iNeuron

Phishing Domain Detection

(Machine Learning)

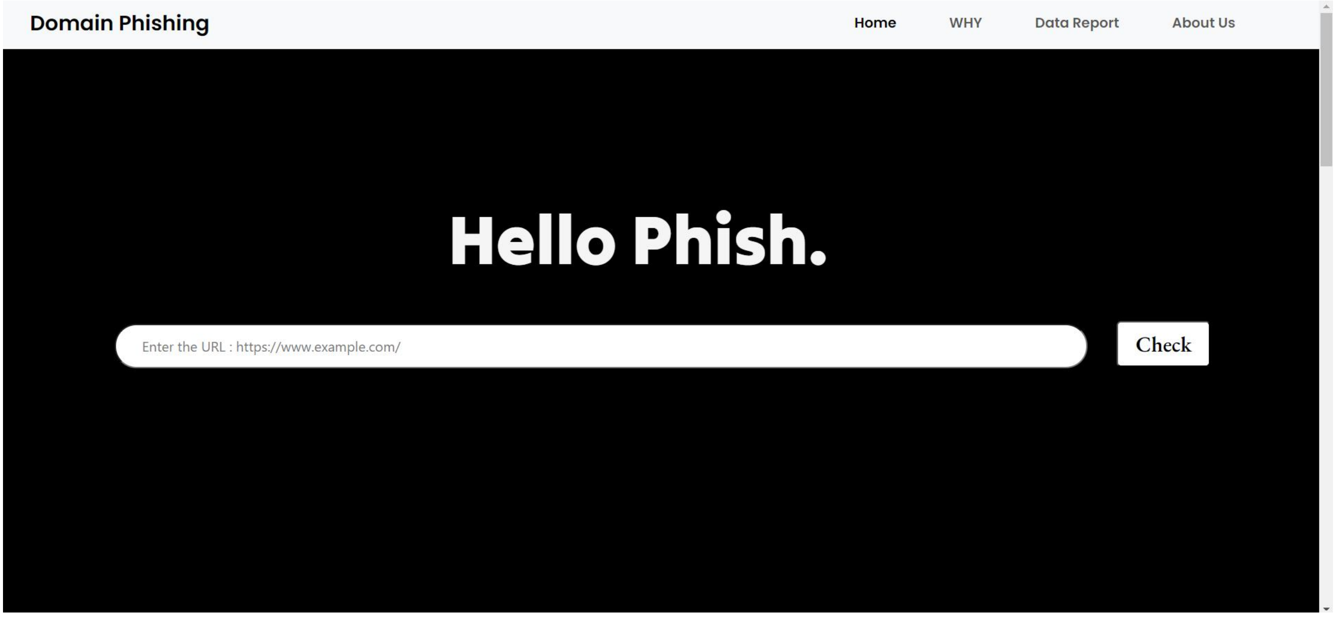
Wireframe Document

Project Members:

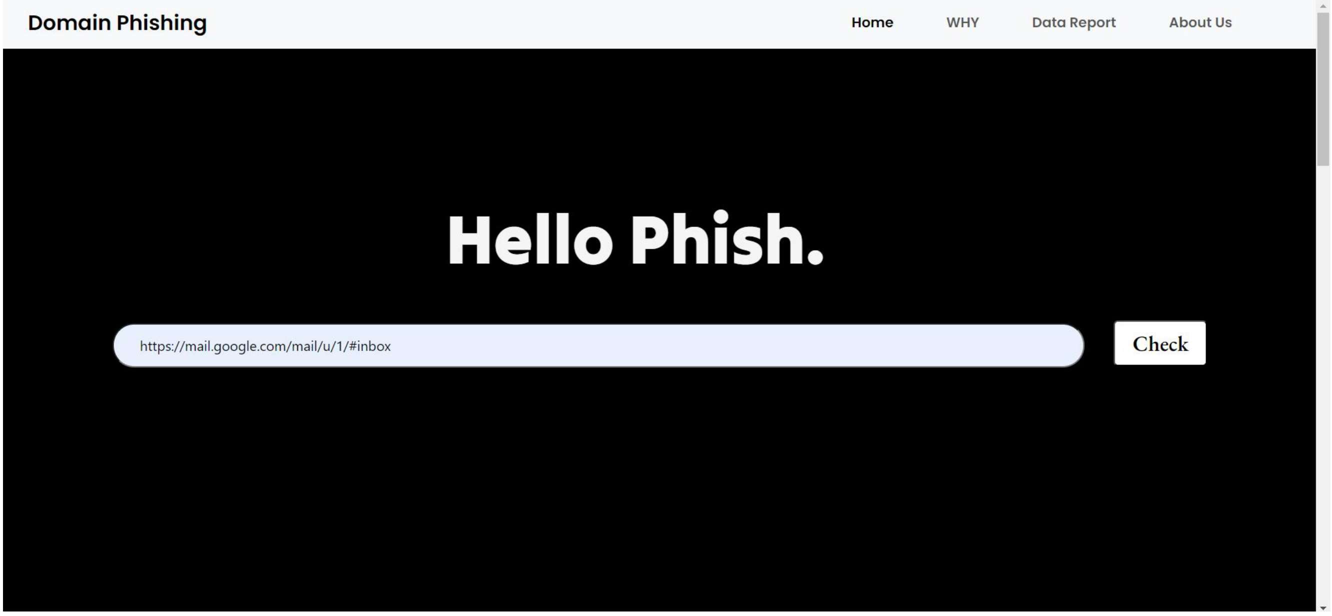
1. Rishabh
2. Shivansh Srivastava
3. Ashish Diwakar

INeuron.ai

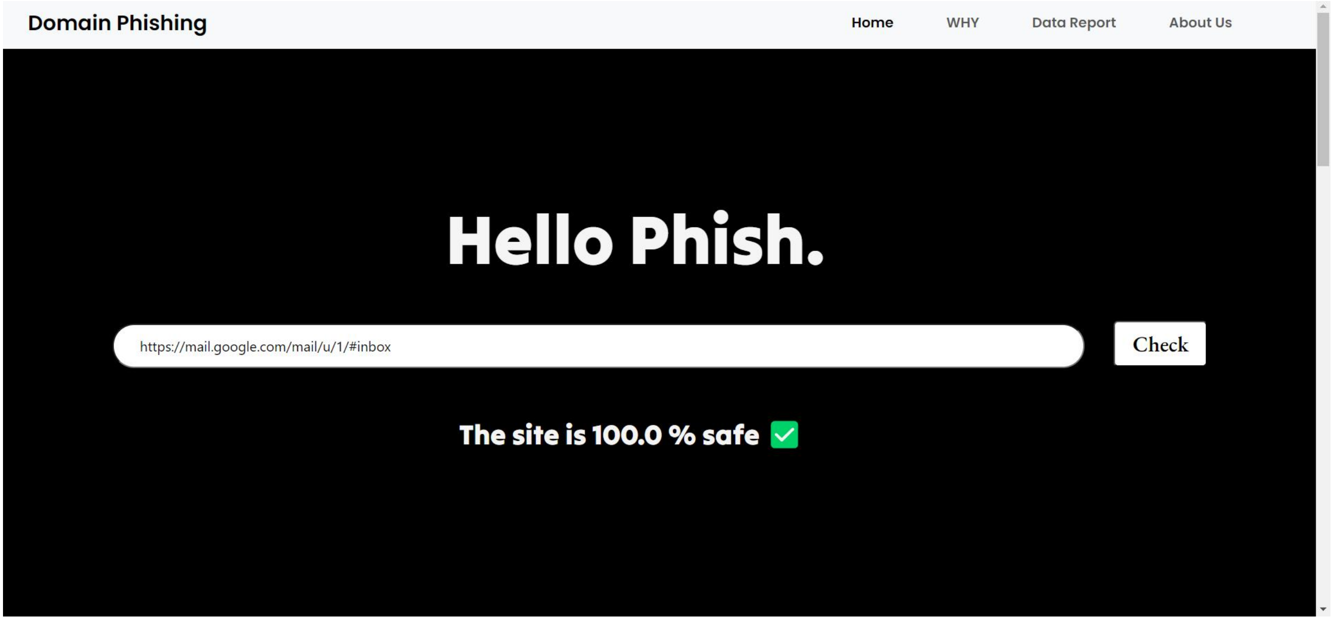
- This is the First page which will be shown on the window,in which 4 option will be shown such as Home, WHY, Data Report and About Us



- The user can enter the link in the search bar as shown below.



- This is the result displayed.

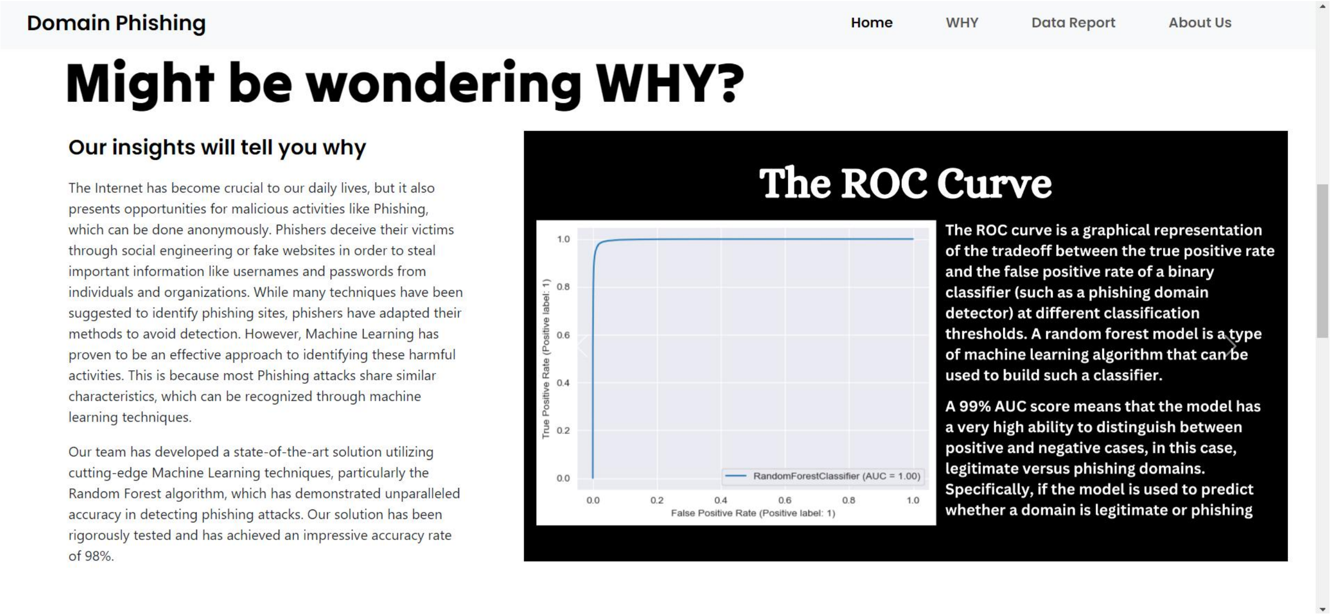


Hello Phish.

https://mail.google.com/mail/u/1/#inbox

The site is 100.0 % safe ✓

- Site Objective Report



Domain Phishing
Home
WHY
Data Report
About Us
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate (Positive label: 1)
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate (Positive label: 1)
RandomForestClassifier (AUC = 1.00)

Might be wondering WHY?

Our insights will tell you why

The Internet has become crucial to our daily lives, but it also presents opportunities for malicious activities like Phishing, which can be done anonymously. Phishers deceive their victims through social engineering or fake websites in order to steal important information like usernames and passwords from individuals and organizations. While many techniques have been suggested to identify phishing sites, phishers have adapted their methods to avoid detection. However, Machine Learning has proven to be an effective approach to identifying these harmful activities. This is because most Phishing attacks share similar characteristics, which can be recognized through machine learning techniques.

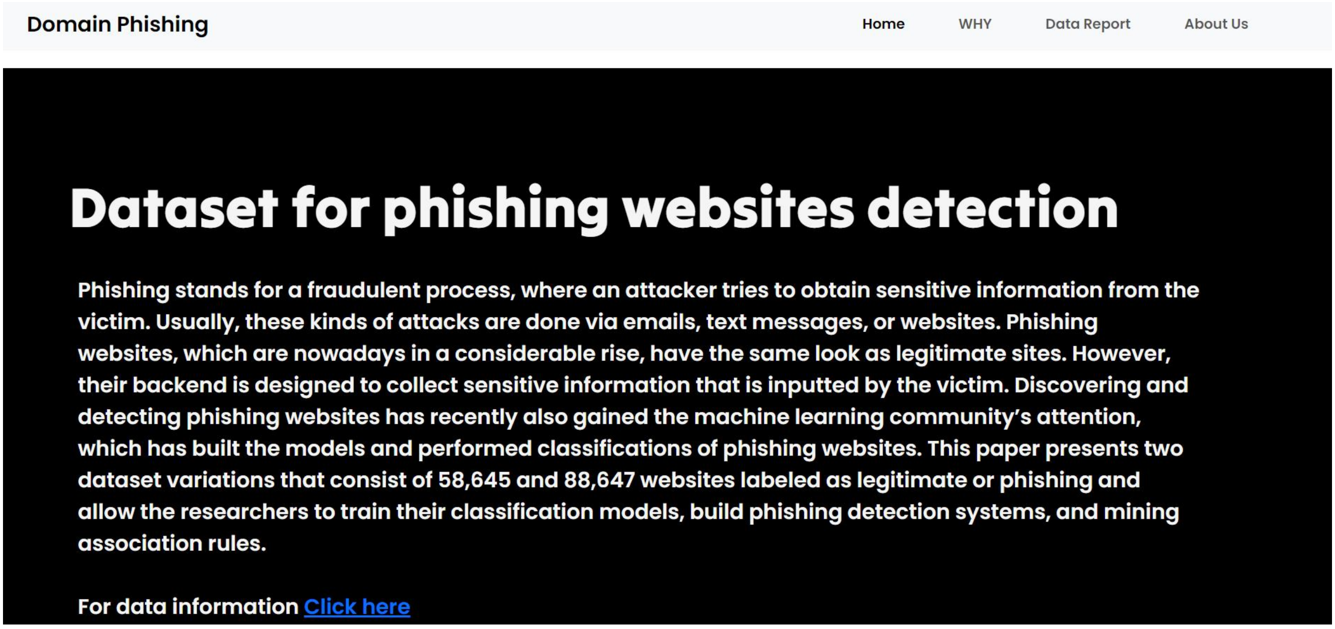
Our team has developed a state-of-the-art solution utilizing cutting-edge Machine Learning techniques, particularly the Random Forest algorithm, which has demonstrated unparalleled accuracy in detecting phishing attacks. Our solution has been rigorously tested and has achieved an impressive accuracy rate of 98%.

The ROC Curve

The ROC curve is a graphical representation of the tradeoff between the true positive rate and the false positive rate of a binary classifier (such as a phishing domain detector) at different classification thresholds. A random forest model is a type of machine learning algorithm that can be used to build such a classifier.

A 99% AUC score means that the model has a very high ability to distinguish between positive and negative cases, in this case, legitimate versus phishing domains. Specifically, if the model is used to predict whether a domain is legitimate or phishing

- Data Report Page



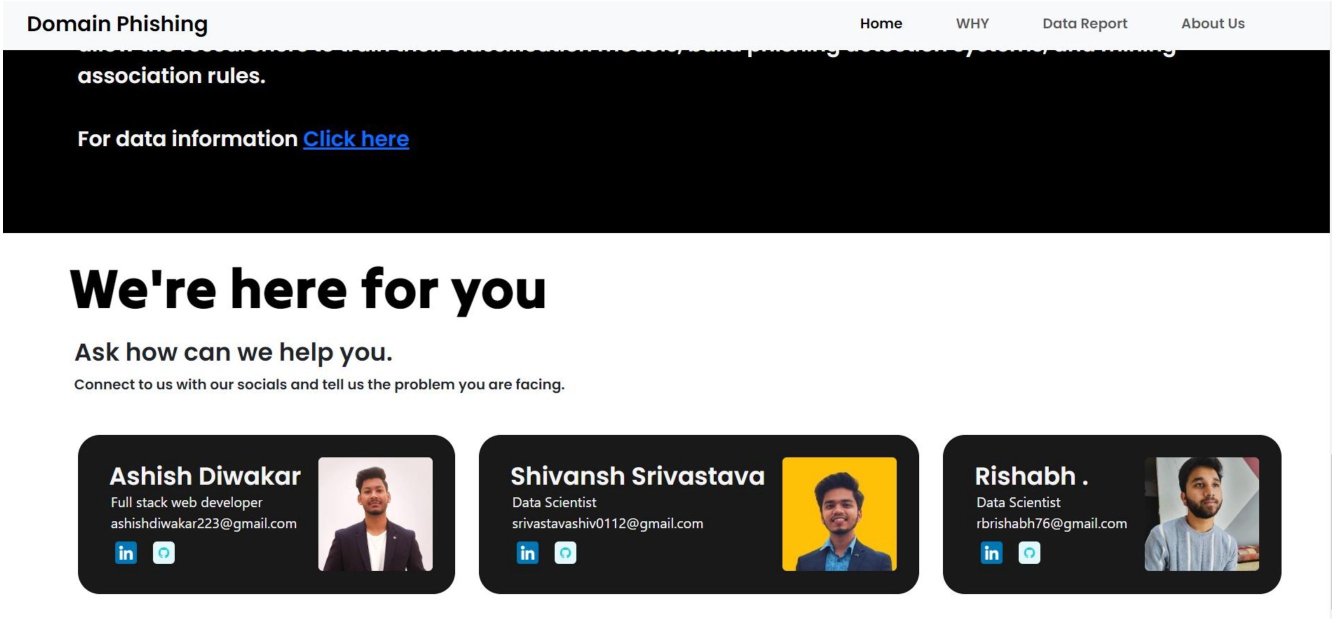
Domain Phishing

Dataset for phishing websites detection

Phishing stands for a fraudulent process, where an attacker tries to obtain sensitive information from the victim. Usually, these kinds of attacks are done via emails, text messages, or websites. Phishing websites, which are nowadays in a considerable rise, have the same look as legitimate sites. However, their backend is designed to collect sensitive information that is inputted by the victim. Discovering and detecting phishing websites has recently also gained the machine learning community's attention, which has built the models and performed classifications of phishing websites. This paper presents two dataset variations that consist of 58,645 and 88,647 websites labeled as legitimate or phishing and allow the researchers to train their classification models, build phishing detection systems, and mining association rules.

For data information Click here

- About Us



Domain Phishing

association rules.

For data information Click here

We're here for you

Ask how can we help you.

Connect to us with our socials and tell us the problem you are facing.

Ashish Diwakar
Full stack web developer
ashishdiwakar223@gmail.com

Shivansh Srivastava
Data Scientist
srivastavashiv0112@gmail.com

Rishabh .
Data Scientist
rbrishabh76@gmail.com

THANK YOU