Phishing Detection of Malicious URL using Machine Learning

ANIMESH DWIVEDI¹,CHIRAG TILWANI²,HARDIK SINGH³, LAVKESH PRASAD4, Prof. PRIYANKA M R5

*Department of Information Science and Engineering, Sapthagiri College of Engineering #14/5, Chikkasandra, Hesaraghatta main road, Bangalore-57 India*

*¹animes.dwivedi04@gmail.com*

*²chiragtilwani08@gmail.com*

*³*[*shrivastwautkarsh005@gmail.com*](mailto:madhuravani026@gmail.com)

*⁴lavkeshprasad@gmail.com*

[*5gayathrir@sapthagiri.edu.in*](mailto:5hodcse@sapthagiri.edu.in)

***Abstract*—** **Our study introduces an innovative approach for identifying phishing-linked harmful website links through advanced machine learning. Compared to traditional methods prone to bypass and manual checks, our system enhances accuracy, speed, and effectiveness. By integrating diverse machine learning algorithms with specialized URL analysis techniques, it swiftly detects malicious URLs within extensive online traffic. This breakthrough redefines cybersecurity, bolstering defenses against phishing attacks and fortifying digital security against evolving threats.**

# INTRODUCTION

Within the field of cybersecurity, spotting malicious URLs linked to phishing attacks has typically been done through manual examination, resulting in delays and uncertain results. Nonetheless, a new strategy that uses machine learning technology has recently surfaced to improve this procedure. By utilizing cutting-edge methods like YOLOv8 and Convolutional Neural Networks (CNNs), this technique strives to quickly identify potential phishing URLs with accuracy. Through leveraging CNNs to detect subtle patterns in URLs, even in the midst of intricate web layouts, this approach guarantees exceptional precision in differentiating harmful links from safe ones.We are currently testing and validating our research on different datasets to create a strong system for detecting malicious URLs in various phishing situations. By combining machine learning and URL analysis, we aim to provide cybersecurity professionals with a reliable tool for quickly and accurately detecting phishing attempts. This paper explains our methods, including the algorithms we use, the experiments we conducted, and our initial results, all of which are crucial for improving cybersecurity defenses.

# LITERATURE SURVEY

The rise in financial crimes utilizing technology is a pressing issue that needs immediate attention. However, traditional classification methods heavily rely on prior knowledge of features which can limit their effectiveness. To tackle this, a hybrid model combining Deep Belief Network's deep learning capabilities with Support Vector Machines' machine learning methods is proposed in this paper. The model first extracts features from unidentified URLs through blacklist filtering, including statistical, webpage code, and webpage text features. It then uses a deep learning model to extract deep features for quick classification.In the end, the feature vectors that result from combining statistical features, webpage code features, and webpage text features are inputted into an SVM model for classification. This model was examined using a dataset that included millions of phishing URLs and legitimate URLs. It was able to achieve an accuracy of 99.96%, a precision rate of 99.94%, and a false positive rate of 51.32%. These results demonstrate that the model outperformed other comparison models.[1]

With the rapid expansion of the Internet, people have shifted from traditional shopping to online commerce. Criminals have also adapted by moving their focus from bank and shop robberies to cybercrimes. They now target victims in the virtual world using tactics like phishing and setting up fake websites to steal sensitive information such as account details and passwords. Distinguishing between a genuine website and a phishing site has become increasingly difficult due to the sophisticated tactics used by attackers that exploit the vulnerabilities of computer users.Despite the efforts of software companies to release new anti-phishing products with various technologies like blacklists, heuristics, visual and machine learning-based approaches, these tools are not foolproof in preventing all phishing attacks. This paper introduces a real-time anti-phishing system that utilizes seven different classification algorithms and features based on natural language processing (NLP). This system stands out from other research in the field due to its language independence, extensive use of phishing and legitimate data, real-time capabilities, ability to detect new websites, independence from third-party services, and use of feature-rich classifiers.In order to evaluate how well the system is performing, we created a fresh set of data and ran tests on it. After analyzing the results from various classification algorithms, it was found that the Random Forest algorithm, using only NLP features, achieved the highest accuracy rate of 97.98% for detecting phishing URLs.[2]

Phishing attacks have become a significant threat in our daily lives and online networks. Attackers disguise malicious URLs as legitimate ones in order to trick users into visiting them and divulging private information. It is crucial to find effective ways to detect these phishing websites and minimize the risks they pose. One popular method is using neural networks, which have the ability to learn from large data sets. However, during the training process, the model can become overwhelmed by irrelevant features, leading to overfitting and rendering it ineffective in detecting phishing sites.To solve this issue, this document presents OFS-NN, a smart model for detecting phishing websites using a combination of optimal feature selection and neural network technology. With OFS-NN, we introduce a new metric called feature validity value (FVV) to assess the importance of specific features in identifying phishing websites. Using this FVV metric, we develop an algorithm to pinpoint the best features for detecting phishing websites, helping to prevent the neural network from becoming too specialized. These chosen features are then used to train the neural network, resulting in a precise classifier that can identify phishing websites effectively.To solve this issue, we introduce OFS-NN, a model for detecting phishing websites using optimal feature selection and neural network. With OFS-NN, we introduce a new metric called feature validity value (FVV) to measure the impact of sensitive features on phishing detection. Using FVV, we design an algorithm to select the best features for detecting phishing sites, addressing overfitting in the neural network. These chosen features are then used to train the neural network, creating an accurate classifier for identifying phishing websites.Based on the experiments, the OFS-NN model has proven to be reliable and consistent in identifying various kinds of phishing websites.[3]

# MOTIVATION

The study, titled "Improving Cybersecurity with Machine Learning-based Phishing Detection," tackles important issues in digital security. By using advanced technologies like machine learning and artificial intelligence, the research aims to transform how we detect harmful URLs linked to phishing attempts. This breakthrough has the power to enhance cybersecurity efforts by spotting phishing threats early on and reducing risks for both individuals and organizations.In addition, the AI-powered detection systems introduced in this study improve the precision of identifying threats and make better use of resources, especially in areas where there is a lack of cybersecurity skills. By simplifying the detection process and cutting down on analysis time, these technologies enable cybersecurity experts to quickly determine and address possible dangers, strengthening online protections against increasing cyber threats.In the end, this research aims to enhance cybersecurity results by helping quickly and accurately identify harmful URLs, as well as opening the door for incorporating machine learning in telemedicine programs, especially in underserved areas with limited healthcare resources.

# PROBLEM DOMAIN

In the domain of "Phishing Detection of Malicious URLs using Machine Learning," several challenges exist that researchers must address to develop effective detection systems. Some of the prominent problems include

i. Data Pre-processing: In this stage, we clean and standardize URL datasets, handle missing values, and encode categorical features. We also perform feature engineering to extract useful information from URLs, such as domain age, URL length, and the presence of special characters.

ii. Feature Extraction: This step involves identifying significant features or patterns in URLs that differentiate between legitimate and malicious links. Techniques like tokenization, n-gram analysis, and domain analysis are used to capture characteristics that suggest phishing attempts.

iii. Creating and implementing machine learning algorithms, such as Random Forest, Gradient Boosting, or Deep Learning models like Recurrent Neural Networks (RNNs) or Convolutional Neural Networks (CNNs), to determine if URLs are safe or harmful based on specific characteristics.

iv. Testing and measuring the effectiveness of the developed models using various evaluation methods like accuracy, precision, recall, F1-score, and ROC-AUC curve. Additionally, cross-validation techniques can be used to validate the reliability of the models.

v.Real-world Deployment and Integration: The process of incorporating the created models into current security systems or web browsers to offer immediate protection from phishing attacks. It is vital to guarantee scalability, efficiency, and compatibility with various platforms when deploying.

# PROBLEM DEFINITION

The study is focused on finding ways to detect harmful URLs in phishing attacks using machine learning. The goal is to create strong algorithms that can accurately identify which URLs are safe and which ones are malicious in order to improve cybersecurity. The research is looking at ways to solve challenges like unbalanced data, complex feature engineering, and vulnerability to attacks that currently limit the effectiveness of detection systems.

The research is using new machine learning techniques to improve how phishing detection systems work. This will help make them more accurate, easy to scale, and effective in various situations. By studying real-world data and analyzing it, the goal is to provide fresh ideas and methods that can improve how we detect phishing attacks. This will lead to better cybersecurity solutions that can protect individuals, businesses, and online communities.

# PROBLEM STATEMENT

“The challenge lies in developing machine learning algorithms capable of accurately identifying malicious URLs amidst vast online traffic, addressing the pressing need for proactive cybersecurity measures against phishing attacks.”

# PROBLEM FORMULATION

Building a strong system to identify harmful URLs using image analysis and machine learning involves the following steps:

1. Pre-processing of URLs: The initial processing includes parsing, feature extraction, and normalization to prepare the data for analysis. This step aims to improve the quality of URL data and reduce noise for better model performance.
2. Extracting features: Relevant features like domain characteristics, lexical patterns, and structural attributes are extracted from the URLs and converted into numerical representations. This process helps quantify the differences between benign and harmful URLs.
3. Creating the Machine Learning Model: Use machine learning methods like supervised learning and ensemble techniques to build a classification model with labeled URL data. A convolutional neural network (CNN) or other relevant architectural designs can be used to detect small patterns that suggest malicious behavior.
4. Measuring Performance: Choose the right evaluation metrics like accuracy, precision, recall, and F1 score to evaluate how well the model performs. Thorough validation on new datasets guarantees the trustworthiness and applicability of the detection system.In this organized method, image processing and machine learning techniques are combined to effectively pinpoint harmful URLs, offering a methodical structure for tackling phishing dangers with precise calculations.

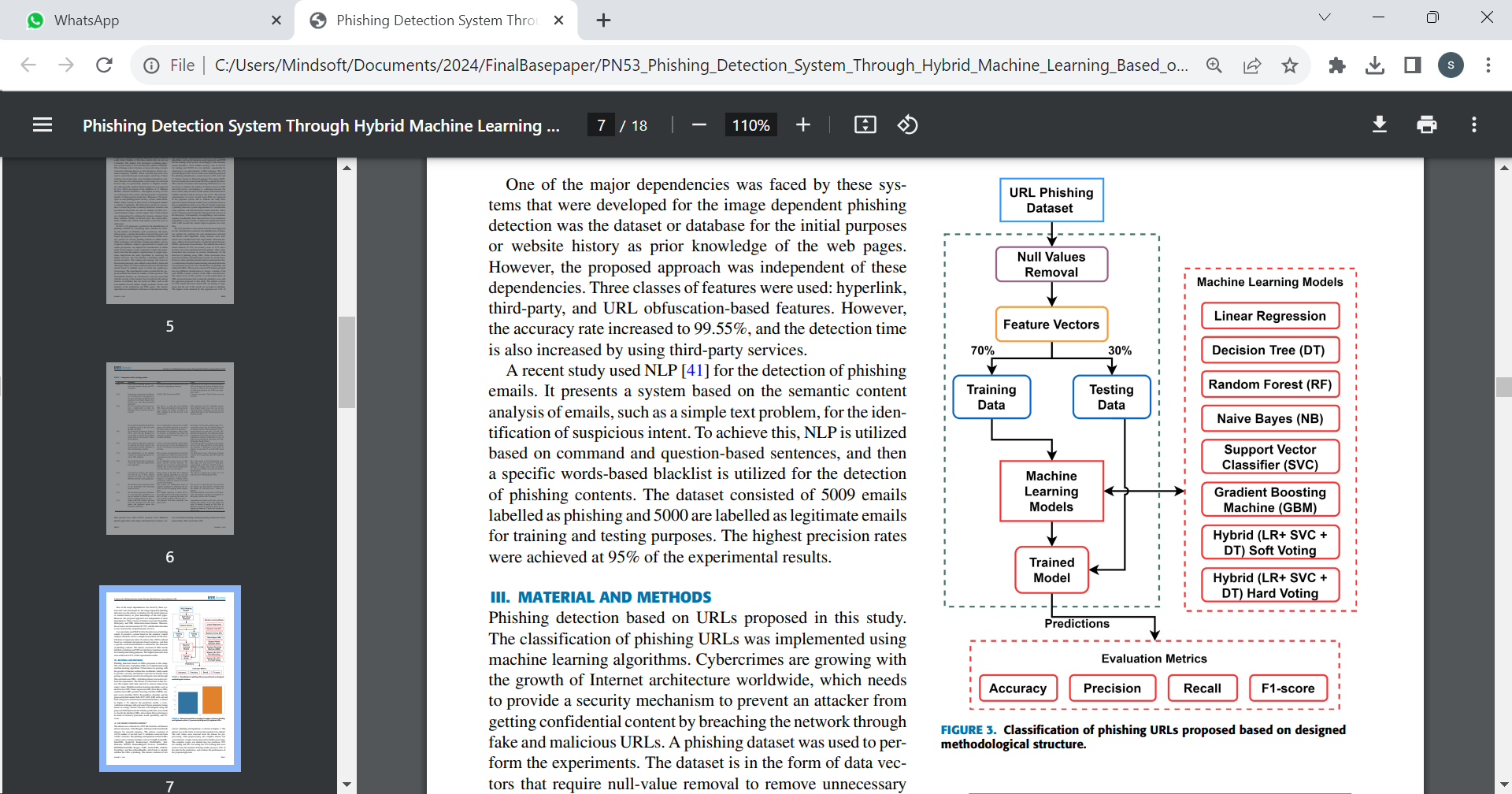


Fig 1:System Architecture

The diagram in Fig. 2 showcases the layout of the detection model for malicious URLs. The system first trains on prepared URL data and then extracts features. These features guide the machine learning model's training, allowing it to differentiate between harmless and harmful URLs. The model that has been trained can subsequently be utilized to spot malicious URLs in live applications, such as those provided by users for assessment.

# PROBLEM SOLVING AND METHODOLOGY

To address the challenge of detecting bone fractures by analyzing images and ML technologies like SMOTE and Voting Classifier, we can consider the following solution approaches:

SMOTE (Synthetic Minority Over-sampling Technique) is a technique used to tackle class imbalance in machine learning data, especially when the minority class is not well-represented. The method involves creating artificial samples for the minority class by using the existing minority samples. This is done by selecting a sample from the minority class and finding its closest neighbors within the same class. A random scaling factor is then chosen, and new synthetic samples are generated by placing points along the lines connecting the original sample to its neighbors, at a specific distance determined by the scaling factor.We keep doing this until we have enough fake samples. SMOTE is better than just copying minority samples because it creates new ones, but it might not be perfect and could make unrealistic samples or not show the whole data pattern.

Voting classifier: The voting classifier is a basic way of merging various classification algorithms, and choosing the right combination rule is crucial for creating classifier ensembles. Voting involves merging the predictions of multiple machine learning algorithms. The average voting approach combines the predictions of three algorithms: random forest classifier, support vector machine classifier, and naive Bayes classifier. These algorithms showed superior accuracy compared to others, so they were chosen for the average voting classification. This method performed effectively, achieving an accuracy of 95.75%.In this study, we utilized a voting classifier with the stacking method, incorporating three different classifiers: support vector machine, random forest, and naive Bayes. Combining these classifiers resulted in a high accuracy rate of around 97.24%, making it the top performer among all algorithms and voting classifiers. The approach involved creating multiple estimators individually and then averaging their predictions, leading to a more accurate overall result compared to using a single estimator.

# RESULT AND SENSITIVITY ANALYSIS

# 

# *RESULT ANALAYSIS*

Fig. 2 and Fig. 3 depicts the result of phishing detection model. This can be furthered described as:

1. Data collection: Machine learning can help detect phishing attacks by collecting a vast amount of phishing and legitimate URLs from sources like public blacklists, web browsing history, and honey pots.
2. Feature Extraction: Features are then extracted from these URLs and websites.
3. Model training: When training an AI model, it is important to use labeled data. The model will then learn to recognize common features in phishing websites.
4. Model evaluation: After training, the model must be evaluated using a separate dataset containing both phishing and legitimate URLs. This evaluation helps to determine the accuracy of the model in detecting phishing websites.
5. Deployment: Once the model has been trained and evaluated, it can be deployed in various ways, such as a web browser extension or security product. This allows the model to assess whether a website is likely to be a phishing site when a user visits it.

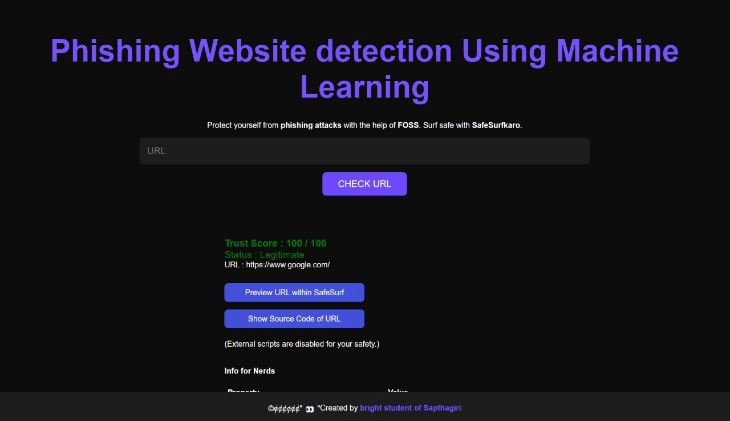


Fig.2: Result of Legitimate URL

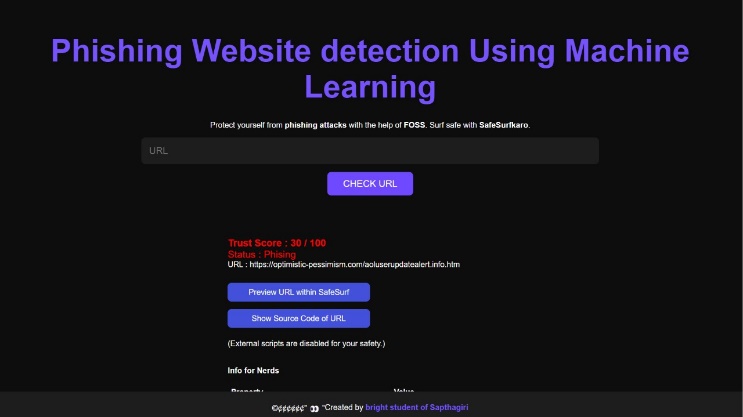


Fig.3: Result of malicious URL

*SENSITIVITY ANALAYSIS*

In comparing the SVM model to the DBN-SVM model, the latter showed a 4% increase in accuracy and false negative rate, as well as a 5% higher precision. This may be attributed to SVM's heavy reliance on labeled training data, which limits performance when there is an inadequate amount of labeled malicious URLs. On the other hand, DBN-SVM combines neural network advantages by extracting deep features and not depending on a single source for features. As for the comparison between the CNN model and DBN-SVM, the latter showed a 3% increase in accuracy, as well as an average of 3% higher false negative rate and precision. This difference could be due to the gradient descent in CNN easily converging to the local minimum rather than the global minimum.When we look at the SVM model versus the DBN-SVM model, the DBN-SVM model displayed a 4% improvement in accuracy and false negative rate, along with a 5% increase in precision.

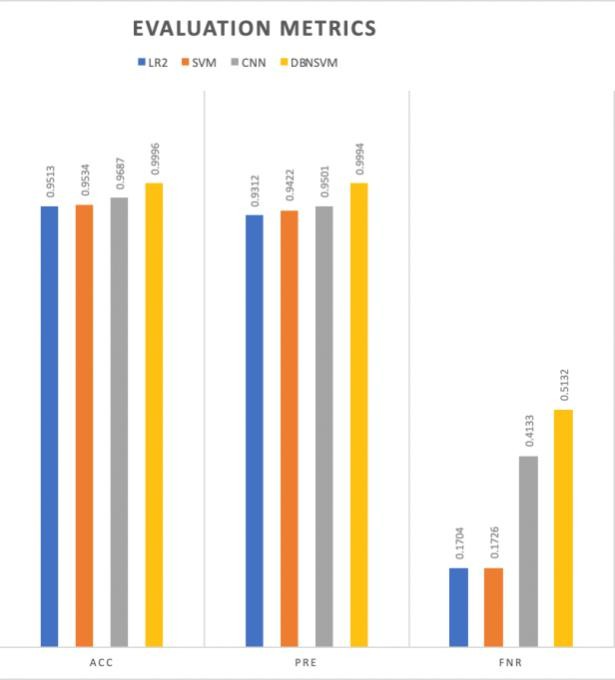


Fig.4: Evaluation metrics of the models.

# COMPARISION OF RESULTS

For comparing the performance of phishing detection models, it can be calculated without URL length and with URL length. It can be calculated for different classifiers using various models, including an ensemble model (SMOTE and MAJORITY VOTING). The results can be presented in a table for easy comparison.

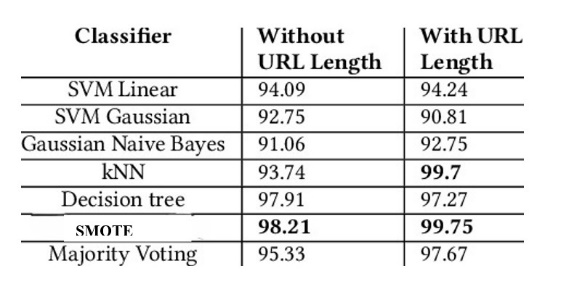


Table I. COMPARISON OF RESULTS

The comparison table of URL lengths for various machine learning classifiers. The classifiers are arranged in order of increasing URL length without URLs. Without URLs, the lengths are all shorter compared to when URLs are present. For instance, SVM Linear has a length of 94.09 without a URL and 94.24 with a URL.

# CONCLUSION

# Our ML-driven system for spotting malicious URLs in phishing attacks is a big step forward in boosting online security. Just like giving medicine for a broken bone, our software quickly deals with potential dangers to protect users from cyber attacks. By constantly improving our model through training, we make sure it's able to catch phishing attempts accurately. This ongoing process not only makes our software more effective, but also shows our dedication to staying ahead of changing online threats.At the heart of our approach is a focus on taking proactive steps to offer strong solutions that reduce risks and create safer online spaces for people and businesses. Our commitment to constantly getting better shows how dedicated we are to moving forward with cybersecurity efforts and making sure digital systems can handle new obstacles.

# FUTURE WORK

* Implementing proactive defense mechanisms: It involves creating sophisticated algorithms that can predict changing phishing tactics and proactively detect new threats to strengthen cybersecurity defenses.
* Integration of real-time analysis::Improving detection systems efficiency by incorporating real-time analysis for continuous monitoring and immediate response to dubious URLs helps decrease response times and minimize risks.

# ACKNOWLEDGMENT

We want to sincerely thank all those who played a part in the research project 'Phishing Detection of Malicious URLs using Machine Learning.' A special shoutout goes to our advisors and mentors for their priceless guidance, as well as healthcare providers and institutions for giving us access to vital datasets and tools. We also recognize the dedication of the research community in image analysis and machine learning, which set the foundation for our project. Lastly, we extend our gratitude to all supporters who helped us improve cybersecurity technology. The success of this project is a result of everyone's collective efforts and support, and we are truly grateful for their contributions

# REFERENCES

1. “Phishing website detection based on Hybrid Model of Deep Belief Network and Support Vector Machine” Xuqiao Yu 2020 IOP Conf. Ser.:Earth Environ. Sci. 602 012001.
2. “Machine learning based phishing detection from URLs” Ebubekir Buber,Banu Diri,January 2019.
3. “OFS-NN: An Effective phishing website detection model based on optimal feature selection and neural network” Erzhou Zhu, Yuyang Chen, Chengcheng Ye,Xuejun Li and Feng Liu, June 18,2019.
4. “A multivocal literature review on growing social engineering based cyber-attacks/threats during the COVID-19 Pandemic: Challenges and Prospective Solutions” Mohammad Hijji,Gulzar Alam,January 2021.
5. “PhiDMA – A phishing detection model with multi-filter approach” Gunikhan Sonowal,K.S. Kuppusamy,July 23,2017.
6. “An Automated Framework for Real-time Phishing URL Detection” Farhan Sadique, Shahriar Badsha,Shamik Sengupta,December 2019.
7. “PhishTime: Continuous Longitudinal Measurement of the Effectiveness of Anti-phishing Blacklists”Adam Oest,Yeganesh Safaei,Penghui Zhang Brad Wardman,Kevin Tyer,August 14 2020.
8. “On Cloaking Behaviors of Malicious Websites”Nayanamana Samarasinghe,Mohammad Mannan.
9. “Exploring A Robust Machine Learning Classigier for Detecting Phishing Domains using SSL Certificates”Akanchha Akanchha,April 2020.
10. “Optimization of URL - Based Phishing Websites Detection through Genetic Algorithm” Muhammad Taseer Sulemana and Shahid Mahmood Awana, April 23,2019.
11. “An Analysis of phishing blacklists: Google Safe Browsing ,openPhish, and PhishTank Simon Bell,February,2020.
12. “Phishing Detection System Through Hybrid Machinge Learning Based on URL” Abdul Karim,Mobeen Shahroz,Khabib Mustofa,Samir Brahim Belhaouari,March 3, 2023.
13. “Phishing Detection Leveraging Machine Learning and Deep Learning: A Review” Dinil Mon Divakaran,16 May 2022.