Project Report

on

DETECTION OF PHISHING WEBSITES USING MACHINE LEARNING

(A Project Report submitted in partial fulfillment of the requirements of Bachelor of Technology in Information Technology of the West Bengal University of Technology, West Bengal)

Submitted by

Arnab Kumar Das
Shubhankar Das
Sayan Banerjee
Sarthak Dey

Under the guidance of
Professor Malabika Sengupta

Dept. of Information Technology

Kalyani Government Engineering College

(Affiliated to West Bengal University of Technology)

Kalyani - 741235, Nadia, WB

2023-2024

Phone: 25826680 (PBX) **e-mail :**
Fax : 2582130

কল্যাণী - ৭৪১ ২৩৫
নদীয়া, পশ্চিমবঙ্গ

Kalyani 741 235
Nadia, West Bengal,India

পত্রাঙ্ক /Ref. No.:
তারিখ / Date :

কল্যাণী গভঃ ইঞ্জিনিয়ারিং কলেজ
Kalyani Government Engineering College
(Govt. of West Bengal)

Certificate of Approval

This is to certify that ……………………………….has done final year project work entitled………………………………………under my direct supervision and he/she has fulfilled all the requirements of relating to the Final Year Project. It is also certified that this project work being submitted, fulfills the norms of academic standard for B. Tech Degree in Information Technology of The West Bengal University of Technology and it has not been submitted for any degree whatsoever by him/her or anyone else previously.

…………………………………………
Head
Department of Information Technology
Kalyani Government Engineering College

……………………………………………..
Supervisor
Department of Information Technology
Kalyani Government Engineering College

…………………………………………
Project Coordinator
Department of Information Technology
Kalyani Government Engineering College

………………………………………………
Examiner

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Malabika Sengupta for her invaluable guidance and unwavering support throughout the course of this project. Her insightful feedback, constructive criticism, and dedication to fostering a spirit of inquiry have been pivotal in shaping and refining our research on "Detecting Phishing Websites using Machine Learning." Her mentorship has not only enriched my understanding of the subject matter but has also inspired a deeper passion for research and innovation.

I would also like to extend my appreciation to the authors of the research papers that served as significant references for our project. The contributions of these researchers have laid a solid foundation for our work, and their insights have been instrumental in shaping our methodology. The following research papers have been particularly influential:

- Görkem Giray, Bedir Tekinerdogan, Sandeep Kumar & Suyash Shukla, “Applications of deep learning for phishing detection: a systematic literature review” in Knowledge and Information Systems [23rd May 2022]

- Ashit Kumar Dutta, “Detecting phishing websites using machine learning technique” in Zhihan Lv, Qingdao University, China October 11, 2021.

Other valuable papers and articles have been enlisted in the References section of the report.

I am thankful for the support and encouragement received from Dr. Sengupta and the authors of the referenced research papers, as their work has played a crucial role in the development and success of our project.

[Date]

1. Arnab Kumar Das - 10200220008-
2. Shubhankar Das - 10200220006 -
3. Sayan Banerjee - 10200220041 -
4. Sarthak Dey - 10200220015 -

Department of Information Technology
Kalyani Government Engineering College

ABSTRACT

Phishing is a type of cyber-attack in which attackers use deceptive tactics to trick individuals into divulging sensitive information such as usernames, passwords, and financial details. Phishing websites pose a significant threat in today's digital landscape. Phishing websites often mimic legitimate websites, making it challenging for users to distinguish between the two. The main purpose of building a phishing website detector is to enhance cybersecurity by identifying and preventing phishing attacks.

Among many key parameters, few parameters as, English efficiency, source year, DNS filter, reviews are used in this work for developing a phishing website detector:

CONTENTS

CHAPTER 1 INTRODUCTION 1

1.1 **Motivation** 1

1.2 **Objective** 2

1.3 **Organization of the Project** 2

CHAPTER 2 BACKGROUND STUDY 3

2.1 **Problem Statement** 4

2.2 **Modes of Phishing** 4-5

2.3 **Workflow of Efficient Phishing Detection Model** 5-6

CHAPTER 3 PROPOSED DETECTION MODEL 7

3.1 **Implementation** 7-8

3.2 **Features** 9

3.3 **Tech Stacks** 9-11

3.4 **Summary of present work** 11-12

3.5 **Future Work** 12

CHAPTER 4 CONCLUSIONS 13

CHAPTER 5 REFERENCES 14

Chapter 1

INTRODUCTION

In the rapidly changing landscape of the internet, cybersecurity is increasingly challenged by sophisticated phishing attacks. Our project addresses this concern by employing machine learning to detect and classify phishing websites accurately, surpassing the limitations of static rule sets commonly used in traditional defense mechanisms.

The primary objective is to design a robust machine learning model capable of discerning patterns in the structure, content, and behavior of phishing sites. Unlike static rule sets, machine learning provides adaptability to the dynamic nature of modern phishing attacks. The model is trained on diverse datasets, allowing it to generalize learning and effectively identify phishing attempts in real-time.

Grounded in the understanding that phishing websites exhibit discernible patterns, our approach draws insights from contemporary research papers, positioning our project at the intersection of cutting-edge research and practical implementation in the field of phishing detection using machine learning. This report comprehensively outlines our research methodology, dataset selection, feature extraction techniques, and evaluation metrics, contributing to the cybersecurity knowledge base and offering a practical solution to the evolving threat of phishing attacks in the digital age.

1.1 **MOTIVATION**

In an era dominated by digital interactions, the surge in phishing attacks demands innovative solutions. Traditional defenses often lag behind, prompting our project's motivation: to leverage advanced machine learning to efficiently identify and categorize suspected phishing websites.

Our platform, centered around a community-sourced repository of potential threats, fosters collective vigilance. We are driven by a commitment to proactively contribute to a safer digital ecosystem, transcending mere detection. Identified phishing websites are shared with organizations like the Anti-Phishing Working Group (APWG), amplifying our impact on a global scale.

Inspired by research insights, we aim to bridge the gap between academic knowledge and practical implementation. Our goal is not just to build a database but to empower users, organizations, and cybersecurity professionals with the tools and intelligence needed to navigate the digital landscape securely. Through this project, we strive to make meaningful contributions to the resilience of the online world, ensuring a safer internet for all.

1.2 **OBJECTIVE**

The upcoming phase of our project focuses on algorithm implementation using Python, integrating the Random Forest and RNN models into our website through APIs. This step translates our research findings into a practical, real-time phishing detection system. Simultaneously, we'll establish a database to systematically store information on identified phishing websites. Future plans involve linking this database to organizations like the Anti-Phishing Working Group (APWG) for wider cybersecurity collaboration. User feedback mechanisms and continuous algorithm refinement will ensure the project remains adaptive and effective in countering evolving phishing threats. The proposed work aims to create a comprehensive and impactful tool that not only detects phishing websites but actively contributes to the global cybersecurity community.

1.3 **ORGANIZATION OF THE PROJECT**

The project's organization revolves around a thorough examination of the phishing landscape, commencing with a detailed exploration of the Problem Statement. This initial phase involves a comprehensive investigation into the various Modes of Phishing, dissecting the techniques employed in these malicious endeavors. Following this, the project delves into elucidating the Workflow of an Efficient Phishing Detection Model, mapping out the sequential steps integral to developing a robust system. Transitioning into the core of the initiative, the Proposed Detection Model encompasses facets such as Implementation, where the practical aspects of model development and deployment are scrutinized. Features of the model are elucidated, highlighting the distinctive functionalities designed to enhance its efficacy. Concurrently, the Tech Stacks employed are detailed, providing insights into the technological foundation of the model. A pivotal component of this phase involves a reflective Summary of Present Work, encapsulating the current state and accomplishments. Looking forward, Future Work outlines the projected trajectory, encompassing planned enhancements and developments intended to fortify the phishing detection model. Through this holistic approach, the project aims to address the multifaceted challenges posed by phishing activities.

Chapter 2

BACKGROUND STUDY

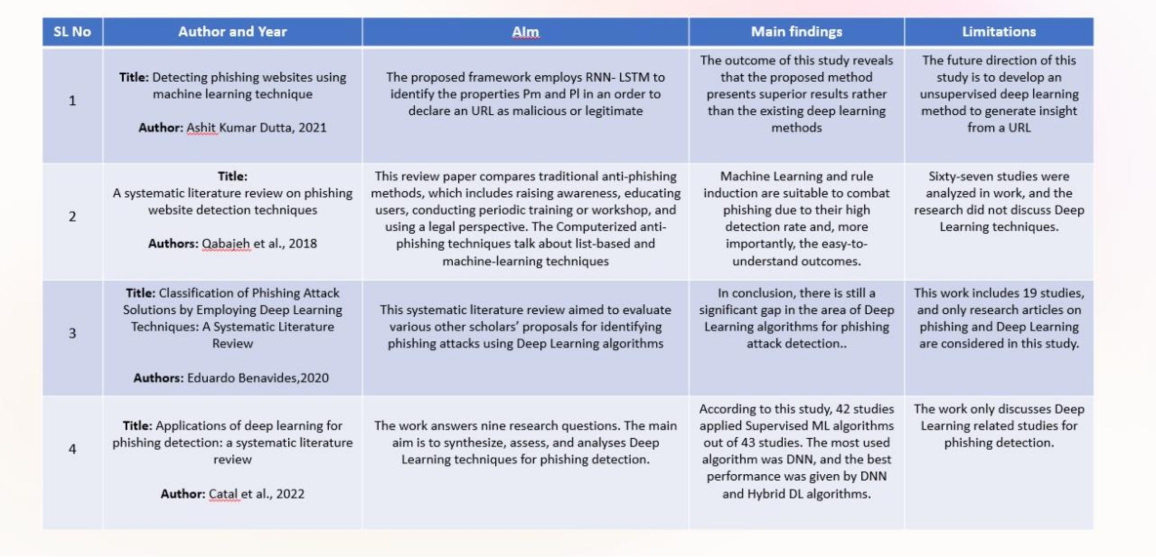
The literature survey conducted for this project provides a comprehensive understanding of the current state-of-the-art techniques, methodologies, and advancements in the domain of phishing detection using machine learning.

The survey encompasses a wide range of research papers, articles, and journals that contribute valuable insights into the intricacies of phishing attacks and the application of machine learning algorithms for their detection. The references to the papers and articles are attached in the References section.

The survey delves into feature extraction methodologies relevant to phishing detection. It investigates how researchers identify discriminative features from web content, structure, and behavior. Commonly used features include URL structures, HTML content analysis, and behavioral patterns, which contribute to the effectiveness of machine learning models.

The survey critically examines the challenges and limitations faced by existing phishing detection systems. It sheds light on areas such as false positives, adversarial attacks, and the need for continuous adaptation to new phishing tactics.

After thoroughly researching the published journals and articles, we found that the past works in this field mainly focused on machine learning algorithms such as Random Forest algorithm, Support Vector Machine, etc. but didn't include Deep Learning to detect the Phishing Websites. We will delve into this aspect and include deep learning in our model to increase its efficiency and accuracy.



SL No	Author and Year	AIm	Main findings	Limitations
1	**Title:** Detecting phishing websites using machine learning technique **Author:** Ashit Kumar Dutta, 2021	The proposed framework employs RNN- LSTM to identify the properties Pm and Pl in an order to declare an URL as malicious or legitimate	The outcome of this study reveals that the proposed method presents superior results rather than the existing deep learning methods	The future direction of this study is to develop an unsupervised deep learning method to generate insight from a URL
2	**Title:** A systematic literature review on phishing website detection techniques **Authors:** Qabajeh et al., 2018	This review paper compares traditional anti-phishing methods, which includes raising awareness, educating users, conducting periodic training or workshop, and using a legal perspective. The Computerized anti-phishing techniques talk about list-based and machine-learning techniques	Machine Learning and rule induction are suitable to combat phishing due to their high detection rate and, more importantly, the easy-to-understand outcomes.	Sixty-seven studies were analyzed in work, and the research did not discuss Deep Learning techniques.
3	**Title:** Classification of Phishing Attack Solutions by Employing Deep Learning Techniques: A Systematic Literature Review **Authors:** Eduardo Benavides,2020	This systematic literature review aimed to evaluate various other scholars' proposals for identifying phishing attacks using Deep Learning algorithms	In conclusion, there is still a significant gap in the area of Deep Learning algorithms for phishing attack detection..	This work includes 19 studies, and only research articles on phishing and Deep Learning are considered in this study.
4	**Title:** Applications of deep learning for phishing detection: a systematic literature review **Author:** Catal et al., 2022	The work answers nine research questions. The main aim is to synthesize, assess, and analyses Deep Learning techniques for phishing detection.	According to this study, 42 studies applied Supervised ML algorithms out of 43 studies. The most used algorithm was DNN, and the best performance was given by DNN and Hybrid DL algorithms.	The work only discusses Deep Learning related studies for phishing detection.

Fig. 1: Literary Survey of the past works.

2.1 **PROBLEM STATEMENT**

Title: Phishing URL Detection System

Problem Statement:
Rising Cyber Threats: The surge in phishing attacks poses a severe threat to online security, demanding a sophisticated solution to accurately detect and prevent malicious URLs.

Objectives:

1. **High Accuracy**: Develop a machine learning model with precise phishing URL classification.
2. **Real-time Detection**: Implement a system for instant analysis, preventing access to harmful websites.
3. **Feature Extraction**: Identify key features for robust analysis of URLs and webpage content.
4. **Behavioral Analysis**: Integrate dynamic behavioral analysis to adapt to evolving phishing tactics.
5. **User-Friendly Interface**: Design an intuitive interface for seamless user interaction.
6. **Scalability**: Ensure efficient handling of a large volume of URL requests without compromising performance.

Outcome:
A proactive defense against phishing attacks, enhancing online security for end-users and professionals.

2.2 **MODES OF PHISHING**

After conducting the literary survey we found the following phishing techniques by which users get scammed-

Email Phishing: Deceptive emails, often impersonating trusted entities, aim to trick recipients into revealing sensitive information through fraudulent links or attachments. Characteristics include spoofed sender addresses and urgent requests.

SMS Phishing (Smishing): Text messages, claiming legitimacy from sources like banks, prompt individuals to click malicious links or disclose sensitive information. Characteristics include urgent messages and requests for personal information.

Call Phishing (Vishing): Voice phishing involves phone calls where attackers, posing as trusted entities, manipulate individuals into divulging sensitive information. Characteristics include manipulative tactics, urgent claims, and requests for information over the phone.

Website Phishing: Fraudulent websites imitate legitimate ones to deceive users into entering sensitive information. Often spread through email or social engineering, these sites aim to capture login credentials, financial details, or personal data.

Social Media Phishing: Cybercriminals exploit social media platforms to deceive users into clicking on malicious links or sharing personal information. Impersonation of trusted contacts and the spread of fake content are common tactics.

Credential Phishing: Attackers use various modes, including emails and fake websites, to trick individuals into divulging login credentials. This information is then exploited for unauthorized access to accounts.

In this project we have primarily focused on Website, Social Media and Email phishing. We also look forward to address the other modes

2.3 **WORKFLOW OF EFFICIENT PHISHING DETECTION**

Successful implementation of machine learning algorithms can automate the phishing detection process, reducing the time required for manual inspection. This efficiency translates to significant time savings for users and organizations dealing with potential threats.

Automated Reporting to APWG:
The automated submission of identified phishing websites to organizations like the Anti-Phishing Working Group (APWG) streamlines the reporting process. This automation not only accelerates the dissemination of threat intelligence but also minimizes the time and effort needed for manual reporting.

Quick User Feedback Loop:
Real-time phishing detection, facilitated by the integration of algorithms and APIs, shortens the feedback loop for users submitting links. Swift responses empower users to make informed decisions promptly, enhancing their ability to navigate the internet securely.

Reduced Investigation Time:
The creation of a structured database allows for efficient storage and retrieval of information on phishing websites. This organized data repository can reduce the time spent on investigations, enabling quicker responses to emerging threats.

Minimal Resource Requirements:

Leveraging open-source technologies and libraries like scikit-learn, TensorFlow, and PostgreSQL can contribute to cost savings by minimizing licensing expenses. Additionally, the collaborative nature of open source often results in efficient problem-solving without the need for extensive financial resources.

User-Driven Threat Intelligence:

The community-driven aspect of the platform, where users contribute links, harnesses collective intelligence. This user-driven model can lead to faster identification and categorization of emerging threats, potentially saving time in the overall threat detection process.

Adaptability for Future Threats:

The scalable and adaptable nature of the project ensures that it remains effective against evolving phishing tactics. This adaptability can result in long-term cost savings by reducing the frequency and scale of necessary updates and modifications.

Reduced Downtime and Losses:

By providing users with a real-time phishing detection tool, the project minimizes the risk of falling victim to phishing attacks. This proactive approach can result in reduced downtime and financial losses that may be incurred due to security breaches.

Educational Impact with Time-Efficient Learning:

The educational impact of the project, in terms of raising awareness about phishing threats, can be achieved more efficiently. Users can quickly grasp and apply cybersecurity best practices, resulting in time-efficient learning and increased resilience against online threats.

Positive ROI on Security Investments:

The project's time and cost-efficient approach, coupled with its potential positive impact on online security, can lead to a positive return on investment (ROI) for organizations investing in cybersecurity measures. This outcome positions the project as a strategic and cost-effective addition to the security landscape.

Chapter 3

PROPOSED DETECTION MODEL

Our project envisions a powerful phishing detection system that combines the versatility of Random Forest and the sequential analysis capabilities of Recurrent Neural Networks (RNN). Users will interact with a user-friendly website, submitting URLs for analysis. The model, implemented in Python, will seamlessly integrate with the website through APIs, ensuring real-time and accurate phishing detection. The frontend, developed using HTML, CSS, JavaScript, and JQuery, offers an intuitive user experience.

A structured database will store information on identified phishing websites, fostering continuous improvement. Detected phishing sites will be systematically reported to organizations like the Anti-Phishing Working Group (APWG), contributing to broader cybersecurity efforts. User feedback mechanisms will enhance the model's adaptability. In essence, our project amalgamates advanced algorithms, collaborative cybersecurity efforts, and user-centric design to fortify our digital landscape against phishing threats.

3.1 **IMPLEMENTATION**

Below are the steps of how we plan on approaching the problem statement:

1. **Define Objectives:** We are building a website where we will collect links of suspected websites. Then we will check the sites using the machine learning algorithms. If the site is a phishing website, we'll add it in our database and then submit the report to organizations like apwg.

2. **Data Collection:** We collected our training and test data from the UCI phishing dataset that is publicly available

3. **Feature Extraction:** Identify relevant features from the URLs that can help distinguish between phishing and legitimate websites. Features might include URL length, presence of HTTPS, domain age, and other relevant characteristics.

4. **Data Preprocessing**: Clean and preprocess the dataset. This involves handling missing values, encoding categorical variables, and scaling numerical features.

5. **Split Data:** Divide the dataset into training and testing sets. This allows you to train the model on one subset and evaluate its performance on unseen data.

6. **Model Selection:** Choose the Random Forest classifier as your machine learning algorithm. Random Forest is effective for classification tasks and can handle a diverse set of features.

7. **Feature Selection:** The difficulty arises when we must determine what are the most relevant features from a set and what combination of features give us near perfect classification accuracies. From the 30 features, we identified five subsets. These were grouped as shown below.

8. **Training:** Train the Random Forest classifier using the training dataset. The model will learn to distinguish between phishing and legitimate websites based on the provided features.

9. **Testing:** Evaluate the model's performance on the testing dataset. Use metrics such as accuracy, precision, recall, and F-score to assess its effectiveness.

10. **Hyper parameter Tuning:** Optimize the performance of the Random Forest by tuning its hyper parameters. This involves adjusting parameters such as the number of trees and tree depth.

11. **Validation:** Perform additional validation, such as cross-validation, to ensure the model's generalizability and robustness.

12. **Deployment:** Once satisfied with the model's performance, deploy it for real-time detection. This could involve integrating it into a web application, browser extension, or network security system.

13. **Monitoring and Updates:** Regularly monitor the model's performance in real-world scenarios and update it as needed to adapt to evolving phishing techniques.

3.2 **FEATURES**

Language Correctness

Measure the English proficiency level of the website content, as phishing sites often contain grammatical errors and awkward phrasing.

Source Year

Evaluate the age of the website, as recently registered domains are more likely to be associated with phishing attempts.

DNS Filter

A DNS filter is a vital component in our phishing website detection strategy. It analyses and categorizes domain names, blocking access to known phishing sites by cross-referencing them with a comprehensive database of malicious entities. This proactive defence mechanism enhances real-time threat mitigation, preventing users from accessing fraudulent websites based on historical associations with phishing activities. Integrating DNS filtering into our system adds a crucial layer of defines, bolstering the effectiveness of our cybersecurity measures.

Reviews

Consider the reputation and feedback from users and security experts to determine the legitimacy of the website.

3.3 **TECH STACKS**

Frontend Development:

HTML: For structuring the content and layout of the website.
CSS: For styling and visual presentation, ensuring an engaging user interface.
JavaScript: To add interactivity and dynamic features to the website.
JQuery: A fast and lightweight JavaScript library for simplifying client-side scripting.

Backend Development:

Node.js: A JavaScript runtime for executing server-side code.
Express.js: A web application framework for Node.js, simplifying the creation of robust APIs.
PostgreSQL: A powerful, open-source relational database management system for efficient data storage and retrieval.

Algorithm Development (Python):

Random Forest Algorithm: Implemented using Python's scikit-learn library, a popular machine learning toolkit.
Recurrent Neural Network (RNN) Algorithm: Developed using Python with TensorFlow or PyTorch, prominent deep learning frameworks.

NumPy: A fundamental package for scientific computing with Python, essential for numerical operations.
Pandas: A data manipulation and analysis library, beneficial for handling structured data.
scikit-learn: A machine learning library that includes tools for classification, regression, clustering, and more.
TensorFlow or PyTorch: Deep learning frameworks for building and training neural networks.

Database Management:

pgAdmin: A comprehensive management tool for PostgreSQL databases, facilitating database administration tasks.
Sequelize: An ORM (Object-Relational Mapping) for Node.js that simplifies database interactions and migrations.
Version Control:

Git: A distributed version control system for tracking changes in the codebase and facilitating collaborative development.
GitHub: Platforms for hosting and managing Git repositories, enabling version control and collaboration.

Text Editor/IDE: Visual Studio Code, Atom, or Sublime Text: Feature-rich text editors suitable for coding, providing a smooth development experience.

API Development:

RESTful API: Building APIs to facilitate communication between the frontend and backend components.

Swagger/OpenAPI: For documenting and testing APIs, ensuring clarity and consistency. By integrating these technologies, the project can leverage a powerful and efficient stack for developing a robust, user-friendly, and effective phishing detection system.

3.4 **SUMMARY OF PRESENT WORK**

The initial phase of our project has been dedicated to laying a robust foundation, encompassing both the development of the user interface and a comprehensive exploration of existing research.
The frontend of our website is now functional, featuring an intuitive input field where users can submit links of websites they suspect to be fraudulent.

Link to the website- https://arnab-batsy.github.io/Detect-Phishing-sites/

Our research phase has been extensive, drawing insights from reputable journals and articles in the field of phishing detection. This literature review has been instrumental in shaping our approach and finalizing the algorithms we plan to implement. We have carefully selected and defined the machine learning algorithms that will power our system, ensuring a potent combination of accuracy and adaptability.

Looking ahead, our immediate future work involves the implementation of these algorithms using Python. This development phase will bring our envisioned machine learning models to life, incorporating the intricacies identified during our research phase. We plan to seamlessly integrate these algorithms into our website using APIs, enabling users to experience real-time phishing detection capabilities.

As a pivotal step, we will establish a database to systematically store information on identified phishing websites. This database will not only serve as a repository for our internal use but will also facilitate the submission of reports to organizations dedicated to combating phishing, such as the Anti-Phishing Working Group (APWG). The integration of a database ensures data persistence and allows for future analyses and refinements.

3.5 **FUTURE WORKS**

Moving forward, our roadmap includes the following key milestones:

Algorithm Implementation: Finalize the development of machine learning algorithms for phishing detection using Python.

API Integration: Seamlessly link the algorithms to the website through APIs, ensuring a user-friendly and responsive interface.

Database Creation: Establish a robust database structure to store information on identified phishing websites, enhancing data management and analysis capabilities.

Submission to APWG: Implement a systematic process to submit reports on identified phishing websites to organizations like APWG, contributing to the broader cybersecurity community.

User Feedback Integration: Incorporate mechanisms for user feedback to continuously improve and refine the performance of our phishing detection system.

Through these planned future works, we aim to transform our research and planning into a functional and impactful tool. By marrying technological innovation with community engagement, we strive to create a comprehensive solution that not only detects phishing websites effectively but actively contributes to the global fight against cyber threats.

CONCLUSION

Chapter 4

In the culmination of our project's initial phases and the roadmap for its future, a cohesive and robust framework for detecting phishing websites using machine learning has taken shape. The completion of the frontend development, coupled with an extensive literature survey, has laid the groundwork for a user-friendly platform that fosters community-driven threat intelligence. The integration of machine learning algorithms, specifically the Random Forest and RNN models, promises to enhance the accuracy and adaptability of our phishing detection system.

The research journey, guided by insights from reputable journals, has deepened our understanding of phishing attacks and their evolving tactics. The literature survey has informed our approach, ensuring that the project aligns with current best practices and stays at the forefront of advancements in cybersecurity.

Looking forward, the implementation phase beckons, where the chosen algorithms will be brought to life using Python, seamlessly integrated into our website through APIs. This critical stage represents the bridge between theory and practical application, translating our research findings into a functional tool capable of real-time phishing detection.

The establishment of a structured database and the planned submission of identified phishing websites to organizations like the Anti-Phishing Working Group (APWG) underscore the project's commitment to actively contributing to the global fight against cyber threats. These future works not only enhance the project's practical impact but also position it within the broader landscape of collaborative cybersecurity initiatives.

In conclusion, our project strives not only to create a sophisticated machine learning-based solution for phishing detection but also to cultivate a community-driven, resilient defense against cyber threats. As we progress into the implementation and refinement phases, we remain dedicated to the overarching goal of creating a safer digital environment, where users can navigate with confidence, shielded from the insidious threat of phishing attacks. The fusion of technological innovation, community collaboration, and a commitment to ongoing improvement defines the essence of our project, embodying a proactive and collective approach to cybersecurity in the digital age.

REFERENCES

Chapter 5

The Published journals and articles that have helped us in gathering information regarding the past works in this topic are enlisted as follows-

1. Görkem Giray, Bedir Tekinerdogan, Sandeep Kumar & Suyash Shukla, "Applications of deep learning for phishing detection: a systematic literature review" in *Knowledge and Information Systems* 23rd May 2022.
2. Eduardo Benavides, Walter Fuertes, Sandra Sanchez & Manuel Sanchez, "Classification of Phishing Attack Solutions by Employing Deep Learning Techniques: A Systematic Literature Review" in *Part of the Smart Innovation, Systems and Technologies book series (SIST,volume 152)* 14th June 2019.
3. Issa Qabajeh, Fadi Thabtah, Francisco Chiclana, "A recent review of conventional vs. automated cybersecurity anti-phishing techniques" in *Centre for Computational Intelligence, De Montfort University, Leicester, UK* 15 June 2018.
4. Ashit Kumar Dutta, "Detecting phishing websites using machine learning technique" in *Zhihan Lv, Qingdao University, China* October 11, 2021.

The online web-links to the above references are as follows-

- https://link.springer.com/article/10.1007/s10115-022-01672-x
- https://link.springer.com/chapter/10.1007/978-981-13-9155-2_5
- https://www.sciencedirect.com/science/article/abs/pii/S1574013717302010?via%3Dihub
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0258361