ISA Journal of Multidisciplinary (ISAJM)

Homepage: https://isapublisher.com/isajm/

Volume 3, Issue 1, Jan-Feb, 2026 ISSN: 3049-1851

Evaluation of Machines Learning Models for Email Spams Detection and Classification

Ikuomola A. J. and Akinbulejo Elijah Obadare

Olusegun Agagu University of Science and Technology, Okitipupa

Received: 10.12.2025 | **Accepted:** 30.12.2025 | **Published:** 02.01.2026

***Corresponding Author:** Akinbulejo Elijah Obadare

DOI: *10.5281/zenodo.18133518*

Abstract **Review Article**

Email plays vital role in the world of digital communication, still it has faced a huge challenge of harmful and malicious spam attacks. In spite a lot of measures and research contributions to guide against spam attacks, yet the spammers are adopting dynamic approaches to evade detection through content manipulation, text obfuscation and removal of hypertext tag which render the traditional methods such rule-based, blacklist ineffective. This research work embarks on evaluation of five widely used machine learning models- Naïve Baye's (NB), Support Vector Machine (SVM), Decision Tree (DT), Multi-Layer Perceptron (MLP) and K-Nearest Neighbour (KNN) as well as their ensemble model which combined all these models in a single model for a better performance. The ensemble model performed better than each model.

Keywords: Email spam detection, Machine learning models, Ensemble classification, Text obfuscation, Cybersecurity.

INTRODUCTION

Rapid adoption of internet as a major part of human daily activities which has greatly influenced the process of communication electronically of which electronic mail (email) acts as one of the most widely used media of communication across the globe. As there is wider acceptance of emails usage among users so also spam emails increase proportionally. In the space of electronic communication, spam email is a major threat that catches the attention of both the researchers, organisation and individual users. The impacts of Spam emails cannot be overemphasized as they consume a lot of time, resources and result to slow internet speed, spread of malwares, phishing, DoS attacks and stealing of confidential information and this pose a severe security challenge to our society. Detection accuracy is a major challenge faced by different researchers using different classification methods.Email is widely used in almost all fields of life for daily transactions or exchange of valuable information. Spam is a significant problem in the world of electronic communication. Meanwhile spam email portrays a huge threat to the millions users daily as it does not only contribute huge to financial losses but also posed a security challenge to sensitive data through malicious links that can cause phishing, DoS and other attacks.Therefore, spam email detection and classification is a critical task in the security of

emails as it assists in protecting both users and organization from such illegal attacks.

In this research work, we will use five different machine learning models such as Naïve Baye's(NB), Support Vector Machine(SVM), Decision Tree(DT), Multi-Layer Perceptron(MLP) , K-Nearest Neighbour (KNN) and their ensemble model to be able to reduce significantly high false negatives and false positive issues that are faced in email classification.

2.0 LITERATURE REVIEW ON EXISTING WORK

Hamsapriya *et al*.,(2018) deployed Naive Baye's classifiers, Multilevel Perceptron (MLP) Classifier and J48 classifier for classifying spam from ham and they concluded that the combined use of different algorithms lead to correction of individual uncorrected errors. Hassan *et al*.(2019) discussed the features extraction methods and used Naive Baye's and Support Vector Machine (SVM) for classifying spam and ham based on the features extracted. Ikuomola *et al*(2020) developed malicious contents detection system using browser extension. Jazzar *et al*.,(2021) compared the performance of support Vector Machine (SVM), Artificial Neutral Network (ANN), Baye's and J48 with the conclusion that Support Vector Machine performed most optimally. Aju and Adedeji (2022) developed spam emails filtering model using ensemble of Decision Tree (DT), Support Vector Machine (SVM) and Multilayer Perceptron (DT, SVM, MLP) with the conclusion that the combined techniques performed better than individual algorithm. Ghogare *et al*., (2024) discussed the impacts of efficient data preprocessing on the overall performance of the machine learning models. Anitha *et al*.,(2023) deployed Natural Language Processing with Machine Learning for spam emails detection and used Naive Baye's, Support Vector Machine (SVM), K-Nearest Neighbour and Random Forests but Naive Baye's has the highest prediction accuracy. Kasturi *et al*.,(2024) explored combination of machines learning and Natural Language Processing for efficient spam detection and accuracy. It also used SVM, KNN, Naive Bayes and Decision tree algorithms and that DT performed most optimally. Thompare *et al* (2023) proposed a comparative machine learning framework for classifying email messages using SVM, RF, KNN, DT. Sultana *et al* (2020) proposed a machine learning base detection primarily using Naive Baye's classifier. Switalski & Kopowka(2019) conducted comparative evaluation of traditional Machine Learning Algorithms for spam detection and classification. Cletus *et al* (2024) proposed a Machine Learning based Malware Classifier

3.0 METHODOLOGY

The method used in this research work includes, Data collection of 10,000 emails samples from each of the three different datasets (kaggle, spam assassin and entron datasets), contents extraction, preprocessing, data transformation, training machine learning classifiers and models performance evaluation.

3.1 Architecture of a Machine Learning Model for Email Spam Detection and Classification

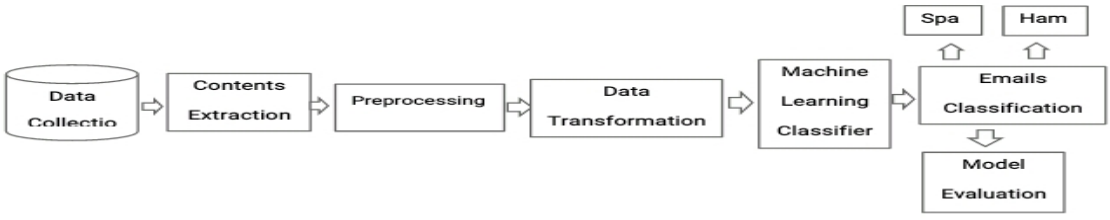


Figure 1: Architectural Design of the Model

3.2 Components of the Architecture Design

I. Data Collection: Data are collected from three different datasets. An unbiased 10,000 email data samples are collected from each of the datasets, Kaggle dataset, enron dataset and SpamAssasin, giving room for a robust dataset across the three datasets with each dataset consists 10,000 emails samples.

ii. Content Extraction: The contents are extracted and cleansed

iii. Preprocessing: Thorough preprocessing' is done on the extracted contents.

iv. Data Transformation: The data are transformed to digital format that can be executed by the machine learning models.

v. Model Training (Machine Learning Classifiers): Machine Learning models are trained using the already properly preprocessed datasets.

vi. Classification: Classification is done while testing the models

vii. Performance Evaluation: Their performance are evaluated thereafter using some performance evaluation metrics.

3.3 MACHINE LEARNING TECHNIQUES

There are five widely spread classification algorithms in machine learning which were selected: Naive Bayes, K-nearest Neighbors, Support Vector Machine, Decision tree and Multi-Layer Perceptron (MLP)

I. Naïve Baye's: Naive Bayes is a probabilistic algorithm that does a good job of classifying spam. It is called "naive‟ because it ignores possible dependencies or correlations among inputs and reduces a multivariate problem to a group of univariate problems. It serves as an important probabilistic method and allows us to exploit ethical grey areas by manipulating the odds of the method's predictions. It's high accuracy in conducting binary classification and straightforward implementation

II. Support Vector Machine (SVM) is a linear classifier that is equivalent to finding the hyperplane separating the classes with maximum indentation. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on . Support Vector Machine's accuracy and precision in classifying different classes of data have led to its widespread adoption as it is suitable for classifying both linear and nonlinear datasets.

III. Decision tree: Decision Tree(DT) is a popular technique for classifying data since the solution it produces is both interpretable and straightforward. It is structured like a tree with central hub, branches and leaves. The terminal node, or leaf node represents a class attribute, and the other nodes represent potential solutions.

IV. Multilayer Perceptron: Multilayer Perceptron are groups of simple processing units (artificial neurons) which are interconnected and communicate with one another by means of sizable number of weighted connections. Each processing unit can accept input from neighbor. Next, the output value is calculated passed to other neighbors. The connection between units are weighted.

V. K-Nearest Neighbors (KNN) is a metric classification method, that is, objects are represented as points in space and distances are calculated between them. Then, it enters a learning phase when training data points are iteratively assigned to a cluster which center is located at the nearest distance.The algorithm is unstable to outliers and not working correctly with a large amount of features.

VI. Ensemble model (NB+SVM+DT+MLP+KNN): Combined the strengths and diversity of all the models for an improved performance.

4.0 Implementation and Results

Implementation of this work was done using Python in connection with some libraries and datasets.The models were built and trained using Scikit-learn library while NLTK helps language progressing.The results from the evaluation of their performance across different datasets are stated as follows:

4.1 Evaluation of the performance of the ML Models on three different datasets

Table 1, 2, 3 shows the performance of the machine learning models on kaggle datasets, enron dataset and spamassasin dataset respectively while figure 2, 3, 4 shows the Bar charts of the performance of the machine learning models on kaggle datasets, enron dataset and spamassasin dataset respectively

Table 1: Kaggle

Model/Metric	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)	ROC-AUC (%)
Naive Baye's (NB)	72.3	71.6	72.7	72.3	72
Support Vector Machine (SVM)	74.4	74.4	74.8	74.5	74
Decision Tree (DT)	71.8	71.3	72.3	71.8	71
Multilayer Perception (MLP)	73.8	73.3	74.3	74.0	73
K-Nearest Neighbour (KNN)	72.3	71.8	71.8	72.3	72
Ensemble (NB+SVM+DT+MLP+KNN)	77.8	75.7	76.3	75.8	77

Figure 2: Chart for Kaggle results

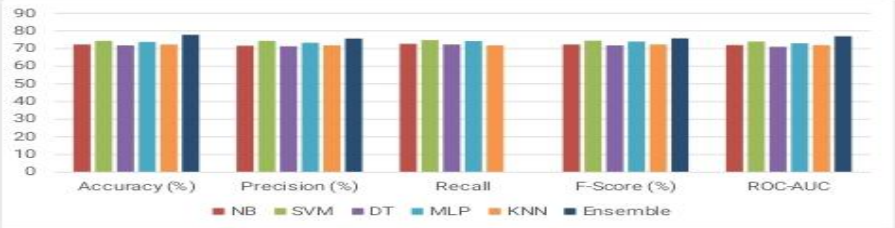


Table 2: Entron

Model/Metric	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)	ROC-AUC (%)
Naive Baye's (NB)	70.7	66.6	70.3	69.5	70
Support Vector Machine (SVM)	71.8	71.3	72.3	69.5	70
Decision Tree (DT)	68.8	67.8	68.8	68.0	68
Multilayer Perception (MLP)	70.8	70.3	71.3	71.8	70
K-Nearest Neighbour (KNN)	66.8	71.3	;72.3	71.8	66
Ensemble (NB+SVM+DT+MLP+KNN)	73.3	72.8	73.8	73.3	73

Figure 3:Chart for Enron Results

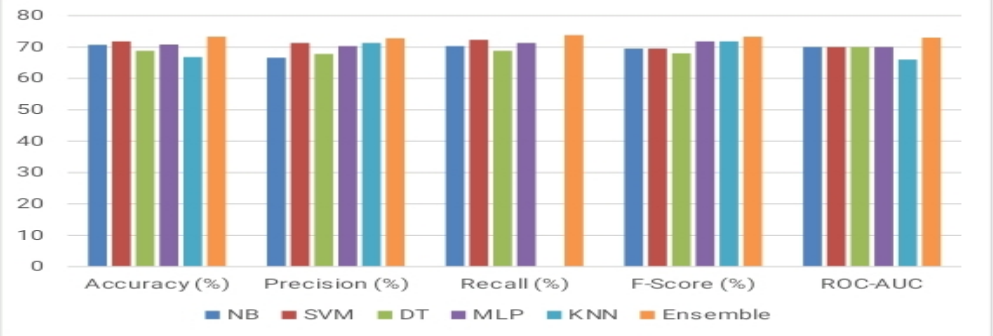
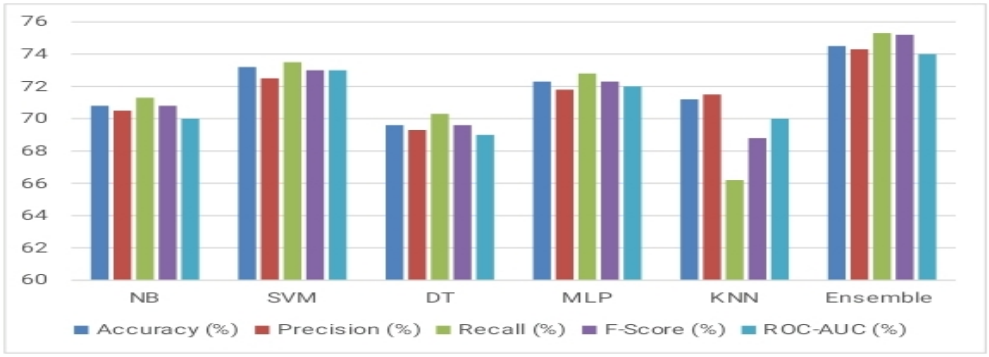


Table 3: SpamAssasin

Model/Metric	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)	ROC-AUC (%)
Naive Baye's (NB)	70.8	70.5	71.3	70.8	70
Support Vector Machine (SVM)	73.2	72.5	73.5	73.0	73
Decision Tree (DT)	69.6	69.3	70.3	69.6	69
Multilayer Perception (MLP)	72.3	71.8	72.8	72.3	72
K-Nearest Neighbour (KNN)	71.2	71.5	66.2	68.8	70
Ensemble (NB+SVM+DT+MLP+KNN)	74.5	74.3	75.3	75.2	74

Figure 4: Chart for Spamassasin Results



5.0 Discussion and Conclusion

It is clear from the table 1, 2,3 that the Ensemble model gives better /higher accuracy, precision, recall, F1-score and ROC/AUC than the other five classifiers/models across all datasets. Also SVM is the best individual model, with high accuracy, precision, recall, F1-score, and AUC across all datasets. All the models perform best on the Kaggle dataset, with the highest accuracy, precision, recall,

F1-score, and AUC compare to the two other datasets (i.e. Enron and SpamAssasin). The models perform worst on the Enron dataset, with the lowest accuracy, precision, recall, F1-score, and AUC.

REFERENCES

Aju O. G. and Adedeji A. J. (2022). An Email Filtering Model Using Ensemble of Machine Learning Techniques. International Journal of Computer Applications Technology and Research Vol. 11, pp 66-71

Cletus A., Opoku A. A., Weyori B.A., (2024), An Evaluation of Current Malwares Trends and Defense Techniques. Journal of Advances in Information Technology Vol. 15, pp 649 - 671 DOI: 10.12.72.0/jait.15.5.6469-671

Dada, E. G., Bassi, J. S., Chiroma, H., Adetunmbi, A. O., & Ajibuwa, O. E. (2019). Machine learning for email spam filtering: review, approaches and open research problems. Heliyon, 5(6), e01802. https://doi.org/10.1016/j.heliyon.2019.e01802

Ghogare P. P., Dawood H., Patil M. P. (2024), Enhancing Spam Email Classification Using Effective Preprocessing Strategies and Optimal Machine Learning Algorithms. Indian Journal of Science and Technology

Hamsapriya T., Renuka D. K., Chakkaravarthi M. R. and Surya P. L., "Spam Classification Based on Supervised Learning Using Machine Learning Techniques," 2018 International Conference on Process Automation, Control and Computing, Coimbatore, 2018, pp 1 - 7, doi: 10.1109/PACC.2011.5979035

Hassan M. A. and Mtewa N. (2019). Feature Extraction and Classification of Spam emails. International Conference on Soft Computing & Machine Learning Intelligence DOI: 10.1109/ISCMI.2018.8703222

Ikuomola A. J., Ogunbamerun A. and Nwanze M.N(2020). Development of a Malicious Website Detection System Using Web Browser Extension. Coast Journal of the Facualty of Science 2(2): pp 482 – 489.

Kasturi K., Rohini K., and Vijayaprathaban (2024) E-Mail Spam Classification using Machine Learning Algorithms and Natural Language Processing, Research Gate Publication, https://www.researchgate.net/publication/378797444

Kontsewaya Y., Antonov E. and Artamonov A. (2020). Evaluating the effectiveness of Machine Learning Methods for Spam Detection. International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence. 10.1016/j.procs.2021.06.056

Jazzar M., Yousef R. F. and Eleyan D. (2021). Evaluation of Machine Learning Techniques for Email Spam Classification. International Journal Education and Management Engineering vol. 4 pp 35 - 42. DOI: 10.5815/ijeme.2021.04.04

Manoj Sethi, Sumesha Chandra, Vinayak Chaudhary, andYash (2021). EmailSpam Detection using Machine Learning and NeuralNetworks. International Research Journalof Engineering and Technology (IRJET). pp 349 - 355

Moutafis I., Anreators A. and Stefaneas P. (2023). Spam Email Detection Using Machine Learning Techniques. European Conference on Cyber Warfare and Security

Reddy Anitha, Kanthala Harivardhan Reddy, A. Abhishek, Myana Manish, G. Viswa Sai Dattu, and Noor Mohammad Ansari(2023),Email Spam Detection Using MachineLearning Journal of Survey in Fisheries Science 10(1), pp 2658-2664

Sharma A. and Arjun N. (2023). Spam Detection Using Using Machine Learning Techniques .

International Journal of Research Publication and Reviews Vol.4, No 7, pp 2478 - 2488

Sultana T., Sapnaz K. A., Sana F. and Najath J. (2020). Email based Spam Detection. International Journal of Engineering Research & Technology (IJERT) Vol.9, pp 135 - 139

Swtalski P., Kopowka M., (2019), Machine Learning Methods in E-mail Spam Classification. Journal of System and Information Technology. DOI: 10.34739/si.2019.23.04

Thombare R. Y., Patil P., Shelter P., Dhakane O. and Chakor K. (2023). Email Classification Using Machine Learning. International Journal of Advanced Research in Science, Communication and Technology Vol3 , pp 347 - 352 , DOI: 10.48175/568