International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 23s,2025
https://theaspd.com/index.php

Improving Email Spam Detection And Classification Through Data Balancing And Ensemble Machine Learning-Based Boosting Approaches

Garima Mishra[1], Dr.Parth Gautam[2]
[1]Research Scholar,Department of Computer Science and Application , Mandsaur University,Mandsaur , Madhya Pradesh (India), mishragarima2708@gmail.com
[2]Supervisor and Assistant Prefessor, Department of Computer Science and Application ,Mandsaur University,Mandsaur , Madhya Pradesh India, parth.gautam@meu.edu.in

***Abstract*—**Email is among the most used and effective internet communication and data or messages sharing method. With the importance and high usage of emails, spam mail has also grown at a great rate. Email systems are faced with the huge and complicated challenges of detecting and filtering spam. The use of traditional methods of identification such as blocklists, real-time blackhole listing and content-based methods is limited. As a result of these constraints, more advanced machine learning (ML) tools have been developed to ameliorate accuracy of spam detection. The current work deals with the problem of email spam identification which is on the rise, i. e., it is a relevant issue in the sphere of digital communications security in which undesired or malicious email messages could be employed to infringe upon the privacy and integrity of user databases. The overall aim was to create an effective classification system that could easily differentiate between spam and legitimate messages. On the Spambase data of the UCI data repository preprocessing was used using feature labeling, splitting training and testing sets and using SMOTEENN to balance the classes to reduce skewness. The three ensemble boosting models AdaBoost, Gradient Boosting (GBC), and CatBoost models were implemented and stringently tested in terms of confusion matrix, classification report, sensitivity, specificity, ROC curves and precision-recall curves. The results were high, with all models showing a similar level of performance, approximately 97.66% for AdaBoost, 97.92% for GBC, and 98% for CatBoost. Notably, CatBoost slightly exceeded the others. The comparative analysis proved that boosting-based models exhibit great resilience to misclassification of spam and non-spam and can be effectively utilized in real-life application. The significance of this work is that it integrates hybrid resampling with the most recent boosting techniques which ensures high performance with unequal data together with the highest possible detection. The values of the research performance measure indicate that ML models have the potential to enhance the adoption of cybersecurity solutions in combating email spam attacks.

Keywords*–Cybersecurity, Email, Spam emails, machine learning, deep learning, balancing, boosting, ensemble models*.**

INTRODUCTION

The internet's technological innovations have completely changed our world by making it possible for knowledge to be shared instantly around the world.[1]. But there is a serious downside to this advancement: the increase in spam in our inboxes. Email spam is a widespread problem that affects consumers worldwide and includes a wide range of unsolicited messages. Email has certainly made communication easier, but it has also led to a rise in spam. It's now essential to effectively separate spam from valid emails[2]. Unsolicited commercial email (UCE), one of the most well-known types of spam, bombards receivers with unsolicited sales spiels for goods or services they never agreed to accrue. Furthermore, emails are dangerous because they might fool receivers into divulging private information by imitating official correspondence from reliable sources[3]. Emails have made it easier for people to work together because they are a quick and affordable way to communicate. They have greatly expedited communication and information exchange on personal as well as professional grounds[4]. However, because of escalating usage and reliance on emails, consumers are now more susceptible to cybersecurity risks encompassing malware infections, spam attacks, and sundry types of wringing. Emails are essential for users as well, since they continue to be crucial in

many fields. An estimated USD 355 million is lost annually as a result of spam triage for the four types of spam[5].

As businesses implement secure email procedures and strong security measures to fend against new attacks[6]. Cybercriminals use email channels as a springboard for attacks that could cause consequential harm to individuals and organisations. In fact, it is said that up to 90% of cyberattacks are caused by emails[7][8]. There are still weaknesses in email security despite efforts to strengthen it. Attackers employ a range of strategies, encompassing social engineering, email account hacking, and creating phoney emails, in order to take advantage of companies and compromise their systems[9]. Since social engineering campaigns aim to deceive staff, get unauthorised access, reveal private information, spread malware, and interfere with necessary operations, they are among the most deceptive of these strategies[10]. Therefore, it is imperative to strengthen cybersecurity protection measures and take action against these expanding email-based threats[11]. There are weaknesses in email networks that bad actors frequently take advantage of. Phishing emails and spam are the most popular attack techniques used by these people. Emails have largely facilitated communication and connections, but a major issue is that spam emails are constantly being sent to recipients. As a result, separating genuine emails from unsolicited spam has become crucial[12].

Additionally, because receivers are forced to reply to unsolicited messages, spam violates their privacy. This privacy barrier can be broken with a simple click[13]. According to studies, over 200 million mobile users receive dreadfully inadequate amounts of spam SMS messages each day, making mobile phones the most often targeted device for hacks[14]. As majority of these spam communications include phishing attempts advertising sundry items, services, or events, it is clear that each spam email incurs a financial penalty to the recipient[15]. Even though there are a number of ways to tell the difference amidst spam and licit messages, spammers usually modify their strategies to get users' blandishment, hindering to tell the difference[16]. Email is still the most widely used and economical way to share information on electronic devices. It is used for sundry purposes, encompassing financial operations, advertising, health information delivery, recruitment, and business-to-business and inter-business communication. It is also used for both functional and non-functional communications both inside and outside of organisations.[17].

Researchers have looked into a number of ways to ameliorate spam detection's efficacy to solve the difficulties that come with email communication ML algorithms have been widely used in a variety of email spam during the past ten years due to significant advancements in the field of AI. Amidst commonly used ML techniques[18] for spam stratification are supervised learning methods, such as Random forest[19], SVMs[20], XGBoost[21], Naive Bayes [22], and ANNs [23]. With their potentiality to grasp data patterns, attributes, and email structures, these incredibly adaptable algorithms offer encouraging results when it comes to spotting phishing emails.[24]. Additionally, by including balance and selection procedures that are normally carried out manually by human experts, machine learning approaches may effectively classify email spam[25]. Building email spam detection models using ML approaches is the driving force behind this project, which aims to accurately identify spam emails from legitimate emails. The following research contributions of this work are:

The study systematically applied and compared three ensemble boosting algorithms (AdaBoost, Gradient Boosting, and CatBoost) on the UCI Spambase dataset for effective spam identification.

- The work addressed dataset imbalance by using the SMOTEENN technique, which combines oversampling and undersampling to improve classification fairness across spam and non-spam classes.
- Multiple evaluation measures, including “accuracy” (acc), sensitivity, “precision” (prec), ROC-AUC, “recall” (rec), specificity, “F1-score” (f1-score), and precision-recall curves, were employed for a holistic performance assessment.
- The results demonstrate that boosting-based machine learning approaches can provide reliable, interpretable, and computationally efficient solutions for real-world email spam filtering.

The significance of this investigation is rooted in its potentiality to strengthen email security by providing an efficient and reliable spam detection framework, addressing one of most persistent threats in digital communication. Unlike conventional approaches that often struggle with imbalanced data and limited evaluation, this study applies hybrid to effectively balance the dataset and ensure fair classification across both spam and non-spam classes. The innventiveness of the work lies in comparative investigation of multiple ensembles boosting models—AdaBoost, Gradient Boosting, and CatBoost—on the same dataset, offering valuable insights into their relative strengths. By combining advanced resampling techniques,

diverse evaluation metrics, and boosting algorithms, this work presents a comprehensive and practical methodology for ameliorating the accuracy, interpretability, and adaptability of real-world email spam filtering systems.

A. Organization of the paper

The remnant article is arranged as follows: The relevant works for email spam detection are covered in Section 2. Using flowcharts and models, Section 3 delineates the process. Section 4 delineates experimental results and a discussion of comparison. Finally, Section 5 delineates conclusion and limitations with future work.

LITERATURE REVIEW

An overview of relevant work is given in this section. Sundry studies have scrutinize email spam detection using both ML and DL techniques. Research has progressed over time from conventional classifiers to ones that target various spam identification issues.

According to the literature, Saleem et al. (2025) have successfully created a hybrid DL model that utilizes GRU (Gated Recurrent Unit) and LSTM (Long Short-Term Memory) separately to detect spam emails. Even though the other models—CNNs, ensemble ML classifiers, LSTM, or GRU—have been used separately in earlier research, the current study has added to the body of literature because it has been able to amalgamate computational efficiency of GRU with benefits of LSTM in capturing long-term dependency. Important problems that are typically faced when implementing standalone deep learning, like the vanishing gradient problem and excessive resource use, have been addressed in this hybridisation. Furthermore, our suggested model has a 90% detection accuracy, which is excellent. Transformer-based models are substantially lighter and can be employed in applications that operate in real time, but they require a lot of computing power[26].

Complementing this, Algimantas Venčkauskas (2024), The study suggests a new domain-specific ontology and approach for emails that are suitable to email CTI sharing solutions and only require the sharing of email message information in order to protect privacy. For propound email domain-specific ontology, a new semantic parser was developed in order to produce a dataset and populate email data while preserving privacy. Using the newly generated dataset, experiments were carried out to detect and categorise spam communications and analyse machine learning techniques. ANOVA (analysis of variance), Chi-squared, Kruskal-Wallis tests, and feature-ranking algorithms were utilized. The kernel naïve Bayes model produced results that were satisfactory in every experiment. The newly constructed semantic parser, proposed domain-specific ontology, and the generated metadata dataset produced the best results for spam email message recognition, with an f-meaasure of 95.92% and the greatest acc of 92.28%.[27].

Other studies compared traditional classifiers directly. N. RamojiRao et al. (2024), study focused on analyzing email spam stratification using two classification algorithms. The research was conducted using a dataset of 5,172 emails. The goal was to compare efficacy of these two ML techniques for classifying emails. The analysis involved evaluating the performance of SVM and a novel Naive Bayes Classifier. For the experiment, a sample size of 25 emails per group was determined using a G power analysis with 80% power. The results stipulated that while SVM achieved an accuracy of 89.78% with a relatively low mean error, the novel NB Classifier outperformed it with an accuracy of 94.09%. This indicates that NB Classifier is more effective at accurately classifying emails as spam or not spam compared to SVM. The statistical significance of the results was confirmed with a p-value of 0.036($p<0.05$), highlighting that Naive Bayes Classifier's superior accuracy is statistically significant. In summary, the study demonstrates that NB Classifier provides better potentiality in email spam detection than the Support Vector Machine[28].

Expanding on this, M. Alsuwit et al. (2024), the study suggests using DL and ML methods to efficiently categorise spam emails. To build reliable models for precise spam identification, techniques including Random Forest (RF), Logistic Regression (LR), Naïve Bayes (NB), and Artificial Neural Networks (ANNs) are used. By combining these methods, goal is to ameliorate spam detection's accuracy and efficacy, helping email and Internet of Things service providers lessen the negative impacts of spam. Promising results were found when the suggested models were evaluated. The remarkable 97% accuracy rate attained by LR, RF, and NB demonstrates their effectiveness in correctly recognising spam emails[29]. Similarly, V. Dharani et al. (2023), suggested a model based on ML algorithms to help people avoid falling for scammers' tricks. The NB method and term frequency-inverse document frequency vectorizer are utilize to execute suggested model acquired the Kaggle dataset and used it to train the model. This model includes a local host webpage

that may be accessed using the PyCharm IDE. The findings obtained indicate that the model's accuracy is 95%[30].

On the other hand, ensemble learning has also gained traction. Temidayo et al. (2023) developed RF and XGBoost(extreme gradient boost) ensemble algorithms using the Enron1 dataset to determine and rank spam emails. The grid-search cross-validation method was subsequently employed to improve the generated ensemble models by exploring hyperparameter spectrum for optimal values. Both algorithms' tuned and baseline (un-tuned) models' performances were assessed and contrasted. We also looked at how hyperparameter adjustment affected both models. Comparing both models to the baseline models, the experimental study's results showed that the hyperparameter adjustment enhanced their efficacy. The accuracy of the XGBoost and tuned RF models was 97.78%. The RF model did not perform as well as the XGBoost model. The XGBoost model that was created is successful and efficient at detecting spam emails[31].

Deep learning remains significant, particularly when applied with NLP preprocessing. K. Debnath et al. (2022), Deep learning models are created using LSTM and BERT to detect and categorise new email spam using the Enron email dataset. The email's text was analysed and data preprocessed using an NLP technique. The outcomes are contrasted with those of earlier email spam detection methods. Using LSTM, the suggested DL method attained best accuracy of 97.15%[32]. And other ML based, Tasnia Toma et al. (2021), Using supervised ML on an existing dataset for email classification, investigating Naïve Bayes, SVM, and RF. We displayed additional efficacy metrics, such as prec, rec, and f-measure, in addition to the algorithms' accuracy. With regard to Multinomial NB, Bernoulli Naïve Bayes, Gaussian NB, RF, and SVM, we obtained high accuracy rates of 97.6%, 91.5%, and 97.8%, respectively[33].

TABLE I. SUMMARY OF RELATED WORK ANALYSIS OF EMAIL SPAM DETECTION USING ML TECHNIQUES

Reference	Methodology	Dataset	Results	Advantages	Limitations	Recommendations
Saleem et al. [26]	Hybrid deep learning (LSTM + GRU)	Not specified (general email dataset)	90% accuracy	Combines LSTM's ability for long-term dependency with GRU's efficiency; addresses vanishing gradient problem	Computationally intensive compared to classical ML; Transformers can be more resource-efficient	Explore integration with Transformer-based architectures for real-time applications
Algimantas Venčkauskas [27]	Domain-specific ontology + semantic parser + Kernel Naïve Bayes	Custom metadata-based dataset	92.28% accuracy, 95.92% F1	Privacy-preserving (uses metadata only); strong performance	Limited to metadata (may miss semantic email content)	Extend ontology for multilingual or cross-domain spam detection
N. RamojiRao et al. ([28]	SVM vs. novel Naïve Bayes Classifier	5,172 emails	SVM: 89.78%; Naïve Bayes: 94.09% (statistically significant, p=0.036)	Demonstrates Naïve Bayes' superiority; statistical validation	Small dataset size; limited feature set	Apply on larger datasets; extend feature engineering for robustness
M. Alsuwit et al. [29]	ML + DL (LR, NB, RF, ANN)	Not specified	LR, RF, NB ≈ 97% accuracy	Shows ML still competitive	ANN performance not	Integrate hybrid ML-DL models;

				vs. DL; robust model comparison	highlighted; dataset not specified	apply to IoT email services
Dharani et al. [30]	Naïve Bayes + TF-IDF, implemented in local host (PyCharm IDE)	Kaggle dataset	95% accuracy	Simple and effective; real-world implementation on localhost	Limited to single ML algorithm; no ensemble/DL comparison	Extend to hybrid/ensemble approaches; deploy on cloud platforms
Temidayo et al. [31]	Ensemble (Random Forest, XGBoost) + Grid Search CV	Enron1 dataset	Tuned RF/XGBoost up to 97.78% (XGBoost best)	Ensemble with hyperparameter tuning outperforms baselines	Requires extensive tuning; higher training cost	Apply on streaming data; explore federated ensemble learning
K. Debnath et al. [32]	Deep learning (LSTM, BERT) + NLP preprocessing	Enron dataset	LSTM: 97.15% accuracy	Leverages sequential modeling; strong performance vs. older models	BERT computationally heavy; dataset limited to Enron	Apply transformers (DistilBERT, RoBERTa) for efficiency; expand datasets
Tasnia Toma et al. [33]	NB (Multinomial, Bernoulli, Gaussian), RF , SVM	Public email dataset	NB (97.6%), RF (97.8%), SVM (91.5%)	Clear comparison across classifiers; RF/NB highly accurate	Performance varies by NB type; lacks DL comparison	Incorporate DL or hybrid methods; evaluate robustness on noisy datasets
Saleem et al. [26]	Hybrid deep learning (LSTM + GRU)	Not specified (general email dataset)	90% accuracy	Combines LSTM's ability for long-term dependency with GRU's efficiency; addresses vanishing gradient problem	Computationally intensive compared to classical ML; Transformers can be more resource-efficient	Explore integration with Transformer-based architectures for real-time applications
Algimantas Venčkauskas [27]	Domain-specific ontology + semantic parser + Kernel Naïve Bayes	Custom metadata-based dataset	92.28% accuracy, 95.92% F1	Privacy-preserving (uses metadata only); strong performance	Limited to metadata (may miss semantic email content)	Extend ontology for multilingual or cross-domain spam detection

B. ***Research gaps/Problems***

In spite of significant progress in email spam detection, existing approaches face several limitations. First, many prior works rely on limited or domain-specific datasets, which restrict generalizability of their findings across diverse email environments. Second, class imbalance remains a major challenge, as spam datasets are often skewed toward non-spam emails, leading to biased classifiers and reduced detection of minority spam instances. Others achieve high accuracy but fail to address class imbalance, leading to biased predictions toward the majority (non-spam) class. Deep learning models, while effective, are computationally expensive and unsuitable for lightweight, real-time deployment. Similarly, metadata-based or single-model approaches often overlook semantic content or lack robustness against evolving spam patterns. Although several studies use single old machine learning or deep learning models, there is limited exploration of multiple ensembles techniques to systematically compare their effectiveness for spam detection. Addressing these gaps, the proposed work leverages the Spambase dataset, applies hybrid sampling techniques for effective class imbalance handling, and evaluates multiple boosting-based ensemble algorithms to identify the most robust and scalable model.

METHODOLOGY

The goal of this investigation is to utilized ML approaches to create an adept approach for detecting spam emails. The procedure entails gathering the Spambase dataset, preparing it, performing exploratory data analysis (EDA), using SMOTEENN to address class imbalance, and dividing dataset into training and testing sets. The best-performing model is then determined by applying boosting models (AdaBoost, Gradient Boosting, and CatBoost) and evaluating their performance using metrics including the confusion matrix, classification report, sensitivity, specificity, ROC curve, and precision-recall curve. Figure 1 illustrates the step-by-step workflow of the suggested methodology.

The flowchart of propound email spam detection system outlining the implementation steps with machine learning models is presented below:

C. ***Data Collection***

The process of obtaining unprocessed data from trustworthy sources in order to train and assess ML models is known as dataset collection. The dataset used in this work is Spambase dataset from UCI ML Repository. It cosists of 4600 rows and 58 columns, where 57 features represent different characteristics of emails, such as frequency of words and characters, and the 58th column is the target variable (0 = not spam, 1 = spam). This dataset is widely used in research for spam detection tasks because it provides a good mix of numerical features and a realistic distribution of spam and non-spam emails.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 23s,2025
https://theaspd.com/index.php

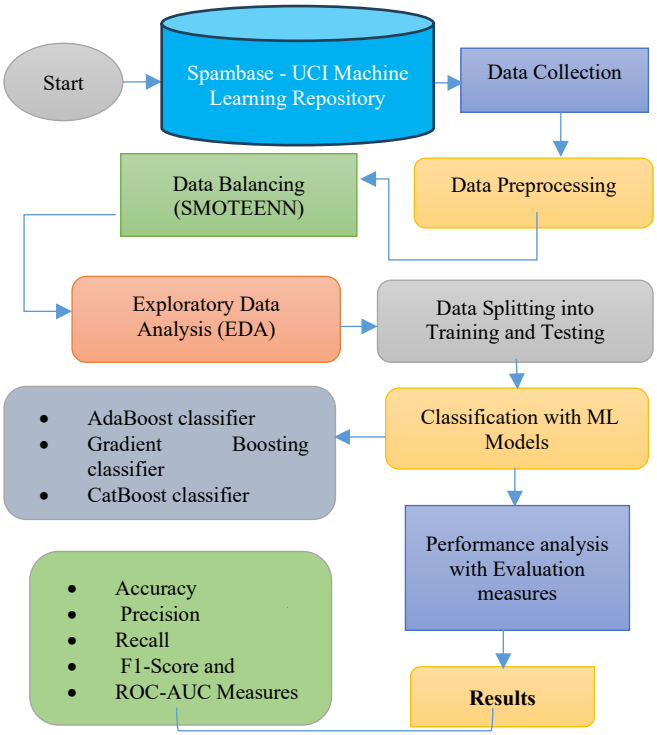

Fig. 1. Flowchart of proposed methodology for email spam identification system using ML techniques

D. Data Preprocessing

Data cleaning and pre-processing were the next steps, which ensured the data's quality and dependability and prepared it for appropriate analysis. The process of transforming raw data into a format suitable for ML is known as data preprocessing. To prepare data for model building, this step is crucial. Since the dataset does not contain predefined column names, meaningful names are manually assigned for better interpretability. The following preprocessing steps are performed:

- Checking dataset shape and structure to ensure data consistency.
- Data type verification using .info() to confirm correct numerical formats.
- Statistical summary analysis using. describe() to observe feature ranges, mean, and variance.
- **Separating features (X) and target (y):** All columns except the target column are taken as predictors, while the target column is used for classification.
- **Checking missing values and anomalies:** The dataset is clean and does not contain null values, making it ready for further processing.

This ensures the dataset is well-arranged and ready for further processing.

Fig. 2. Boxplots of dataset visualization

Figure 2, displays boxplots for each of the 57 features of the Spambase dataset. Using a five-number summary—the minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum—box-and-whisker charts are a standardised way to display a dataset's distribution. In Q1 through Q3, box itself delineates interquartile range (IQR), which encompasses the middle 50% of the data. Inside the box, the line indicates the median. The whiskers show the rest of the data distribution and extend from the box; points outside of them are considered outliers. The boxplots in the figure show that most of the features have a high number of outliers, with the median value for many features being close to zero.

E. Handling Imbalanced Data (SMOTEEN)

Datasets are frequently unbalanced, with a class having much more samples contrary to other, in many real-world classification tasks, such as email spam detection. When there are substantially more samples in one class than the other, this is known as class imbalance and can skew models[34]. A hybrid resampling technique known as SMOTEENN (Synthetic Minority Oversampling Technique + Edited Nearest Neighbours) was used to rectify this mismatch[35]. Initially, the dataset was imbalanced with 2788 non-spam (class 0) and 1812 spam (class 1) samples. To fix this, the SMOTEENN technique was applied.

SMOTE: In order to add additional data points without creating duplicates, it creates synthetic data by interpolating amidst minority class data samples that are currently available and their closest neighbour. Additionally, because minority class instance are expanded without duplication, overfitting may be avoided. It uses the following equation (1) to create synthetic data "x_new":

$$x_{new} = x_i + \lambda \times (x_{nn} - x_i) \quad (1)$$

Where: x_i= minority class, ⋋ = a random value amidst [0, 1]. and ⋋ = a random value amidst [0, 1].

SMOTE+ENN: This hybrid method ameliorates quality of synthetic data generated by SMOTE by removing instances of closest neighbour being incorrectly classified using Edited closest Neighbour (ENN)[36].

ENN cleaning following equation (2):

$$\left| \begin{matrix} if\ x_i\ is\ misclassified\ by\ its\ k\ nearest \\ neighbours, remove\ x_i \end{matrix} \right| \quad (2)$$

Equation (3) can be used to illustrate this method:

$$S_{balanced} = ENN\ (SMOTE(S_{minority}, S_{majority})) \quad (3)$$

- SMOTE(Sminority,Smajority) generates xnew.
- The ENN eliminates noisy data.

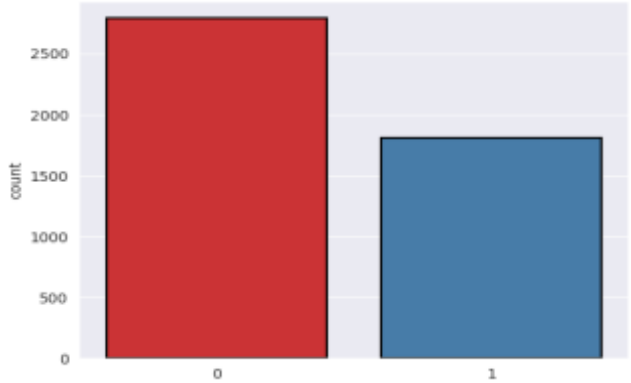

Fig. 3. Bar graph of data distribution classes after balancing

After applying SMOTEENN, the dataset was effectively balanced with around 2046 spam (class 1) and 1802 non-spam (class 0) emails, shows in bar graph figure 3. This balancing ensured that the boosting models received a fair representation of both classes during training, which improved their potentiality to correctly classify spam emails while maintaining accuracy for non-spam.

F. Data Splitting

After preprocessing the data, balancing it, and grouping it in terms of features and target labels, the second step was to divide the data into training and test sets. ML models were built and trained using the training set and then tested using the testing set. Of various techniques used to split a dataset, a fixed split strategy

was used in this study whereby the dataset was segregated into 70 percent training and 30 percent testing. This procedure demonstrates that a sufficient quantity of data was employed to train models, but at the same time, there were enough unseen samples to reliably assess the potentiality of the models, as well as the fact that SMOTEENN balancing was used to ensure fairness of data distribution amidst spam and non-spam classes in both subsets.

G. Classification ML Models Implementation

Model implementation is the process of applying ML algorithms to the prepared dataset to perform classification. Three ensemble boosting-based algorithms were implemented (AdaBoost, Gradient Boosting, and CatBoost). All three models were assessed using 10-fold cross-validation on both training and test datasets to ensure robust and unbiased performance measurement. These models discussed below:

1) AdaBoost

The family of algorithms known as "boosting" has a number of variations, the most well-known of which being AdaBoost. It is an ensemble learning technique that creates a strong classifier by amalgamating several weak classifiers, typically decision stumps. It functions by giving misclassified samples larger weights so that later classifiers can concentrate more on challenging instances[37]. The final classifier is given by eq. (4):

$$F(x) = sign(\sum_{m=1}^{M} \alpha_m \, h_m(x)) \quad (4)$$

Where:

- $h_m(x)$ = weak learner at iteration m.
- α_m= 1n $\left(\frac{1-\epsilon_m}{\epsilon_m}\right)$ is the weight of the weak learner, based on error ϵ_m.
- $F(x)$= final strong classifier

In the email spam detection task, AdaBoost was applied to iteratively improve classification performance by focusing on emails that were incorrectly stratified in earlier rounds. This helped increase sensitivity of the model toward detecting spam emails.

2) Gradient Boosting

Gradient Boosting is an advanced boosting algorithm where models are built sequentially, and each new learner attempts to minimize the errors (residuals) of prior model using gradient descent[38]. Unlike AdaBoost, which adjusts sample weights, Gradient Boosting fits new learners to the negative gradient of the loss function is given by eq. (5). At each iteration m:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (5)$$

Where:

- $F_m(x)$ = updated model after iteration m.
- $h_m(x)$= weak learner trained on residuals (errors) γ_m.
- γ_m = learning rate controlling contribution of $h_m(x)$.

For email spam detection, Gradient Boosting improved classification by systematically reducing residual errors from prior iterations. This sequential optimization allowed model to capture perplex patterns in the dataset, leading to better generalization on unseen test data.

3) CatBoost

The gradient boosting technique CatBoost was created especially to manage categorical information effectively and avoid overfitting. It uses ordered boosting and permutation-driven methods to reduce bias and variance[39]. The general CatBoost model follows the gradient boosting principle is given by eq. (6):

$$F_m(x) = F_{m-1}(x) + \eta . h_m(x) \quad (6)$$

Where:

- η =learning rate (in this work, 0.73)
- $F_m(x)$ = updated model after iteration m ((here, 300 iterations).
- $h_m(x)$= weak learner (decision tree of depth 4 in this case)

In this work, CatBoost was applied with tuned hyperparameters (iterations = 300, learning rate = 0.73, depth = 4). Its ordered boosting strategy minimized overfitting while improving classification accuracy. This

made CatBoost particularly effective compared to traditional boosting approaches in classifying spam and non-spam emails.

H. Performance Matrix

We computed confusion matrix, acc, rec, prec, and f-measure to assess email spam classification models. Additionally, we computed the ROC AUC to assess the efficacy of oversampling technique as our dataset was initially unbalanced. By offering a thorough comparison of AdaBoost, Gradient Boosting, and CatBoost, these assessment metrics aid in determining which model is best for email spam detection. Each measure's concept and computation are explained in depth in equations 7 through 11:

1) Accuracy

It is ratio of appropriately stratified instances (both spam and non-spam) among all predictions. While commonly used, it can sometimes be misleading in imbalanced datasets because it may favor the majority class.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (7)$$

2) Precision

The percentage of accurately anticipated positive cases (spam emails) out of all occurrences projected as positive is called precision, sometimes referred to as positive predictive value. It measures how well the model predicts whether an email is spam or not.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (8)$$

3) Recall (Sensitivity/TPR)

Recall is the rate of correctly recognized actual positive cases (spam emails) by the model. It delineates that model could not send spam emails away.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (9)$$

4) Specificity (FPR)

Specificity The percentage of true negative cases (non-spam mails) which are correctly detected. It is a sign of the accuracy of model in identifying the valid emails and not triggering a false spam flag.

$$FPR = \frac{FP}{FP+TN} \quad (10)$$

5) F1-Score

The harmonic mean between prec and rec is called the F1-score. It is a measurement that includes both FP and FN, which is especially beneficial when dealing with uneven data, such as spam.

$$F1 - Score = \frac{2*Precision*Recall}{(Precision+Recall)} \quad (11)$$

Where:

- **TP (true positives)** delineates aggregate count of emails that are appropriately categorised as spam,
- **FP (false positives)** delineates aggregate count of emails that were mistakenly categorised as spam emails,
- **FN (false negatives)** delineates aggregate count of emails that were mistakenly categorised as non-spam, and finally
- **TN (true negatives)** delineates aggregate count of emails that have been correctly classified as non-spam.

The confusion matrix brings out efficacy of the model. FP, FN, TP, and TN numbers are presented. Finally, the probability of a randomly sampled phishing email being rated higher than a randomly sampled authentic email is called the receiver operating characteristic area under the curve (ROC AUC) score. This metric can be computed by area under ROC curve.

Drawing on primary concerns that should be addressed to reach the objective of building credible spam detection system, in this work such factors like the imbalance of data, noise, and the need to categorize the data correctly are outlined. The chosen boosting algorithms have good capability of learning, and hybrid balancing procedures improve the impartiality among classes. The method provides not only accuracy but also strength in differentiating spam and legitimate emails by authenticating the performance using several evaluation measurements

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 23s,2025
https://theaspd.com/index.php

RESULT ANALYSIS AND DISCUSSION

This segment delineates the results analysis of ML models with experimental setup given first. All tests were done on a regular workstation equipped with an i7-core Intel/AMD processor (3.0+ GHz) with 16GB RAM and SSD storage, and without graphic acceleration. Python 310 with libraries NumPy, Pandas, scikit-learn, imbalanced-learn (SMOTEENN), CatBoost, Matplotlib, and Seaborn were used in the software environment. Also, cross-validation 10-fold was done on cross_val_predict. Table II shows performance of three boosting-based ML models of AdaBoost, Gradient Boosting (GBC), and CatBoost on the email spam classification problem. Their results were very high, with all models reaching above 97 per cent, proving the models to be effective in stratifying mails that are spam and those who are not. CatBoost was marginally better and showed the highest accuracy (98%), f-measure (98.15), and recall (97.91) which should reflect a better balance amidst prec and sensitivity. Gradient Boosting also showed competitive results with an acc of 97.92% and f-measure of 98.07%, while AdaBoost, though slightly lower, still maintained robust performance with 97.66% acc. Across all models, ROC-AUC reached 98%, highlighting their excellent discriminative ability. Overall, the results confirm that all three ensemble methods are reliable for spam detection, with CatBoost offering a marginal advantage in most evaluation metrics.

TABLE II. MACHINE LEANING MODELS' PERFORMANCE WITH MEASURES FOR EMAIL SPAM CLASSIFICATION

Measures	AdaBoost	GBC	CatBoost
Accuracy	97.66	97.92	98
Precision	98.06	98.38	98.38
Recall	97.59	97.75	97.91
F1-score	97.83	98.07	98.15
Sensitivity	97.59	97.75	97.91
Specificity	97.74	98.11	98.11
Roc-AUC	98	98	98

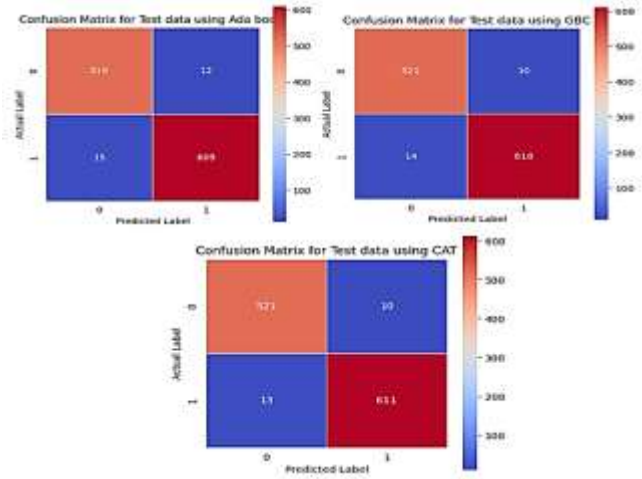

Fig. 4. Confusion matrix of proposed models for email spam

Figure 4 illustrates the confusion matrices of AdaBoost, Gradient Boosting (GBC), and CatBoost for email spam detection, highlighting their performance on both positive (spam) and negative (non-spam) classes. All three models demonstrate high accuracy in classifying non-spam emails (TN) and spam emails (TP), with only a small number of misclassifications. For the negative class (non-spam), AdaBoost correctly predicted 519 out of 531 cases, while GBC and CatBoost improved slightly to 521, showing stronger control over false alarms (false positives). For the positive class (spam), AdaBoost correctly classified 609 out of 624 cases, GBC classified 610, and CatBoost achieved the best with 611, thereby minimizing missed detections (false negatives). Since false negatives (missed spam) generally have a more serious impact in spam filtering systems compared to false positives, CatBoost's superior handling of the positive class provides an added advantage. Overall, these results indicate that while all three boosting models are

reliable, CatBoost demonstrates the most balanced performance by maximizing spam detection while minimizing false alarms.

Classification Report of Test data using Ada boost

	precision	recall	f1-score	support
0	0.97	0.98	0.97	531
1	0.98	0.98	0.98	624
accuracy			0.98	1155
macro avg	0.98	0.98	0.98	1155
weighted avg	0.98	0.98	0.98	1155

Classification Report of Test data using GBC

	precision	recall	f1-score	support
0	0.97	0.98	0.98	531
1	0.98	0.98	0.98	624
accuracy			0.98	1155
macro avg	0.98	0.98	0.98	1155
weighted avg	0.98	0.98	0.98	1155

Classification Report of Test data using CAT

	precision	recall	f1-score	support
0	0.98	0.98	0.98	531
1	0.98	0.98	0.98	624
accuracy			0.98	1155
macro avg	0.98	0.98	0.98	1155
weighted avg	0.98	0.98	0.98	1155

Fig. 5. Classification report of proposed models for email spam

Figure 5 presents the classification reports of AdaBoost, Gradient Boosting (GBC), and CatBoost models on the test dataset. All three models achieved nearly identical results with 98% overall accuracy, supported by consistently high prec, rec, and f-measure for both spam (class 1) and non-spam (class 0). AdaBoost slightly lagged in precision for non-spam class (0.97), while GBC and CatBoost achieved a balanced 0.98 across all metrics, showing stronger stability. These results confirm that the boosting models, particularly CatBoost, deliver reliable and consistent performance in distinguishing spam from legitimate emails.

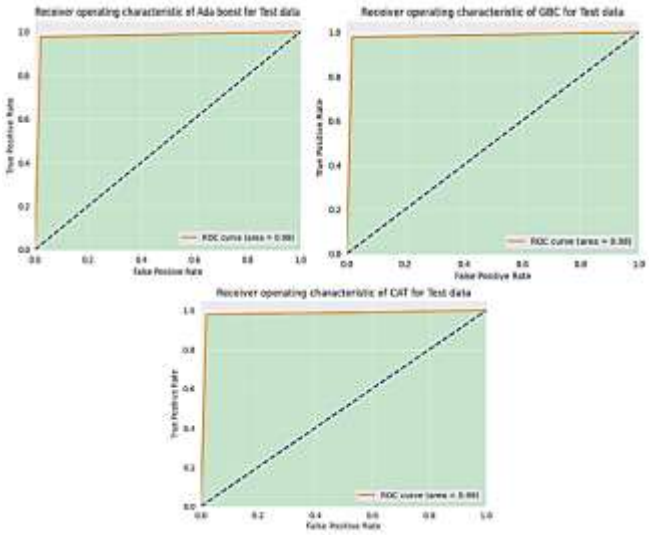

Fig. 6. ROC curve of proposed models for email spam

The ROC (Receiver Operating Characteristic) curves for AdaBoost, Gradient Boosting, and CatBoost demonstrate the strong discriminative power of the models shows in figure 6, with all achieving an AUC of 0.98. This indicates that the models effectively separate spam (positive class) from non-spam (negative class) emails across various thresholds. The curves are steep toward the top-left corner, highlighting very high TP rates with minimal FP, which is crucial for spam identification, as it minimizes the risk of letting spam slip into the inbox while maintaining reliable classification of legitimate emails.

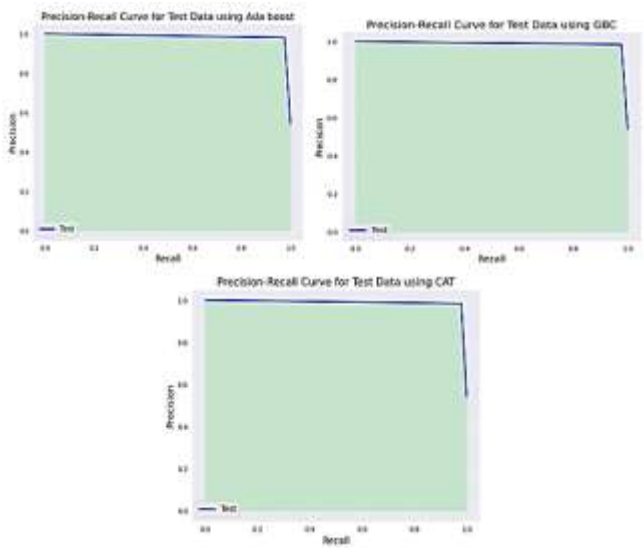

Fig. 7. PR curve of proposed models for email spam

The Precision–Recall (PR) curves further emphasize the robustness of the models in handling class imbalance, shows in figure 7. All three boosting models maintain high precision and recall values close to 1.0, with only minor drops at extreme thresholds. This means the models are highly capable of correctly identifying spam emails (high recall) while also ensuring that most emails labeled as spam are indeed spam (high precision). Compared to ROC, PR curves are more informative in imbalanced datasets, and the near-perfect PR curves here validate that the proposed models perform exceptionally well in separating emails that are spam from those that are not.

The results clearly justify the efficacy of the propound boosting models for email spam startification, as evidenced by high performance across all evaluation metrics. The confusion matrices highlight minimal misclassifications, with CatBoost showing the best balance between detecting spam and avoiding false alarms. The classification reports confirm consistent prec, rec, and f-measure near 0.98, ensuring reliability for both spam and non-spam classes. Furthermore, ROC curves with an AUC of 0.98 demonstrate strong discriminative power, while the PR curves validate robustness in handling class imbalance, maintaining near-perfect precision and recall. Collectively, these findings establish CatBoost as the most effective model, while AdaBoost and GBC also deliver competitive and reliable results.

DISCUSSION

The comparative scrut in Table II clearly highlights the superiority of suggested boosting models over several widely used existing approaches for email spam detection. Traditional models like LR (Logistic Regression) and DT (Decision Trees) deliver relatively good accuracy but fall short in precision and F1-score compared to boosting methods, reflecting their limitations in handling complex patterns and imbalanced data. DL methods such as DistilBERT and CNN show promising potential, with CNN performing competitively, yet they require significantly higher computational resources and may face challenges with smaller datasets. In contrast, the proposed ensemble boosting models—AdaBoost, Gradient Boosting, and particularly CatBoost, which achieved the best overall performance with an acc of 98%, prec of 98.38%, rec of 97.91%, and f-measure of 98.15%—consistently outperform existing baselines across all key evaluation measures. This demonstrates not only the robustness and adaptability of boosting algorithms but also their efficiency in achieving state-of-the-art efficacy without the heavy computational overhead of DL, making them highly practical for real-world spam filtering systems.

TABLE III. PROPOSED ML MODELS 'COMPARISON WITH EXISTING MODELS FOR EMAIL SPAM DETECTION

Models	Measures	Accuracy	Precision	Recall	F1-score
Proposed Models	AdaBoost	97.66	98.06	97.59	97.83
	GBC	97.92	98.38	97.75	98.07

	CatBoost	98	98.38	97.91	98.15
Existing Models	LR[26]	97	97.50	97.50	97.50
	DistilBERT [40]	93	93	92	92
	DT[41]	92.6	89	94	92
	CNN[42]	96.60	95.84	96.87	96.35

The main advantage of this work is the demonstration that ensemble boosting models, especially CatBoost, provide highly accurate and unwavering solutions for email spam stratification contrary to conventional ML and certain DL approaches. By tackling the class imbalance challenge with SMOTEENN, the system ensures balanced learning across spam and non-spam classes, thereby reducing false classifications. Moreover, the approach achieves excellent performance without requiring heavy computational resources or extensive feature engineering, making it both efficient and practical for real-world deployment. Overall, the study shows that boosting algorithms are a good option for efficient spam filtering systems due to their scalability, flexibility, and resilience.

CONCLUSION

Spam is a real nuisance to email-users as spam usually interferes with their work or leisure. Machine learning methodologies are often employed as the driving force behind spam detection solutions since they are effective and typically have a high classification accuracy level. The paper illustrates that boosting-based machine learning models are very effective in spam email detection and CatBoost has come out as the best performer. After applying SMOTEENN balancing and evaluating across multiple metrics, CatBoost achieved 98% acc, 98.38% prec, 97.91% rec, 98.15% f-measure, 97.91% sensitivity, 98.11% specificity, and 0.98 ROC-AUC. Compared to AdaBoost and Gradient Boosting, CatBoost maintained slightly higher stability across precision, recall, and specificity, reducing both FP and FN more effectively. These findings highlight the robustness of CatBoost in handling imbalanced data and complex feature interactions, making it highly apt for real-world email filtering applications. Overall, results emphasize advantage of advanced boosting algorithms in building secure, adaptive, and reliable spam detection systems.

This study delineates efficacy of boosting models for spam detection but has certain limitations, including dependancy on a single dataset, use of individual models, and limited feature engineering. Future work could address these gaps by testing on diverse datasets such as Enron, Ling-Spam, or TREC Public Spam Corpus, applying advanced NLP techniques like Word2Vec, GloVe, BERT, or RoBERTa for richer feature extraction, and exploring hybrid models that amalgamating boosting algorithms with DL (e.g., CNN–CatBoost, LSTM–XGBoost) or ensemble frameworks blending multiple boosting and transformer-based models. Such improvements would enhance generalizability, robustness against evolving spam patterns, and scalability for real-time email filtering systems.

REFERENCES

[1] A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti, and M. Alazab, "A comprehensive survey for intelligent spam email detection," 2019. doi: 10.1109/ACCESS.2019.2954791.

[2] 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*, vol. 12, pp. 143627–143657, 2024, doi: 10.1109/ACCESS.2024.3467996.

[3] P. Sharma and U. Bhardwaj, "Machine learning based spam E-mail detection," *Int. J. Intell. Eng. Syst.*, vol. 11, no. 3, pp. 1–10, 2018, doi: 10.22266/IJIES2018.0630.01.

[4] T. Gangavarapu, C. D. Jaidhar, and B. Chanduka, "Applicability of machine learning in spam and phishing email filtering: review and approaches," *Artif. Intell. Rev.*, vol. 53, no. 7, pp. 5019–5081, 2020, doi: 10.1007/s10462-020-09814-9.

[5] S. Larabi-Marie-Sainte, S. Ghouzali, T. Saba, L. Aburahmah, and R. Almohaini, "Improving spam email detection using deep recurrent neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 25, no. 3, pp. 1625–1633, 2022, doi: 10.11591/ijeecs.v25.i3.pp1625-1633.

[6] S. S. S. Neeli, "Critical Cybersecurity Strategies for Database Protection against Cyber Attacks," *J. Artif. Intell. Mach. Learn. Data Sci.*, vol. 1, no. 1, pp. 2102–2106, 2023, doi: 10.51219/JAIMLD/sethu-sesha-synam-neeli/461.

[7] S. B. Shah, "Machine Learning for Cyber Threat Detection and Prevention in Critical Infrastructure," *Dep. Oper. Bus. Anal. Inf. Syst. (OBAIS*, vol. 2, no. 2, pp. 1–7, 2025, doi: 10.5281/zenodo.14955016.

[8] R. Patel, "Automated Threat Detection and Risk Mitigation for ICS (Industrial Control Systems) Employing Deep Learning in Cybersecurity Defence," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, pp. 584–591, Dec. 2023, doi: 10.14741/ijcet/v.13.6.11.
[9] A. R. Bilipelli, "Forecasting the Evolution of Cyber Attacks in FinTech Using Transformer-Based Time Series Models," *Int. J. Res. Anal. Rev.*, vol. 12, no. 3, pp. 1–7, 2023.
[10] H. Kali, "The Future Of Hr Cybersecurity: Ai-Enabled Anomaly Detection In Workday Security," *Int. J. Recent Technol. Sci. Manag.*, vol. 8, no. 6, pp. 80–88, 2023.
[11] V. Prajapati, "Enhancing Threat Intelligence and Cyber Defense through Big Data Analytics: A Review Study," *J. Glob. Res. Math. Arch.*, vol. 12, no. 4, pp. 1–10, 2025.
[12] D. Goel and A. K. Jain, "Mobile phishing attacks and defence mechanisms: State of art and open research challenges," 2018. doi: 10.1016/j.cose.2017.12.006.
[13] S. Zavrak and S. Yilmaz, "Email spam detection using hierarchical attention hybrid deep learning method," *Expert Syst. Appl.*, vol. 233, 2023, doi: 10.1016/j.eswa.2023.120977.
[14] M. F. A. Kadir, A. F. A. Abidin, M. A. Mohamed, and N. A. Hamid, "Spam detection by using machine learning based binary classifier," *Indones. J. Electr. Eng. Comput. Sci.*, 2022, doi: 10.11591/ijeecs.v26.i1.pp310-317.
[15] Nirav Kumar Prajapati, "Federated Learning for Privacy-Preserving Cybersecurity: A Review on Secure Threat Detection," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 4, pp. 520–528, Apr. 2025, doi: 10.48175/IJARSCT-25168.
[16] F. Jáñez-Martino, R. Alaiz-Rodríguez, V. González-Castro, E. Fidalgo, and E. Alegre, "A review of spam email detection: analysis of spammer strategies and the dataset shift problem," *Artif. Intell. Rev.*, 2023, doi: 10.1007/s10462-022-10195-4.
[17] N. Prajapati, "The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications," *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, pp. 16–22, 2025, doi: 10.56472/25832646/JETA-V5I2P103.
[18] N. Ahmed, R. Amin, H. Aldabbas, D. Koundal, B. Alouffi, and T. Shah, "Machine Learning Techniques for Spam Detection in Email and IoT Platforms: Analysis and Research Challenges," 2022. doi: 10.1155/2022/1862888.
[19] M. Fayaz, A. Khan, J. U. Rahman, A. Alharbi, M. I. Uddin, and B. Alouffi, "Ensemble machine learning model for classification of spam product reviews," *Complexity*, 2020, doi: 10.1155/2020/8857570.
[20] S. A. Khamis, C. F. M. Foozy, M. F. A. Aziz, and N. Rahim, "Header Based Email Spam Detection Framework Using Support Vector Machine (SVM) Technique," in *Advances in Intelligent Systems and Computing*, 2020. doi: 10.1007/978-3-030-36056-6_6.
[21] I. Basyar, Adiwijaya, and D. T. Murdiansyah, "Email spam classification using gated recurrent unit and long short-term memory," *J. Comput. Sci.*, 2020, doi: 10.3844/JCSSP.2020.559.567.
[22] A. D. Wibisono, S. Dadi Rizkiono, and A. Wantoro, "FILTERING SPAM EMAIL MENGGUNAKAN METODE NAIVE BAYES," *TELEFORTECH J. Telemat. Inf. Technol.*, 2020, doi: 10.33365/tft.v1i1.685.
[23] S. Sinha, I. Ghosh, and S. C. Satapathy, "A study for ann model for spam classification," in *Advances in Intelligent Systems and Computing*, 2021. doi: 10.1007/978-981-15-5679-1_31.
[24] J. Lee, F. Tang, P. Ye, F. Abbasi, P. Hay, and D. M. Divakaran, "D-Fence: A flexible, efficient, and comprehensive phishing email detection system," in *Proceedings - 2021 IEEE European Symposium on Security and Privacy, Euro S and P 2021*, 2021. doi: 10.1109/EuroSP51992.2021.00045.
[25] K. Agarwal, P. Uniyal, S. Virendrasingh, S. Krishna, and V. Dutt, "Spam Mail Classification Using Ensemble and Non-Ensemble Machine Learning Algorithms," in *Lecture Notes in Networks and Systems*, 2021. doi: 10.1007/978-981-15-7106-0_18.
[26] S. Saleem, Z. U. Islam, S. S. U. Hasan, H. Akbar, M. F. Khan, and S. A. Ibrar, "Spam Email Detection Using Long Short-Term Memory and Gated Recurrent Unit," *Appl. Sci.*, vol. 15, no. 13, p. 7407, Jul. 2025, doi: 10.3390/app15137407.
[27] A. Venčkauskas, J. Toldinas, N. Morkevičius, and F. Sanfilippo, "An Email Cyber Threat Intelligence Method Using Domain Ontology and Machine Learning," *Electronics*, vol. 13, no. 14, p. 2716, Jul. 2024, doi: 10.3390/electronics13142716.
[28] N. RamojiRao, S. Anusuya, Harshavardhini, and C. B. Sembuli, "Analysis of Email Spam Detection Using Naive Bayes and Support Vector Machine Classification Algorithms," in *2024 IEEE 9th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, IEEE, Nov. 2024, pp. 1–5. doi: 10.1109/ICETAS62372.2024.11120193.
[29] M. H. Alsuwit, M. A. Haq, and M. A. Aleisa, "Advancing Email Spam Classification using Machine Learning and Deep Learning Techniques," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 4, pp. 14994–15001, Aug. 2024, doi: 10.48084/etasr.7631.
[30] V. Dharani, D. Hegde, and Mohana, "Spam SMS (or) Email Detection and Classification using Machine Learning," in *Proceedings - 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023*, 2023. doi: 10.1109/ICSSIT55814.2023.10060908.
[31] T. O. Omotehinwa and D. O. Oyewola, "Hyperparameter Optimization of Ensemble Models for Spam Email Detection," *Appl. Sci.*, vol. 13, no. 3, p. 1971, Feb. 2023, doi: 10.3390/app13031971.
[32] 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 2022*, 2022. doi: 10.1109/COM-IT-CON54601.2022.9850588.
[33] T. Toma, S. Hassan, and M. Arifuzzaman, "An Analysis of Supervised Machine Learning Algorithms for Spam Email Detection," in *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, IEEE, Jul. 2021, pp. 1–5. doi: 10.1109/ACMI53878.2021.9528108.

[34] F. S. Alsubaei, A. A. Almazroi, and N. Ayub, "Enhancing Phishing Detection: A Novel Hybrid Deep Learning Framework for Cybercrime Forensics," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3351946.
[35] C. Kumar, T. S. Bharti, and S. Prakash, "A hybrid Data-Driven framework for Spam detection in Online Social Network," in *Procedia Computer Science*, 2022. doi: 10.1016/j.procs.2022.12.408.
[36] Y. Zhu, Y. Hu, Q. Liu, H. Liu, C. Ma, and J. Yin, "A Hybrid Approach for Predicting Corporate Financial Risk: Integrating SMOTE-ENN and NGBoost," *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3323198.
[37] A. Ghourabi and M. Alohaly, "Enhancing Spam Message Classification and Detection Using Transformer-Based Embedding and Ensemble Learning," *Sensors*, 2023, doi: 10.3390/s23083861.
[38] S. Alrefaai, G. Ozdemir, and A. Mohamed, "Detecting Phishing Websites Using Machine Learning," in *HORA 2022 - 4th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, 2022. doi: 10.1109/HORA55278.2022.9799917.
[39] A. Odeh, Q. A. Al-Haija, A. Aref, and A. A. Taleb, "Comparative Study of CatBoost, XGBoost, and LightGBM for Enhanced URL Phishing Detection: A Performance Assessment," *J. Internet Serv. Inf. Secur.*, vol. 13, no. 4, pp. 1–11, 2023, doi: 10.58346/JISIS.2023.I4.001.
[40] T. Liu, S. Li, Y. Dong, Y. Mo, and S. He, "Spam Detection and Classification Based on DistilBERT Deep Learning Algorithm," *Appl. Sci. Eng. J. Adv. Res. Peer Rev. Ref. J. ISSN*, no. 3, pp. 6–10, 2024, doi: 10.5281/zenodo.11180575.
[41] P. Malhotra and S. Malik, "Spam Email Detection Using Machine Learning and Deep Learning Techniques," *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4145123.
[42] M. A. Wani, M. ElAffendi, and K. A. Shakil, "AI-Generated Spam Review Detection Framework with Deep Learning Algorithms and Natural Language Processing," *Computers*, vol. 13, no. 10, p. 264, Oct. 2024, doi: 10.3390/computers13100264.