A MACHINE LEARNING-BASED FRAMEWORK FOR PHISHING DETECTION AND PREVENTION

Ibrahim Maalej, Anthony Polak, Emanuel Ruiz

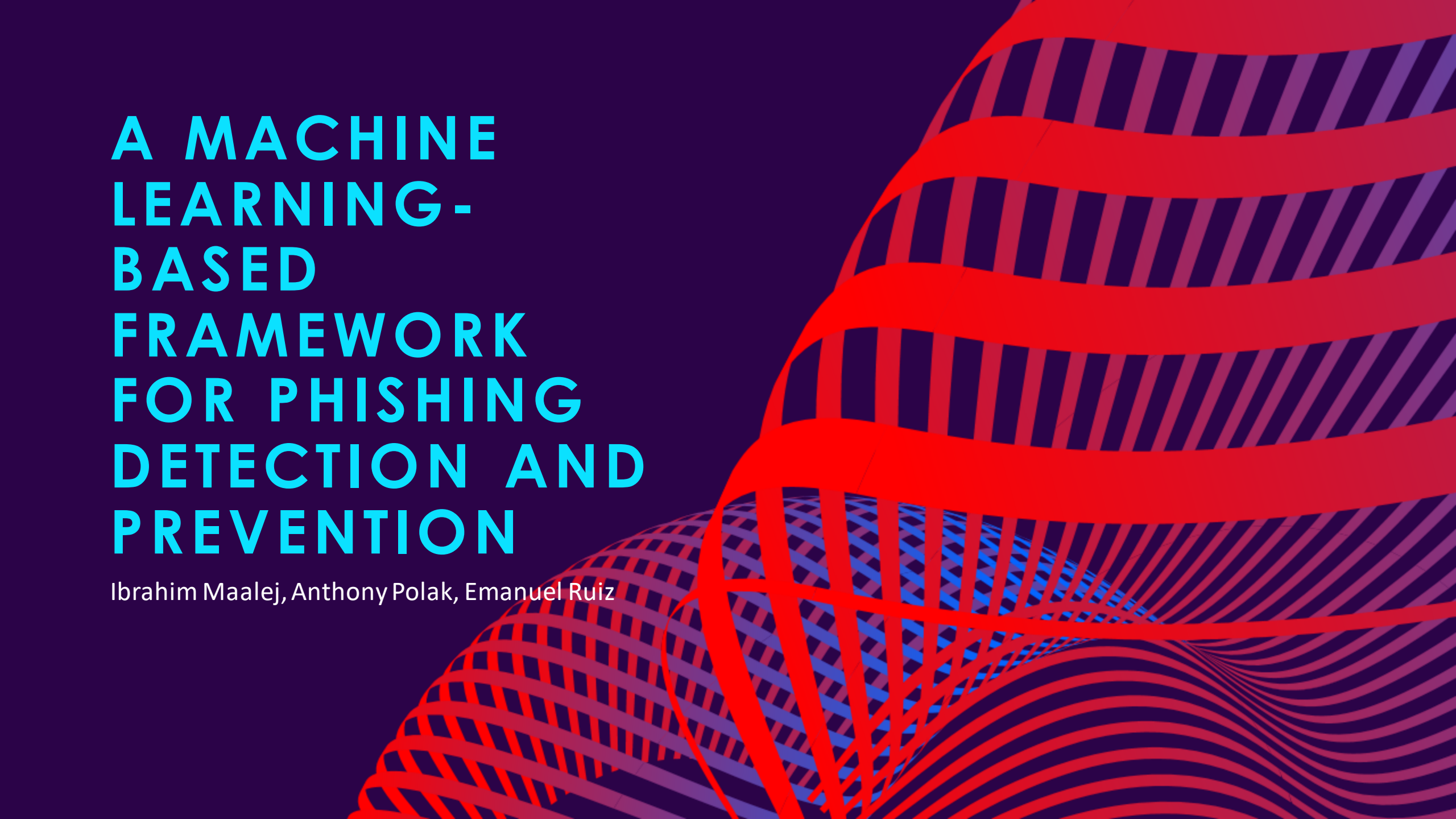
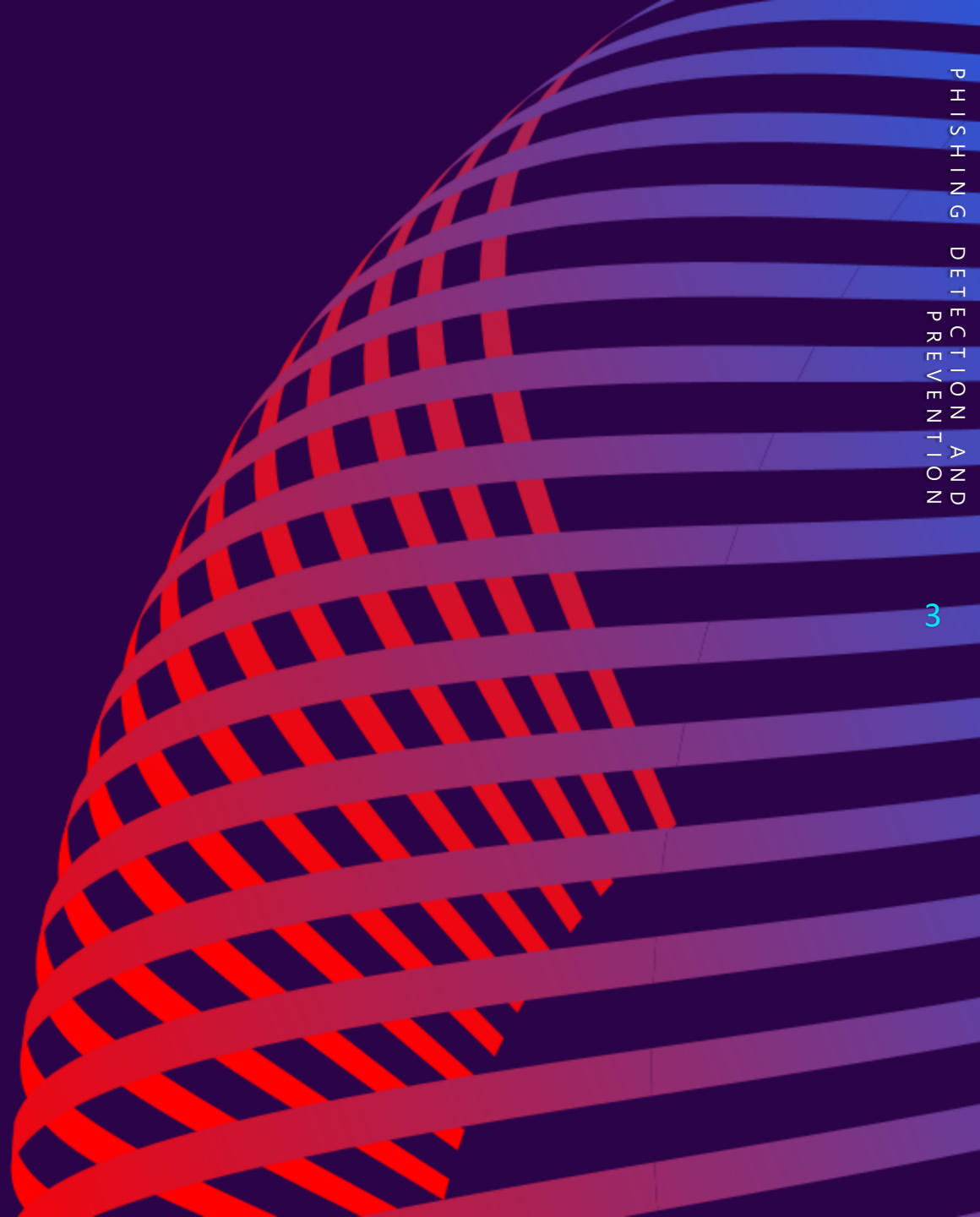


Table of contents

INTRODUCTION



WHAT IS MACHINE LEARNING?

Definition:

- A branch of artificial intelligence focused on enabling computers to learn from data.

Purpose:

- Automate decision-making and predictions based on patterns in data.

Types:

- Supervised, unsupervised, and reinforcement learning.

WHAT IS PHISHING?

Definition:

- A type of cyber-attack where attackers attempt to deceive users into providing sensitive information.

Methods:

- Often involves fraudulent emails, messages, or websites.

Impact:

- Can lead to identity theft, financial loss, or unauthorized access to data.

DANGERS OF PHISHING

22%

FBI'S 2021 INTERNET CRIME REPORT SHOWS PHISHING ATTACKS ACCOUNTED FOR 22% OF ALL DATA BREACHES

$44.2 M

$44.2 MILLION WAS STOLEN BY CYBER CRIMINALS THROUGH PHISHING ATTACKS IN 2021. (AAG IT)

97%

INTEL STUDY SHOWS 97% OF PEOPLE FAIL AT IDENTIFYING PHISHING EMAILS FROM GENUINE EMAILS

OUR SOLUTION

The overarching problem addressed by this project is the persistent vulnerability of individuals and organizations to phishing threats.

We set out to design a sophisticated AI model-based tool aimed at enhancing cybersecurity and protecting users through the detection of link-related phishing attacks in web content

METHODOLOGY

Research the problem

Understand the intricacies of phishing attacks, their indicators, and machine learning models

Find the data

Gather comprehensive datasets containing benign and phishing links, ensuring that the data is diverse and up-to-date.

Build the model

Utilize machine learning algorithms to construct a classifier that can accurately distinguish between safe and phishing links.

Develop the tool

Integrate the model into a user-friendly application, providing real-time phishing detection to assist users in staying secure online.

DATASET

DATASET ACQUISITION AND PREPARATION

- FINAL DATASETS BUILT UPON:
  - GREGA VRBANČIČ'S HTTPS://DOI.ORG/10.1016/J.DIB.2020.106438
  - ABDELHAKIM HANNOUSSE AND SALIMA YAHIOUCHE'S HTTPS://DOI.ORG/10.17632/C2GW7FY2J4.3
- DATASETS ARE CLEANED, TRANSFORMED, ENRICHED, MERGED, AND TRIMMED USING PANDAS.
- 65.36% NON-PHISHING LINKS, 34.64% PHISHING LINKS
- SHAPE: (40,000, 116)
  - 40,000 INSTANCES
  - 116 FEATURES

URL FEATURIZATION

- Length of the URL and its sub-components
- Quantity of special characters and their ratios in different parts of the URL
- IP address as domain
- Redirections
- Suspicious keywords within sub-components
- HTTPS as the protocol or a part of another section of the URL

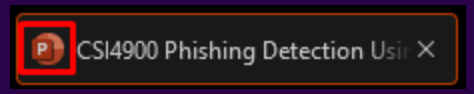
DOMAIN FEATURIZATION

- Indexed by Google
- Time response
- TLS/SSL Certification
- Page Rank (Tranco)
- DNS Record
- Domain Age
- Quantity of Resolved IPs
- Quantity of servers
  - MX servers
  - Name servers

ROOM FOR IMPROVEMENT

HTML/JAVASCRIPT ELEMENTS FROM URL PAGES

- Existence of Favicon
- Javascript Obfuscation
- Extraction of HTML tags from the page
  - <script>, <meta>, <link>, <a>



(A)

```
function setText(data) {
document.getElementById("myDiv").innerHTML = data;
}
```

(B)

```
function ghds3x(n) {
h = "\x69\u006En\u0065r\x48T\u004DL";
a="s c v o v d h e . n i";x=a.split(" ");b="gztxleWentBsyf";
r=b.replace("z",x[7]).replace("x","E").replace("s","").replace("f","I")
["repl" + "ace"]("W","m")+"d";
c="my"+String.fromCharCode(68)+x[10]+"v";
s=x[5]+x[3]+x[1]+"um"+x[7]+x[9]+"t";d=this[s][r](c);if(+!![])
{ d[h]=n; } else { d[h]=c; } }
```

MODEL

MODEL DEVELOPMENT

Research

Researched many potential methods to handle the phishing classification problem

Model Selection

Numerous models pitted against each other to find best performance

Featurization & Analysis

Features chosen and analyzed for maximum impact and coverage

Final Implementation

Implement & combine everything together to construct model ensuring reliable and efficient detection

RESEARCH

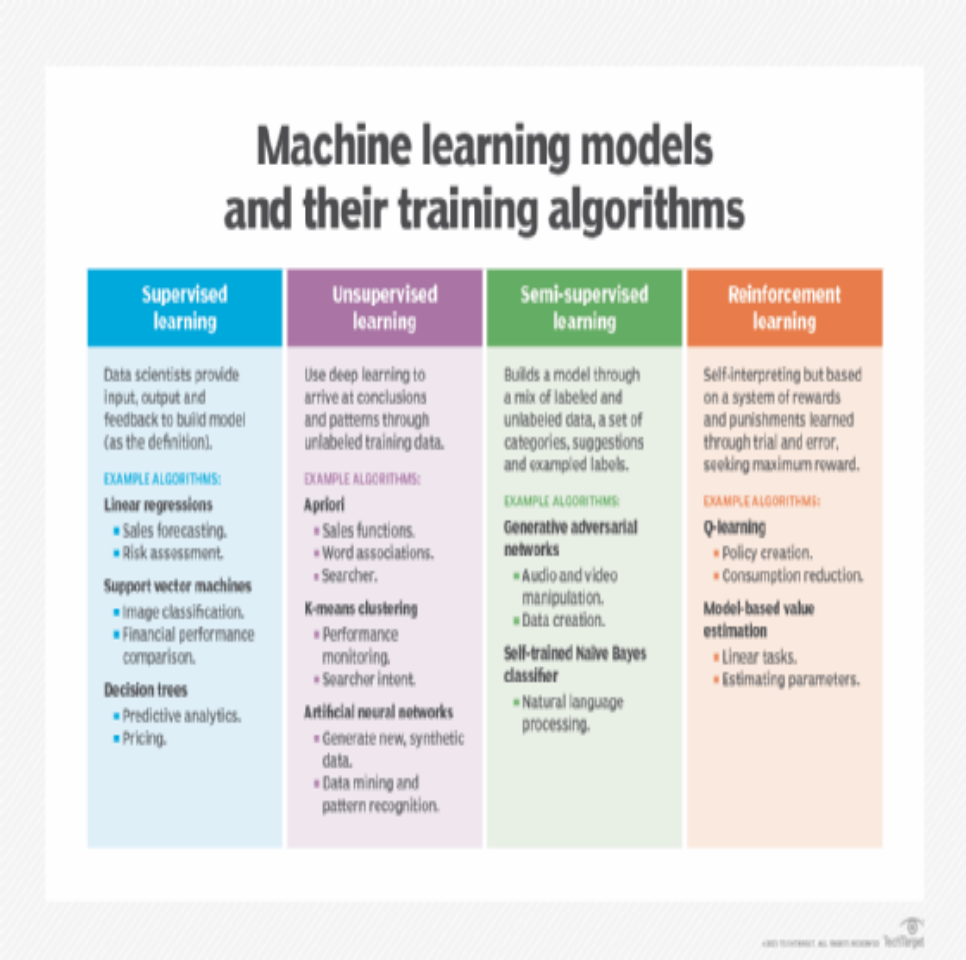


Fig 1. The four main types of machine learning and their most common algorithms.

Languages and Packages:

- Many languages and Packages to choose from
- Python using scikit-learn as main foundation found to be most suitable

Feature Types:

- Various possible types such as text-based, URL-based, HTML-Based, signature-based, etc.
- Chose URL-based with some signature-based mixed in

Classification Methods:

- Numerous methods such as supervised, unsupervised, etc.
- A Supervised method chosen to compliment URL-based features

Phishing Indicators

- Many types of attacks
- What signals a possible URL-based attack?
- Common metrics like checking google page rank or URL shortening services

MODEL SELECTION

	PRECISION	RECALL	F1-SCORE	TEST SCORE	TRAIN SCORE
Random Forest	0 | 0.97 1 | 0.94	0 | 0.97 1 | 0.95	0 | 0.97 1 | 0.94	0.961	0.999
Light GBM	0 | 0.97 1 | 0.94	0 | 0.97 1 | 0.95	0 | 0.97 1 | 0.94	0.960	0.977
Logistic Regression	0 | 0.95 1 | 0.87	0 | 0.93 1 | 0.91	0 | 0.94 1 | 0.89	0.922	0.924
Decision Tree	0 | 0.95 1 | 0.92	0 | 0.96 1 | 0.91	0 | 0.96 1 | 0.92	0.944	1
Naïve Bayes	0 | 0.92 1 | 0.80	0 | 0.89 1 | 0.86	0 | 0.91 1 | 0.83	0.879	0.879
Support Vector Machine	0 | 0.70 1 | 0.81	0 | 0.98 1 | 0.20	0 | 0.81 1 | 0.31	0.707	0.707
Multilayer Perceptron	0 | 0.83 1 | 0.95	0 | 0.98 1 | 0.63	0 | 0.90 1 | 0.75	0.859	0.862
BEST	**Random Forest**	**Random Forest**	**Random Forest**	**Random Forest**	**Decision Tree**

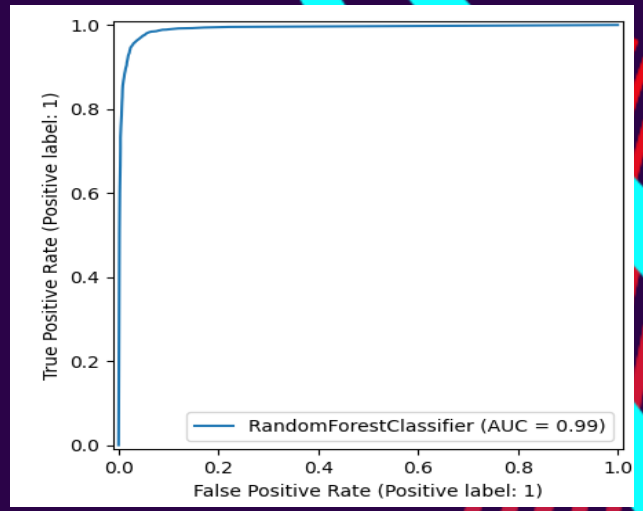


Fig 2. Random Forest AUC/ROC Display

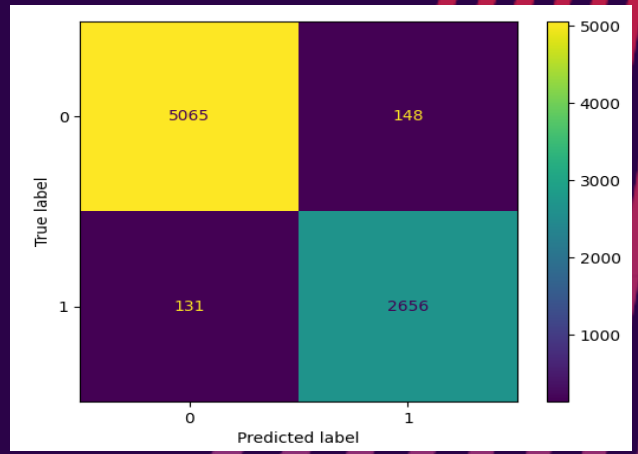


Fig 3. Random Forest Confusion Matrix

FEATURE ANALYSIS

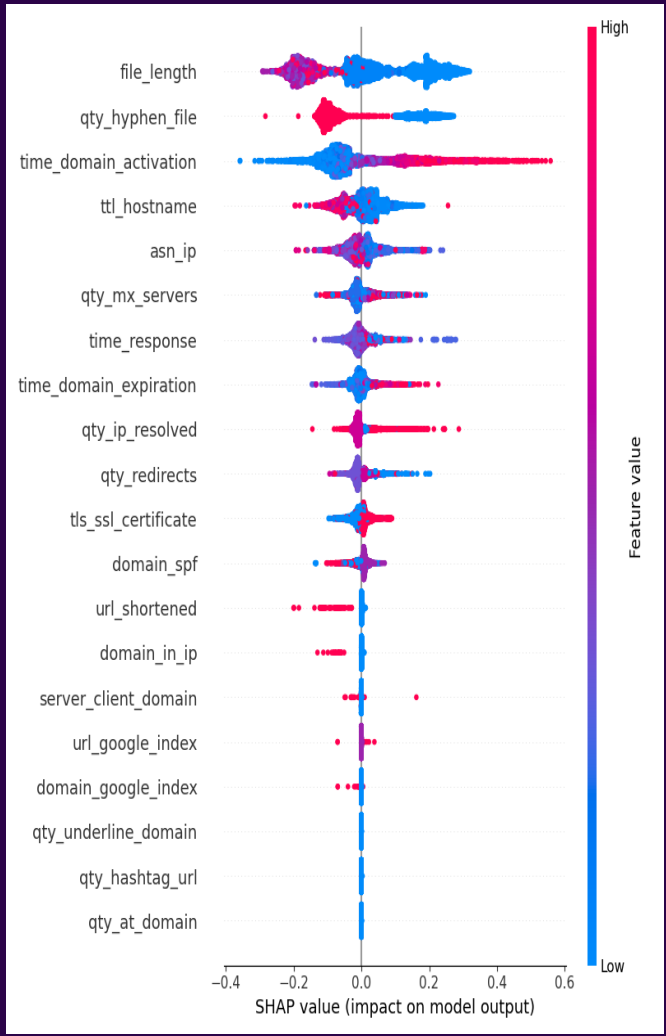


Fig 4. SHAP Summary Graph

Variance Inflation Factor (VIF):

- Assess multicollinearity among features to ensure that each feature provides unique information to the model.

SHAP:

- values to understand each feature's contribution to the model's predictions and their impact on classification outcomes.

Permutation Importance:

- Measure the importance of each feature by observing the change in model performance when the feature's values are randomly permuted.

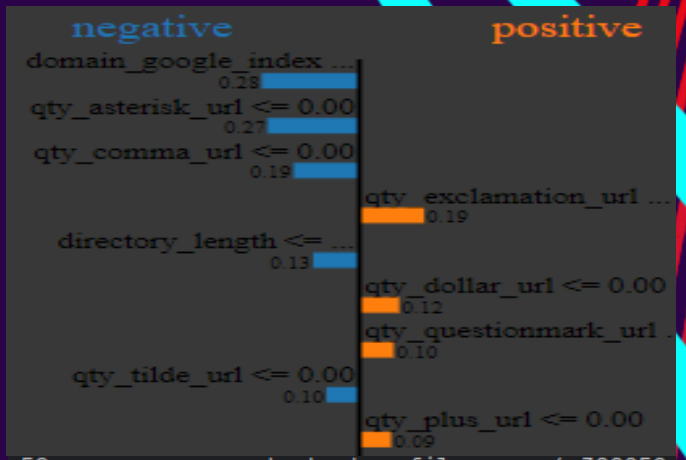
LIME

- provide interpretable insights into the model's predictions for specific instances by approximating the model with a simpler, interpretable model.

directory_length	0.180 +/- 0.003
time_domain_activation	0.062 +/- 0.002
length_url	0.016 +/- 0.001
qty_dot_domain	0.010 +/- 0.001
ttl_hostname	0.007 +/- 0.001
qty_nameservers	0.006 +/- 0.001
asn_ip	0.006 +/- 0.001
time_response	0.005 +/- 0.001
qty_mx_servers	0.003 +/- 0.000
qty_ip_resolved	0.003 +/- 0.000
time_domain_expiration	0.003 +/- 0.001
domain_spf	0.002 +/- 0.000

Fig 5. Permutation Importance

Fig 7. LIME Feature Insights



59	qty_hyphen_file	6.389052
103	qty_nameservers	5.676059
75	file_length	4.568873
100	time_domain_activation	3.396401
102	qty_ip_resolved	3.006201
106	tls_ssl_certificate	2.424170
104	qty_mx_servers	2.304309
109	domain_google_index	1.661424
108	url_google_index	1.658722
101	time_domain_expiration	1.648173
99	asn_ip	1.516875
107	qty_redirects	1.422552
105	ttl_hostname	1.374359
97	time_response	1.338245
38	domain_in_ip	1.155511
110	url_shortened	1.128012
39	server_client_domain	1.028662
98	domain_spf	1.022042

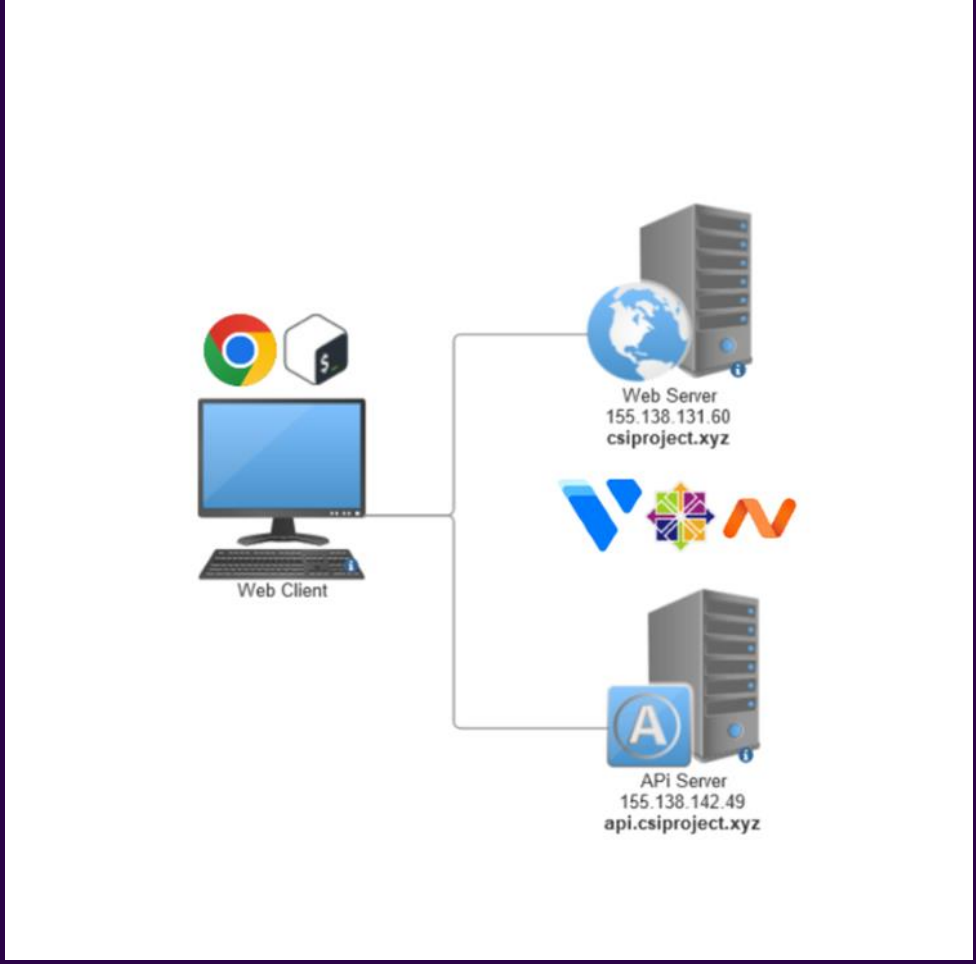
Fig 6. VIF Inflation Factor

FINAL IMPLEMENTATION

- IMPLEMENTED USING PYTHON, INCLUDING LIBRARIES AND PACKAGES SUCH AS SCIKIT-LEARN, PANDAS, ETC.
- 900+ LINES OF CODE WITH FEATURE IMPLEMENTATION
- TRAINED ON FINAL PROCESSED QUALITY DATASET
- 115 FEATURES, 1 TARGET
- HYPERPARAMETER TUNING FOR BEST RESULTS
- 10-FOLD CROSS VALIDATED, 0.91 MATTHEWS COEFFICIENT
- 96% ACCURACY, F1-SCORE, PRECISION, RECALL ON AVERAGE
- STRONG, RELIABLE AND EFFICIENT MODEL FOR DETECTION

WEB INFRASTRUCTURE

OUR SERVERS



VULTR Provides:

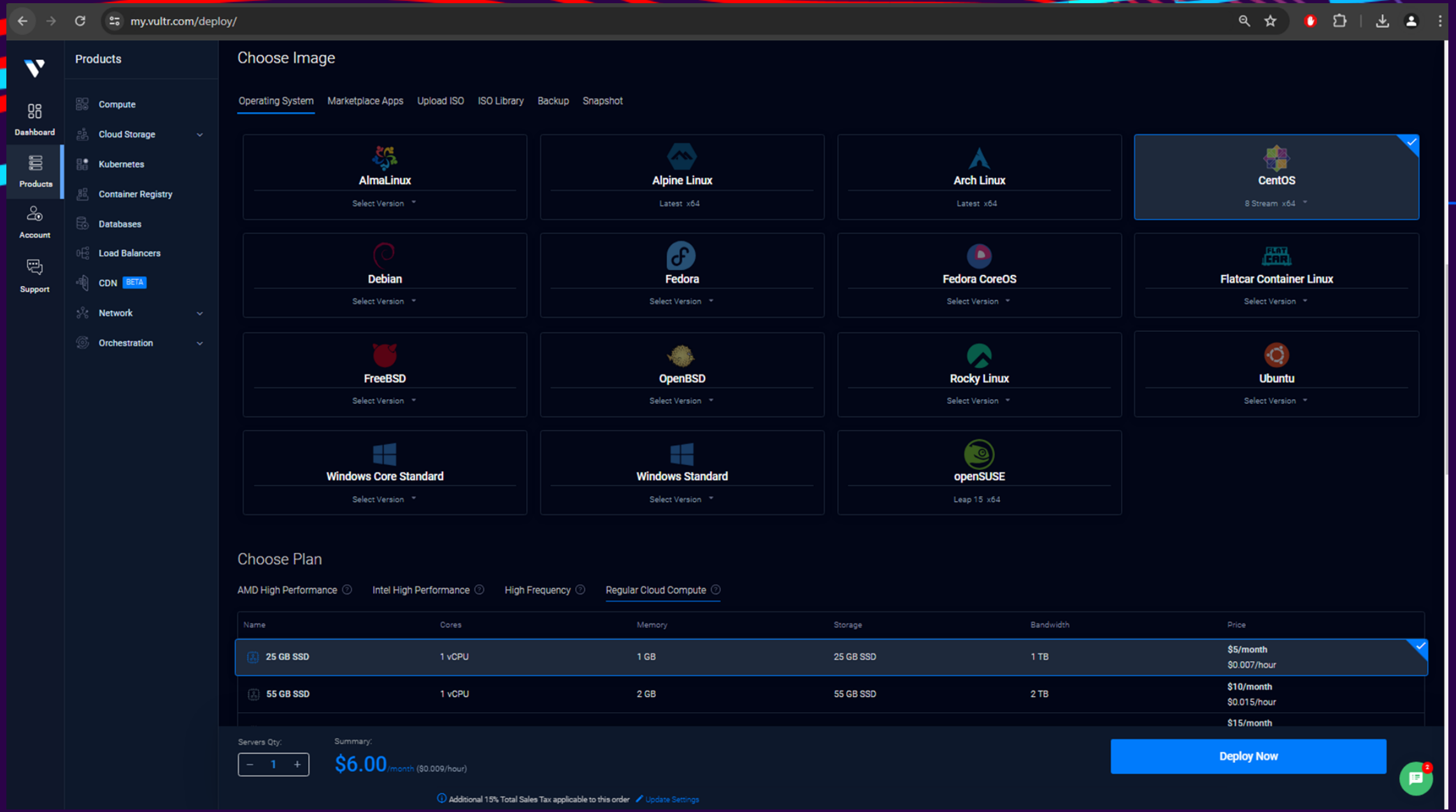
- Hardware Resources for computation
  Toronto DC - Regular Cloud Compute 1 vCPU, 1024 MB RAM, 25 GB SSD, 1.00 TB Transfer (5$/mo)
- DNS (supporting SLD)
- Firewall
- Reserved IPs
- Snapshots and Optional Upgrades

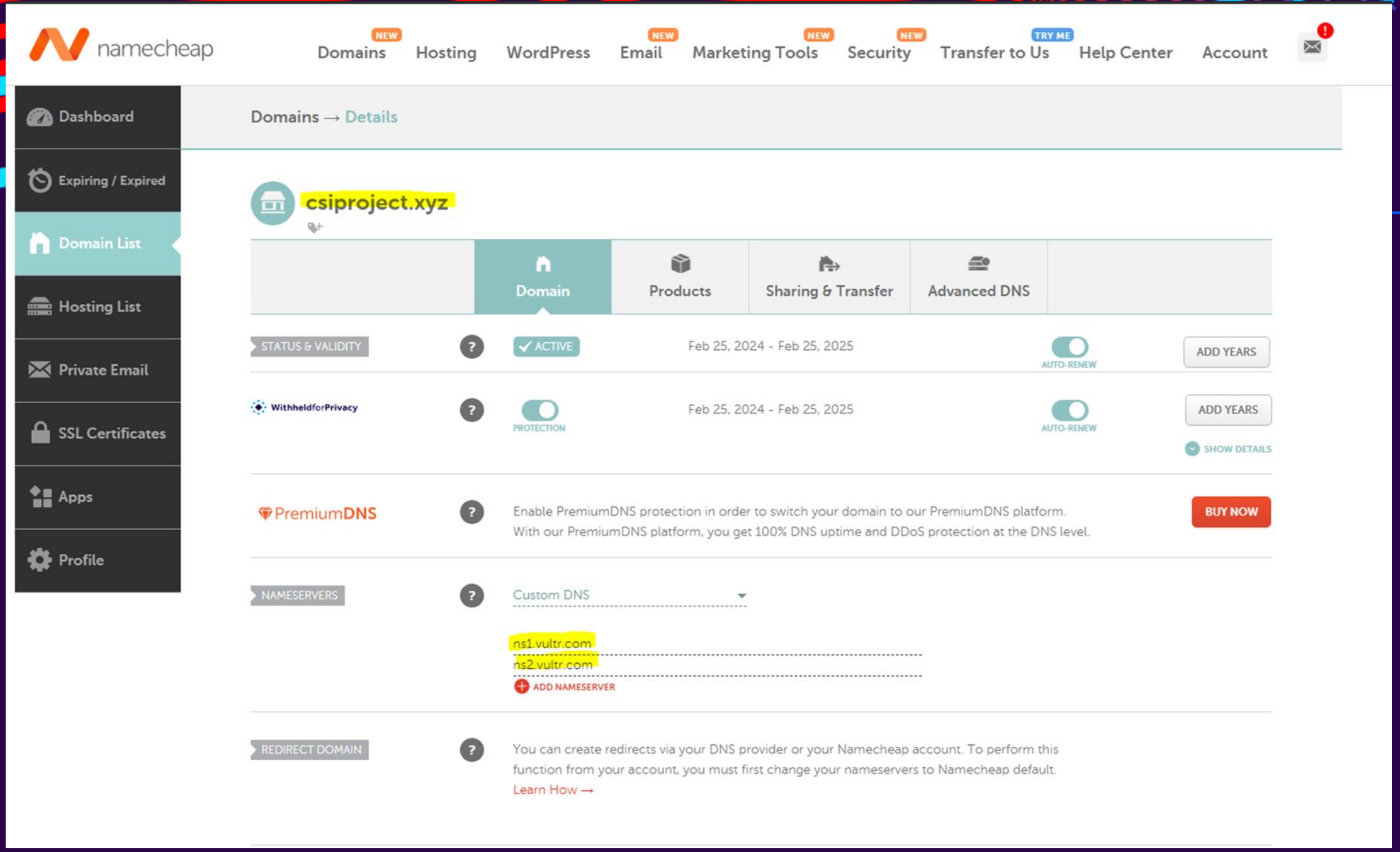
Namecheap provides:

- TLD – csiproject.xyz

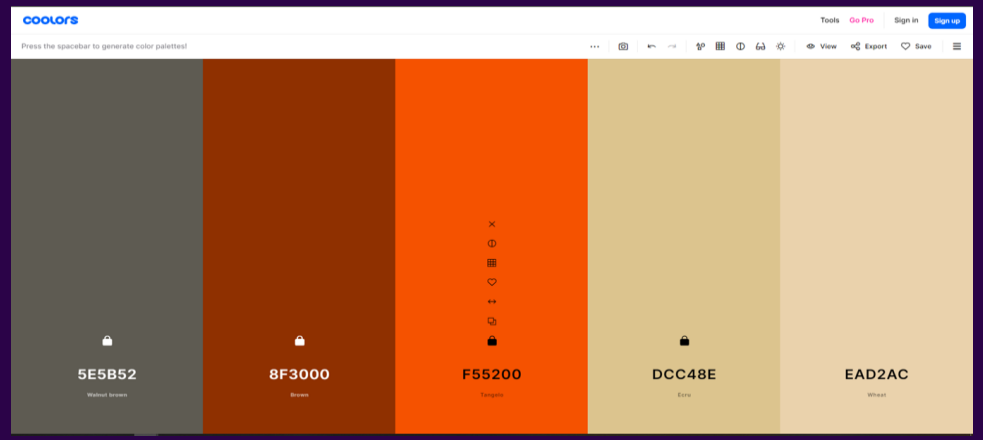
Shared configurations

- Centos Stream 9
- Certbot certificates
- Firewall ports opened

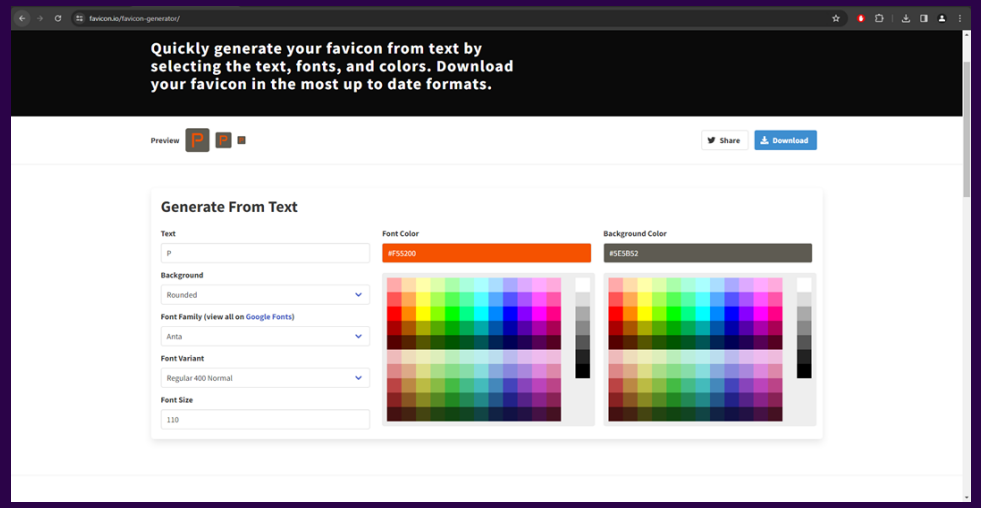

my.vultr.com/deploy/
Products
Dashboard
Products
Account
Support
Compute
Cloud Storage
Kubernetes
Container Registry
Databases
Load Balancers
CDN BETA
Network
Orchestration
Choose Image
Operating System
Marketplace Apps
Upload ISO
ISO Library
Backup
Snapshot
AlmaLinux
Select Version
Alpine Linux
Latest x64
Arch Linux
Latest x64
CentOS
8 Stream x64
Debian
Select Version
Fedora
Select Version
Fedora CoreOS
Select Version
Flatcar Container Linux
Select Version
FreeBSD
Select Version
OpenBSD
Select Version
Rocky Linux
Select Version
Ubuntu
Select Version
Windows Core Standard
Select Version
Windows Standard
Select Version
openSUSE
Leap 15 x64
Choose Plan
AMD High Performance
Intel High Performance
High Frequency
Regular Cloud Compute
Name
Cores
Memory
Storage
Bandwidth
Price
25 GB SSD
1 vCPU
1 GB
25 GB SSD
1 TB
$5/month
$0.007/hour
55 GB SSD
1 vCPU
2 GB
55 GB SSD
2 TB
$10/month
$0.015/hour
$15/month
Servers Qty:
1
Summary:
$6.00/month ($0.009/hour)
Deploy Now
Additional 15% Total Sales Tax applicable to this order
Update Settings

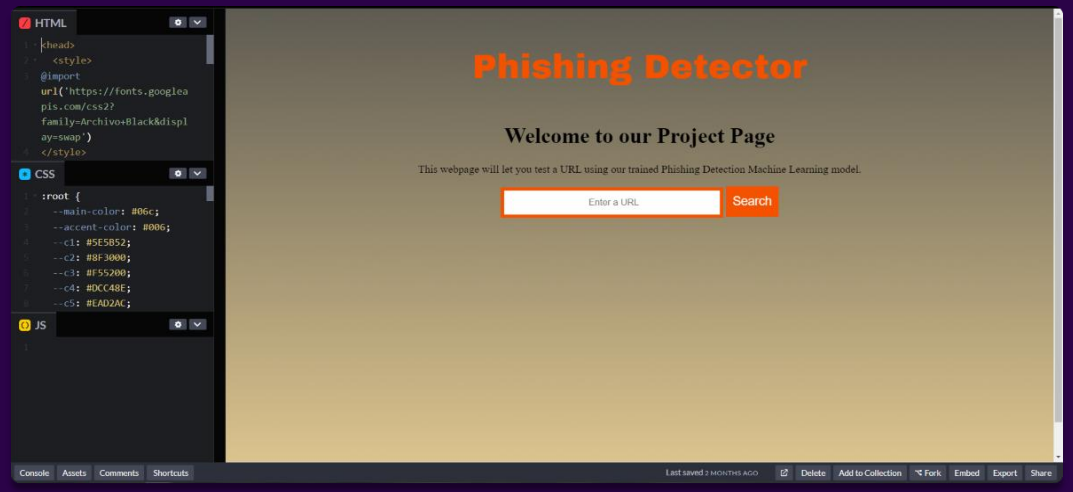

namecheap
Domains NEW
Hosting
WordPress
Email NEW
Marketing Tools NEW
Security NEW
Transfer to Us TRY ME
Help Center
Account
Dashboard
Expiring / Expired
Domain List
Hosting List
Private Email
SSL Certificates
Apps
Profile
Domains → Details
csiproject.xyz
Domain
Products
Sharing & Transfer
Advanced DNS
STATUS & VALIDITY
ACTIVE
Feb 25, 2024 - Feb 25, 2025
AUTO-RENEW
ADD YEARS
WithheldforPrivacy
PROTECTION
Feb 25, 2024 - Feb 25, 2025
AUTO-RENEW
ADD YEARS
SHOW DETAILS
PremiumDNS
Enable PremiumDNS protection in order to switch your domain to our PremiumDNS platform.
With our PremiumDNS platform, you get 100% DNS uptime and DDoS protection at the DNS level.
BUY NOW
NAMESERVERS
Custom DNS
ns1.vultr.com
ns2.vultr.com
ADD NAMESERVER
REDIRECT DOMAIN
You can create redirects via your DNS provider or your Namecheap account. To perform this function from your account, you must first change your nameservers to Namecheap default.
Learn How →

DESIGN



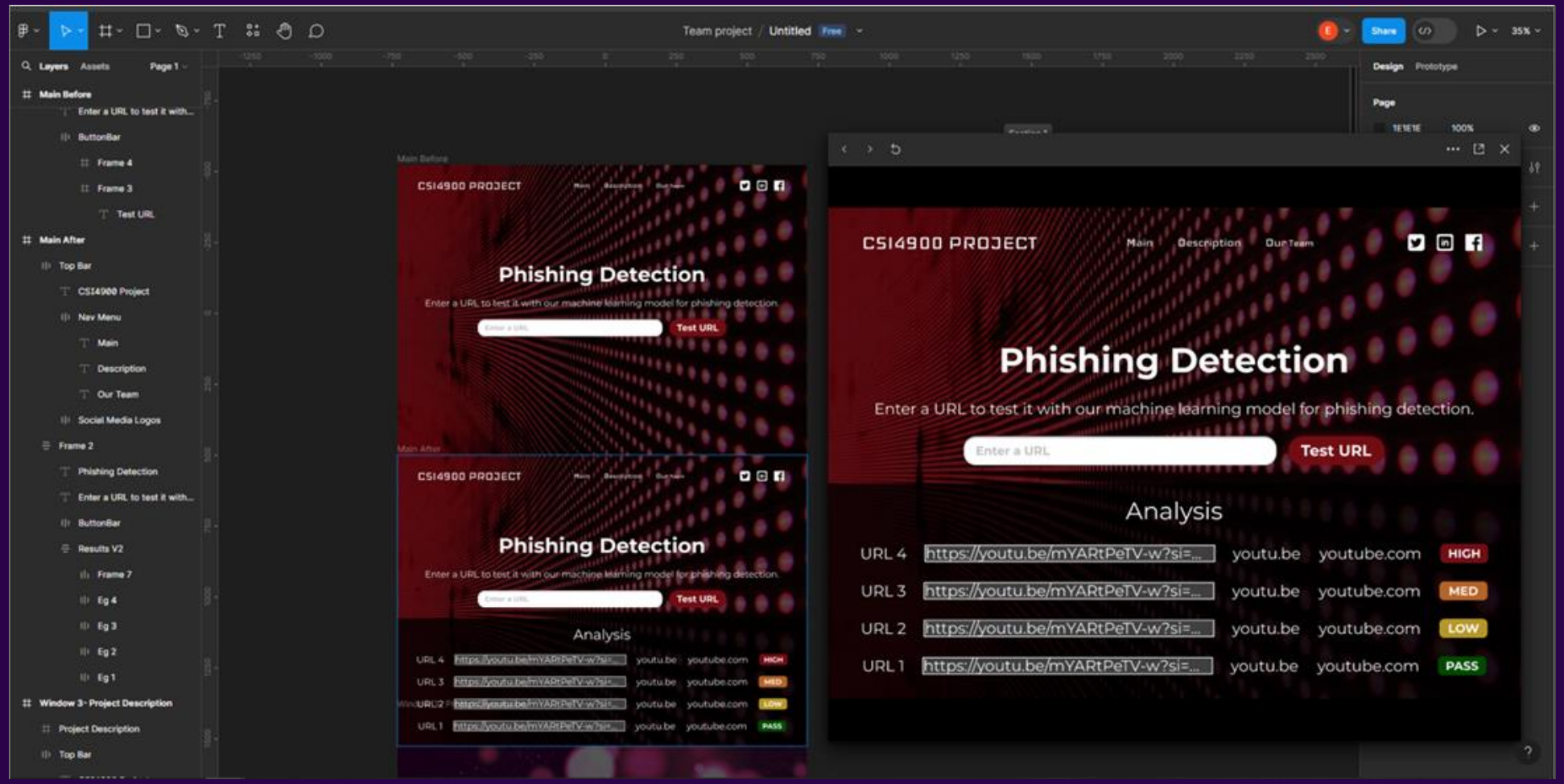
- Start with basic HTML, CSS, JS
- Keep it simple and consistent
- Introducing more elements adds complexity



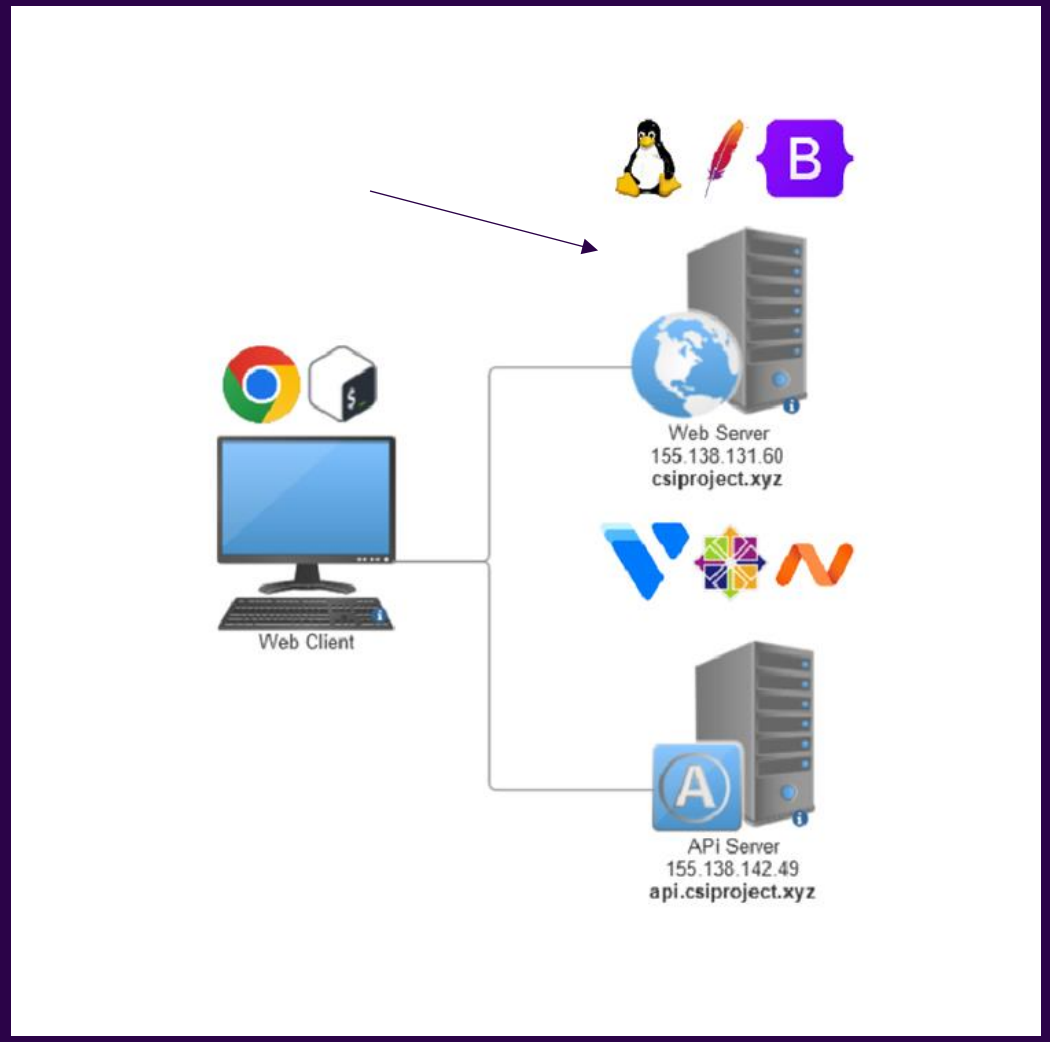


DESIGN

Figma for prototyping



WEB SERVER DETAILS

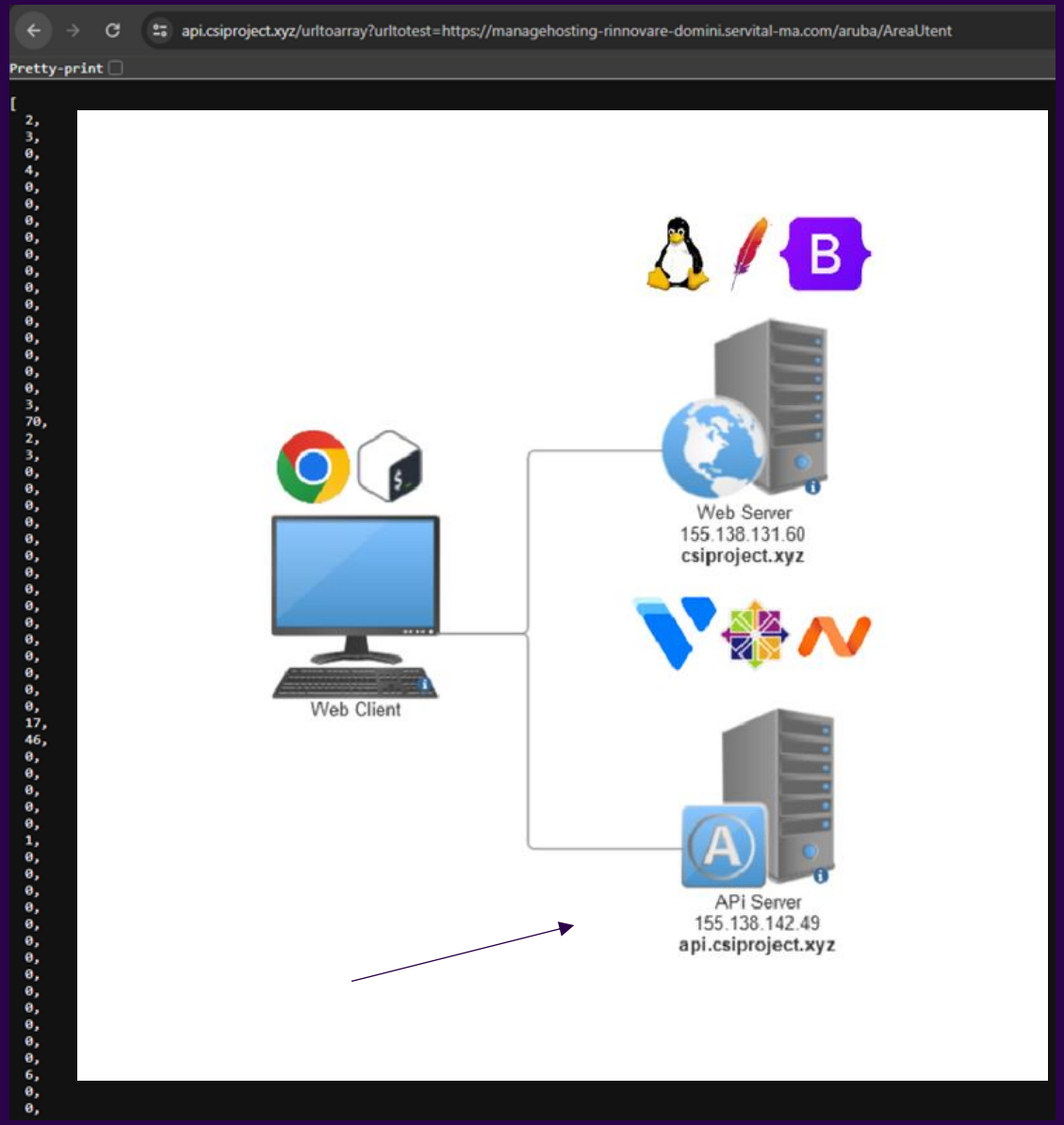


- Using Apache as the Web Server app
- Sending our webpage files which now incorporate the Bootstrap framework
- Apache configured to use HTTPS (port 443) and 303 Redirect (http -> https)
- Self-signed certificate -> Certbot

These are important for modern websites as web browsers will often complain

```
[root@csiprojwebserver ~]# systemctl  status httpd
● httpd.service - The Apache HTTP Server
     Loaded: loaded (/usr/lib/systemd/system/httpd.service; enabled; preset: disabled)
     Active: active (running) since Thu 2024-03-21 00:47:54 UTC; 3 weeks 5 days ago
```

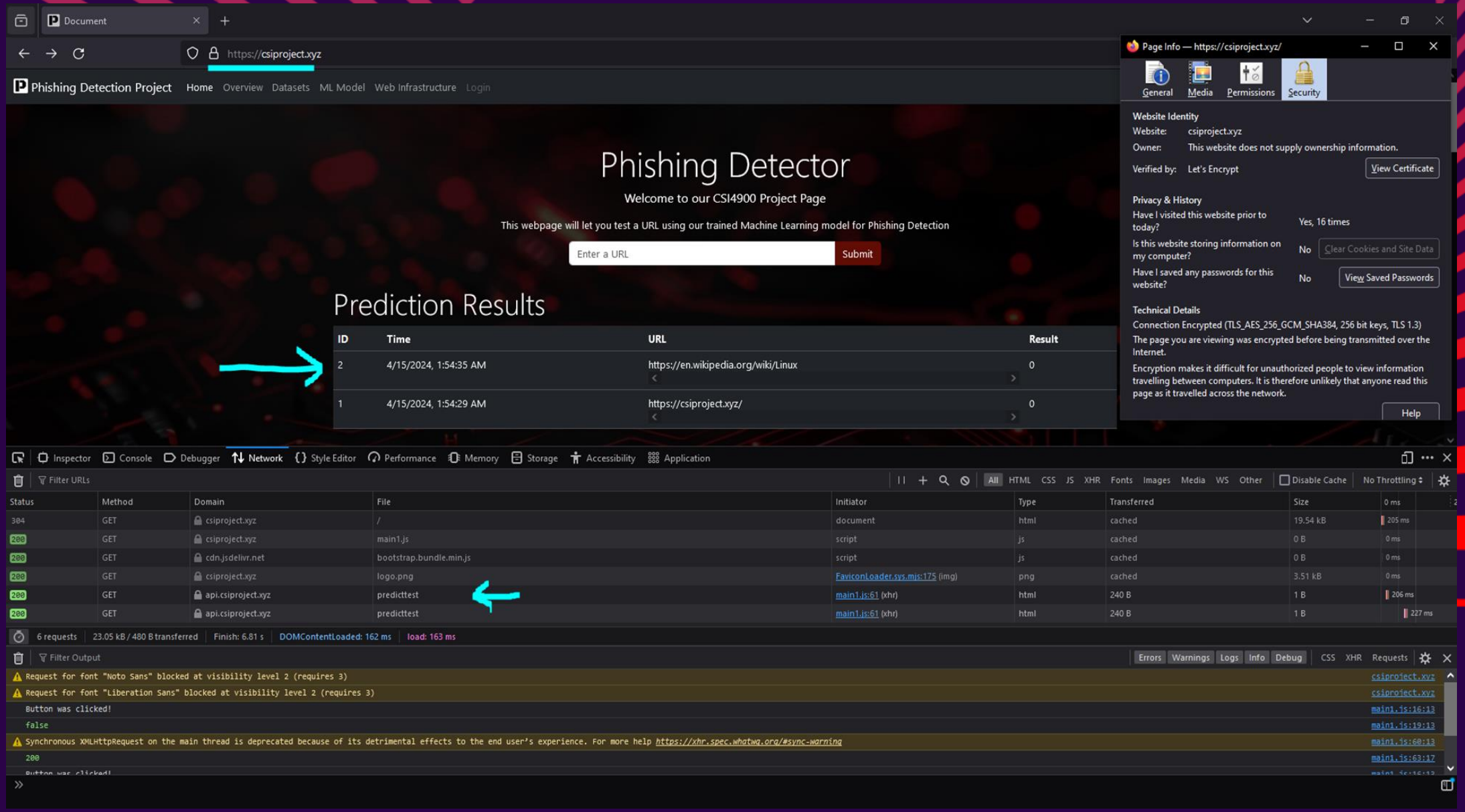
API SERVER DETAILS



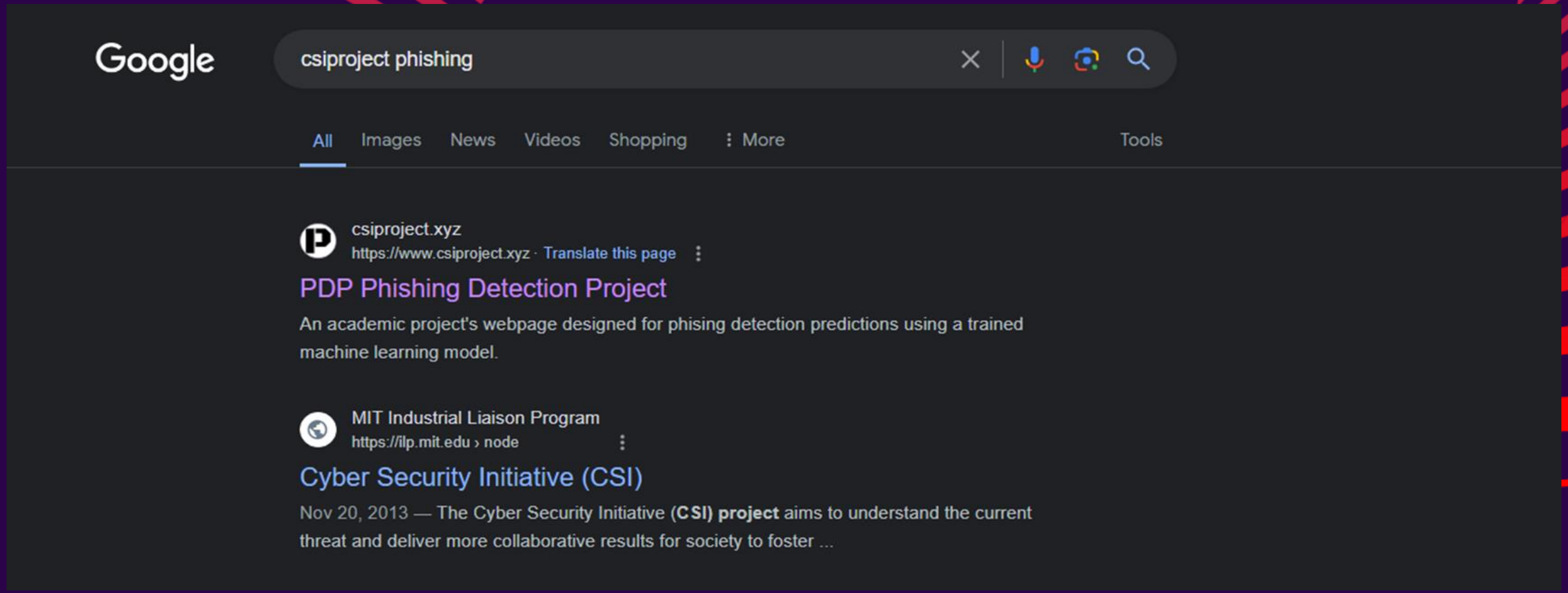
- Using Flask – Python web framework that can be used for APIs
- Using PIP to install all our Python packages
- Code functionality of API endpoints with Python
- Storing trained model as joblib
- Set up Certbot certificate

```
[root@csiprojapi pythonproj]# python api.py
 * Serving Flask app 'api'
 * Debug mode: on
INFO:werkzeug:WARNING: This is a development server.
 * Running on all addresses (0.0.0.0)
 * Running on https://127.0.0.1:443
 * Running on https://155.138.142.49:443
INFO:werkzeug:Press CTRL+C to quit
```

RESULTS


Document
https://csiproject.xyz
Phishing Detection Project
Home
Overview
Datasets
ML Model
Web Infrastructure
Login
Phishing Detector
Welcome to our CSI4900 Project Page
This webpage will let you test a URL using our trained Machine Learning model for Phishing Detection
Enter a URL
Submit
Prediction Results
ID
Time
URL
Result
2
4/15/2024, 1:54:35 AM
https://en.wikipedia.org/wiki/Linux
0
1
4/15/2024, 1:54:29 AM
https://csiproject.xyz/
0
Page Info — https://csiproject.xyz/
General
Media
Permissions
Security
Website Identity
Website: csiproject.xyz
Owner: This website does not supply ownership information.
Verified by: Let's Encrypt
View Certificate
Privacy & History
Have I visited this website prior to today? Yes, 16 times
Is this website storing information on my computer? No
Clear Cookies and Site Data
Have I saved any passwords for this website? No
View Saved Passwords
Technical Details
Connection Encrypted (TLS_AES_256_GCM_SHA384, 256 bit keys, TLS 1.3)
The page you are viewing was encrypted before being transmitted over the Internet.
Encryption makes it difficult for unauthorized people to view information travelling between computers. It is therefore unlikely that anyone read this page as it travelled across the network.
Help
Inspector
Console
Debugger
Network
Style Editor
Performance
Memory
Storage
Accessibility
Application
Filter URLs
All
HTML
CSS
JS
XHR
Fonts
Images
Media
WS
Other
Disable Cache
No Throttling
Status
Method
Domain
File
Initiator
Type
Transferred
Size
304 GET csiproject.xyz / document html cached 19.54 kB
200 GET csiproject.xyz main1.js script js cached 0 B
200 GET cdn.jsdelivr.net bootstrap.bundle.min.js script js cached 0 B
200 GET csiproject.xyz logo.png FaviconLoader.sys.mjs:175 (img) png cached 3.51 kB
200 GET api.csiproject.xyz predicttest main1.js:61 (xhr) html 240 B 1 B
200 GET api.csiproject.xyz predicttest main1.js:61 (xhr) html 240 B 1 B
6 requests
23.05 kB / 480 B transferred
Finish: 6.81 s
DOMContentLoaded: 162 ms
load: 163 ms
Filter Output
Errors
Warnings
Logs
Info
Debug
CSS
XHR
Requests
Request for font "Noto Sans" blocked at visibility level 2 (requires 3)
Request for font "Liberation Sans" blocked at visibility level 2 (requires 3)
Button was clicked!
false
Synchronous XMLHttpRequest on the main thread is deprecated because of its detrimental effects to the end user's experience. For more help https://xhr.spec.whatwg.org/#sync-warning
200

RESULTS


Google
csiproject phishing
All
Images
News
Videos
Shopping
More
Tools
csiproject.xyz
https://www.csiproject.xyz · Translate this page
PDP Phishing Detection Project
An academic project's webpage designed for phising detection predictions using a trained machine learning model.
MIT Industrial Liaison Program
https://ilp.mit.edu › node
Cyber Security Initiative (CSI)
Nov 20, 2013 — The Cyber Security Initiative (CSI) project aims to understand the current threat and deliver more collaborative results for society to foster ...

CONCLUSION

Through the design and implementation of a machine learning model and tool for phishing detection, we aimed to help combat the ever-evolving cyber security threats that impact individuals and businesses alike.

THANK YOU

Are there any questions?