# ANALYSING MACHINE LEARNING TECHNIQUES FOR PHISHING DETECTION: A COMPARATIVE STUDY

**ADITYA SAXENA**

STUDENT, UG

NATIONAL POST GRADUATE COLLEGE

COLLEGE, LUCKNOW, UTTAR PRADESH, INDIA

PRADESH, INDIA

adityasaxena2003@gmail.com

**MR. MAHESH KUMAR TIWARI**

ASSISTANT PROFESSOR

NATIONAL POST GRADUATE, LUCKNOW, UTTAR PRADESH, INDIA

maheshyogi26@gmail.com

**KEYWORDS**

Phishing attacks, Online security, Machine learning models, Phishing detection, Diverse dataset, Traditional classifiers, Advanced techniques, Neural networks, Ensemble methods, Features, Website

## ABSTRACT

With the advent of digital world, there are several threats that have become a cause of grave security and privacy concern and amongst them is phishing. Phishing attacks pose a significant threat to individuals, organizations, and online security. As the sophistication of phishing campaigns continues to evolve, the need for effective detection methods becomes paramount. This abstract provides an overview of a comprehensive study that compares various machine learning models for the task of phishing detection. In this research, a diverse dataset of phishing and legitimate websites is employed to assess the performance of different machine learning

content, URL structure, User behaviour, Evaluation metrics, Accuracy, Precision, Recall, F1score, AUC-ROC, Adversarial attacks, Rea-time detection, Robustness, Scalability.

algorithms. The models considered in this study include traditional classifiers such as logistic regression and decision trees, as well as more advanced techniques such as neural networks and ensemble methods. Features derived from website content, URL structure, and user behaviour are used to train and evaluate these models. The evaluation metrics used for comparison encompass accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

Additionally, the study explores the robustness of the models against adversarial attacks and their scalability in real-time detection scenarios. The findings of this research provide valuable insights into the strengths and weaknesses of various machine learning approaches for phishing detection. Such insights can guide the selection of appropriate models for specific use cases and contribute to the ongoing efforts to enhance online security. Ultimately, the goal is to improve the accuracy and efficiency of phishing detection systems to mitigate the risks posed by these malicious activities in an ever-evolving digital landscape.

## 1. INTRODUCTION

In today`s virtual landscape, the proliferation of threats has turn out to be a distinguished concern, with phishing rising as a giant threat to safety and privacy. These malicious assaults gift bold dangers to individuals and organizations usually evolving in complexity and sophistication. This creation gives a top-level view of a complete look at aimed toward comparing diverse device gaining knowledge of fashions for phishing detection. Features derived from internet site content, URL structure, and consumer conduct shape the muse for education and assessing those fashions. Evaluation metrics, together with accuracy, precision, recall, F1-score, and the place below the receiver working feature curve (AUC-ROC), are hired for comparison.

## 2. LITERATURE REVIEW

- Dr. R. Dhanalakshmi et. el [1] had introduced various machine learning techniques used for phishing detection, including decision trees, neural networks, and ensemble techniques. Discuss the suitability of these techniques for detecting phishing attempts.
- I. Kaur et. el [2] provides a detailed description of the datasets used in the study, including sources and characteristics. We provide an overview of their suitability for this particular task. This document focuses on phishing attacks, their prevalence, and the risks they pose to individuals and organizations.
- K. Abdul Haleem et. el [3] provided an overview of phishing attacks and the complexities involved in classifying and distinguishing between legitimate and phishing entities. Explain how specific machine learning models or techniques have been developed to classify phishing attacks and how these models aim to address the challenge of accurately identifying phishing attempts.
- S. AKhila et. el [4] gave a comprehensive review and classification of various machine learning approaches and algorithms for identifying phishing websites. The paper basically evaluates and compare the effectiveness, strengths, and weaknesses of different machine learning techniques applied in different studies.
- Dhiman Sarma et. el [5] gave the effectiveness of gadget learning (ML) in comparison to standard strategies in detecting phishing threats. Traditional strategies, encompassing rule-primarily based totally structures and blacklists, have barriers in adapting to evolving phishing tactics.
- Moham0med Hazim Alkawaz et. el [6] states the function of gadget learning (ML) in fostering phishing schooling and recognition. As cyber threats, especially phishing attacks, maintain to escalate, ML gives modern answers to decorate customers` information and vigilance.
- Andei Paleyes et. el [7] investigates the prison demanding situations related to deploying gadget learning (ML) for phishing prevention. As ML technology end up vital in cybersecurity, worries associated with privacy, facts protection, and compliance with current legal guidelines emerge.
- Ashit Kumar Dutt [8] investigates delves into the moral concerns surrounding the utility of system learning (ML) for phishing detection. As ML technology play an increasing number of crucial functions in cybersecurity, knowledge and addressing moral implications are crucial.
- Anuraag Velamati [9] explores the integration of Natural Language Processing (NLP) and machine learning (ML) for the classification of phishing emails,

presenting a comparative study of their effectiveness. Phishing attacks often leverage sophisticated linguistic tactics to deceive users, prompting researchers to examine the synergy between NLP and ML in email security.

- Panagiotis Bountaka et. el [10] states that comparative examine of gadget studying fashions withinside the evaluation of phishing campaigns.
- Dhiman Sarma et. el [11] states that evaluation specializes in exploring the resilience of system learning (ML) fashions withinside the context of phishing assaults. Phishing threats constantly evolve, annoying adaptive and resilient protection mechanisms.

**TABLE 1: PAPERS WITH PROBLEMS FACED.**

| **S. no** | ***Paper*** | ***Author*** | **Method used** | **Problem** |
|---|---|---|---|---|
| 1 | A Machine Learning Approach to Phishing Detection and Defence | Dr. R. Dhanalakshmi Dr. K. Kavitha | Logistic Regression<br>• Decision Trees<br>• Random Forests<br>• Support Vector Machines (SVM)<br>• Neural Networks<br>• Natural Language Process (NLP) | Imbalanced dataset<br>• Real-time detection<br>• Feature engineering<br>Detection of sophisticated attack |
| 2 | Phishing Websites Detection Using Machine Learning Techniques | I. Kaur, A. Rani, and A. Pannu, | • Feature Extraction<br>Logistic Regression<br>• Decision Trees<br>• Random Forests<br>• Support Vector Machines (SVM)<br>• Neural Networks | Imbalanced Dataset<br>Generalization to new threats<br>• Feature engineering<br>• Real time Detection<br>Adversarial attacks |

| | | | | |
|---|---|---|---|---|
| 3 | Machine Learning Techniques for Phishing Detection and Classification | K. Abdul Haleem and T. Kavitha, | • F1 score<br>• Feature Extraction<br>Logistic Regression<br>• Decision Trees<br>• Random Forests<br>• Support Vector Machines (SVM) | Imbalanced Dataset<br>Generalization to new threats<br>• Feature engineering<br>• Real time Detection<br>Adversarial attacks |
| 4 | A Review on Phishing Website Detection using Machine Learning Techniques | S. Akhila and Dr. A. Vadivel | Logistic Regression<br>• Decision Trees<br>• Random Forests<br>• Support Vector Machines (SVM) | Vulnerability in machine learning algorithms<br>Datasets are less than what is needed |
| 5 | Comparative Analysis of Machine Learning and Traditional Methods in Phishing Technology | Dhiman Sarma, Tanni Mittra, Rose Mary Bawm and Sohrab Hossain | • Blacklist<br>• Heuristic<br>• Domain Key identified mail<br>Comparative Analysis | • Ethical Consideration<br>• Feature Selection<br>Imbalanced datasets |
| 6 | Machine Learning for Phishing Education and | Mohammed Hazim Alkawaz, Stephanie Joanne Steven, | • Feature extraction<br>• Deep Learning<br>• Natural Language Process (NLP) | Data quality and availability<br>• Ethical Consideration |

| | | | | |
|---|---|---|---|---|
| | Awareness: An Analysis | Omar Farook Mohammad | | |
| 7 | Legal Challenges in Deploying Machine Learning for Phishing Prevention | Andei Paleyes, Raoul-Gabrief Urna and Neil D. Lawrence | • Natural Language Process (NLP)<br>• data protection regulations (e.g., GDPR), privacy laws, and anti-spam regulations. | Transparency<br>Privacy and protection<br>• Legal liability |
| 8 | Ethical Considerations in Machine Learning for Phishing Detection | Ashit Kumar Dutta. | Feature Engineering<br>• Natural Language Process (NLP)<br>Ensemble Methods<br>Ethical Frameworks | • Bias and fairness<br>• Privacy concern<br>Security risk |
| 9 | Comparative Study of Machine Learning Models for Analysing Phishing Campaigns | Anuraag Velamati | • Text processing<br>• Ensemble method<br>• Natural Language Process (NLP) | Real world deployments<br>• Feature selection |
| 10 | NLP and Machine Learning for Phishing Email Classificati | Panagiotis Bountakas, Konstantinos Koutroumpouchos and | • Text processing<br>• Ensemble method<br>• Evaluated matrix using f1 score, accuracy, recall | • Feature extracting<br>Imbalanced data<br>• Semantic dataset |

| | | | | |
|---|---|---|---|---|
| | on: A Comparative Study | Christos Xenakis | | • Handling unstructured datasets |
| 11 | Phishing Resilience: A Comparative Study of Machine Learning Models | Dhiman Sarma, Tanni Mittra, Rose Mary Bawm and Sohrab Hossain | Feature Engineering in emails<br>• Deep Learning | Adapt new algorithms<br>• No generalized datasets |

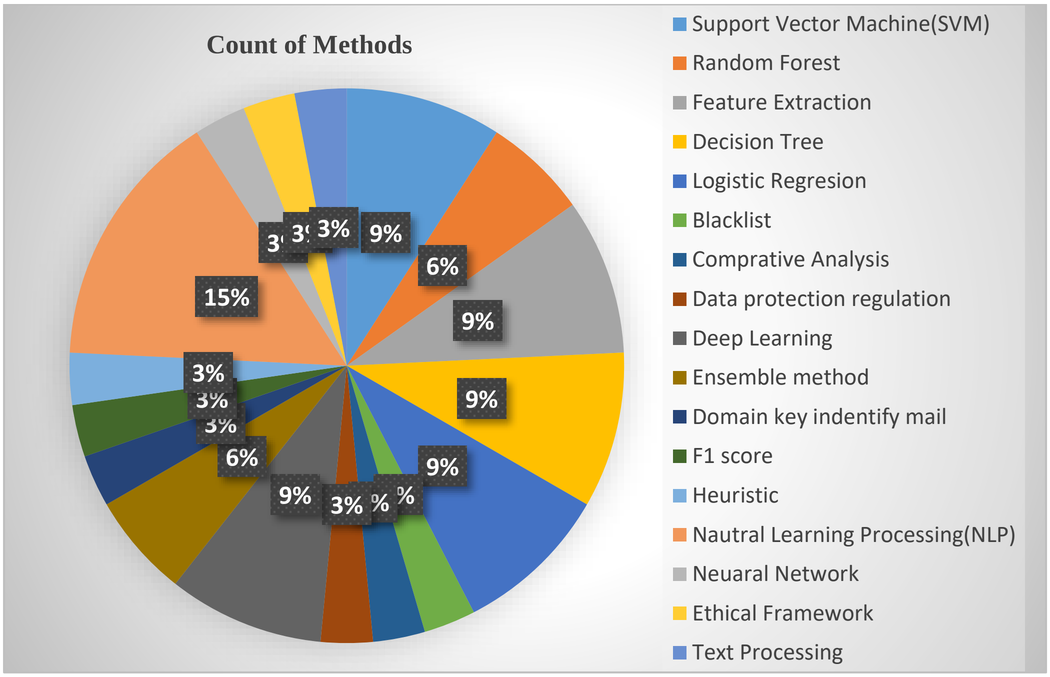

FIGURE1: GRAPH SHOWING METHODS WEIGHTAGE.

## 3. INTRODUCTION TO MACHINE LEARNING

In the field of artificial intelligence, machine learning focuses on creating algorithms that let computers analyse, interpret, and forecast data in order to make judgments or predictions.[1] These algorithms search through huge datasets for patterns and relationships using statistical methods rather than expressly programming them for a given job. Machine learning models are useful tools for jobs like image identification, natural language processing, recommendation systems, and autonomous cars because they can continuously learn from data through iterative learning.[2] This has resulted in a transformation of how we engage with technology in our daily lives and driven innovation.[4] There are mainly 3 type of machine learning namely supervised learning, unsupervised and reinforcement learning.

- Supervised Learning: Supervised mastering entails schooling a version on a categorised dataset, wherein every enter statistics factor is related to a corresponding goal label. The version learns the mapping among inputs and outputs primarily based totally at the furnished examples. In the context of phishing detection, supervised mastering algorithms may be educated on a dataset of recognised phishing URLs, wherein every URL is categorised as both valid or phishing.
- Unsupervised Learning: Unsupervised mastering entails schooling a version on an unlabelled dataset, wherein the set of rules attempts to discover hidden styles or systems withinside the statistics without specific guidance.
- Reinforcement Learning: Reinforcement mastering entails schooling a version to make sequences of selections in surroundings to maximise a few perceptions of cumulative reward. The version learns via trial and error, receiving remarks from the surroundings withinside the shape of rewards or penalties. In the context of phishing detection, reinforcement mastering may be implemented in adaptive protection structures wherein the version constantly learns to conform its detection techniques primarily based totally at the remarks it gets from customers or different additives of the system.

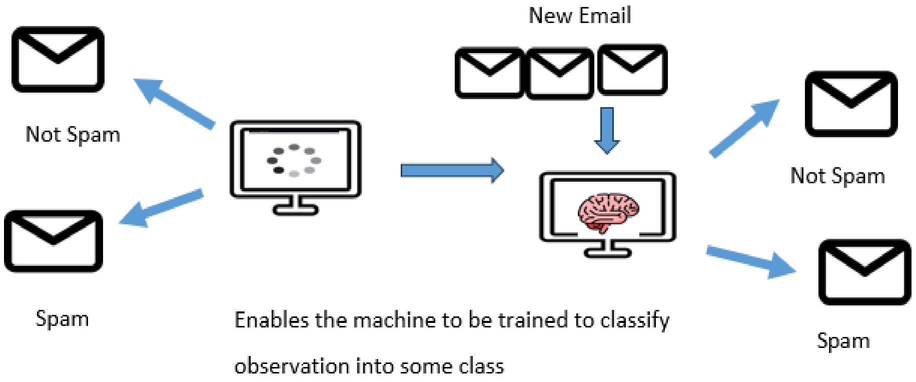
New Email
Not Spam
Spam
Not Spam
Spam
Enables the machine to be trained to classify
observation into some class

**FIGURE 2 MACHINE LEARNING HELPS TO IDENTIFY SPAM MAILS**

## 4. CONSEQUENCE

The consequences of falling victim to such an attack can range from financial loss to serious compromise of sensitive data, with far-reaching implications for both. Essentially, phishing attacks exploit people's trust and gullibility, often masquerading as legitimate and trustworthy sources [6].

- Detection Improvement: The contrast among device getting to know algorithms and phishing strategies fosters the improvement of greater state-of-the-art detection methods. By reading the styles and traits of phishing attempts, device getting to know fashions may be delicate to higher become aware of and thwart phishing emails, websites, and different malicious sports.
- Adversarial Innovation: The evolution of device getting to know strategies for detecting and stopping phishing assaults might also additionally set off malicious actors to innovate their techniques to bypass those defence.
- Enhanced Protection: The contrast of numerous devices getting to know methods to phishing detection can result in the implementation of greater sturdy safety measures.

User Awareness and Education: The contrast among device getting to know and phishing underscores the significance of person recognition and schooling in thwarting a hit assay

## 5. ROLE OF MACHINE LEARNING

- Pattern Recognition: Machine gaining knowledge of algorithms can examine styles and traits of phishing emails, websites, or different malicious sports to perceive not unusual place capabilities indicative of phishing tries.[7]
- Feature Extraction: Machine gaining knowledge of fashions can mechanically extract applicable capabilities from phishing-associated information, along with electronic mail headers, content, hyperlinks, and sender information.[8]
- Classification: Once applicable capabilities are extracted, gadget gaining knowledge of fashions classify incoming emails or messages as both valid or phishing tries. [1]
- Anomaly Detection: Machine gaining knowledge of also can locate anomalies in person conduct or community site visitors that can imply phishing [6]
- Ensemble Methods: Ensemble methods, which integrate predictions from a couple of gadgest gaining knowledge of fashions, are frequently used to enhance the robustness and accuracy of phishing detection systems. [10]

- Continuous Learning: Machine gaining knowledge of fashions may be skilled constantly on new information to conform to evolving phishing strategies and rising threats.[11]
- Feedback Loops: Feedback loops permit gadget gaining knowledge of fashions to study from their errors and enhance over time. [4]
- Integration with Security Systems: Machine gaining knowledge of-primarily based totally phishing detection answers may be included with present safety systems, along with electronic mail gateways, firewalls, or endpoint safety platforms [9]

User Awareness and Education: Machine gaining knowledge of also can help person recognition and schooling tasks with the aid of using reading person conduct and supplying personalised schooling or steering on figuring out and responding to phishing tries effectively.[3]

## 6. MODELS USED IN PHISHING DETECTION

### 6.1 LOGISTIC REGRESSION

Logistic regression is a statistical technique used for binary class responsibilities, in which the aim is to expect the opportunity that a statement belongs to one in all training. Despite its name, logistic regression is a class set of rules in preference to a regression set of rules.[1]

In the context of phishing detection, logistic regression may be used to categorise emails or web sites as both valid or phishing tries primarily based totally on numerous functions extracted from the data. Logistic regression works with inside the context of phishing detection:[2]

- Feature Extraction: Before making use of logistic regression, applicable functions want to be extracted from the data. These functions can consist of traits of the e-mail header (e.g., sender address, situation line), content (e.g., textual content evaluation for phishing-associated key phrases or patterns), and URLs (e.g., area reputation, presence of suspicious key phrases). [3]
- Model Training: Once the functions are extracted, the logistic regression version is skilled the use of a categorised dataset. In this dataset, every example (e.g., e-mail or internet site) is categorised as both valid (terrible elegance) or phishing (high-quality elegance).[4]
- Probability Estimation: Logistic regression fashions output chances in preference to discrete elegance labels. The output of logistic regression is a

opportunity rating among 0 and 1, which represents the chance of an example belonging to the high-quality elegance (phishing).[5]

- Decision Boundary: Logistic regression separates the function area into areas similar to the 2 training the use of a selection boundary. This selection boundary is a hyperplane that maximizes the chance of efficiently classifying the schooling data. [6]
- Deployment: Once the logistic regression version is skilled and evaluated, it is able to be deployed in a real-global phishing detection system. Incoming emails or internet site requests may be processed with the aid of using the version to estimate the opportunity of phishing, and suitable movements may be taken primarily based totally at the expected chances (e.g., blocking off suspicious emails or alerting users).[7]

### 6.2 SUPPORT VECTOR MACHINES (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm used for classification or regression tasks. [1] SVM is a powerful and versatile supervised learning algorithm used for both classification and regression tasks. The main goal of a classification problem is to find a hyperplane that optimally partitions a data set into classes. It is especially effective for binary classification [2]. As part of phishing detection, SVM is used to classify websites or digital content as phishing or legitimate. This works by finding a hyperplane that separates two classes. By mapping input data into a high-dimensional feature space, SVM identifies optimal boundaries between classes and allows new data points to be classified based on their position relative to those boundaries.[7]

We use SVM for phishing technology.

- High Accuracy: SVMs are recognised gain excessive accuracy in category responsibilities. In the case of phishing technology, correctly distinguishing among valid and malicious websites, emails, or different kinds of communique is vital for stopping customers from falling sufferer to phishing attacks.[8]
- Robustness to Overfitting: SVMs are much less vulnerable to overfitting in comparison to different device gaining knowledge of algorithms, making them appropriate for phishing detection responsibilities wherein the education information can be constrained or noisy. [9]
- Effective Handling of High-Dimensional Data: Phishing detection regularly entails studying information with a big wide variety of functions, including URLs, content, sender information, etc. [10]

- Flexibility with Kernel Functions: SVMs can make to deal with non-linear relationships among functions. This flexibility permits them to seize complicated choice boundaries, making them powerful in detecting state-of-the-art phishing tries which could contain diffused versions in functions.[11]

### 6.3 DECISION TREE

A decision tree is a hierarchical model that makes decisions based on various characteristics. A decision tree is a tree-like structure used for both classification and regression. Create a model that partitions a dataset into subsets based on the values of various features. This process continues recursively, eventually forming a tree structure where each internal node represents a feature or attribute, each branch represents a decision rule, and each leaf node represents an outcome or classification. [10]

We use decision tree for phishing technology.

- URL structure: Analysing the structure of URLs to come across suspicious styles inclusive of misspellings or uncommon characters. [9]
- Domain age and recognition: Checking the age and recognition of the area web website hosting the internet site. [8]
- Presence of HTTPS: Determining whether or not the internet site makes use of steady HTTPS encryption. [6]
- Presence of login forms: Detecting whether or not the internet site activates customers to go into touchy information. [5]
- Source of e mail: Analysing the sender`s e-mail deal with and evaluating it to recognized valid sources. [7]
- Content analysis: Examining the content material of emails or web sites for suspicious language or requests [4]

### 6.4 NEURAL NETWORKS

Neural networks are a set of algorithms inspired by the structure and function of the human brain [1]. They are made up of layers of interconnected nodes (neurons) that can learn to recognize patterns, understand relationships in data, and make predictions [3]. Neural networks are widely used in phishing detection because they can model complex, nonlinear relationships within data sets [5]. It can process and learn from large amounts of information to recognize patterns that may indicate a phishing attempt, including identifying anomalies in website content, URL structure, and user behaviour. [6]

#### 6.4.1 STRENGTHS

- Nonlinear relationships: Neural networks can handle nonlinear relationships between features, allowing for effective modelling of complex patterns in data.[2]
- Performance on large datasets: It performs well on large and diverse datasets, making it suitable for tasks where the data is large and diverse. [5]
- Data requirements: Neural networks typically require large amounts of data to train effectively. Insufficient data can lead to overfitting and under generalization. [6]
- Computing power: Training neural networks can be computationally intensive, especially for deep networks and large data sets. [7]
- Interpretability: For complex architectures, it can be difficult to understand how the network reached a particular conclusion, which can affect its interpretability. [9]

### 6.4.2 EXPLANATION OF KEYWORDS USED

- **Phishing attacks**: Phishing attacks are fraudulent attempts to obtain sensitive information by impersonating a trusted organization [1].

- **Cyber criminals:** Cyber criminals use various techniques such as fraudulent emails or websites to trick individuals into sharing their personal information, financial information, and login credentials. [2]

- **Online security:** This includes a variety of practices and technologies that protect against unauthorized access, data breaches, malware, and other cyber-attacks [3]. It includes the use of firewalls, encryption, multifactor authentication, and continuous security updates to reduce potential risks. [4]

- **Machine learning models:** These are algorithms that allow computers to learn from data and make decisions and predictions without being explicitly programmed [5]. For phishing detection, machine learning models use historical and real-time data to identify patterns and characteristics of phishing attacks.[6]

- **Phishing Detection:** The process of identifying and preventing phishing attacks using a variety of techniques and technologies. Phishing detection relies on machine learning to analyse various data sources such as website content, URL structure, and user behaviour to distinguish between legitimate and phishing incidents. [8]

- **Traditional classifiers:** These classifiers are often simpler and easier to interpret, but may not be complex enough to handle the nuances of phishing patterns. [10]
- **Advanced techniques:** Advanced machine learning models such as neural networks and ensemble techniques used for phishing detection. They provide greater accuracy and the ability to analyse complex relationships in your data, contributing to improved performance. [11]
- **Neural networks:** A class of machine learning models inspired by the structure and function of the human brain. [5]
- **Ensemble method:** A method that combines multiple machine learning models to improve accuracy and overall performance. [8]

## 7. CONCLUSION

Phishing attacks represent a formidable threat to individuals, businesses, and the overall landscape of online security. These fraudulent tactics, designed to deceive and exploit trust, can lead to a spectrum of severe consequences, including substantial financial losses, data breaches, compromised credentials, and threats to online security. As phishing attacks evolve and grow more sophisticated, their impact continues to pose a considerable risk to the digital community. The consequences of falling victim to phishing attacks range from profound economic loss to the serious compromise of sensitive data, each carrying far-reaching implications for both individuals and organizations. Additionally, the increasing complexity of detecting phishing attacks makes it significantly more difficult to detect those using traditional means. Using sophisticated social engineering techniques, email and domain spoofing, sophisticated phishing websites, multiple attack vectors, and personalized tactics, these attacks go beyond traditional security measures and are difficult to prevent and identify. In summary, deploying machine learning is critical to mitigating the risks posed by advanced and evolving phishing attacks. The ability to detect and analyse subtle changes in patterns, continuously learn from new data, and make decisions in near real-time is critical to strengthening online security and protecting against the multifaceted nature of phishing threats.

## 8. REFERENCES

- "A Machine Learning Approach to Phishing Detection and Defence" by Dr. R. Dhanalakshmi, Dr. K. Kavitha, published in the International Journal of Advanced Research in Computer and *Communication Engineering.
- "Phishing Websites Detection Using Machine Learning Techniques" by I. Kaur, A. Rani, and A. Pannu, published in the International Journal of Advanced Research in Computer Science.
- "Machine Learning Techniques for Phishing Detection and Classification" by K. Abdul Haleem and T. Kavitha, published in the International Journal of Scientific & Technology Research.
- "A Review on Phishing Website Detection using Machine Learning Techniques" by S. Akhila and Dr. A. Vadivel, published in the International Journal of Computer Science and Mobile Computing.
- Comparative Analysis of Machine Learning and Traditional Methods in Phishing Technology by Dhiman Sarma, Tanni Mittra, Rose Mary Bawm and Sohrab Hossain published in International Conference of Inventive Computation and Information Technologies, Scopus.
- Machine Learning for Phishing Education and Awareness: An Analysis by Mohammed Hazim Alkawaz, Stephanie Joanne Steven, Omar Farook Mohammad published in IEEE.
- Legal Challenges in Deploying Machine Learning for Phishing Prevention by Andei Paleyes, Raoul-Gabrief Urna and Neil D. Lawrence published in ACM Journal.
- Ethical Considerations in Machine Learning for Phishing Detection by Ashit Kumar Dutta.

- Comparative Study of Machine Learning Models for Analysing Phishing Campaigns by Anuraag Velamati in International Journal of Engineering Applied Sciences and Technology.
- NLP and Machine Learning for Phishing Email Classification: A Comparative Study by Panagiotis Bountakas, Konstantinos Koutroumpouchos and Christos Xenakis published in ACM Journal.