagavalli, & G.Suganya. (2010). Email Spam Filtering using Supervised Machine Learning Techniques. *International Journal on Computer Science and Engineering*, *2*.

[36] Yang, X., Song, Z., King, I., & Xu, Z. (2023). A Survey on Deep Semi-Supervised Learning. *IEEE Transactions on Knowledge and Data Engineering*, *35*(9), 8934–8954. https://doi.org/10.1109/TKDE.2022.3220219