International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

IDENTIFICATION OF PHISHING EMAILS WITH MACHINE LEARNING ALGORITHMS FOR IMPROVED CYBERSECURITY

Srinivasa Rao Singireddy
Architect, The Clorox Services Company, Pleasanton, CA.
srinivasarao.singireddy@gmail.com

Abstract

Phishing is a social engineering tactic primarily designed to acquire personal or private information, potentially causing damage to the targeted individual or organisation in many ways. Consequently, there is an imperative need for precise phishing detection. This project aims to create an efficient and dependable system for an early identification or classification of emails as phishing or benign, using two datasets from the Nazario Fraudulent Corpus and the Apache SpamAssassin website, including 4,600 phishing and benign emails. A classification framework is constructed using several ML models, including DT, SVM, multilayer perceptrons (MLP), and LSTM. The data has been divided into an 80:20 ratio for training and testing these models. The performance of these models was evaluated in the present study using classification metrics such as accuracy, recall, precision, and F1-score. A comparison analysis reveals that LSTM models surpass others, attaining the best recall(98.56%), accuracy(98.89%), precision(98.87%), and F1-score(98.87%), while decision tree, SVM, and MLP recorded lower accuracy rates of 98.06%, 96.90%, and 98.90%, respectively. These conclusions show how beneficial ML approaches are in improving abilities to distinguish between genuine and phishing URLs.

Keywords: *Phishing Emails Identification, Email Classification, Machine Learning, Cybersecurity.*

I. INTRODUCTION

Phishing e-mail messages are forged and are sent by the perpetrators with an intention to induce the receivables to surrender their identity information. In a phishing assault, the target user is tricked into divulging sensitive information, such as their passwords and credit card numbers, by promising to take them to a safe website or file [1]. The phisher may send out the mails to thousands of people and even though few of them will be deceived, the sender may make large profits. A most commontype of cybercrime is Phishing and this involves encouraging avictims to reveal their accounts details, passwords and banking details.

Some possible cyber threats that have adverse effects on an organization include: Email phishing is another attack that may cause loss of face, loss of funds and identity theft[2]. Phishing schemes have grown tremendously popular over the past decade with a minimum of millions of people falling for it each month. The alarm for organisations is that the task to protect them against this growing threat is becoming increasingly challenging due to this expansion. Criminally, the detection and prevention of email phishing has been made more challenging in recent time. Thus, phishers never run out of ideas since the work of security solutions and law enforcement agencies is to counter their schemes. To guard against this threat organisations require adaptable attack detection and prevention resources[3].

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

People are still needed for the identification of typical phishing techniques to inspect the body of the email message, the title, and sender. Phishing assaults are becoming more sophisticated, however, so current methods remain insufficient[4][5]. In recent times, DL and ML-based solutions have shown capable of surpassing the shortcomings of conventional phishing detection algorithms. Phishing email detection models may be trained using ML algorithms. These algorithms are capable of learning phishing patterns and traits from extensive datasets on the subject. Important elements connected to phishing activity must be determined before training can begin. For effective detection algorithms, this often calls for domain knowledge and meticulous feature selection.

1.1 Contribution of study

The research attempts to provide a ML framework for an identification of phishing emails, with an emphasis on enhancing cybersecurity in digital communication. Here are the key contributions of the study:

- To collect phishing email dataset for Phishing Emails with Machine Learning.
- Conduct data preprocessing that includes tokenisation, stemming with the Porter Stemmer, and the removal of stop words.
- The research emphasises the utilisation of ML classification models (DT, SVM, MLP, and LSTM) to identify phishing emails.
- Evaluated model efficiency with accuracy, F1-score, recall, and precision.

1.2 Structure of paper

A following paper organzaed as: Section I provide the topic overview with contribution. Then Section II and III provide the literature review on this topic and proposed methodology with each step. Section IV provide the experimental results and discussion of implemented AI models with comparative analysis. At last section V providesa conclusion and future work.

II. LITERATURE REVIEW

To better understand the previous work on email phishing detection. This section provides a literature review on phishing email identification.Alsosummry of the literature review discussed in table 1 below:

This study Zannat et al. (2023) strives to use deep learning techniques for email categorisation in the Bangla language. After looking at a number of algorithms and creating a dataset, the researchers in this study found that the Bi-LSTM technique could identify Bangla phishing emails with the highest accuracy (97%)[6].

In this paper, Pallavi and Jayarekha, (2023) provide a way to use ML algorithms to identify spam emails. Analysis of the Kaggle-obtained content-based filtering email dataset allowed for the engineering of the necessary features for training the ML models. They analysed the dataset's performance after testing several ML techniques. The outcomes prove that the proposed method is comparatively effective in the identification of spam mailing lists with a best accuracy extent of 99.8% and the Rmse of 0.2. More specifically, they applied a number of ML classifier algorithms

such as DT, Voting classifiers, RF, LR, etc on presented dataset to inspect which of them shows the highest level of accuracy. In using this technology, the email clients and servers will be able to independently recognise the spam emails and enhance the ability of their identification[7].

In this study, Debnath and Kar, (2022) with an aim of establishing models on email spam emails with the use of DL and ML to distinguish between spam and genuine communications. In this paper, the DL models that are LSTM and BERT are used to classify newly-identified email spam in the Enron email dataset. A method based on NLP was used to examine and prepare the email's content for data analysis. The outcomes are contrasted with the earlier models used for detecting spam in email. The proposed DL approach reached a maximum accuracy of 99.14% with BERT, 98.34% with BiLSTM, and 97.15% with LSTM. All implementations use Python[8].

In this study, Thakur et al. (2022) the goal of spam detection is to provide individuals with relevant emails while identifying and removing spam emails. Just about every email provider offers spam detection, but it isn't always accurate; in fact, it occasionally labels legitimate emails as spam. The comparative analysis technique is the main emphasis of this study. It involves applying several ML models to the same dataset. They compared the various ML models using the accuracy and precision metrics. The accuracy rate achieved by SVM is 98.09%[9].

In this paper, Toma, Hassan and Arifuzzaman, (2021) They investigate NB, SVM, and RFC using an existing email classification dataset and the supervised ML approach. Accuracy wasn't the only performance statistic shown; additional metrics including F1 score, precision, and recall were as well. In each technique, they achieved a high rate of accuracy, such as 98.8% for MNB, 97.6% for BNB, 91.5% for GNB, 97.8% for RFC, and the same for SVM, respectively [10].

In this study, Sonowal, (2020) The current body of knowledge fails to adequately handle the challenge of sound-alike terms, which is a critical limitation. This research thus suggests a paradigm called SPEDAS (Sounds-alike contents). Screen reader users may now get assistance in detecting phishing emails. This model can identify phishing emails by checking for material that sounds similar. The experimental findings show that the proposed model was 83% accurate.[11].

In this study, Octaviani et al. (2020) in order to compare the MNBC, SVM, and RNN algorithms in order to find the one that estimates spam in emails the most accurately. Using the Classification Report, we can see how each method performed according toprecision, accuracy, memory, and f1 score. This research highlights data showing that the SVM algorithm outperforms other methods for spam email categorisation with 96% accuracy, 0.92 precision, 0.96 recall, and 0.94 f1-score [12].

In this study, Niu et al. (2018) presents a model proposal called Cuckoo Search SVM abbreviated as CS-SVM. Also, the hybrid classifier is constructed from the 23 characteristics determined by the CS-SVM. This hybrid classifier opens a function of optimising the parameter selection of the RBF through the integration of Cuckoo Search (CS) to SVM. Among them 20,071 non phishing emails in total and 1,384 emails that are phishing are used in the experimentation. The research compares the suggested technique to a simple SVM classifier with default parameters and finds that it outperforms it in terms of phishing email categorisation.A max accuracy of 99.52% is achieved using the CS-SVM classifier [13] Table 1 gives a comprehensive summary of the related works that are discussed below.

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

TABLE I. SUMMARY OF RELATED WORK ON PHISHING EMAIL IDENTIFICATION USING MACHINE LEARNING

Ref	Methodology	Performance	Limitations	Future Work
Niu *et al.*, [13]	CS-SVM (Cuckoo Search integrated with SVM to optimize RBF parameters), extracts 23 features from phishing emails.	Accuracy: 99.52%	Focuses only on optimizing SVM parameters using Cuckoo Search. Limited to phishing email detection.	Exploring other hybrid optimization techniques for SVM parameter tuning. Applying model to other cybersecurity threats.
Debnath and Kar, [8]	LSTM and BERT-based models, NLP applied for text preprocessing, used Enron email dataset.	Accuracy: BERT 99.14%, BiLSTM 98.34%, LSTM 97.15%	Limited dataset (Enron), high computational resources required for deep learning models.	Explore ensemble approaches combining deep learning models for better performance and reduced resource usage.
Thakur *et al.*, [9]	Applied various machine learning models (SVM, Naïve Bayes, etc.), comparing performance metrics.	SVM Accuracy: 98.09%	Lacks deep learning comparisons. Limited features used for evaluation.	Expand analysis to include deep learning models and additional feature engineering.
Sonowal [11]	SPEDAS Model focused on detecting phishing emails with sound-alike content to assist screen readers.	Accuracy: 83%	Lower accuracy compared to other models, limited to sound-alike phishing detection.	Improve feature extraction techniques and extend the model to cover more types of phishing content.
Zannat *et al.*, [6]	Developed a new dataset for Bangla language phishing emails, used Bi-LSTM for classification.	Accuracy: 97%	Focused on Bangla language only, dataset is not widely available.	Extend the research to other languages and improve dataset availability for further research.
Octaviani *et al.*, [12]	Compared MNBC, SVM, RNN on spam email classification using Classification Report metrics.	SVM Accuracy: 96%, Precision: 0.92, Recall: 0.96, F1-Score: 0.94	Limited feature set, basic ML algorithms without deep learning models.	Test more advanced models like DL and ensemble techniques for better accuracy.
Pallavi and Jayarekha [7]	Used Decision Trees, Voting Classifiers, Random Forests, Logistic Regression. Features engineered from Kaggle dataset.	Accuracy: 99.8%, RMSE: 0.2	High performance achieved, but feature engineering could be further optimized.	Test other machine learning algorithms and extend feature selection techniques.
Toma, Hassan and Arifuzzaman [10]	Compared Naïve Bayes (Multinomial, Bernoulli, Gaussian), Random Forest, SVM using supervised learning techniques.	Naïve Bayes: 98.8%, 97.6%, 91.5%. Random Forest: 97.8%. SVM: 98.5%	Limited analysis on deep learning models, did not consider ensemble methods.	Incorporate deep learning models and ensemble approaches for better spam detection accuracy.

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

III. METHODOLOGY

In order to create a system that is effective as well as reliable in identifying and classifying fraudulent emails, this study gathered phishing email dataset. This dataset comprise phishing and benign emails, respectively. During the preprocessing phase, the email header, body, and text have been analysed to extract features. The concentration has been on URL-related attributes, including hexadecimal URLs, domain count, and IP-based URLs. Using the Porter Stemmer, text features have been parsed, tokenised, and stemmed. HTML elements and attachments have been processed, and tokens have been normalised by removing inflectional endings. In order tooptimise the model's functionality, stop phrases such as "the" and "then" have been eliminated. WordNet has been employed to apply semantic text processing to tokens, thereby enhancing the accuracy of classification and the identification of semantic relationships by incorporating conceptually related words. DT, SVM, MLP, and LSTM models were trained and validated using data that was split 80:20. Confusion matrices, accuracy, precision, and recall have been used to evaluate these models for email fraud classification. Following figure 1 provides proposed system architecture.

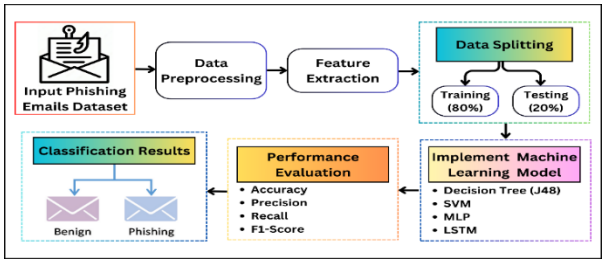

Fig. 1. Flowchart for Phishing Emails Identification

Each step of the flowchart for phishing email identification is listed below:

3.1 Dataset Description

The dataset used for phishing email detection comprises 4,600 authentic email messages, equally divided among 2,300 phishing emails from the Nazario Phishing Corpus and 2,300 benign emails obtained from the Apache SpamAssassin. Each email entry has elements such as "Website Index," denoting the email's unique identity, "Subject," encapsulating the email's title, "Content," elaborating on the body of the email, and "Content-Type," specifying the email's format. In order to determine how well the suggested model distinguishes between legitimate and phishing emails, this dataset is used for both training and evaluation purposes.

3.2 Data Pre-Processing

An important consideration in machine learning research is data preprocessing. The preparation of dependable input data sets is done prior to the building of learning models. The data preprocessing part of this research include removing elements from the email body, text, and header with an emphasis on URL-related information such IP URLs, domain counts, and hexadecimal URLs. Text features are extracted using parsing, tokenization, and stemming (via Porter Stemmer), where HTML tags and attachments are processed, and tokens are normalized by removing inflectional endings. Stop words like "the" and "then" are eliminated to enhance model performance. Additionally, semantic text processing is applied using WordNet to enrich tokens

with conceptually related words, improving the identification of semantic relationships and enhancing classification accuracy.

3.3 Data splitting

Cyber threat detection required partitioning the whole dataset into training and testing subgroups. Data used for testing makes up the remaining 20%, whereas 80% is used for training.

3.4 Machine Learning Classification Models

To understand the fundamental functioning and characteristics of the ML models used in this research. This section briefly describes the DT, SVM, MLP, and LSTM classifiers.

1) Decision Tree (J48)

A DT is a model that uses a tree to show many ways a choice may be made and the possible results of each one[14]. A characteristic is represented by a node, a decision by a branch, and a result (class or decision) by a leaf in a decision tree.

2) Support Vector Machine (SVM)

The speed and efficacy of SVMs have made them a popular supervised tool for text classification[15]. A two-dimensional line called a hyperplane is created using the given training data to effectively divide the categories. This hyperplane has been called the decision boundary by some.

3) Multilayer Perceptron (MLP)

MLPs are a subset of feedforward ANNs in which each layer is completely linked. The term "MLP" is used interchangeably with "ANN" when referring to these networks more generally, and "multiple layers of perception" when describing them more accurately[16][17].

4) Long Short-Term Memory (LSTM)

Because RNN designs like LSTM can capture long-term dependencies, they are frequently utilised in voice recognition, NLP, and time-series forecasting. Nonetheless, the diminishing gradient presents a problem for traditional RNNs. LSTM may express sequential data with long-term dependencies by storing and retrieving information from previous time steps using a memory cell [18]. The four primary parts of LSTM are a memory cell, three component gates (Input, Output, and Forget), and the neural network itself. The input gate controls how new inputs are allowed into a memory cell. An amount of information that a memory cell should erase from its prior state is determined by the forget gate. Lastly, the data storage and communication function is handled by the memory cell; the states that the cell outputs are controlled by the output gate. These LSTM gates are trained using the sigmoid and tanth activation functions. To indicate whether the gates are open or closed, the sigmoid function produces values between zero and one. Memory cell state strength is represented by the tanh function, which gives values between -1 and 1[19].
The optimal values for the parameters of the LSTM model are determined during training by minimising the loss function via backpropagation across time. With the output layer and weight matrices and bias vectors of every gate included in aparameters. The size of the gradient may be limited during training via gradient clipping. LSTMs can nevertheless experience disappearing or bursting gradients, depending on the gradient size. LSTMs may be used to forecast future prices

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

by detecting trends in past price data. The model predicts future prices by analysing a set of previous prices. While training the LSTM, it is necessary to optimise the loss function that quantifies the discrepancy between the actual and projected prices [20].
The following equations characterise the forward training procedure of an LSTM network [21]:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \dots\dots\dots\dots (1)$$
$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \dots\dots\dots\dots (2)$$
$$= f_t . c_{t-1} + i_t . tanh(W_c[h_{t-1}, x_t] + b_c) \dots\dots\dots\dots (3)$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \dots\dots\dots\dots (4)$$
$$h_t = O_t . \tanh(c_t) \dots\dots\dots\dots (5)$$

Time steps t indicate the input (xt), hidden state (ht), cell state (ct), input gate (it), forget gate (ft), and output gate (ot). W is used to represent the weight matrices, while b is used to represent the bias vectors. The output is bound between 0 and 1 using the sigmoid function and between -1 and 1 using the hyperbolic tangent function tanh, respectively.

3.5 Model Evaluation

Moreover, model evaluation is a critical component in the introduction of an ML system. Thus it helps us in judging how well an ML model works also gives us an idea on what is strength and what is the weakness of using this specific model. In this work, the performances of the ML models have been evaluated using f1-score, accuracy, precision, recall, and confusion matrix. A confusion matrix is one of the classification model performance metrics measurements that indicate its accuracy. It shows the two cases of false positive, false negative, true positive and lastly true negative. This matrix enables analysis of the accuracy of the model; identify cases of overhead and error in the classification and improvement of the forecasting error. For enhanced understanding of the confusion matrix, please refer to Figure 2 below.

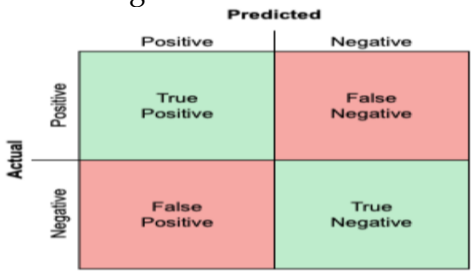

Fig. 2. Confusion Matrix

Accuracy: The accuracy may be expressed as a ratio of a number of accurate prediction (true positives and true negatives) as a percentage of the total forecasts. This can be formulated mathematically using the following formula Equation(6)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots\dots (6)$$

Precision: Precision of detection of genuine positives is determined using the percentage while precision quantifies the proportion of anticipated positive cases that actually are positive. This may be defined mathematically using Equation (7).

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

$$Precision = \frac{TP}{TP + FP} \ldots . (7)$$

Recall: The recall measures how accurate the percentage of true positive results is with reference to the total obtained number of positive results. When the positive class is crucial or when you want to minimise false negatives, this measure is beneficial. It can be defined mathematically using equation (8).

$$Recall = \frac{TP}{TP + FN} \ldots \ldots . (8)$$

F1-Score: The F1-score, which is an effective statistic when the dataset's class distribution is not uniform, is determined by taking the harmonic mean of the recall and precision. It can be defined mathematically using equation (9).

$$F1 - Score = \frac{2 * (Precision * Recall)}{Precision + Recall} .. (9)$$

These metrics assess how well ML models perform in identifying email phishing.

IV. RESULT ANALYSIS AND DISCUSSION

The ML models' experimental outcomes that were applied to the dataset of phishing emails. In order to facilitate comprehension of the experimental outcomes, the subsequent section implemented an assortment of graphs, charts, and tables.

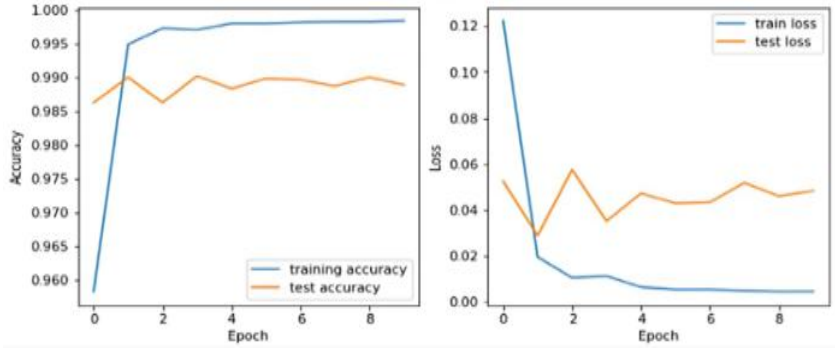

Fig. 3. Line Graph for the LSTM Model's Training and Testing Accuracy and Loss

Figure 3 shows the data used to evaluate the LSTM model's performance, including the training and testing accuracy and loss line graphs. The epochs are shown on the x-axis of the graph, while the accuracy and loss are shown on the y-axis for each epoch.The line graph demonstrates that the training accuracy increases significantly and stabilises around 100%, whilst the test accuracy shows minor oscillations but regularly exceeds 98%. The loss graph reveals a significant decline in training loss, signifying enhanced model performance, whilst the test loss exhibits minor fluctuations but mostly stays low, showing effective generalisation.

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

TABLE II. CLASSIFICATION REPORT OF LSTM MODEL FOR PHISHING EMAIL IDENTIFICATION

Classification Report of LSTM Model				
	Precision	**Recall**	**F1-Score**	**Support**
0	0.99	0.99	0.99	3081
1	0.99	0.99	0.99	2331
Micro Avg	0.99	0.99	0.99	5412
Macro Avg	0.99	0.99	0.99	5412
Weighted Avg	0.99	0.99	0.99	5412
Samples Avg	0.99	0.99	0.99	5412

A classification report evaluating the LSTM model's efficacy in identifying phishing emails is shown in Table II. The classification report includes information on the 0 and 1 classes in terms of f1-score, recall, accuracy, precision, and support. The results of the testing phase showed that with 3081 and 2331 support counts, respectively, the LSTM model consistently achieved 99% recall, accuracy, precision, and f1-score.

4.1 Comparative analysis

In this section, provide the ML models like DT[22], SVM [23], MLP[24], and LSTM comparison for phishing email detection for cybersecurity. The comparison of models is based on phishing email dataset and performance matrix including F1-score, recall, accuracy, and precision.

TABLE III. COMPARISON ANALYSISOF DIFFERENT ML MODELS FOR PHISHING EMAIL DETECTIONUSING CLASSIFICATION METRICS

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Decision Tree (J48)	98.06	98.20	97.90	98.10
SVM	96.90	97.0	96.90	96.90
MLP	98.37	98.90	97.83	98.36
LSTM	98.89	98.87	98.56	98.87

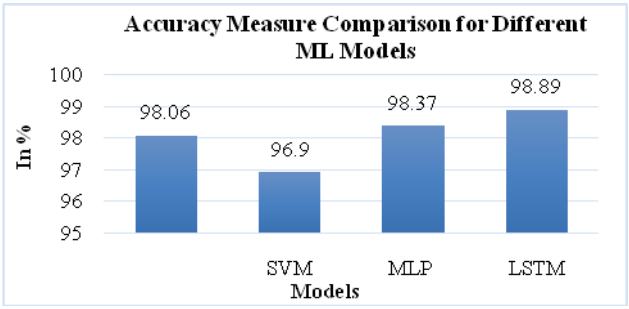

Fig. 4. Comparison of Accuracy Measures for Different ML Models in Phishing Emails Identification

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

Figure 4 and Table III present a comparison of accuracy measures by using different machine learning model's accuracy scores to determine an optimal model for identifying phishing emails. In the figure, the x-axis represents the ML models including decision tree, SVM, MLP, and LSTM, while a y-axis shows all these models' scores as a percentage. The graph clearly depicts that the LSTM model outperforms other models and achieved the highest accuracy score of 98.89%, while the other model's decision tree, SVM, and MLP obtained accuracy scores of 98.06%, 96.9% and 98.37%, respectively, throughout the testing phase. In conclusion, the comparison demonstrates that the LSTM is the most effective model for classifying phishing emails.

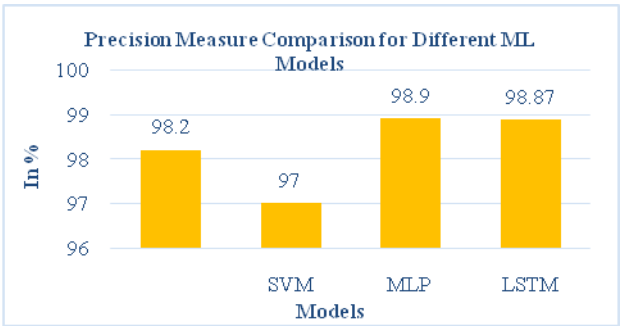

Fig. 5. Comparison of Precision Measures for Different ML Models in Phishing Emails Identification

Figure 5 and Table III present a comparison of precision measures by using different machine learning model's precision scores to determine an optimal model for identifying phishing emails. The graph clearly depicts that the MLP and LSTMmodels outperform others and achieved the highest precision scores of 98.9% and 98.87, while the other model's decision tree and SVMobtained lower precision scores of 98.2% and 97%, respectively. Hence, having analyzed all the earlier discussion one can state that MLP and LSTM can be considered as the most effective for being applied in the process of classification of phishing emails.

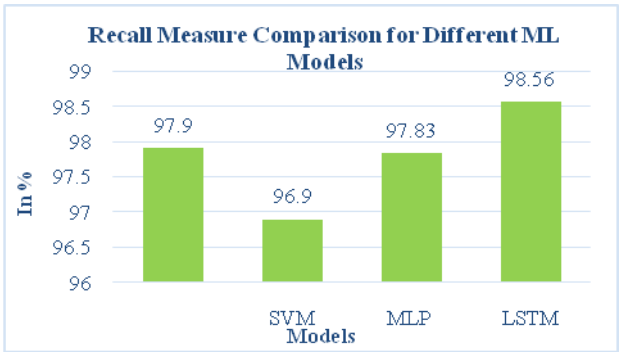

Fig. 6. Comparison of Recall Measures for Different ML Models in Phishing Emails Identification

The optimum model for spotting phishing emails was determined by comparing the recall scores of several models, as shown in Figure 6 and Table III. The LSTM model received the highest recall score of 98.56% while the rest of the classifiers gave out accuracy of 97.9% in decision tree, 96.9% in SVM, and 97.83% in MLP as presented in the above graph. Finally, it can be shown that the LSTM outperforms all other models in the classification of phishing emails.

International Journal of Core Engineering & Management
Volume-7, Issue-11, 2024 ISSN No: 2348-9510

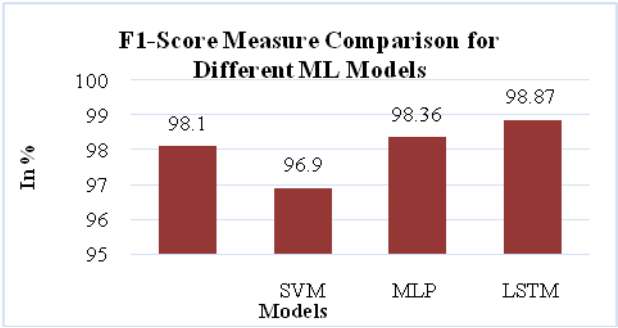

Fig. 7. Comparison of F1 Measures for Different ML Models in Phishing Emails Identification

The F1 metrics in Table III and Figure 7 are used to evaluate several models in order to find the one that is best at identifying fraudulent emails. As it can be seen from the graph below, LSTM model has the highest f1-score of 98.87% making it the best model. However, the decision tree model, SVM, and MLP model provided an f1 score of 98.1%, 96.9% and 98.36 respectively. Last but not least, the findings indicated that there is a superiority of LSTM when classify the phishing emails in contrast with the other models.

V. CONCLUSION AND FUTURE SCOPE

The most accurate description of a cybercrime that involves an average citizen of a nation sending emails or other forms of electronic contact to a victim with intent to fleece him /her is phishing attack. An attacker is an individual with an inherent criminal nature; he or she is capable of launching an attack through emails containing links with a risky content or through emails containing a payload that is designed to capture login details among other personal data. Such links in these emails cause all round losses in form of emotional losses, financial losses and physical losses such as loss of property and cash through identity theft and banking fraud. By specific contrast, this research sought only to verify or debunk the maliciousness of each given email using DT, SVM, LSTM, and MLP models. The F1-score, precisely, recall, and accuracy are evaluation criteria. Among the models tested, the LSTM model had the best results for detecting phishing emails (98.89% accuracy, 98.87% precision, 98.56 recall, and 98.87% F1-score). On the basis of the above-discussed results, LSTM could be confidently referred to as the approach of choice for addressing the problem of detecting phishing emails. Thus, to enhance its ability to respond to new threats, future research can focus on the expansion of the number of phishing approaches in the dataset. Ensemble techniques and other sophisticated machine learning algorithms might potentially enhance classification performance even more by combining the best features of many models. Other ways to improve the user experience include adding the ability to detect in real-time and creating interfaces that are easy to use for email clients. The system may be made more resistant to complex phishing efforts if behavioural analysis, such as user engagement with emails, were to be investigated for inclusion. Maintaining the system's efficacy in the ever-changing world of cyber threats requires continuous upgrades and modifications.

REFERENCES

1. P. Saraswat and M. Singh Solanki, "Phishing Detection in E-mails using Machine Learning," Proc. Int. Conf. Technol. Adv. Comput. Sci. ICTACS 2022, vol. 12, no. 7, pp. 420–424, 2022, doi: 10.1109/ICTACS56270.2022.9987839.
2. S. Atawneh and H. Aljehani, "Phishing Email Detection Model Using Deep Learning," Electron., 2023, doi: 10.3390/electronics12204261.
3. A. Safi and S. Singh, "A systematic literature review on phishing website detection techniques," J. King Saud Univ. - Comput. Inf. Sci., vol. 35, no. 2, pp. 590–611, 2023, doi: https://doi.org/10.1016/j.jksuci.2023.01.004.
4. O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," Expert Syst. Appl., 2019, doi: 10.1016/j.eswa.2018.09.029.
5. V. Kumar, A. Sahoo, and F. W. Liou, "Cyber-enabled product lifecycle management: A multi-agent framework," in Procedia Manufacturing, 2019. doi: 10.1016/j.promfg.2020.01.247.
6. R. Zannat, A. A. Mumu, A. Rahman Khan, T. Mubashshira, and S. R. Mahmud, "A Deep Learning-Based Approach for Detecting Bangla Spam Emails," in International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2023, 2023. doi: 10.1109/ICECCME57830.2023.10252671.
7. N. Pallavi and P. Jayarekha, "Efficient Spam Email Classification Using Machine Learning Algorithms," in 7th IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions, CSITSS 2023 - Proceedings, 2023. doi: 10.1109/CSITSS60515.2023.10334171.
8. 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.
9. P. Thakur, K. Joshi, P. Thakral, and S. Jain, "Detection of Email Spam using Machine Learning Algorithms: A Comparative Study," in 2022 8th International Conference on Signal Processing and Communication, ICSC 2022, 2022. doi: 10.1109/ICSC56524.2022.10009149.
10. 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 2021, 2021. doi: 10.1109/ACMI53878.2021.9528108.
11. G. Sonowal, "A model for detecting sounds-alike phishing email contents for persons with visual impairments," in Proceedings of the International Conference on e-Learning, ICEL, 2020. doi: 10.1109/econf51404.2020.9385451.
12. N. L. Octaviani, E. Hari Rachmawanto, C. A. Sari, and D. Rosal Ignatius Moses Setiadi, "Comparison of multinomial naïve bayes classifier, support vector machine, and recurrent neural network to classify email spams," in Proceedings - 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, iSemantic 2020, 2020. doi: 10.1109/iSemantic50169.2020.9234296.
13. W. Niu, X. Zhang, G. Yang, Z. Ma, and Z. Zhuo, "Phishing emails detection using CS-SVM," in Proceedings - 15th IEEE International Symposium on Parallel and Distributed Processing with Applications and 16th IEEE International Conference on Ubiquitous

Computing and Communications, ISPA/IUCC 2017, 2018. doi: 10.1109/ISPA/IUCC.2017.00160.
14. N. N. Kumar, "Model of Decision Tree for Email Classification," Int. J. Sci. Res., vol. 11, no. 7, pp. 1502–1505, 2022, doi: 10.21275/SR22722110223.
15. S. Rawal, B. Rawal, A. Shaheen, and S. Malik, "Phishing Detection in E-mails using Machine Learning," Int. J. Appl. Inf. Syst., vol. 12, no. 7, pp. 21–24, 2017, doi: 10.5120/ijais2017451713.
16. A. Odeh, "PHISHING WEBSITE DETECTION USING MULTILAYER PERCEPTRON," Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., vol. 24, no. December, pp. 194–198, 2021, doi: 10.32628/cseit217354.
17. H. Sinha, "Predicting Employee Performance in Business Environments Using Effective Machine Learning Models," Int. J. Nov. Res. Dev., vol. 9, no. 9, pp. 875–881, 2024.
18. K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Trans. Neural Networks Learn. Syst., 2017, doi: 10.1109/TNNLS.2016.2582924.
19. Z. Li et al., "Towards binary-valued gates for robust LSTM training," in 35th International Conference on Machine Learning, ICML 2018, 2018.
20. M. A. Istiake Sunny, M. M. S. Maswood, and A. G. Alharbi, "Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model," in 2nd Novel Intelligent and Leading Emerging Sciences Conference, NILES 2020, 2020. doi: 10.1109/NILES50944.2020.9257950.
21. P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," Fractal Fract., 2023, doi: 10.3390/fractalfract7020203.
22. A. Y. Daeef, R. Badlishah Ahmad, Y. Yacob, N. Yaakob, and M. N. Bin Mohd. Warip, "Phishing email classifiers evaluation: Email body and header approach," J. Theor. Appl. Inf. Technol., vol. 80, no. 2, pp. 354–361, 2015.
23. A. Yasin and A. Abuhasan, "An Intelligent Classification Model for Phishing Email Detection," Int. J. Netw. Secur. Its Appl., vol. 8, no. 4, pp. 55–72, 2016, doi: 10.5121/ijnsa.2016.8405.
24. S. A. Dalia, A. A. A. A. Hanan, and I. Abbas, "Effective Phishing Emails Detection Method," Turkish J. Comput. Math. Educ., vol. 12, no. 14, pp. 4898–4904, 2021.