IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

Exploit URL Detection Using Supervised Learning Algorithms

Mrs.G.Rekha[1] M Thejaswini[2] M swamy[3] M Hemanth[4] V Aswini[5]

Associate Professor, Department of Computer Science and Engineering, Kuppam Engineering College, Kuppam, Andhra Pradesh, India[1]

UG Student, Department of Computer Science and Engineering, Kuppam Engineering College, Kuppam, Andhra Pradesh, India [2345]

ABSTRACT: Phishing remains a significant concern in our evolving digital world, where cybercrime thrives on unauthorized access to private information through computer systems. Phishing via deceptive URLs is a common tactic, aiming to steal user data when individuals visit malicious websites. Detecting such URLs is challenging, despite existing blacklisting methods used by the web security community. These methods, often relying on manual reporting and heuristic analysis, can miss newly emerging or incorrectly evaluated malicious sites. To address this, we propose employing machine learning algorithms like Decision Trees, Random Forests, Multi-Layer Perceptron (MLP), XGBoost (XGB), and Support Vector Machines (SVM) to identify malicious URLs. This involves extracting features from URLs and applying the models to determine their maliciousness.

I. INTRODUCTION

In today's digital age, machine learning stands as a transformative force, shaping the landscape of technology, business, and society at large. At its core, machine learning is a subset of artificial intelligence (AI) that enables computers to learn from data, identify patterns, and make decisions or predictions without explicit programming. This capability empowers machines to perform tasks and solve problems that were once the exclusive domain of human intelligence.
The power of machine learning lies in its ability to extract insights and knowledge from vast amounts of data, enabling computers to generalize from past experiences and adapt to new situations. Traditional programming paradigms involve explicitly instructing a computer on how to perform a task through a set of rules or algorithms. In contrast, machine learning algorithms learn from examples and iteratively improve their performance over time through experience.

One of the defining characteristics of machine learning is its versatility and applicability across a wide range of domains and industries. From healthcare and finance to marketing and entertainment, machine learning finds applications in diverse fields, revolutionizing processes, and driving innovation. For example, healthcare, machine learning algorithms can analyse medical imaging data to detect diseases early or assist clinicians in making more accurate diagnoses. In finance, predictive analytics powered by machine learning can forecast market trends, optimize investment strategies, and detect fraudulent transactions.

Machine learning algorithms can be broadly categorized into three main types: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the algorithm is trained on a labelled dataset, where each example is associated with a corresponding target or outcome. The algorithm learns to map input data to output labels, enabling it to make predictions on new, unseen data. Common supervised learning tasks include classification, regression, and ranking.

Unsupervised learning, on the other hand, involves training algorithms on unlabelled data, where the objective is to uncover hidden patterns or structures within the data. Clustering, dimensionality reduction, and anomaly detection are typical unsupervised learning tasks.

Reinforcement learning, inspired by psychology, involves training agents to interact with an environment and learn optimal decision-making strategies through trial and error. This approach has applications in robotics, game playing, and autonomous systems.

Machine learning represents a paradigm shift in how we approach problem-solving and decision-making in the digital

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

age. By harnessing the power of data and algorithms, machine learning enables us to unlock new insights, automate tasks, and drive innovation across various domains. As we continue to push the boundaries of AI, it is imperative to do so responsibly, ensuring that the benefits of machine learning are balanced with ethical considerations and societal welfare.

In our rapidly advancing digital age, machine learning (ML) has seamlessly integrated into various aspects of our daily lives, significantly impacting how we interact with technology and navigate the online landscape. At its core, machine learning is a subset of artificial intelligence (AI) that empowers computers to learn from data, recognize patterns, and make decisions autonomously without explicit programming.

Everyday scenarios such as personalized recommendations on streaming platforms, predictive text suggestions on smartphones, and even facial recognition features in social media apps are all manifestations of machine learning in action. These applications demonstrate the essence of machine learning: the ability to analyze vast amounts of data, extract meaningful insights, and deliver tailored experiences to users.

While machine learning finds numerous applications across diverse domains, one area where its impact is particularly pronounced is in cybersecurity, specifically in the detection and mitigation of phishing attacks. Phishing attacks represent a pervasive threat in the digital realm, wherein malicious actors impersonate legitimate entities to deceive users into divulging sensitive information like passwords, financial details, or personal data.

II. EXISTING METHOD

The study delves into the increasing threat of phishing attacks in cyberspace and proposes a novel approach utilizing machine learning algorithms for effective detection.

The authors explore various features extracted from URLs and employ different machine learning classifiers to discern between legitimate and phishing URLs. They conduct experiments to evaluate the performance of their proposed system, comparing it with existing methods, and demonstrate promising results in terms of accuracy, precision, recall, and F1-score.

The findings suggest that their approach can effectively contribute to the mitigation of phishing threats by accurately identifying malicious URLs, thus enhancing cybersecurity measures. Overall, the paper provides valuable insights and contributions to the field of cybersecurity by leveraging machine learning techniques for phishing URL detection.

For the machine learning classification phase, the authors experiment with several popular algorithms such as Decision Trees, Random Forest, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Logistic Regression. They train these classifiers on the training set and then evaluate their performance using metrics like accuracy, precision, recall, and F1-score on the testing set.

Additionally, the authors employ techniques such as cross-validation to ensure the robustness of their models and to mitigate issues like overfitting.

Disadvantages of Existing System

- High Complexity
- Low efficiency compared to other models
- Limited interpretability

III. PROPOSED SYSTEM

The proposed system for URL classification integrates machine learning algorithms and feature engineering techniques to differentiate legitimate and malicious URLs. It utilizes a diverse dataset comprising domain information, structural features, and content-based characteristics. Employing decision trees, random forests, MLP, XGBoost, and SVM, the system identifies optimal classification methods. With hyperparameter tuning and cross-validation, it enhances model performance. Additionally, it features a user-friendly interface for seamless deployment, bolstering cybersecurity and threat detection capabilities. Through rigorous testing, the system aims to achieve high accuracy in URL classification, fortifying internet security.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

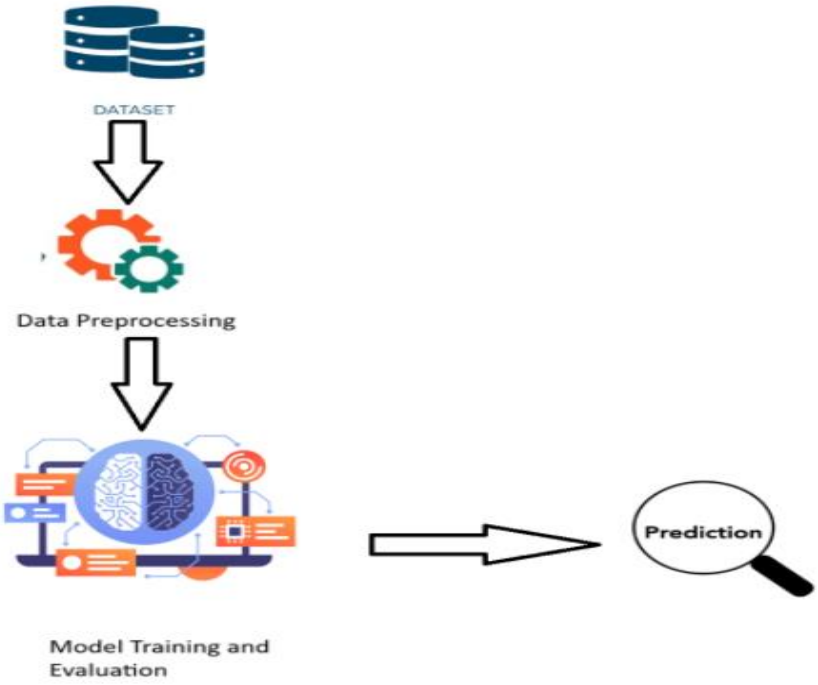
|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

Advantages of Proposed System

- Highest accuracy
- Reduces time complexity.
- Easy to use
- Automatic feature extraction.
- Real-time detection

SYSTEM ARCHITECTURE



IV. LITERATURE SURVEY

- Some of the relevant work reported by authors in reputed journals are discussed here (Shrivas & Suryawanshi, 2017) collected phishing data set from UCI repository and implemented the phishing data set on rapid miner tool and compared decision tree, random tree, random forest algorithms for method of phishing detection involved fixed black and white listing databases. But, these methods are not efficient because a duplicate website can be developed very fast.
- So most of these methods cannot make an accurate decision dynamically on whether the new website is phishing or legitimate. Hence, number of new phishing websites may be classified as legitimate website. In this situation, it is preferred to develop guidelines to extract specific features from websites and then use them to predict the type of web page.
- classification of phishing and non-phishing. Accuracy obtained from decision tree was 91.8% which was the best as compared to other algorithms random tree 66.7%, random forest 78.8% and decision stump 84%. *(Subasi et al., 2017)* compared various algorithms like random forest, support vector machine, decision tree on WEKA open source tool. This work reported that Random Forest is quicker, robust and more accurate as compared to KNN, SVM Rotation forest and Decision Tree. Random forest yielded best results here based on accuracy 97.36%.
- *(Hodžić* and Kevrić, *2016)* compared the algorithms multilayer perceptron (MLP), decision tree, random forest, C4.5, rotation tree (REP tree) etc. All the experiments were conducted in WEKA tool and this work reported that Rotation tree achieved overall best accuracy of 89.1% compared to other algorithms.
- In their paper (Kalaiselvan et. al) collected phishing dataset from the Phistank website and compared the algorithms C4.5, SVM, Naïve and ZeroR, to classify phishing dataset into phishing and legitimate. Here the accuracy of developed methods was assessed after applying the 10 fold cross-validation and Naïve Bayes

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

algorithm was found to perform better than other algorithm. In their research paper Support Vector Machine, Gaussian and NMC classifiers have been employed by (Kaur et al, 2015) along with fuzzy logic. Fuzzy based detection system provides effective aid in detecting phishing websites. It successfully resulted in low false positive and high true positive for classifying phishing websites. A methodology to detect phishing website based on machine learning classifiers is presented in (Ali et. al, 2017) which uses a wrapper features selection method. Some common supervised machine learning techniques have been used by authors here to accurately detect phishing websites and they found that wrapper based algorithm performs better as compared to normal method**.**

- (Mohammad et al, 2014) in his research proposed structuring neural networks) based intelligent model for predicting phishing attacks. The authors were able to atomise phishing website detection with frequent change in phishing websites using 17 different features**. (**Aburrous et al., 2010) purposed an Intelligent system to detect phishing in e-banking where they combined fuzzy logic model with machine learning algorithms to detect phishing websites. They differentiated between different types of phishing websites using 10 fold cross validation and achieved 86.38% accuracy, which is very low.
- (Chen et al, 2010) purposed method for concocted spoof detection using different algorithms as Bayesian Network, C4.5, Logit Regression, Naïve Bayes, Neural Network and SVM (linear composite, linear, polynomial, RBF kernels). The authors achieved an accuracy of 92.56% among 900 legit concocted, and spoof e-commerce websites.

V. METHODOLOGY

1. Dataset

The dataset comprises of 10,000 URL tests, each portrayed by 17 highlights pointed toward recognizing genuine and malicious URLs for network safety purposes. These features incorporate markers, for example, the presence of IP addresses, "@" images, URL length, redirection, HTTPS use, and that's only the tip of the iceberg. The target of this dataset is to work with the improvement of AI models able to do precisely characterizing URLs into genuine or pernicious classifications. By utilizing these models, network protection experts can improve their capacity to recognize and battle potential dangers presented by malevolent URLs, in this way reinforcing the security of online clients and frameworks. Preprocessing steps, including dealing with missing qualities and encoding all out factors, are fundamental prior to preparing machine learning models.

The Data preprocessing phase by loading the dataset and conducting exploratory data analysis (EDA) to understand its structure and characteristics. This includes checking the dataset's shape, listing its features, and examining its information to identify any missing values or inconsistencies. Furthermore, the code shuffles the dataset's rows to ensure an even distribution when splitting it into training and testing sets. Additionally, the code removes the "Domain" column, which likely serves as an identifier and doesn't contribute to the classification task. Finally, preprocessing involves splitting the dataset into features (X) and the target variable (y), which are then further divided into training and testing subsets using a conventional 80-20 split. These preprocessing steps lay the foundation for training and evaluating machine learning models effectively, enabling accurate classification of URLs into legitimate and malicious categories for cybersecurity purposes.

a Data preprocessing, including merging the data and a major challenge when attempting to add a dataset to the machine learning model is null values. Because of this, all null values are removed before adding the dataset to the machine learning model.

b Finally, we apply the machine learning model to all the features generated by the feature extraction module with the help of machine learning algorithms such as Decision Tree, Random Forest Classifier, Multilayer Perceptron, XGB, SVM.

2.Feature Extraction

As part of this step, we extract features from the URL dataset. The extracted features are classified into Address Bar based Features and Domain based Features and a total of 18 features are taken into consideration.

- Domain: The domain name of the URL, indicating the website's identity.
- Have_IP: Binary indicator (0 or 1) denoting whether the URL contains an IP address instead of a domain name.
- Have_At: Binary indicator (0 or 1) indicating the presence of the "@" symbol in the URL, often used in email addresses.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

- URL_Length: Categorical variable representing the length of the URL, which can provide insights into its complexity and potential for obfuscation.
- URL_Depth: Categorical variable indicating the depth of the URL path, revealing how nested the URL structure is.
- Redirection: Binary indicator (0 or 1) signifying whether the URL involves redirection, which can be indicative of phishing attempts.
- https_Domain: Binary indicator (0 or 1) indicating if the URL uses HTTPS protocol for secure communication.
- TinyURL: Binary indicator (0 or 1) denoting the usage of TinyURL for URL shortening, which can obscure the actual destination.
- Prefix/Suffix: Binary indicator (0 or 1) indicating the presence of prefix or suffix in the URL, which might be indicative of phishing attempts.
- DNS_Record: Binary indicator (0 or 1) representing the existence of DNS records for the domain, which verifies its legitimacy.
- Web_Traffic: Binary indicator (0 or 1) denoting the presence of web traffic to the URL, indicating its popularity or activity level.
- Domain_Age: Binary indicator (0 or 1) representing the age of the domain, with older domains often considered more trustworthy.
- Domain_End: Binary indicator (0 or 1) indicating the end of the domain, which might influence its legitimacy.
- iFrame: Binary indicator (0 or 1) signifying the presence of iframe in the URL, which can be used for embedding external content.
- Mouse_Over: Binary indicator (0 or 1) indicating the presence of mouse-over events in the URL, potentially revealing hidden links or actions.
- Right_Click: Binary indicator (0 or 1) denoting the ability to right-click on the URL, which might be disabled in phishing attempts.
- Web_Forwards: Binary indicator (0 or 1) representing web forwards in the URL, indicating potential redirection or forwarding mechanisms.
- Label: Target variable denoting the classification of the URL as legitimate (0) or malicious (1), which is the main focus of the classification task.

Machine Learning Algorithms

Decision Tree algorithm:

The Decision Tree classifier is a flexible and interpretable calculation for both characterization and relapse errands. It segments the component space in light of the upsides of info highlights, making a tree-like design where each inside hub addresses an element, each branch addresses a choice in view of that component, and each leaf hub addresses a class name. In this code, a Decision Tree classifier with a most extreme profundity of 5 is started up and prepared on the preparation dataset. This profundity impediment forestalls overfitting by limiting the intricacy of the tree. The model is assessed on both the preparation and testing datasets to evaluate its precision in foreseeing whether a URL is real or pernicious.

$$LH = \sum fi(1 - fi)$$

Random Forest algorithm:

The Random Forest classifier is a troupe learning technique in light of choice trees. It develops numerous choice trees during preparing and yields the method of the classes anticipated by individual trees. This gathering approach mitigates overfitting and further develop speculation execution. In this code, an Irregular Woodland classifier with a most extreme profundity of 5 is launched and prepared on the preparation dataset. Like the Decision Tree, the Random Forest model's precision is assessed on both the preparation and testing datasets to evaluate its adequacy in ordering URLs.

$$RFfii = \frac{\sum j \in alltreesnormfi_{ij}}{T}$$

$$fii = \sum sjCj$$

j:nodes j splits on feature i

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

Multilayer Perceptron:

The Multilayer Perceptron classifier is a kind of fake brain network described by different layers of interconnected hubs (neurons). It can learn complex examples and connections in information by changing the loads and predispositions of associations between neurons during preparing. In this code, a MLP classifier with three secret layers, each containing 100 neurons, is launched and prepared on the preparation dataset. The model's exactness is then assessed on both the preparation and testing datasets to quantify its exhibition in URL characterization.

Extrem Gradient Boosting(XGB):
The XGBoost (Extreme Gradient Boosting) classifier is a strong slope supporting calculation known for its speed, versatility, and exactness. It successively assembles a group of powerless students (decision trees) and consolidates their expectations to work on generally execution. In this code, a XGBoost classifier with a learning pace of 0.4 and most extreme profundity of 7 is launched and prepared on the preparation dataset. Like different models, its exactness is assessed on both the preparation and testing datasets to evaluate its viability in arranging URLs.

Support Vector Machine:
The Support Vector Machine classifier is a broadly involved managed learning calculation for characterization undertakings. It finds the hyperplane that best isolates the classes in the component space, augmenting the edge between the classes. In this code, a SVM classifier with a direct bit and regularization boundary C=1.0 is launched and prepared on the preparation dataset. Its exactness is then assessed on both the preparation and testing datasets to decide its presentation in recognizing genuine and malevolent URLs. SVMs are especially powerful in high-layered spaces and can deal with both direct and non-straight connections between highlights.

VI. SIMULATED RESULT

The results section encapsulates the performance of various machine learning models employed for URL classification. Each model's accuracy on both training and test datasets is outlined, providing a comprehensive overview of their predictive capabilities. The Decision Tree classifier achieved an accuracy of 81.6% on the training set and 80.4% on the test set, indicating its ability to generalize reasonably well to unseen data.

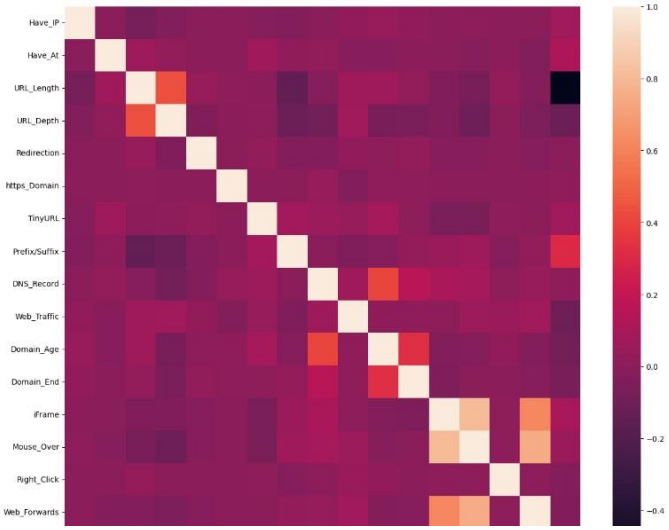
- Following closely, the Random Forest classifier exhibited slightly higher performance with accuracies of
- 82.1% on the training data and 81.2% on the test data. Notably, the Multilayer Perceptrons (MLP) model displayed the highest accuracy among all models, boasting 86.9% on the training data and 85.5% on the test data. Similarly, the XGBoost classifier showcased strong performance with accuracies of 87.0% and 85.1% on the training and test datasets, respectively.
- Conversely, the Support Vector Machine (SVM) classifier demonstrated slightly lower accuracy, achieving 80.5% on the training data and 79.0% on the test data. These detailed results offer valuable insights into the efficacy of each model, with the MLP and XGBoost models emerging as the top performers in accurately classifying URLs as legitimate or malicious.
- Among the models evaluated, the Multilayer Perceptrons (MLP) and XGBoost classifiers emerged as the top performers in accurately classifying URLs as legitimate or malicious. The MLP model exhibited the highest accuracy on both the training and test datasets, with 86.9% and 85.5%, respectively. Similarly, the XGBoost classifier demonstrated strong performance with accuracies of 87.0% on the training data and 85.1% on the test data. These models outperformed others due to their ability to capture complex patterns and relationships within the dataset, thanks to their inherent flexibility and ensemble learning techniques. The MLP model, with its multiple hidden layers and neurons, excels at learning intricate features from the data, while XGBoost leverages the collective wisdom of multiple decision trees to achieve superior classification accuracy. Additionally, both models are known for their robustness to overfitting and generalization capability, contributing to their exceptional performance in this task. These results underscore the effectiveness of MLP and XGBoost classifiers in addressing the challenges of URL classification for cybersecurity applications.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

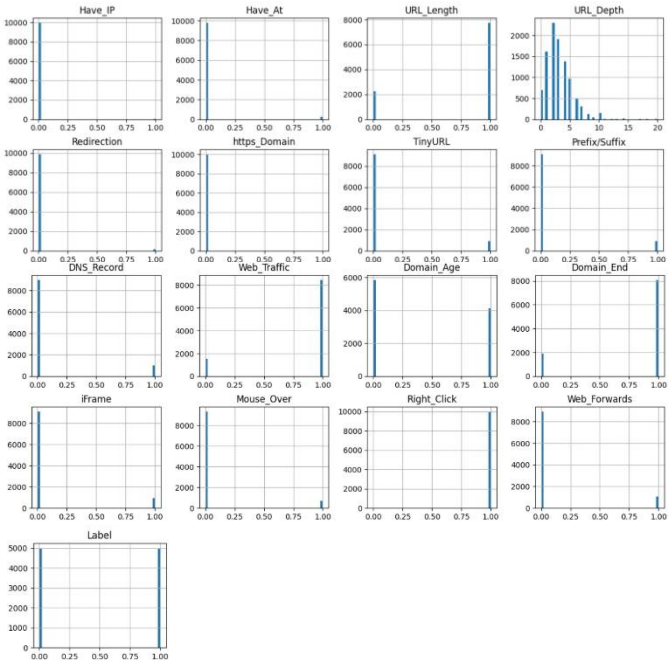
| DOI: 10.15680/IJIRCCE.2024.1204247 |



The heatmap of the dataset used is produced above.

The below are the visualizations of the features used in the dataset.

We present a simulated experimental assessment of the proposed method using an actual dataset:



VII. CONCLUSION

This work presents Security keyword-based data discovery using HT-HDFS for improved speed. This is an expansion of the homomorphic cryptography-based hash tree framework, or HT-HDFS, for processing picture data. The total time needed to complete the job is reduced by using the homomorphic cryptography approach. Fast keyword search without compromising the semantic security of the encrypted keywords is made feasible with Searchable Public-Key Ciphertext with Hidden Structures Algorithm.All keyword-searchable ciphertext in HT-HDFS is organised by hidden relations. A search algorithm is guided to find all matching ciphertexts efficiently by obtaining the minimum information of the relations through the search trapdoor (Keyword encrypted with Homomorphic-Cryptography) corresponding to a keyword. We create an HT-HDFS scheme from the ground up with a concealed star-like structure in the ciphertext.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204247 |

VIII. FUTURE WORK

The project on detecting phishing websites holds substantial future scope for advancement and expansion. Some potential avenues for further development and improvement include:

1. Perceptron Enhanced Feature Engineering Perceptron: Continuously refining and expanding the set of features used for detecting phishing websites can improve the accuracy and robustness of the machine learning models. This could involve incorporating additional contextual information, such as website content analysis, user interaction patterns, or real-time data sources, to enhance the detection capabilities.

2. Perceptron Advanced Machine Learning Techniques Perceptron: Exploring advanced machine learning techniques, such as deep learning architectures (e.g., convolutional neural networks or recurrent neural networks), ensemble methods, or semi-supervised learning approaches, may further improve the performance of the phishing detection system. These techniques could help capture more intricate patterns and relationships in the data, leading to better generalization and adaptability.

3. Perceptron Real-Time Detection and Response Perceptron: Integrating real-time detection capabilities into the system can enable proactive identification and mitigation of phishing threats as they emerge. Implementing mechanisms for continuous monitoring of website behavior, anomaly detection, and automated response actions can enhance the system's effectiveness in combating evolving phishing attacks.

REFERENCES

1. Aburrous, M., Hossain, M. A., Dahal, K., & Thabtah, F. (2010). Intelligent phishing detection system for e-banking using fuzzy data mining. *Expert systems with applications*, *37*(12), 7913-7921.
2. Abu-Nimeh, S., Nappa, D., Wang, X., & Nair, S. (2007, October). A comparison of machine learning techniques for phishing detection. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit* (pp. 60-69). ACM.
3. Ali, W. (2017). Phishing Website Detection based on Supervised Machine Learning with Wrapper Features Selection. *International Journal of Advanced Computer Science and Applications*, *8*(9), 72-78.
4. Anti-Phishing Working Group, "Phishing Activity Trends Report 1 Quarter," *Most*, no. March, pp. 1–12, 2010.
5. Barraclough, P. A., Hossain, M. A., Tahir, M. A., Sexton, G., & Aslam, N. (2013). Intelligent phishing detection and protection scheme for online transactions. *Expert Systems with Applications*, *40*(11), 4697-4706.
6. Chen, H., Vasardani, M., & Winter, S. (2017). Geo-referencing Place from Everyday Natural Language Descriptions. *arXiv preprint arXiv:1710.03346*.
7. Hodžić, A., and Kevrić, J., (2016) Comparison of Machine Learning Techniques in Phishing Website Classification. International Conference on Economic and Social Studies (ICESoS'16), 249–256.
8. Kahksha & Naaz, S. (2018) Machine Learning Algorithm to Predict Survivability in Breast Cancer Patients. *Int. J. Comput. Sci. Eng.*, 10(4), 97–101.
9. Kalaiselvan, O. & Edwinraja, S. (2015) Predicting Phishing Websites using Rule Based Techniques. *International Journal of Emerging Technology and Innovative Engineering*, I(4), 180–185.
10. Kaur, S., & Sharma, S., (2015). Performing Efficient Phishing Webpage Detection. *International Journal of Computer Sciences and Engineering*, 3(7), 52–56.
11. Mohammad, R. M., Thabtah, F., & McCluskey, L. (2014). Intelligent rule-based phishing websites classification. *IET Information Security*, *8*(3), 153-160.
12. Namdev, N., Agrawal, S., & Silkari, S. (2015). Recent advancement in machine learning based internet traffic classification. *Procedia Computer Science*, *60*, 784-791.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details