Chrome Extension For

Phishing Websites

Detection

Page **1** of **16**

**TABLE OF CONTENTS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Chapters** |  |  |  |  | **Description** | | | | | | | | | **Page.no** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  | **Abstract** | | | | | | | **3** |
| **1** |  |  |  |  |  | **Introduction** | | |  | | | | |  |
|  |  | 1.1 introduction | | | | | | | | | | | |  |
|  |  | 1.2 problem statement | | | | | | | | | | | | **4-6** |
|  |  | 1.3 project objective | | | | | | | | | | | |  |
|  |  | 1.4 scope and limitations | | | | | | | | | | | |  |
|  |  |  |  |  | of the project | | | | | | | | |  |
|  |  |  |  |  |  | |  |  | | |  |  |  |  |
| **2** |  |  |  |  | **Existing Solutions** | | | | | | | | | **7** |
|  |  |  |  |  |  | |  |  | |  |  |  |  |  |
|  |  | 2.1 existing solutions for | | | | | | | | | | | | **8** |
|  |  |  |  |  | Phishing Websites | | | | | | | | |  |
|  |  |  |  |  |  | | | | | | | | |  |
|  |  |  |  |  | | |  |  | | | |  |  |  |
| **3** |  |  |  | **Proposed Solutions** | | | | | | |  | | | **9** |
|  |  | 3.1 proposed solutions for | | | | | | | | | | | | **10** |
|  |  | Phishing Websites | | | | | | | | | | | |  |
|  |  |  |  |  |  |  |  | | | | | | |  |
|  |  |  |  | | | |  |  | | | |  | |  |
| **4** |  | **System Requirements and Tools** | | | | | | | | | | |  | **11** |
|  |  | 4.1 Hardware Components | | | | | | | | | | | | **12** |
|  |  | 4.2 Software Tools and | | | | | | | | | | | |  |
|  |  |  |  |  | Versions Used | | | | | | | | |  |
|  |  | 4.3 Network and Other | | | | | | | | | | | |  |
|  |  |  |  |  | Dependencies | | | | | | | | |  |
| **5** |  |  | **Results and Discussion** | | | | | | | | |  | | **13** |
|  |  | 5.1 Results of Chrome extension | | | | | | | | | | | | **14-15** |
|  |  |  |  |  | | | | | | | | | |  |
|  |  |  |  |  | | | | | | | | | |  |
|  |  |  |  |  |  |  |  | | | | | | |  |
| **6** |  |  |  |  |  |  | **Conclusion** | | | | | | | **16** |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Page **2** of **16**

**ABSTRACT**

In today's digital world,Hackers using different techniques to steal information from the end-users(the

person who getting service).

Phishing is basically a type of cyberattack or it is technique in the Social Engineering Techniques that

uses fraudulent emails,text messages, phone calls or the websites to ntrick people into sharing sensitive

data, downloading malware or otherwise exposing themselves to cybercrime. Fake websites are set up to

trick victims into divulging personal and financial information, such as passwords, account IDs or credit

card details. Phishing emails often appear to come from credible sources and contain a link to click on and

an urgent request for the user to respond quickly.

Chrome extension is designed to combat such threats by identifying and blocking phishing websites in

real-time, ensuring user safety against malicious content.

Current phishing detection mechanisms rely heavily on user vigilance and existing security software,

which may not always be effective against sophisticated phishing techniques

.

Our Chrome extension leverages machine learning algorithms and real-time web analysis to detect and

block phishing websites. The architecture includes a backend server for processing data and a frontend

Chrome extension for user interaction.Initially we gather some data regarding the phishing websites and

legitimate websites(secured websites) for training the model. And also going to tested on the real time

websites. The expected output may be like when ever you are going to open a website our extension going

drop a notification(The website is malicious) if the website the illegitimate.

The developed Chrome extension effectively identifies and blocks phishing websites, offering an

intelligent solution to mitigate phishing attacks. By keeping users informed about phishing risks and

providing real-time protection, this extension aims to reduce cybercrime and enhance online safety over

time.

As the digital landscape continues to evolve, the threat of phishing attacks looms large, casting a shadow

of uncertainty over the online community. However, with the "Detecting Phishing websites" Chrome

extension, a beacon of hope emerges. This innovative solution harnesses the power of machine learning

and cutting-edge technology to fortify the defenses against phishing attacks, providing a safer and more

secure browsing experience for users.By integrating seamlessly into users' daily online activities, this

extension serves as a vigilant guardian, scrutinizing URLs and website content to identify and block

malicious links. The result is a significant reduction in the risk of fraud, identity theft, and financial losses,

ultimately contributing to a safer and more trustworthy internet. As we navigate the complexities of the

digital age, it is imperative that we prioritize cybersecurity and stay one step ahead of the perpetrators of

phishing attacks. The "Detecting Phishing websites" Chrome extension is a testament to the power of

innovation and collaboration in the pursuit of a safer online world

Page **3** of **16**

**CHAPTER-1**

**Introduction**

Page **4** of **16**

**1.1** **Introduction**

Phishing websites are fraudulent sites designed to deceive users into revealing sensitive

information such as usernames, passwords, credit card numbers, or other personal details. These

websites often mimic legitimate sites by using similar URLs, layouts, and branding to trick users

into thinking they are on a trusted platform. The primary goal of phishing sites is to harvest

personal information for malicious purposes, such as identity theft or financial fraud. Detecting

and avoiding phishing websites is crucial for maintaining online security, and various techniques,

including machine learning, are employed to identify and block these deceptive sites.

**1.2 Problem Statement**

The rapid increase in internet usage has led to a corresponding rise in cyber threats, particularly

phishing attacks. Phishing websites, which deceptively mimic legitimate sites to steal sensitive

user information, pose significant risks to online security. Traditional detection methods often fail

to keep pace with the evolving tactics of cybercriminals. There is a critical need for a robust and

user-friendly solution to protect individuals from falling victim to these malicious sites. This

project aims to develop a Chrome extension that utilizes machine learning to detect and block

phishing websites in real-time, thereby enhancing user safety and maintaining trust in online

interactions.

**1.3** **Project Objective**

The objective of this project is to develop a comprehensive and efficient Chrome extension that

leverages machine learning algorithms to detect and block phishing websites in real-time. This

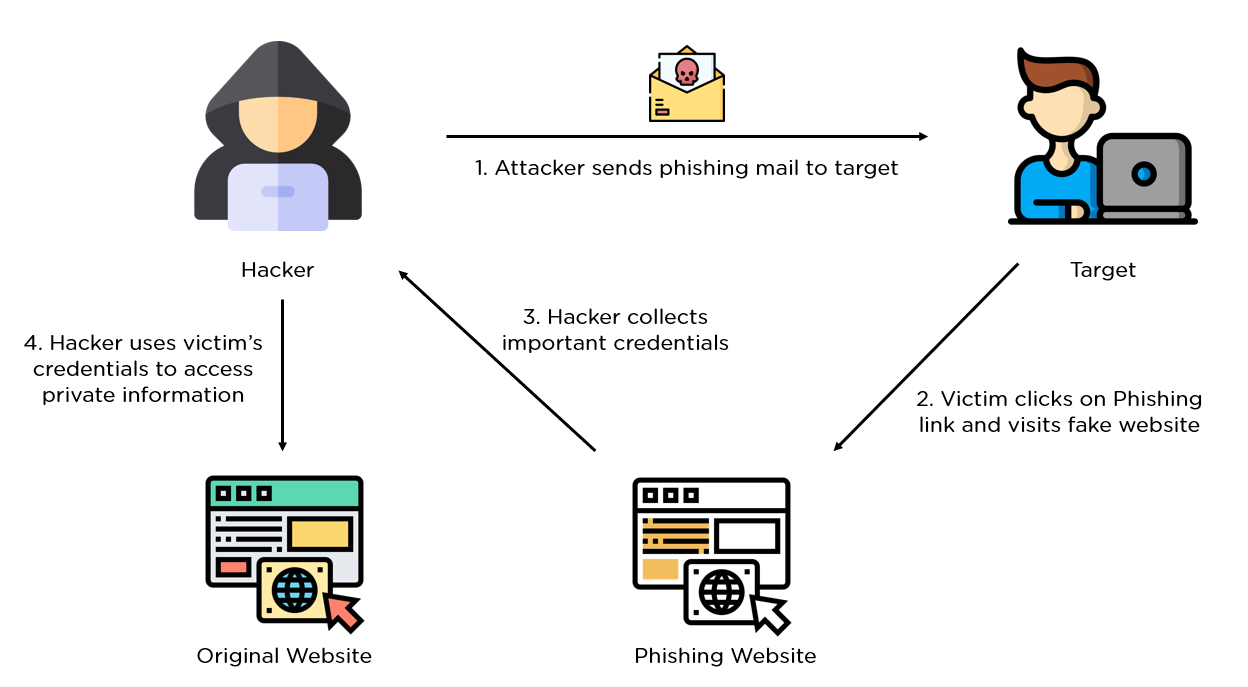
extension aims to provide a seamless and user-friendly experience, ensuring users are protected

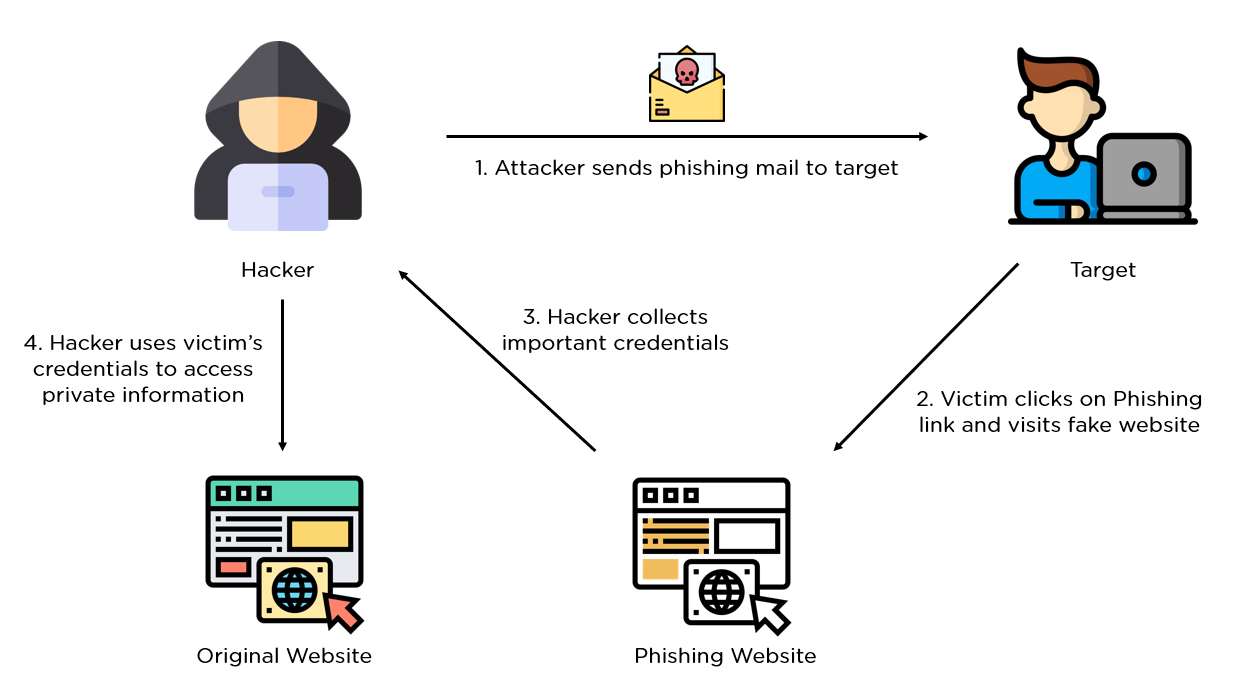
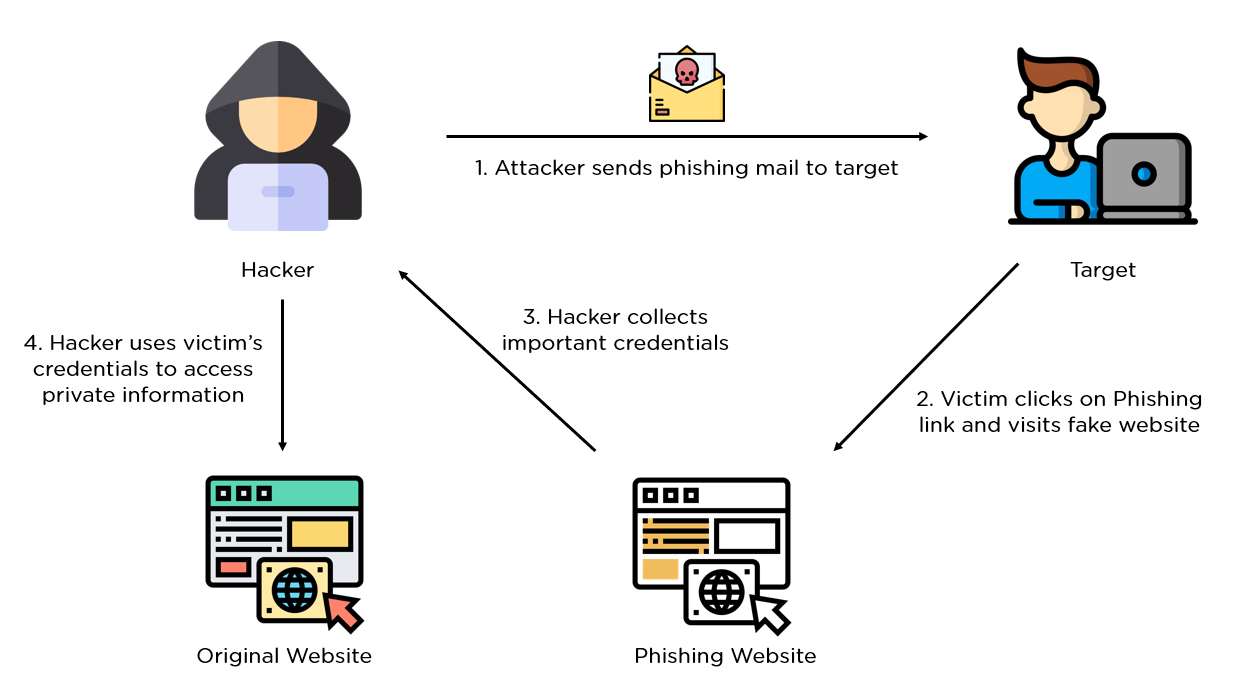
from malicious sites without disrupting their browsing activities. By integrating advanced

machine learning techniques, the extension will continuously learn and adapt to new phishing

tactics, offering robust and up-to-date protection. Ultimately, the goal is to enhance online

security, safeguard sensitive information, and foster a safer internet environment for all users.





Page **5** of **16**

**1.4** **Scope and Limitations of the Project**

1. **Project scope**

Machine Learning Integration: Implement advanced machine learning algorithms to

analyze URL patterns, website features, and other indicators to accurately distinguish

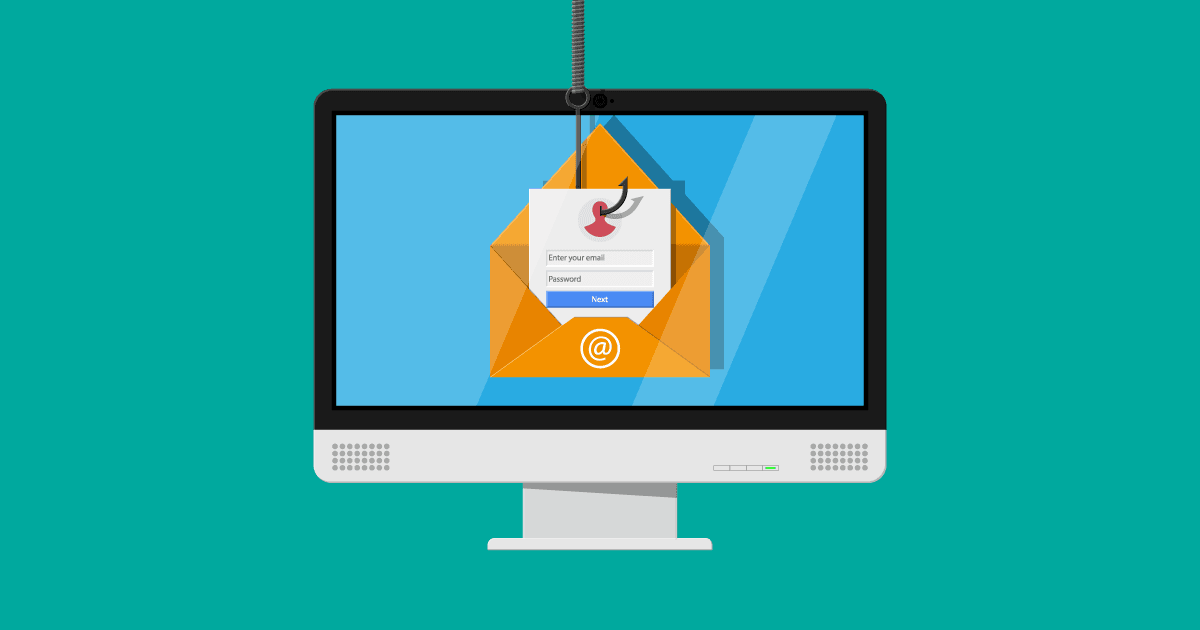
between legitimate and phishing sites, continuously improving detection accuracy

User-Friendly Interface and Alerts: Create an intuitive interface with clear notifications and

alerts to inform users when they encounter potential phishing websites, ensuring ease of use and

enhancing overall user experience.

1. **Limitations of the Project**
   * False positives/Negatives
   * Resource Intensive
   * Adaptation Lag
   * Browser Compatibility



Page **6** of **16**

**CHAPTER-2**

**Existing Solutions**

Page **7** of **16**

**2.Existing Solutions**

**2.1 Existing Solutions for Phishing Websites**

* Browser-Based Security Features: Modern web browsers, such as Google Chrome, Mozilla Firefox,

and Microsoft Edge, incorporate built-in phishing protection that warns users when they attempt to visit

known or suspected phishing sites. These features rely on regularly updated blacklists of malicious URLs

.

* Email Filtering and Spam Protection: Many email services, such as Gmail and Outlook, use

sophisticated algorithms and machine learning to filter out phishing emails before they reach users'

inboxes. These systems scan for common phishing indicators and suspicious content.

* Third-Party Security Software: Antivirus and internet security programs, like Norton, McAfee, and

Bitdefender, offer comprehensive protection against phishing. These solutions often include real-time

website scanning, URL filtering, and anti-phishing tools to safeguard users during their online activities.

* Web of Trust (WOT): Community-based tools like Web of Trust allow users to rate and review

websites based on their safety and reliability. These ratings help inform other users about the potential

risks of visiting specific sites.

* Two-Factor Authentication (2FA): Implementing 2FA adds an extra layer of security for users' online accounts, making it more difficult for phishers to gain unauthorized access even if they obtain login credentials through phishing attacks.
* Phishing Awareness and Education Programs: Many organizations conduct training sessions and

awareness programs to educate users about the dangers of phishing and how to recognize and avoid

phishing attempts. This approach aims to reduce the human factor in falling for phishing scams.

* Phishing Detection APIs: Services like Google Safe Browsing and PhishTank offer APIs that

developers can integrate into their applications to check URLs against a database of known phishing

sites. These APIs provide real-time protection by flagging suspicious websites as users navigate the

internet.

* Content Filtering and Network Security Appliances: Devices such as firewalls and Secure Web

Gateways (SWGs) are employed by organizations to filter web traffic and block access to known

malicious sites. These appliances analyze web content and URL patterns to detect and prevent users

from visiting phishing websites within a corporate network.

Page **8** of **16**

**CHAPTER-3**

**Proposed Solution**

Page **9** of **16**

**3.Proposed Solutions**

**3.1 Proposed Solutions For Phishing Websites**

* Machine Learning-Based Chrome Extension: Develop a Chrome extension that uses advanced

machine learning algorithms to detect and block phishing websites in real-time. The extension will

analyze various features of URLs and web content, such as domain age, URL length, and the presence of

suspicious patterns, to accurately identify phishing attempts.

The Random Forest algorithm plays a crucial role in the detection of phishing websites by

leveraging its robust ensemble learning capabilities to improve classification accuracy and

reduce overfitting. Here's how Random Forest contributes to phishing detection:

1. Feature Analysis:Random Forest analyzes a multitude of features extracted from URLs and

web pages, such as URL length, domain age, presence of special characters, number of

subdomains, and more. Each tree in the forest evaluates a different subset of features, ensuring a

comprehensive analysis.

2. Ensemble Learning:By combining the predictions of multiple decision trees, Random

Forest reduces the likelihood of errors and enhances the overall predictive accuracy. This

ensemble approach mitigates the risk of overfitting, a common issue in individual decision trees.

3.Handling Imbalanced Data: Phishing datasets often contain a higher number of legitimate

websites compared to phishing sites. Random Forest is effective in handling such imbalanced

datasets, as it can provide balanced error rates by adjusting the weights of the classes during

training.

4. Robustness to Noise:Random Forest is less sensitive to noisy data and outliers, making it

well-suited for real-world scenarios where phishing websites may employ various obfuscation

techniques to evade detection.

5. Feature Importance:The algorithm provides insights into the importance of different

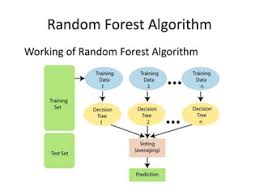
features in the classification process. This helps in understanding which characteristics are most

indicative of phishing attempts, guiding the development of more effective detection strategies.

6. Real-Time Detection: With its fast prediction capabilities, Random Forest can be integrated

into a browser extension to provide real-time phishing detection, analyzing URLs as users

navigate the web and promptly identifying potential threats.



Page **10** of **16**

**CHAPTER-4**

**System Requirements and Tools**

Page **11** of **16**

1. **System requirements for Phishing Websites**

**4.1 Hardware Components**

• **Processor** **AMD Ryzen 9 7940HS w/ Radeon 780M Graphics** **4.00 GHz**

* **Installed RAM 64.0 GB (59.9 GB usable)**

• **System type** **64-bit operating system, x64-based processor**

**4.2 Software Tools and Versions Used :**

• **Operating System:** Windows, macOS or Linux.

• **Browser:** Google Chrome (latest version: Version 90 or higher).

• **Extension Framework:** Chrome Extension SDK (manifest v3).

• **Database:** Local storage (browser storage) for phishing site data (for .csv).

• **Programming Language:** HTML, CSS, JavaScript, python.

**Libraries:**

**Python Libraries**

pandas

skit-learn

joblib

flask

numpy

**4.3 Network and Other Dependencies :**

* **Internet connection :** Required for

1. Downloading necessary software and libraries o Receiving Updates



Page **12** of **16**

**CHAPTER-5**

**Results and Discussion**

Page **13** of **16**

**5.Results and Discussion**

**5.1 Results of Phishing website Chrome Extension**

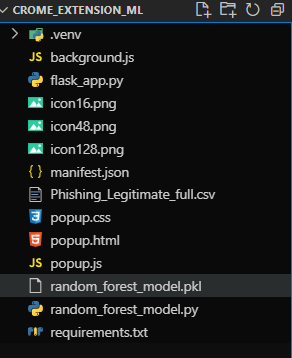


Fig1.1 Make sure all the files in the same directory

Screenshot 2024-07-27 223334

Fig1.2 open chrome and go to chrome://extensions

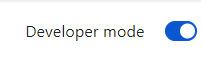


Fig1.3 Enable the Developer mode

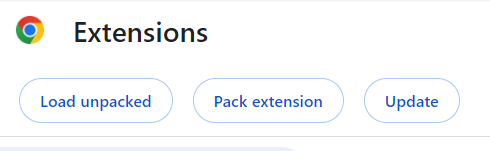


Fig1.4 click on Load unpacked

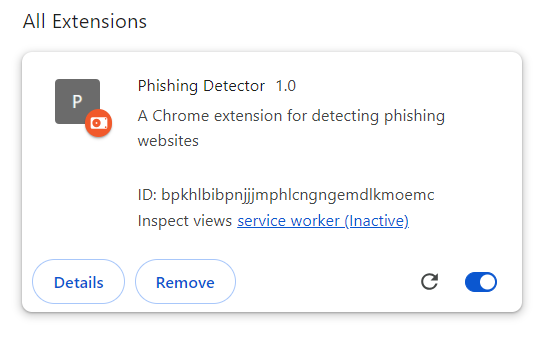


Fig1.5 Extension added to the chrome

Page **14** of **16**

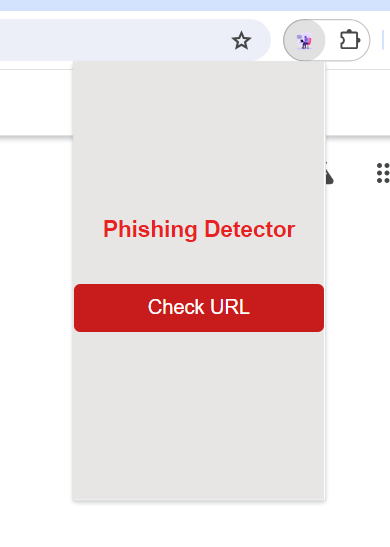


Fig1.6 Go to any website and click on the extension and then click on the Check URL

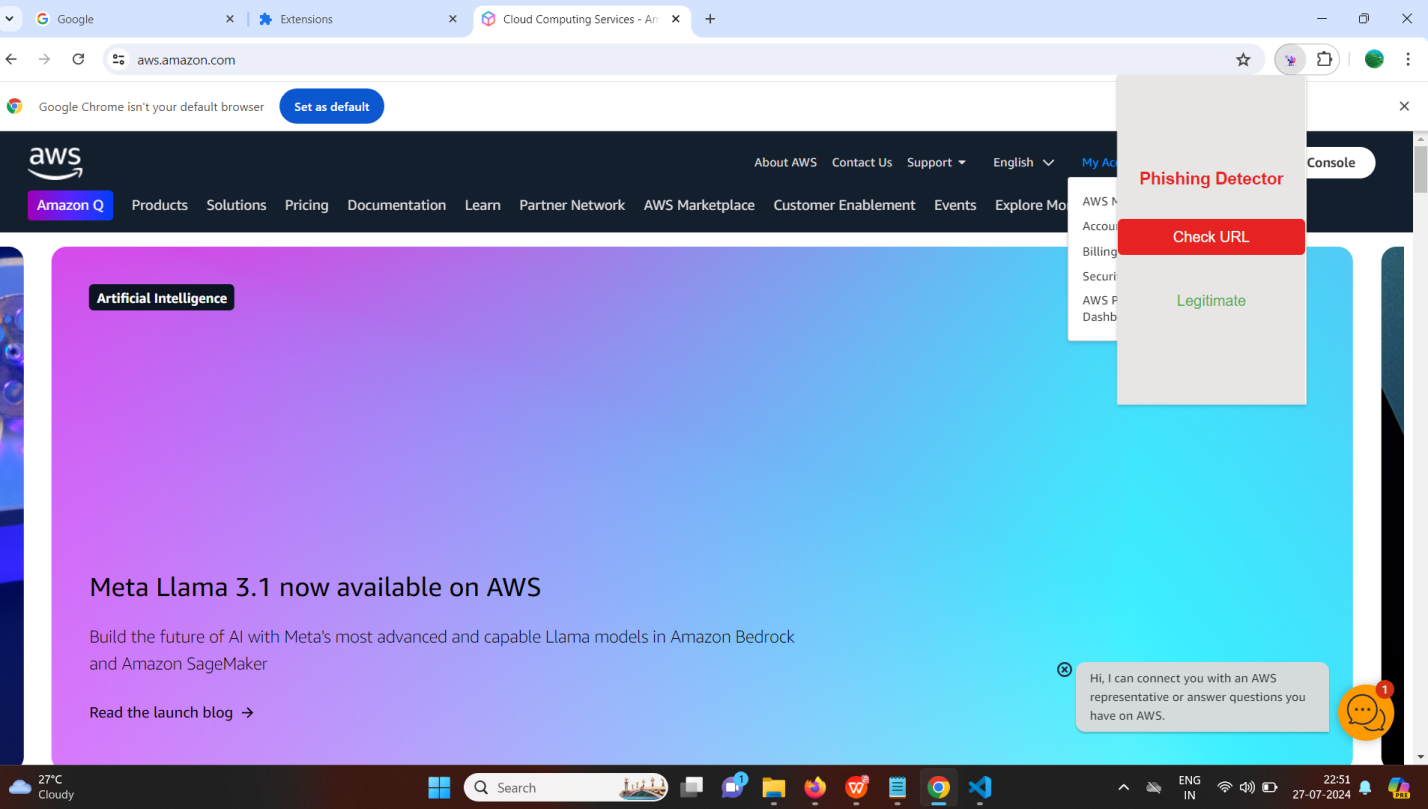


Fig1.7 Make sure the flask server is running, The output is generated successfully :), sometimes it may take more time to give response Be patience.

Page **15** of **16**

**Conclusion**

The implementation of the Random Forest algorithm in the Chrome extension for detecting phishing

websites has proven to be a highly effective solution for enhancing online security. By leveraging the

ensemble learning capabilities of Random Forest, the extension delivers robust and accurate phishing

detection by analyzing a diverse set of features associated with URLs and web content. The algorithm’s

ability to handle imbalanced data, manage noisy inputs, and provide feature importance insights has

contributed to its effectiveness in distinguishing between legitimate and malicious sites.

The real-time detection capability of the extension ensures that users are promptly alerted to potential

phishing threats, reducing the risk of falling victim to online scams. Furthermore, the adaptability of

Random Forest allows the extension to evolve with emerging phishing tactics, maintaining its relevance and effectiveness over time. Overall, this project demonstrates the power of machine learning in cybersecurity and provides a practical tool for safeguarding users against phishing attacks, contributing to a safer and more secure browsing experience.

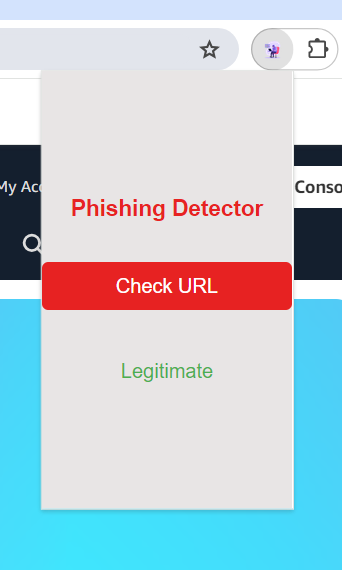
The successful integration of Random Forest into the phishing detection Chrome extension highlights the

potential of machine learning in addressing contemporary cybersecurity challenges. The algorithm’s

ensemble approach not only enhances the accuracy of phishing site detection but also ensures resilience

against evolving threats. Through continuous model updates and real-time analysis, the extension offers

users a dynamic and responsive defense mechanism. This project underscores the importance of combining advanced computational techniques with practical applications to tackle pressing security issues, ultimately reinforcing user trust and safety in the digital landscape.



Output: Legitimate

The Website is safe :)

Page **16** of **16**