Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

# A COMPARATIVE REVIEW OF MACHINE LEARNING TECHNIQUES FOR EMAIL SPAM DETECTION: FROM CLASSICAL MODELS TO MULTI-MODEL DEEP LEARNING

**Haroon Ahmad[*1], Hasnain Abbas[2], Muhammad Mansoor[3], Muhammad Ali Qureshi[4]**

[*1,2,4]*Department of Information and Communication Engineering the Islamia University of Bahawalpur, Pakistan*
[3]*Department of Information and Communications Engineering*

[*1]sm1491844@gmail.com, [3]malikmansoor1237@gmail.com, [4]ali.qureshi@iub.edu.pk

**DOI:** https://doi.org/10.5281/zenodo.18429385

**Keywords**
*Spam Detection, Naive Bayes, Logistic Regression, Random Forest, Machine Learning, Email Forensics, Deep Learning, Comparative Review*

**Article History**
*Received: 30 October 2025*
*Accepted: 18 December 2025*
*Published: 31 December 2025*

**Copyright** @Author
**Corresponding Author: ***
**Haroon Ahmad**

**Abstract**
*Email spam has been a constant menace in the internet space, as it has grown to be more than texts scam- ming but composed of multi-content attacks. Within the last twenty years, spam detection has undergone different eras of development methodologically; starting with rule-based filters, and then evolving to classical machine learning, and most recently to deep learning. In this paper, an analysis of the main techniques has been done and three underlying algorithms, namely Naive Bayes, Logistic Regression, and Random Forest, and their performance, constraints, and their usefulness in real- world settings have been provided. We evaluate their theoretical foundations, feature engineering specifications and their relative accuracy on benchmarks such as SpamAssassin and Enron. In addition, we discuss the way in which the current methods, such as ensemble methods and multi-model deep learning, overcome the limitations of classical models. In this review, we have conducted the synthesis of the 40 recent studies (2013-2025) and concluded that the classical methods are still useful due to their simplicity and interpretability, whereas hybrid methods and transformer-based systems currently represent the state of the art in forensic-grade spam detection.*

## INTRODUCTION

Email spam has remained a major cybersecurity problem. since the dawn of the internet. Current estimates suggest that close to fifty percent of total email traffic in the world is unsolicited or malicious, one of the major agents of phishing, malware. and financial fraud [1]–[3]. The history of spam detection, represents general trends in artificial intelligence: starting with using heuristic rules (e.g. SpamAssassin), moving up. classical statistical learning, and currently evolving towards deep neural architecture [4], [5].

There are three algorithms that are prevalent in the classical approaches the literature: Naive Bayes (NB), logistic regression. (LR), and Random Forest (RF). Each offers distinct accuracy, speed, interpretability, and robustness trade-offs. Naive Bayes is based on the Bayes theorem and indepenen-feature. Its popularity as a result of its efficiency has led to its use as dence assumptions unusual high performance on text data. [6], [7]. Logistic Regression gave a probabilistic model with great success. interpretability by use of coefficient analysis [8], [9].Random Forest is a

collection of decision trees that enhanced robustness by reducing the overfitting and dealing with non-linear feature interactions [10], [11].

These approaches are limited by their historical achievement. lions to the current-day spamming schemes- especially the ones engaging visual materials, adversarial blurring, and multi-modal payloads [12]–[14]. Recent work has thus shifted to the depths of learning, text and BERT-style models. Image transformers Vision, frequently combined in multi-model architectures [15]– [17].

The classical spam is presented in a systematic review in this paper identifies ways of detection, evaluates its performance critically, and places their role into perspective in the times of deep learning. We integrate findings of 40 studies (2013-2025) to respond to: Comparison of NB, LR, and RF in accuracy, efficiency and forensic utility What are the gaps which are covered by modern methods II. Machine learning methods based on classical learning algorithms.

## I. Classical Machine Learning Methods

### A. Naive Bayes

Naive Bayes is still among the most used algo- spam filtering rithms because it is simple and computer-calculable. efficiency. The model presupposes conditional independence between. features assign the label of the class to it, and it is enabled to estimate posterior likelihoods by the Bayes' theorem: P (spam x) = P (x spam) P (spam) P (x) In reality, one uses bag-of- words to represent emails. model, commonly TF-IDF weighted. Despite its strong independence assumption– little of which is generally true in natural language–Naive Bayes always has a 85-93when using normal data sets such as SpamAssassin and Ling-Spam [18], [19]. low training time, minimal are its main strengths. memory usage, and strength to irrelevant features. However, it is challenged by correlated characteristics and fails to represent semantic meaning or surrounding implication, and hence subject to synonym attacks and obfuscated text [2], [20]. Devendran et al. noted that Naive Bayes performs poorly on imbalanced datasets, a common scenario in real-world email traffic [21].

### B. Logistic Regression

Logistic Regression estimates the odds logarithms of an email being spam.

**TABLE I: Comparative Analysis of Classical Spam Detection**

| Property | Naive Bayes | Logistic Regression | Random Forest |
|---|---|---|---|
| Typical Accuracy (%) | 85–93 | 88–94 | 90–95 |
| Training Speed | Very Fast | Fast | Moderate |
| Inference Speed | Very Fast | Fast | Slow |
| Interpretability | Low | High | Medium |
| Handles Feature Correlation | Poorly | Well | Very Well |
| Robust to Irrelevant Features | Yes | No | Yes |
| Requires Feature Scaling | No | Yes | No |

$$\log \frac{P(\text{spam} \mid \mathbf{x})}{1 - P(\text{spam} \mid \mathbf{x})} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i$$

The assumption of feature independence does not hold, as is the case with Naive Bayes, Doing better on correlated text features often. LR normally has accuracies of 88-94 and offers very high. explainable coefficients; an example of this is a large positive bi of the word free suggests high

spam association. [22], [23]. This interpretability is valuable in forensic contexts where analysts must justify decisions [24], [25].However, LR remains a linear Without complex feature interactions in terms of feature interactions can be captured in a model wide reaching feature engineering by hand. It also requires careful preprocessing (e.g. normalization, sparse vectors) and is not insensitive to class imbalance [26]. Maheswari and Bushra showed that LR performs better than the Naive Bayes when and n-gram features but nonetheless is still behind ensemble methods [27].

### C. Random Forest

Random Forest eliminates the linearity problem with LR. through building a combination of decision trees that are not correlated with each other.Training is done on a random subsample of data and features on each tree and the voting of predictions is done through majority voting. This is threaded by nature to deal with non-linear relationships, feature sparse data, and high-dimensional interactions, and high-dimensional interactions, and high-dimensional sparse data without exten- sive. preprocessing [28], [29].Research reports Random Forest accuracy. in the 90-95% range on email datasets and with added benefits of intrinsic features rankings of importance and anti- overfitting [30], [31].Yet, RF sacrifices are not interpretable, though global. they are harder to predict and feature importance is available. to explain than LR coefficients [32].Additionally, RF models are more inference-slow as compared to NB or LR, that is a disadvantage of high-throughput email gateways [**?**], [33].Random Forest was the best at predicting the crystalline miners.Alboloushi et al. pointed out that the best to predict the crystalline miners was Random Forest when used together with header, excels in identifying spoofed email metadata analysis [34].

## II. Comparative Performance and Limitations

Table I highlights some of the main features of the three approaches According to empirical research of 2013 through 2025. The table is purposely small enough to fit in the IEEE column width and maintaining readability..

ll three approaches have one serious weakness in common; they all use it only on the textual information. As spam evolves to embed malicious code in images, PDFs or HTML structures, Threats are not identified in classical models [35]. For example, A spam mail with an image of a is an example with no suspicious text–but as a fake invoice will be classified as legitimate by NB, LR, or RF. This gap has motivated

### Methods

the formation of multi-model methods of analysis that are simultaneous [36], [37]. Shukla et al. demonstrated that text- only filters are easily circumvented by spoofed email using text visual deceit, forensic analysis of images is required. [38].

## III. Evolution Toward Modern Approaches

In order to eliminate the text-only constraint, recent studies examine hybrid and deep learning algorithms. Ensemble techniques mix classics (e.g., stacking NB, LR, and RF) to increase precision, Tian et al. saw 96.2% with the use of a stacking approach [39].Other ones combine heuristic rules with explainable filtering through machine learning [40], [41].

The adoption is the greatest development, though. of deep learning. The authors Hina et al. used LSTM networks on. extract sequential relationships in email text, 94.8 percent accuracy. On the fine-tuning of BERT to contextual, Nasreen et al. narrowed it down to contextual understanding, whereas Liu et al. relied on DistilBERT as an efficient network. Critically, these models still ignore visual content.

Most importantly, these models do not take into account the visual content. Multi-model deep learning has become the frontier. Al-Ghamdi and Alsubait combined text and image characteristics with the basic concatenation [36],but recent studies use cross-attention. mechanisms to dynamically coordinate modalities. Advanced frame- DeBERTa-v3 is used

as a part of works to integrate text and Vision Transformer. In the case of images, more than 97% accuracy is attained on real-datasets world email documents consideration of spam as a multi-content threat [37]. The topic of recurrent architectures was studied by Saleem et al to the dynamics of email sequences across time, and Kyaw et al deep learning on phishing detection reviewed systematically. Abualhaj et al. used Harris to improve detection. Hawks optimization, proving the worth of metaheuristic feature selection [42].

## IV. Discussion and Future Directions

Classical approaches also have a place especially in resource- based restricted conditions or conditions of model transparency. mandated. Many open-source filters are run by Naive Bayes. because to its speed, whereas Logistic Regression is used in forensic settings on the interpretability of it by which it is understood [24], [43].

Random Forest offers a balance of sufficient strength to be used generally Nevertheless, the future of spam detection is admittedly multi-model. Embedded As enemies become more and more reliant on images document-based attacks, and scripts, defenders are left to use sys-content analysers that are

used to analyse any type of content. The future studies ought to concentrate on about lightweight edge deployable multi- model architectures. Meta-data forensics (metadata, headers, timestamps) integrationwith content analysis resistance to adversarial attacks privacy-preserving training to counter- evasion attacks and privacy-preserving training to mitigate data bias. The synergy of was highlighted by Pandey et al. Complete cybercrime forensics: email, web, and disk investigation [5].

## V. Conclusion

his review has followed the history of spam detection since. The classical machine learning to the modern deep learning. Naive Each of the Bayes, Logistic Regression, and Random Forest provides. as well as valuable trade-offs on accuracy, speed and interpretability, and are altogether a grown-up basis of the textual filtering. They are however unable to process visual or multi content emails constrains their effectiveness in dealing with modern threats. Multi-model deep learning- fusing linguistic introduces itself as well as visual intelligence based on architectures such as DeBERTa and ViT–a paradigm shift to holistic, forensic grade spam detection. As the threat environment keeps on changing, so should our methodologies of defense as well, incorporate classical modern architectural innovativeness. [37], [39].

## References

E. H. Tusher, M. A. Ismail, M. A. Rahman, A. H. Alenezi, and M. Uddin, "Email spam: A comprehensive review of optimize detection methods, challenges, and open research problems," *IEEE Access*, 2024.

A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti, and M. Alazab, "A comprehensive survey for intelligent spam email detection," *Ieee Access*, vol. 7, pp. 168 261–168 295, 2019.

D. E. Salhi, M. Rawashdeh, T. Bdair, and M. Al Zamil, "Two-step email spam detection: Comparing machine and deep learning accuracy." *Electrotehnica, Electronica, Automatica*, vol. 73, no. 2, 2025.

E. Altulaihan, A. Alismail, M. Hafizur Rahman, and A. A. Ibrahim, "Email security issues, tools, and techniques used in investigation," *Sustainability*, vol. 15, no. 13, p. 10612, 2023.

B. Pandey, P. Pandey, A. Kulmuratova, and L. Rzayeva, "Efficient usage of web forensics, disk forensics and email forensics in successful investigation of cyber crime," *International Journal of Information Technology*, vol. 16, no. 6, pp. 3815–3824, 2024.

P. Malhotra and S. Malik, "Spam email detection using machine learning and deep learning techniques," in *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*, 2022.

K. S. Ubale and K. A. Shirsath, "Evaluation of classification algorithms for effective spam email detection using spam email dataset," in *Transformative Applied Research in Computing, Engineering, Science and Technology*. CRC Press, 2025, pp. 118–125.

K. Debnath and N. Kar, "Email spam detection using deep learning approach," in *2022 international conference on machine learning, big data, cloud and parallel computing (COM-IT-CON)*, vol. 1. IEEE, 2022,
pp. 37–41.

L. Wang, "Spam email detection using na¨ıve bayes classifier," in *ITM Web of Conferences*, vol. 70. EDP Sciences, 2025, p. 04028.

E. H. Tusher, M. A. Ismail, and A. F. Mat Raffei, "Email spam classification based on deep learning methods: a review," *Iraqi Journal for Computer Science and Mathematics*, vol. 6, no. 1, p. 2, 2025.

B. K. Dedeturk and B. Akay, "A parallel hybrid approach integrating clonal selection with artificial bee colony for logistic regression in spam email detection," *Neural Computing and Applications*, vol. 37, no. 27,
pp. 22 401–22 419, 2025.

S. Yu, "Covert communication by means of email spam: A challenge for digital investigation," *Digital Investigation*, vol. 13, pp. 72–79, 2015.

C. Opara, P. Modesti, and L. Golightly, "Evaluating spam filters and stylometric detection of ai-generated phishing emails," *Expert Systems with Applications*, vol. 276, p. 127044, 2025.

S. Jamal, H. Wimmer, and I. H. Sarker, "An improved transformer-based model for detecting phishing, spam and ham emails: A large language model approach," *Security and Privacy*, vol. 7, no. 5, p. e402, 2024.

M. Hina, M. Ali, A. R. Javed, F. Ghabban, L. A. Khan, and Z. Jalil, "Sefaced: Semantic-based forensic analysis and classification of e-mail data using deep learning," *IEEE Access*, vol. 9, pp. 98 398–98 411, 2021.

G. Nasreen, M. M. Khan, M. Younus, B. Zafar, and M. K. Hanif, "Email spam detection by deep learning models using novel feature selection technique and bert," *Egyptian Informatics Journal*, vol. 26, p. 100473, 2024.

T. Liu, S. Li, Y. Dong, Y. Mo, and S. He, "Spam detection and classification based on distilbert deep learning algorithm," *Applied Science and Engineering Journal for Advanced Research*, vol. 3, no. 3, pp. 6–10, 2024.

F. Hossain, M. N. Uddin, and R. K. Halder, "Analysis of optimized machine learning and deep learning techniques for spam detection," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE, 2021, pp. 1–7.

P. H. Kyaw, J. Gutierrez, and A. Ghobakhlou, "A systematic review of deep learning techniques for phishing email detection," *Electronics*, vol. 13, no. 19, p. 3823, 2024.

M. Alazab, R. Layton, R. Broadhurst, and B. Bouhours, "Malicious spam emails developments and authorship attribution," in *2013 fourth cybercrime and trustworthy computing workshop*. IEEE, 2013, pp. 58–68.

V. K. Devendran, H. Shahriar, and V. Clincy, "A comparative study of email forensic tools," *Journal of Information Security*, vol. 6, no. 2, p. 111, 2015.

Z. B. Siddique, M. A. Khan, I. U. Din, A. Almogren, I. Mohiuddin, and
S. Nazir, "Machine learning-based detection of spam emails," *Scientific Programming*, vol. 2021, no. 1, p. 6508784, 2021.

E. R. Ejirika and T. O. Omotehinwa, "Analysis of machine learning models for spam email detection and real-time integration," in *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*. IEEE, 2024, pp. 1–10.

A. Ghafarian, A. Mady, and K. Park, "An empirical analysis of email forensics tools," *International Journal of Network Security & Its Applications (IJNSA) Vol*, vol.

12, 2020.

S. D. Hamdi and A. M. Radhi, "Digital cyber forensics contribution for email analysis," *Journal of Engineering and Sustainable Development*, vol. 24, no. 4, pp. 9–19, 2020.

S. Alsudani, H. Nasrawi, M. Shattawi, and A. Ghazikhani, "Enhancing spam detection: A crow-optimized ffnn with lstm for email security," *Wasit Journal of Computer and Mathematics Science*, vol. 3, no. 1, pp. 28–39, 2024.

K. U. Maheswari and S. N. Bushra, "Machine learning forensics to gauge the likelihood of fraud in emails," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2021, pp. 1567–1572.

S. Zavrak and S. Yilmaz, "Email spam detection using hierarchical attention hybrid deep learning method," *Expert Systems with Applications*, vol. 233, p. 120977, 2023.

L. Sankaine, J. G. Ndia, and D. Kaburu, "An english-swahili email spam detection model for improved accuracy using convolutional neural networks," *Mesopotamian Journal of CyberSecurity*, vol. 5, no. 2, pp. 590–605, 2025.

K. V. Samarthrao and V. M. Rohokale, "Enhancement of email spam detection using improved deep learning algorithms for cyber security," *Journal of Computer Security*, vol. 30, no. 2, pp. 231–264, 2022.

A. Sheneamer, "Comparison of deep and traditional learning methods for email spam filtering," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 1–6, 2021.

K. Morovati and S. S. Kadam, "Detection of phishing emails with email forensic analysis and machine learning techniques." *International Journal of Cyber-Security and Digital Forensics*, vol. 8, no. 2, pp. 98–108, 2019.

D. K. Yadav, A. Raj, Rajlakshmi, N. Kumar, and R. Kumari, "Enhancing email security: A real-time machine learning-based spam detection system," *Internet Technology Letters*, vol. 8, no. 5, p. e618, 2025.

A. G. AlBoloushi, F. A. Al-Meer, F. Nawshin, and D. Unal, "Network forensics: Techniques, challenges, and incident response," in *Proceedings of Eighth International Conference on Information System Design and Intelligent Applications*. Springer, 2025, pp. 51–62.

I. Riadi, S. Sunardi, and F. T. Fitri, "Spamming forensic analysis using network forensics development life cycle method," *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 6, no. 1, pp. 108–117, 2022.

N. Al-Ghamdi and T. Alsubait, "Digital forensics and machine learning to fraudulent email prediction," in *2022 Fifth National Conference of Saudi Computers Colleges (NCCC)*. IEEE, 2022, pp. 99–106.

V. S. Kadam, S. Pingale, S. R. Biradar, V. M. Rohokale, and K. D. Bamane, "Designing a novel framework of email spam detection using an improved heuristic algorithm and dual-scale feature fusion-based adaptive convolution neural network," *Information Security Journal: A Global Perspective*, vol. 34, no. 4, pp. 286–309, 2025.

S. Shukla, M. Misra, and G. Varshney, "Spoofed email based cyberattack detection using machine learning," *Journal of Computer Information Systems*, vol. 65, no. 2, pp. 159–171, 2025.

Y. Tian, X. Dai, Z. Li, H. Guo, and X. Mao, "Improving the accuracy of cybersecurity spam email detection using ensemble techniques: A stacking approach machine learning for spam email detection," *PLoS One*, vol. 20, no. 9, p. e0331574, 2025.

N. Kaushik, T. S. Rathore, and P. Kumar, "Email traceback: Securing systems from phishing and malicious link prevention," in *2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N)*. IEEE, 2024, pp. 647–652.

L. Xie, Y. Liu, and G. Chen, "A forensic analysis solution of the email network based on email contents," in *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*. IEEE, 2015, pp. 1613–1619.

M. M. Abualhaj, S. N. Alkhatib, A. A. Abu-Shareha, A. M. Alsaaidah, and

M. Anbar, "Enhancing spam detection using harris hawks optimization algorithm," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 23, no. 2, pp. 447–454, 2025.

F. Armknecht and A. Dewald, "Privacy-preserving email forensics,"

*Digital Investigation*, vol. 14, pp. S127–S136, 2015.