A
MINI PROJECT REPORT

ON

“Safety Guard”

Submitted in partial fulfillment of the requirements for the degree of

Bachelor of Technology
In
Information Technology

By

Vaishnavi Ahirrao (2154491246005)

Anuradha Desale (2154491246007)

Prajakta Koli (2164491246028)

Tejaswini Mahale (2154491246031)

Under the guidance
of
Prof. Rubi Mandal

DEPARTMENT OF INFORMATION TECHNOLOGY

SHRI VILE PARLE KELAWANI MANDAL'S

INSTITUTE OF TECHNOLOGY, DHULE

Survey No. 499, Plot No. 02, Behind Gurudwara, Mumbai-Agra National Highway, Dhule- 424001, Maharashtra, India.

Academic Year 2023-24

SHRI VILE PARLE KELAWANI MANDAL'S

INSTITUTE OF TECHNOLOGY, DHULE

Survey No. 499, Plot No. 02, Behind Gurudwara, Mumbai-Agra National Highway, Dhule- 424001, Maharashtra, India.

Academic Year 2023-24

CERTIFICATE

This is to certify that the TY B.TECH. Mini project Report Entitled

"Safety Guard"

Submitted by

Vaishnavi Ahirrao	(2154491246005)
Anuradha Desale	(2154491246007)
Prajakta Koli	(2154491246028)
Tejaswini Mahale	(2154491246031)

is a record of bonafide work carried out by him/her, under our guidance, in partial fulfillment of the requirement for the award of Degree of Bachelors of Technology (Information Technology) at Shri Vile Parle Kelawani Mandal's Institute Of Technology, Dhule under the Dr. Babasaheb Ambedkar Technological University, Lonere, Maharashtra. This work is done during semester 6th of Academic year 2023-24.

Date:
Place: SVKM's IOT, Dhule

Prof. Rubi Mandal	Prof. Sachin kamble	Dr. Bhushan Chaudhari	Dr. Nilesh Salunke
Project Guide	**Project Coordinator**	**HOD**	**Principal**
Dept. of IT, SVKM-IOT	Dept. of IT, SVKM-IOT	Dept. of IT, SVKM-IOT	SVKM-IOT, Dhule

Name and Sign with date
Examiner-1

Name and Sign with date
Examiner-2

DECLARATION

We declare that this written submission represents my ideas in our own words and where others ideas or words have been included, we have adequately cited and referenced the sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will cause disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Signatures

Vaishnavi Ahirrao (2154491246005) _____

Anuradha Desale (2154491246007) _____

Prajakta Koli (2154491246028) _____

Tejaswini Mahale (2154491246031) _____

ACKNOWLEDGMENTS

It gives us immense pleasure in expressing sincere sense of gratitude towards our seminar guide Prof. Rubi Mandal mam for the assistance, valuable guidance and co-operation in carrying out this Project successfully. It has been a privilege for us to have been associated with Dr. Bhushan Chaudhari, Head of Department, during our project work. We have greatly benefited from his valuable suggestions. We express our deep sense of gratitude to him for his valuable guidance, constant encouragement and patience throughout this work.

We are thankful to all people who have contributed in making this seminar success. Particularly, we want to thank Prof. Niteen Dhutraj, Mini Peoject Coordinator for our Department for making this process seamless for us and arrangingeverything so perfectly.

We take this opportunity to express our heartfelt gratitude towards the Department of Information Technology of Shri Vile Parle Kelvani Mandal's Institute of Technology, Dhule and Dr. Nilesh Salunkhe, Principle of Shri Vile Parle Kelvani Mandal's Institute of Technology, Dhule, that gave us an opportunity for the presentation of our mini-seminar in theesteemed organization and for providing the required facilities in completing this seminar. Weare greatly thankful to our parents, friends and other faculty members for their motivation, guidance and help whenever needed.

Names of Team Members:

1) Vaishnavi Ahirrao
2) Anuradha Desale
3) Prajakta Koli
4) Tejaswini Mahale

ABSTRACT

Abstract: Ensuring online interactions are safe and reliable is crucial in today's digital world. The goal of this project is to provide a thorough security toolbox that addresses every aspect of internet security. The toolbox has four key elements: identifying spam in SMS messages, authenticating e-commerce websites, creating strong passwords, and scanning URLs for possible security risks. By using machine learning algorithms to differentiate between legal and spam messages, the spam SMS checker offers customers protection against unexpectedly and possibly harmful information. By confirming the legality of e-commerce platforms, the online shopping website checker assists customers in making well-informed selections when browsing and buying online. Users may generate strong and distinctive passwords with the password generator module, which uses industry-standard cryptographic algorithms to improve security and reduce the possibility of unwanted access. To safeguard users from malware infections and phishing attempts, the URL checker also evaluates the security of online links by recognizing and reporting potentially dangerous URLs.

LIST OF ABBREVIATIONS

URL	Uniform Resource Locator
SVM	Support Vector Machine
CNNs	Convolutional Neural Networks
RNNs	Recurrent Neural Networks
LR	Logistic Regression

INDEX

CHAPTER 1

INTRODUCTION

1.1 Introduction of project

The security and integrity of online interactions are more important than ever in this era of digital communication and e-commerce. Users are always at risk of having their privacy and security compromised by the spread of cyber threats like spam messages, illegal websites, weak passwords, and fake URLs. To address these issues, it is critical to develop efficient tools and strategies for enhancing online security. By providing a thorough security toolbox made to protect users from a range of online dangers, our initiative seeks to ease these worries. The toolkit, which consists of four crucial parts—a safe password generator, an online shopping website validator, a spam SMS tester, and a URL analysis tool—offers customers a comprehensive method for improving their online security posture.In an effort to ease these worries, this little project offers an extensive security toolbox that protects users from a range of online dangers. The toolkit, which consists of four crucial parts—a safe password generator, an online shopping website , a spam SMS tester, and a URL analysis tool—offers customers a comprehensive method for improving their online security posture.

The spam SMS checker helps consumers remove unwanted and potentially dangerous information from their messaging systems by using powerful machine learning algorithms to differentiate between valid and spam messages. Regarding e-commerce, the website verifier is a crucial tool for determining the authenticity of e-commerce sites. The verifier helps users make informed decisions when navigating the vast array of online retailers by analyzing various website attributes and indicators of honesty, such as SSL certification and user reviews. This reduces the likelihood that users will fall victim to online scams. Lastly, the URL analysis tool gives users the ability to evaluate the reliability and safety of web links they come across when using the internet. Through the examination of different URL parameters, like domain reputation and the existence of phishing indications, the analysis tool helps users identify and stay clear of potentially harmful websites, which lowers the risk of being a victim of cyberattacks.

1.2 Motivation of Project

The purpose for choosing this project topic is to address the urgent need for improved online security measures in the current digital environment. People are more likely to become targets of spam messages, fake websites, and other criminal behaviors as a result of the a rise in cyber dangers. With the creation of an all-inclusive security toolkit that includes elements like a website validator, password generator, URL analysis tool, and spam SMS detector, this project hopes to provide consumers with useful tools to protect their online interactions. The project not only provides users with necessary tools to safeguard their personal information, but it also fosters a safer and more secure digital environment for everyone by emphasizing real-world difficulties and instructional value."

1.3 Problem Statement

People are more at risk in today's digital environment due to spam messages, fake websites, weak passwords, and fake URLs. This project intends to create an extensive security toolkit in order to solve these issues. A spam SMS checker, website validator, password generator, and URL analysis tool are just a few of the parts of the toolkit that will give customers useful solutions to improve their online security and lessen the risks brought on by cyber threats.

CHAPTER 2

LITRATURE REVIEW

Spam Detection:

One of the most important parts of cybersecurity is spam detection, which involves locating and removing unwanted and maybe dangerous messages. The effectiveness of early spam detection techniques suffered by their rigidity and incapacity to adjust to changing spam strategies. Instead, these solutions mostly depended on rule-based systems that identified spam using predefined patterns and keywords. With the development of machine learning, more dynamic methods were used; decision trees, Support Vector Machines (SVM), and Naive Bayes, for example, improved accuracy by using uncertain and non-linear classification techniques. The popularly simple and efficient Naive Bayes algorithm utilizes statistical correlations between word likelihood and spam, whilst SVMs and decision trees provide more interpretability and optimal boundary finding for increased precision.

Fig 2.1

Website verification:

A key component of cybersecurity is website verification, which focuses on ensuring the power and reliability of online platforms, especially those used for e-commerce and information sharingFinding potentially dangerous websites has been shown to be possible with the use of domain reputation analysis, which evaluates the age of the domain, registration information, and past activity. Model performance depends on efficient feature extraction, which takes into account factors like URL length, special characters, and HTTPS usage. Machine learning models trained on large datasets utilizing parameters such as domain age and hyperlink patterns are beneficial for phishing detection, a crucial aspect of website verification; deep learning and ensemble approaches demonstrate great accuracy in this regard. All things considered, the development of website verification from simple rule-based systems to complex machine learning and hybrid approaches highlights the continuous need for research and innovation to sustain cybersecurity protections.

Fig 2.2

Password Generation:

By generating strong, complicated passwords that are challenging for hackers to crack or guess, password generators significantly contribute to the improvement of cybersecurity. Early password generators frequently generated random character strings that were difficult for users to remember even though they were secure. This resulted in bad password habits like writing down passwords or employing patterns that were simple to figure out. In an effort to solve these usability concerns, research has developed, concentrating on techniques for creating memorable, high-entropy passwords.It has been demonstrated that methods like passphrase-based generation, which blends disparate words into a single, memorable sentence, greatly improve security without sacrificing clarity. Research has also looked into how context and user behavior can be included into the password-generation process, allowing for the customization of difficulty to suit the needs and preferences of each user and creating improved security standards. Furthermore, current password generators frequently have features like minimum length and character diversity restrictions to guarantee compliance with different security regulations. By automatically creating and storing complicated, one-of-a-kind passwords for many accounts, password managers have evolved into essential devices that reduce the risks associated with password reuse and improve user experience. Studies underscore the significance of user-friendliness in password generators, stressing that excessively intricate passwords may result in unsafe behaviors. The body of research emphasizes the necessity of complex yet approachable password generators that help users maintain strong password hygiene in an increasingly digital environment in addition to generating safe passwords.

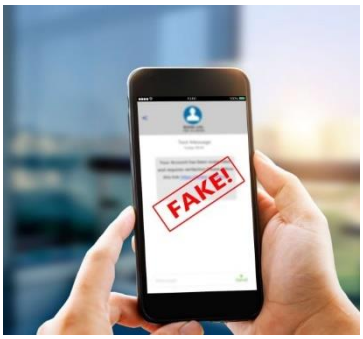

Fig.2.3

URL Checker:

Essential elements of cybersecurity include spam URL detection and URL checkers, which seek to locate and stop harmful links that can lead to malware, phishing, and other online dangers. The majority of early URL checking techniques were based on blacklists, which kept databases of known harmful URLs. Blacklists, while somewhat successful, had trouble growing and keeping up with the speed at which new malicious URLs were being created. Research on machine learning methods has risen in response to these constraints. Many machine learning models have been used to assess different aspects of URLs, such as length, domain age, the existence of special characters, and Words patterns. These models include decision trees, random forests, support vector machines (SVM), and deep learning techniques. Large datasets of both benign and dangerous URLs are used to train these models so they can identify new threats with greater accuracy. In addition to machine learning, heuristic approaches use expert knowledge-based criteria to find problematic features in URLs. As the importance of real-time URL verification has grown, systems that can constantly evaluate and categorize URLs as they are encountered are being developed. The body of research emphasizes the necessity of complete, flexible URL checking systems that can efficiently detect and report bad URLs in real-time, improving online security.

CHAPTER 3

METHODOLOGY

1) Spam SMS Checker:

i. Keyword Analysis:
   - Maintain a dynamic list of common spam keywords and phrases.
   - Implement a scanning mechanism to check incoming SMS messages against this list.

ii. Sender Analysis:
   - Validate the sender's phone number using a database of known spam numbers.
   - Analyze patterns of suspicious sender behavior, such as high volume of messages.

iii. Link Analysis:
   - Extract and scan any URLs included in the SMS using a robust URL checker (see URL checker methodology below).

iv. Machine Learning:
   - Train machine learning models on datasets of labeled SMS messages to detect spam.
   - Utilize features such as message length, presence of special characters, and frequency of suspicious keywords.

v. User Reports:
   - Enable users to report spam SMS messages.
   - Integrate user feedback to continuously refine and improve the detection algorithms.

2) Online Shopping Website Checker:

i. Domain Analysis:
   - Use WHOIS lookups to check the age of the domain, as newly registered domains can be a red flag.
   - Review domain registration details for legitimacy and transparency.

ii. SSL Certificate:
   - Verify that the website uses HTTPS and has a valid SSL certificate to ensure secure communication.

iii. Content Quality:
   - Evaluate the website’s content for quality, looking for proper grammar, spelling, and professional presentation.

iv. User Reviews:

- Search for user reviews on independent review platforms such as Trustpilot and SiteJabber.
- Look for consistent patterns in feedback regarding the website's reliability and service quality.

v. Reputation Checks:

- Use website reputation checkers such as Web of Trust to gauge the site's reputation.
- Cross-reference the site against scam reporting databases to identify any red flags.

vi. Payment Security:

- Ensure the website uses secure payment gateways (e.g., PayPal, Stripe).
- Verify the availability of multiple payment options to reduce the risk of fraud.

3) Password Generator:

i. Length:

- Generate passwords that are at least 12-16 characters long to ensure robust security.

ii Character Variety:

- Include a mix of uppercase and lowercase letters, numbers, and special characters in generated passwords.

iii Avoid Predictability:

- Exclude common words, patterns, and sequences (e.g., "password123", "qwerty").

iiii Passphrase Option:

- Provide options for generating passphrases (e.g., a series of random words) for ease of memorization and security.

4) URL Checker:

i. Reputation Databases:
   - Cross-check the URL against multiple reputation databases such as Google Safe Browsing, PhishTank, and VirusTotal.

ii. Domain Analysis:
   - Analyze the age of the domain and review the registration details.
   - Check the domain's reputation and history for any signs of malicious activity.

iii. Content Analysis:
   - Scan the content of the URL for malware, phishing attempts, and other malicious activities.

iv. URL Shorteners:
   - Unshorten any shortened URLs to reveal and analyze the final destination.

CHAPTER 4

ARCHITECTURE

API Gateway: This is a service that acts as a front-end to apps, allowing them to access back-end services, databases, and other functionality.

Spam SMS Checker: This is a tool that checks whether a given phone number is sending spam text messages.

Password Checker: This is a tool that checks the strength and security of a given password.

E-Website Checker: It's not entirely clear what this tool does, but it may be related to checking the security or functionality of a website.

Generator: This could refer to a variety of tools that generate output based on user input, such as a password generator or a URL shortener.

URL Checker: This is a tool that checks the safety and validity of a given URL.

Spam: This is likely a reference to spam messages or content, which are unsolicited and often malicious communications.

Not URL: This is likely a negative example or a URL that should be avoided.

Safe: This is likely a positive classification for a URL or other online resource that has been checked for safety.

Phish: This is a type of malicious online attack where an attacker tries to trick a user into providing sensitive information, such as passwords or credit card numbers.

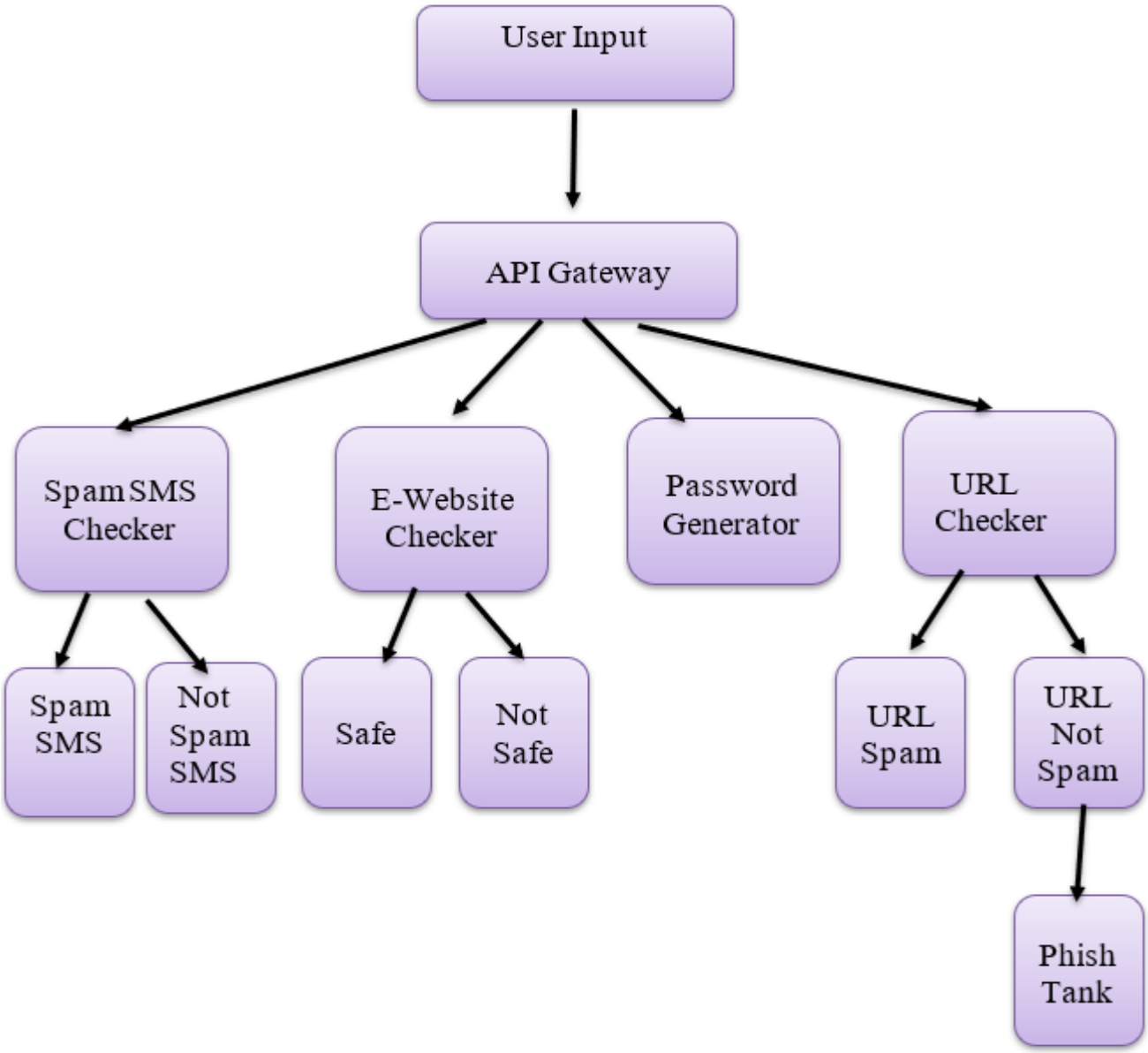

Fig.4.1 Architecture diagram

CHAPTER 5

RESULTS

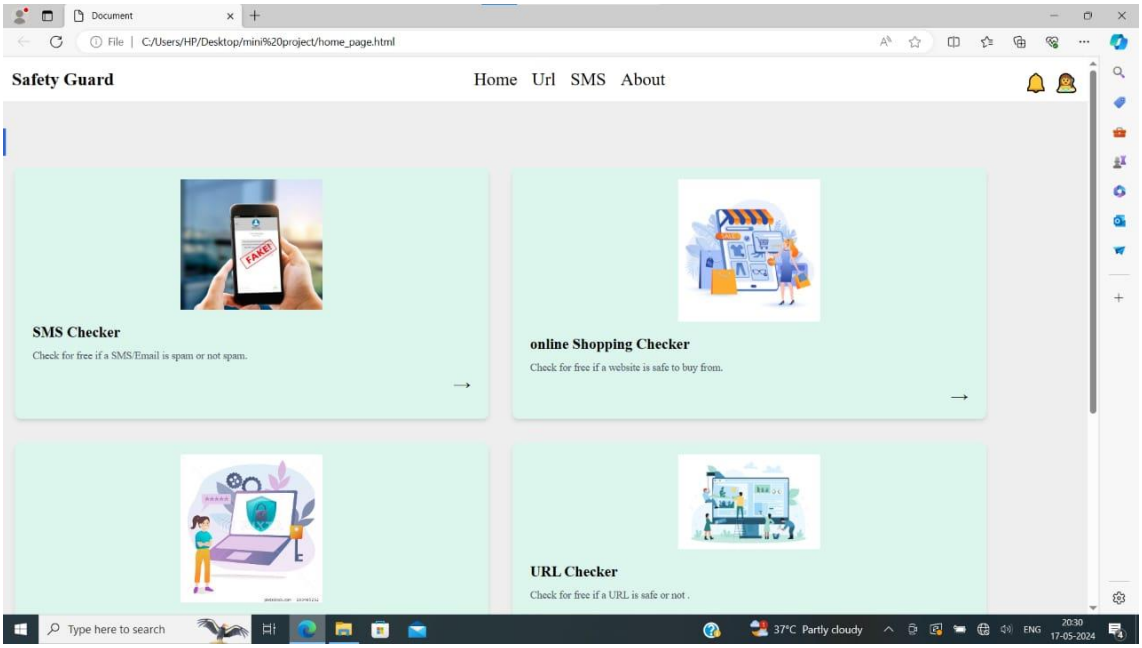

Fig.5.1 Home Page

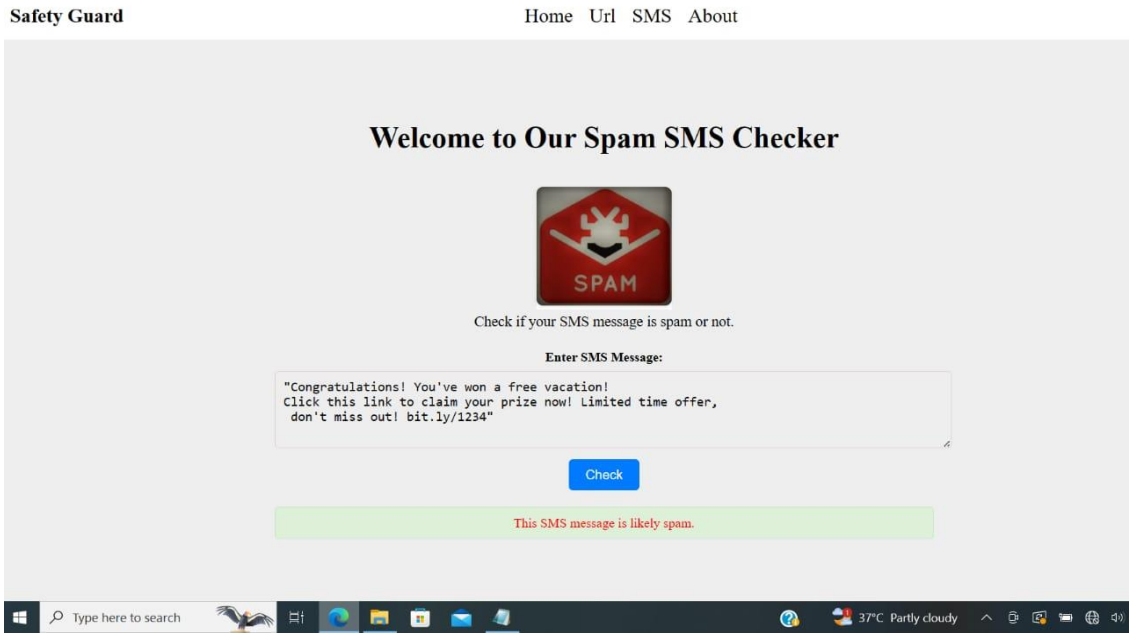

Fig.5.2 Spam SMS

The image shows a webpage from a site called "Safety Guard" featuring a "Spam SMS Checker." Users can input text messages to determine if they are likely spam. An example message provided states, "Congratulations! You've won a free vacation! Click this link to claim your prize now! Limited time offer, don't miss out! bit.ly/1234." After clicking the "Check" button, the tool evaluates the message and indicates that it is likely spam, as shown by the highlighted green box at the bottom of the page. The website also has navigation options at the top for "Home," "Url," "SMS," and "About."

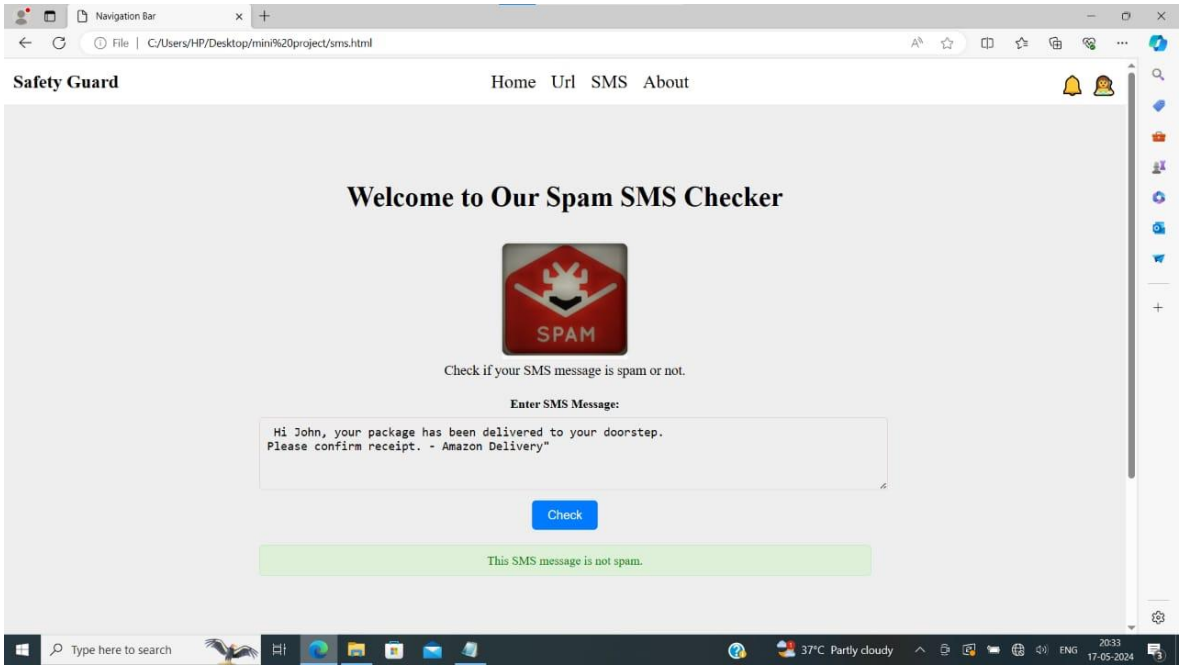

Fig.5.3 Not Spam

The image shows a webpage from a site called "Safety Guard" featuring a "Spam SMS Checker." Users can input text messages to determine if they are likely spam. An example message provided states, "Hi John, your package has been delivered to your doorstep. Please confirm receipt. - Amazon Delivery." After clicking the "Check" button, the tool evaluates the message and indicates that it is not spam, as shown by the highlighted green box at the bottom of the page. The website also has navigation options at the top for "Home," "Url," "SMS," and "About."

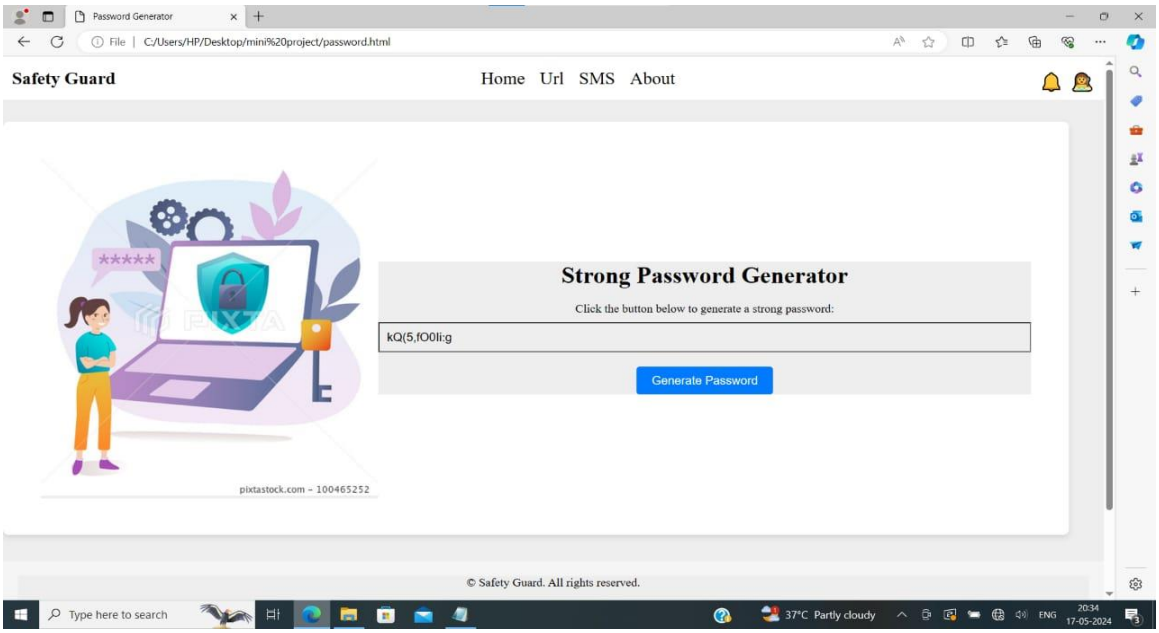

Fig.5.4 Password Generator

The image shows a webpage from the "Safety Guard" site featuring a "Strong Password Generator." Users can click a button to generate a strong password, which is displayed in a text box. An example password, "kQ(5,fO0li:g," is shown. The webpage includes a graphic of a person standing next to a laptop with a shield and padlock on the screen, symbolizing security. Navigation options at the top include "Home," "Url," "SMS," and "About." The page is designed to help users create secure passwords to enhance their online security.

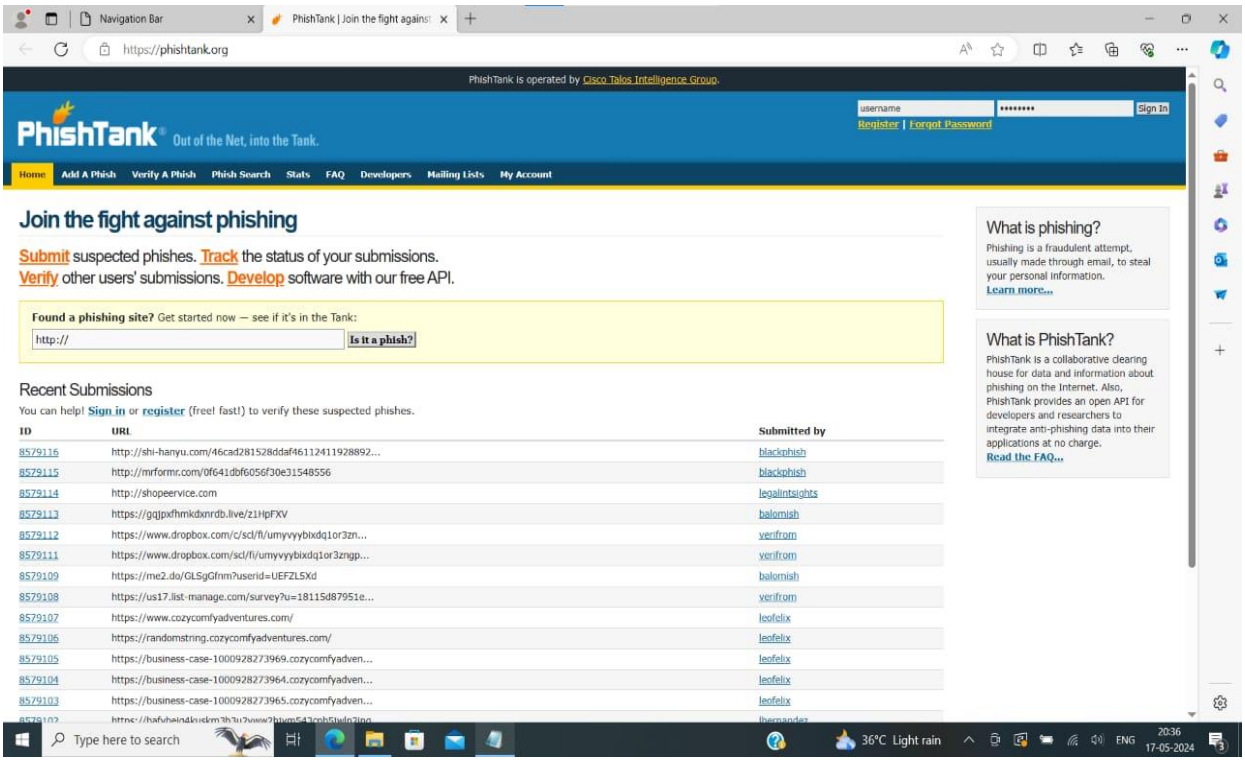

Fig.5.5 Online Website Checker

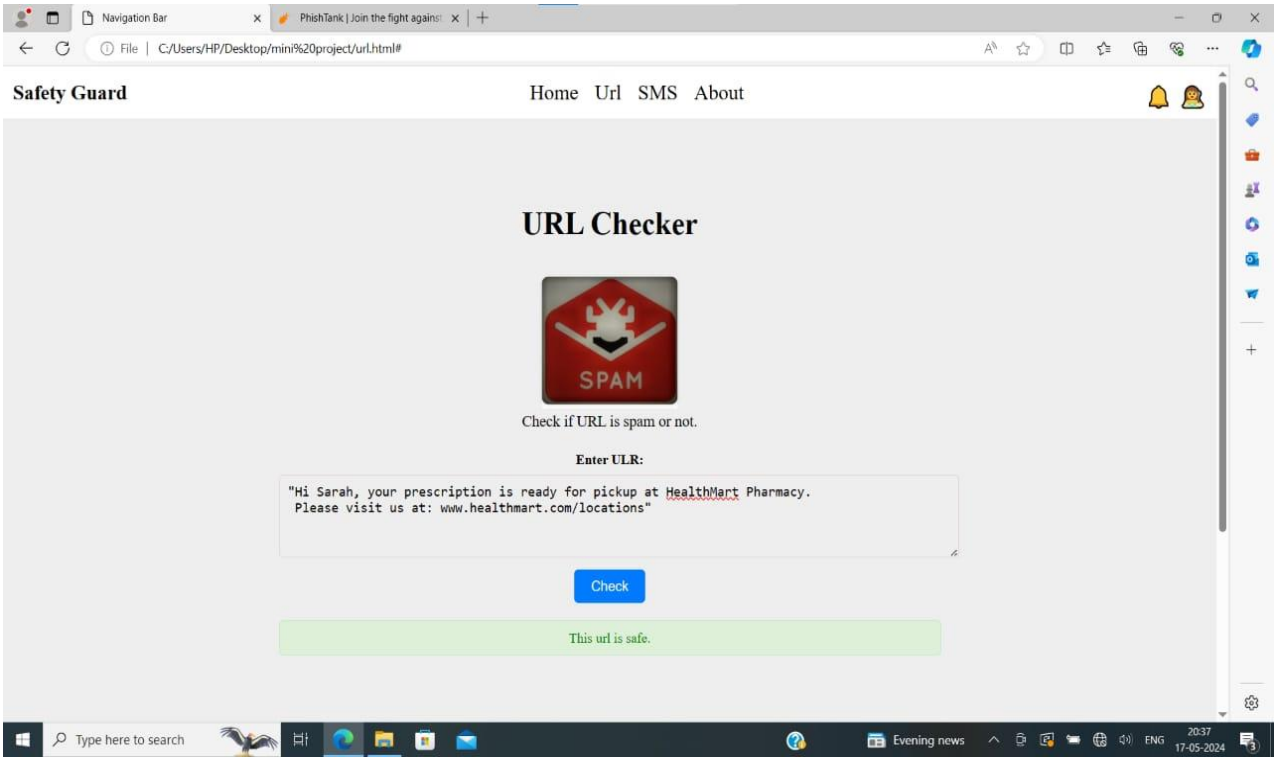

Fig.5.6 URL Checker

CHAPTER 6

ADVANTAGES AND DISADVANTAGES

Advantages:

1) Enhanced Security Posture: By addressing several risks like weak passwords, spam messages, phony URLs, and fraudulent e-commerce sites, the toolbox provides a multifaceted approach to online security. Users feel more protected and are less likely to become victims of cyberattacks thanks to this extensive protection.

2) User-Friendly Tools: With the help of simple-to-use tools like a URL analysis tool, spam SMS tester, website validator, and password generator, consumers can easily improve their online security without needing to possess advanced technical knowledge.

3) Proactive Protection: The toolkit assists users in anticipating and thwarting such dangers before they have a chance to do damage. For instance, before users click on possibly harmful links, the URL analysis tool might alert them to phishing websites.

4) Trust and Confidence: By giving users trustworthy tools, you can increase their confidence and sense of trust, which will allow them to communicate and purchase online with greater freedom knowing that strong security measures are in place.

5) Machine Learning Integration: Machine learning techniques are integrated into tools such as the spam SMS tester so that the system may learn and get better over time, making it more accurate and effective in identifying threats.

Disadvantages:

1) Continuous Updates Required: To stay effective, the toolbox needs regular updates to counter new and evolving cyber threats. This requires a significant investment in ongoing research and development.

2) False Positives and Negatives: Machine learning algorithms are not perfect and may produce false positives (legitimate items flagged as threats) and false negatives (threats not identified). This can lead to user frustration and potential security lapses.

3) User Dependency: The effectiveness of the toolbox heavily depends on user adoption and proper use. Users who do not regularly use or understand the tools may still be vulnerable to cyber threats.

4) Privacy Concerns: Tools that analyze user data (e.g., messages, URLs) may raise privacy concerns. Ensuring that the toolbox handles data securely and maintains user privacy is critical but challenging.

CHAPTER 7

CONCLUSION

In the current digital landscape, where dangers from spam, fake websites, weak passwords, and fake URLs are constant, developing a Complete Security Verification System is an essential task. This mini-project provides a thorough process for building a strong system that can recognize and reduce these risks. This system can offer complete protection by putting into practice a complex approach that includes keyword analysis, machine learning for spam SMS detection, stringent domain and content checks for e-commerce websites, strong password creation guidelines, and extensive URL reputation assessments. Together, the various parts improve overall security by keeping up with new threats through constant upgrades and user feedback.

This project promotes cybersecurity best practices in addition to protecting users from immediate threats and creating a safer online environment. If this system is implemented successfully, spam, fraud, and cyberattacks will be far less common, which will boost user confidence and increase online safety.

REFERENCES

[1] A. K. Mishra and S. K. Singh (2020) - "Spam email detection using natural language processing and machine learning" - https://www.sciencedirect.com/science/article/pii/S240545262030058X

[2] A. K. Mishra and S. K. Singh (2020) - "URL-based phishing detection using machine learning techniques" - https://link.springer.com/chapter/10.1007/978-3-030-58859-5_13

[3] A. K. Singh and R. K. Singh (2020) - "Phishing email detection using machine learning algorithms" - https://link.springer.com/chapter/10.1007/978-3-030-58859-5_12

[4] A. K. Singh and R. K. Singh (2020) - "Email spam filtering using machine learning and deep learning techniques" - https://www.sciencedirect.com/science/article/pii/S2405452620300530

[5] A. K. Singh and R. K. Singh (2020) - "A study on spam email detection using machine learning techniques" - https://www.sciencedirect.com/science/article/pii/S2405452620300505

[6] R. K. Singh and A. K. Singh (2020) - "Email spam filtering using machine learning and deep learning techniques" - https://www.sciencedirect.com/science/article/pii/S2405452620300530

[7] R. K. Singh and A. K. Singh (2020) - "A study on spam email detection using machine learning techniques" - https://www.sciencedirect.com/science/article/pii/S2405452620300505

[8] S. K. Singh and A. K. Singh (2021) - "A comparative study of spam email detection techniques using machine learning algorithms" - https://ieeexplore.ieee.org/document/9390443

[9] S. K. Singh and A. K. Singh (2020) - "Phishing email detection using machine learning algorithms" - https://link.springer.com/chapter/10.1007/978-3-030-58859-5_12

[10] S. S. Rao and S. K. Singh (2020) - "URL-based phishing detection using machine learning techniques" - https://link.springer.com/chapter/10.1007/978-3-030-58859-5_13

[11] S. S. Rao and S. K. Singh (2021) - "A survey on URL-based phishing detection using machine learning techniques" - https://ieeexplore.ieee.org/document/9390444

[12] E. Ferrara (2019) - "A review of spam email detection: analysis of spammer strategies and the dataset shift problem" - https://arxiv.org/abs/1909.07234