



**Đại học Quốc gia TP.HCM
Trường Đại học Bách Khoa
Khoa Khoa học và Kỹ thuật Máy tính**



Đồ án tốt nghiệp

PHÁT HIỆN LỪA ĐẢO BẰNG CÁC PHƯƠNG PHÁP HỌC MÁY

Hội đồng:

Hội đồng 18 KHMT

GVHD:

TS. Nguyễn Lê Duy Lai

GVPB:

TS. Nguyễn Đức Thái

Ngày 12 tháng 06 năm 2023

SINH VIÊN THỰC HIỆN



Nguyễn Đăng Hải

1913254



Phạm Đại Hoàng An

1912539

NỘI DUNG

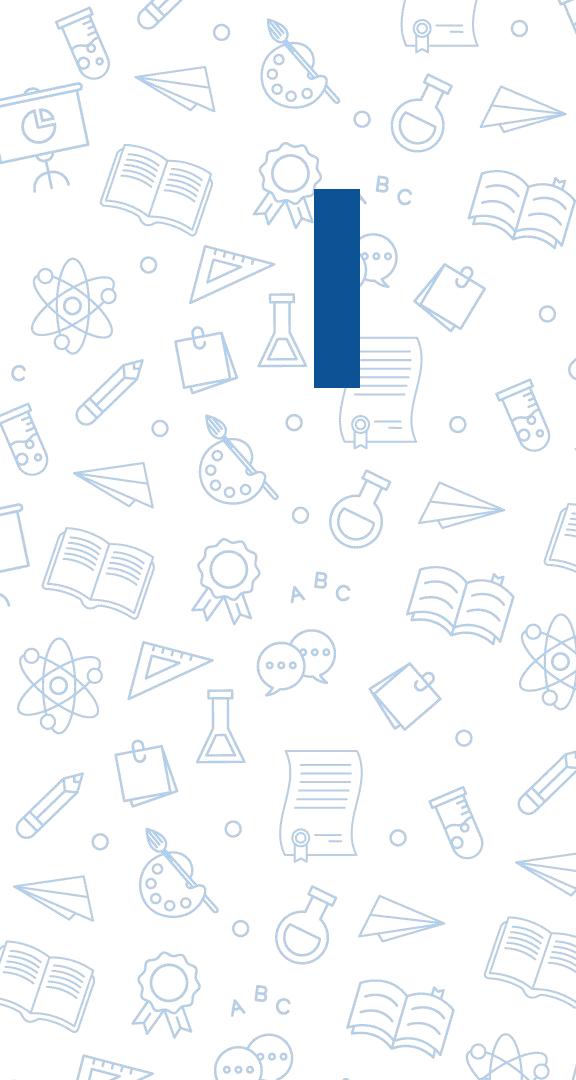
I. TỔNG QUAN ĐỀ TÀI

II. MÔ HÌNH PHÁT HIỆN GIẢ MẠO

III. XÂY DỰNG ĐÁNH GIÁ MÔ HÌNH

IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG

V. ĐÁNH GIÁ VÀ KẾT LUẬN



TỔNG QUAN ĐỀ TÀI

- Tấn công lừa đảo là một vấn đề nghiêm trọng trong thời đại kỹ thuật số.
- Là hành vi tạo ra các trang web giả mạo, giả danh hoặc sao chép các trang web tồn tại để lừa đảo người dùng
- Thường có giao diện và nội dung tương tự như các trang web chính thức

1,025,968

Tổng số vụ lừa đảo qua mạng xảy ra trong Q1/2022

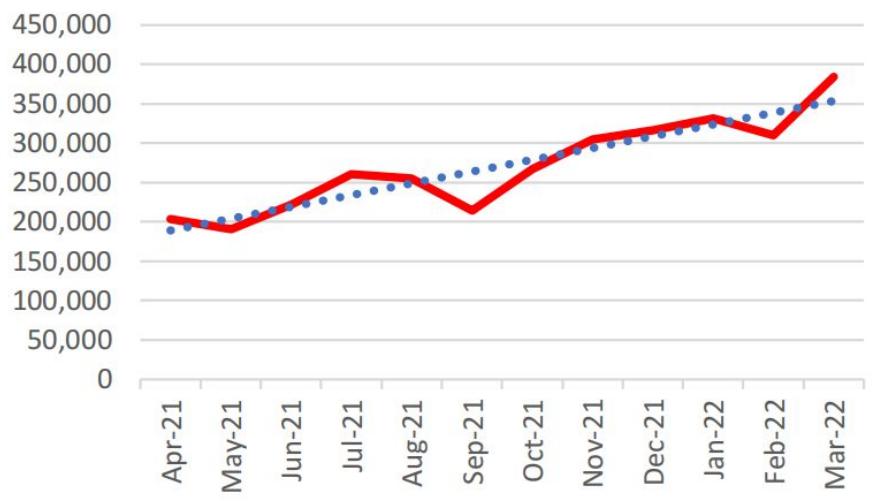
384,291

Số vụ xảy ra trong T3/2022, tháng có số vụ tấn công kỷ lục

23.6%

Phần trăm số vụ xảy ra thiệt hại đến ngành tài chính, ngân hàng

Phishing Attacks, 2Q2021 - 1Q2022



Hình: Số lượng các vụ tấn công lừa đảo từ tháng 4/2021 đến tháng 3/2022

(Nguồn:
https://docs.apwg.org/reports/apwg_trends_report_q1_2021.pdf)

2. Các công trình nghiên cứu, hệ thống liên quan

I. TỔNG QUAN ĐỀ TÀI



ChongLuaDao

★★★★★ 182 ⓘ

| Hỗ trợ tiếp cận | 40.000+ người dùng

★★★★★ 125 ⓘ



AegisWeb3

✓ www.aegisweb3.com ⓘ Nổi bật

★★★★★ 125 ⓘ

| Nâng suất | 100.000+ người dùng



WOT: Website Security & Safety Checker

✓ mywot.com ⓘ Nổi bật

★★★★★ 10.492 ⓘ

| Nâng suất | 900.000+ người dùng

- Nhận thấy ứng dụng có độ chính xác không quá cao.
- Phần lớn ứng dụng dựa vào phương pháp truyền thống, do đó những trang web mới nhất không phát hiện được.

TOWARDS BENCHMARK DATASETS FOR MACHINE LEARNING
BASED WEBSITE PHISHING DETECTION: AN EXPERIMENTAL
STUDY

A PREPRINT

Abdelhakim Hannousse
Department of Computer Science
Université 8 Mai 1945, Guelma
BP 401, Guelma 24000, Algeria
hannousse.abdelhakim@univ-guelma.dz

Salima Yahiouche
Department of Computer Science
LRS laboratory, Badji Mokhtar University
BP 12, Annaba 23000, Algeria
yahiouche.salima@univ-annaba.dz

October 27, 2020

Nguồn:

<https://www.sciencedirect.com/science/article/abs/pii/S0952197621001950>



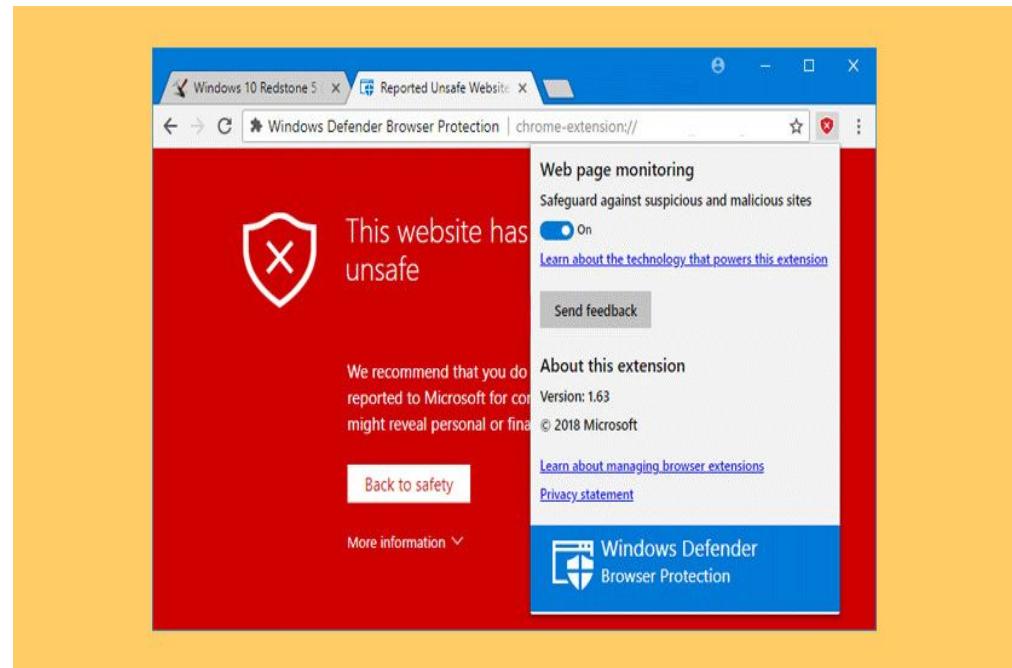
Review

A Survey of Machine Learning-Based Solutions for Phishing Website Detection

Lizhen Tang * and Qusay H. Mahmoud

Nguồn: <https://www.mdpi.com/2504-4990/3/3/34>

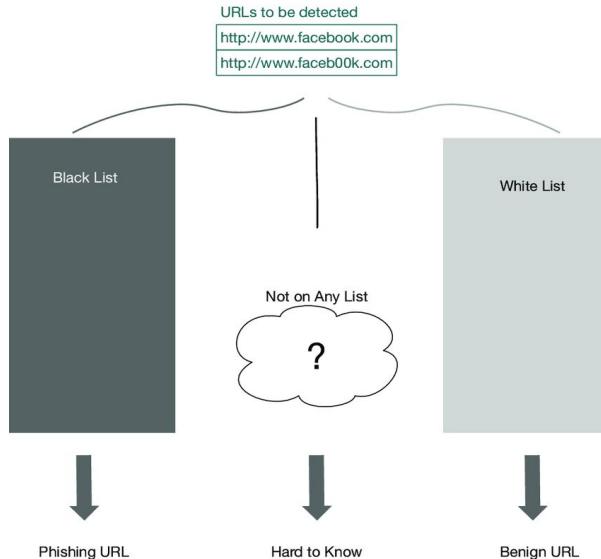
- Phát triển một mô hình học máy tối ưu hơn so với những mô hình đã có sẵn.
- Phát triển một hệ thống phần mềm phát hiện những trang web giả mạo theo thời gian thực.
- Tích hợp phần mềm vào trình duyệt web như một tiện ích mở rộng (extension).



MÔ HÌNH PHÁT HIỆN GIẢ MẠO

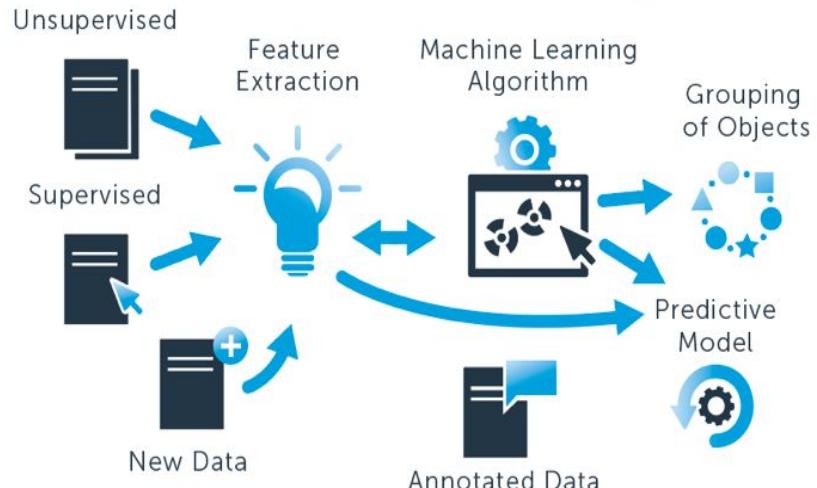
1. Cách thức phát hiện lừa đảo

II. MÔ HÌNH PHÁT HIỆN GIẢ MẠO



List-based detection

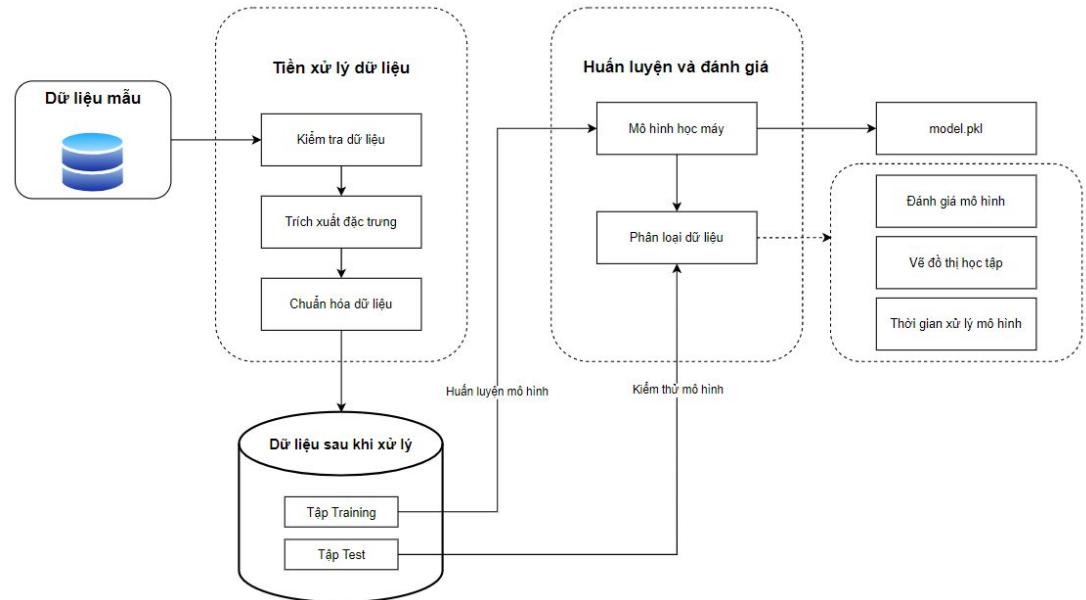
Machine Learning



Machine learning-based detection

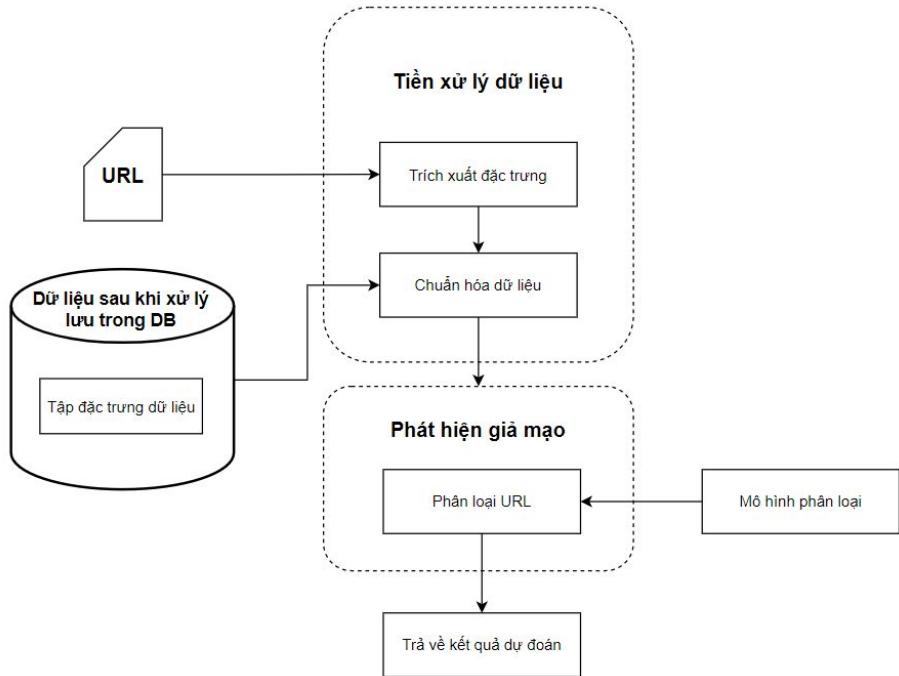
Gồm 2 giai đoạn:

- Giai đoạn tiền xử lý (Pre-Processing)
- Giai đoạn đào tạo (Training & Testing).



Hình: Quy trình huấn luyện mô hình

- Xử lý URL nhận được khi người dùng mở một trang web mới.
- Phân loại bằng cách đưa vào mô hình.

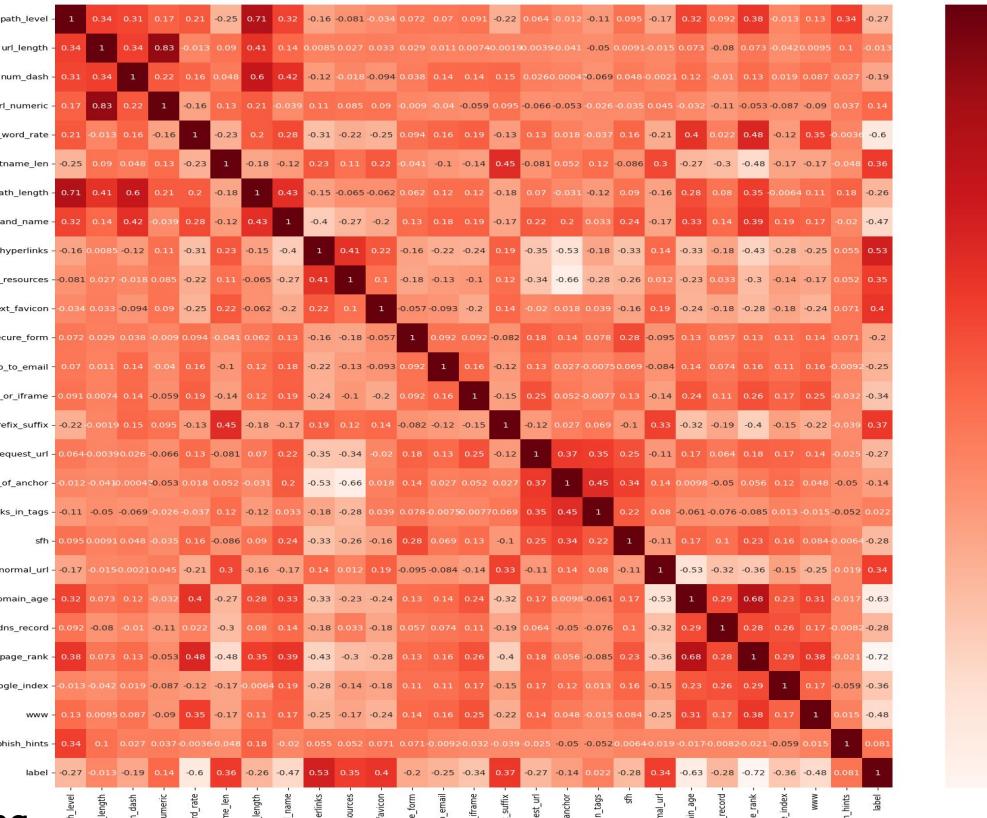


Hình: Quy trình phát hiện giả mạo

3. Các đặc trưng của trang web

II. MÔ HÌNH PHÁT HIỆN GIẢ MẠO

Chọn ra những đặc tính có giá trị tương quan với "label" trên 0.2



Hình: Ma trận tương quan của đặc trưng

PHÂN LOẠI ĐẶC TRƯNG



10

URL-BASED

URL-length, IP,
subdomain, symbol@,..



11

CONTENT-BASE D

Hyperlink, external CSS,
external Favicon,...



5

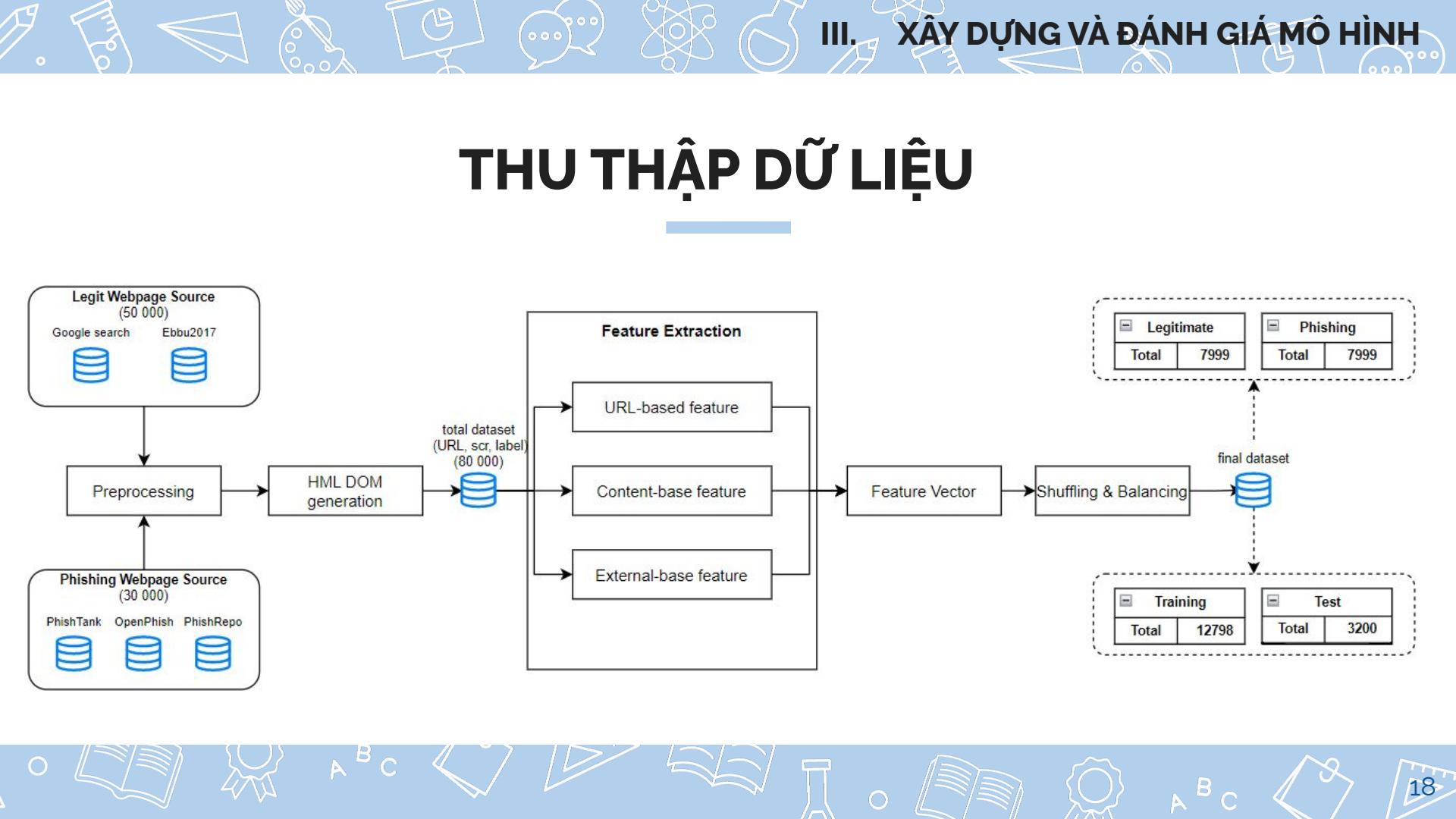
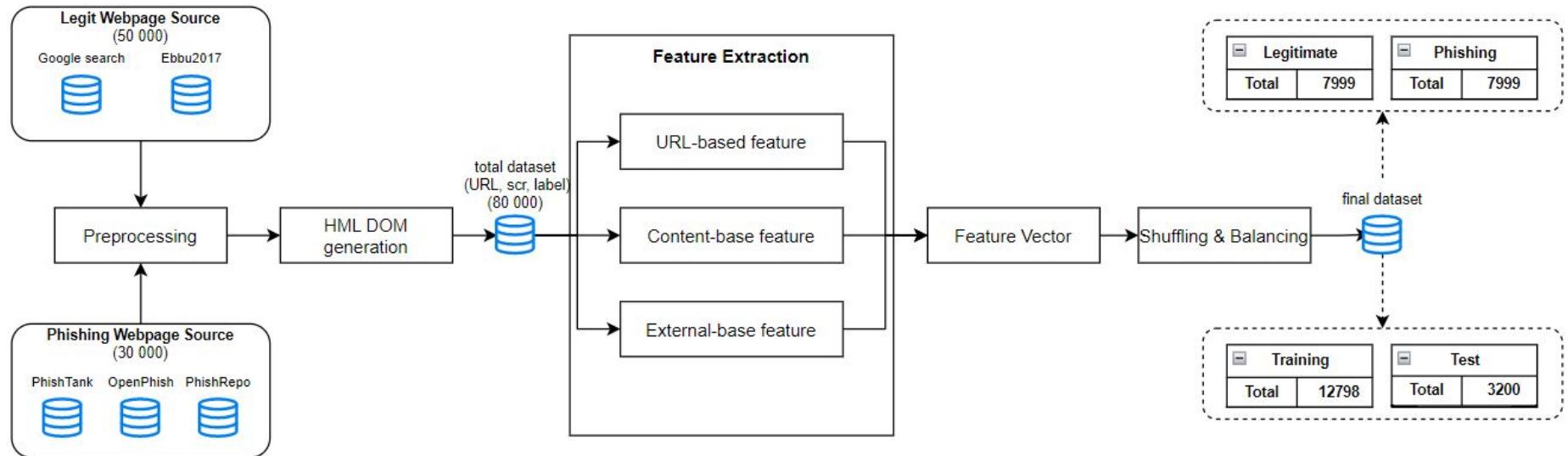
EXTERNAL-BASE D

Domain age, web traffic,
DNS record,...

XÂY DỰNG VÀ ĐÁNH GIÁ MÔ HÌNH

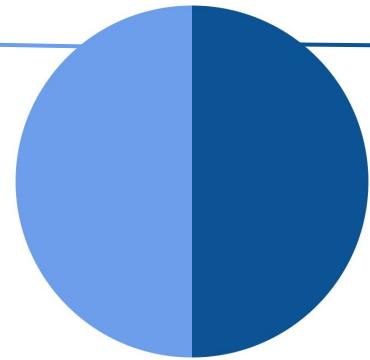


THU THẬP DỮ LIỆU



TỔNG HỢP DỮ LIỆU

50% URL giả
mạo



50% URL hợp lệ



3200

Tập giá trị dùng để test



15998

Tổng dữ liệu



12798

Tập giá trị dùng để training

TIỀN XỬ LÝ DỮ LIỆU

URL	Path_level	URL_length	Num_dash	...	Label
https://diario.elmundo.sv/	0	26	0	...	0
https://tff4ox.webwave.dev/	0	27	0	...	1
https://www.alphabet.com/en-ww	1	30	1	...	0

MÔ HÌNH HỌC MÁY

Model	No.Features	Total data	Macro F1	Accuracy
SVM	30 / 26	3170 / 15998	86% / 79%	86% / 79%
Logistic Regression	30 / 26	3170 / 15998	82% / 89%	83% / 89%
Decision Tree	30 / 26	3170 / 15998	82% / 94%	83% / 94%
kNN	30 / 26	3170 / 15998	85% / 91%	86% / 91%
Random Forest	30 / 26	3170 / 15998	88% / 97%	88% / 97%

ENSEMBLE LEARNING

Model	Estimator-model	Macro F1	Accuracy	Time
Stacked Generalization	Decision Tree	97%	97%	0.014s
Stacked Generalization	kNN	97%	97%	0.024s
Stacked Generalization	SVM	97%	97%	0.023s
Stacked Generalization	Logistic Regression	97%	97%	0.025s
Stacked Generalization	Random Forest	97%	97%	0.021s

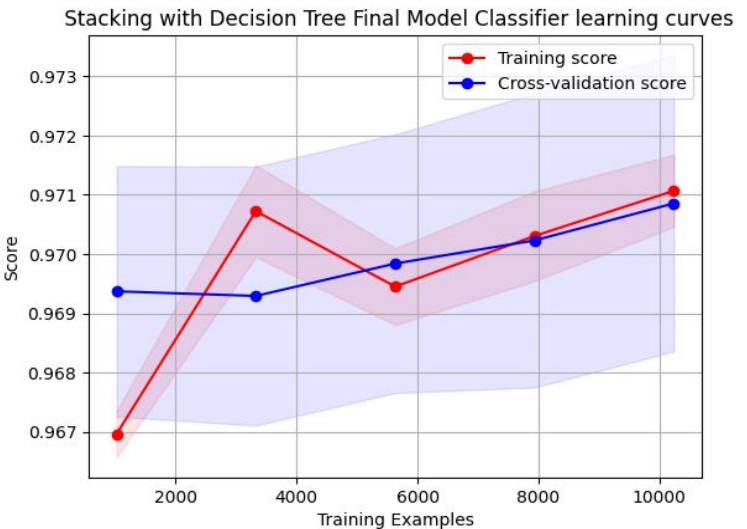
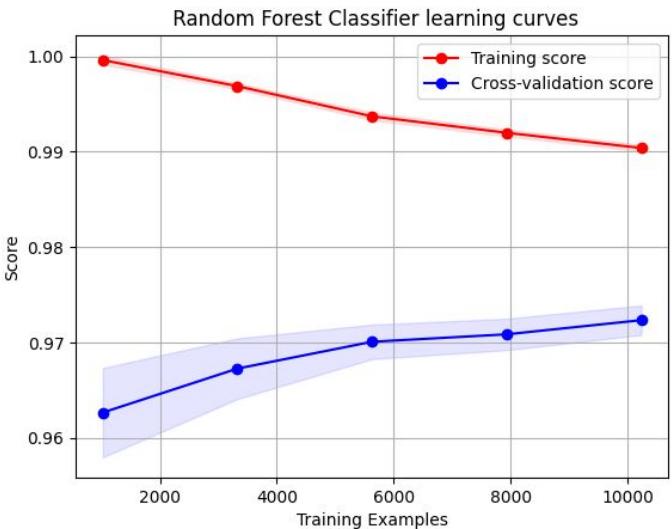
ENSEMBLE LEARNING

Model	Estimator-model	Macro F1	Accuracy	Time
Blending	Decision Tree	97%	97%	0.027s
Blending	kNN	97%	97%	0.025s
Blending	SVM	97%	97%	0.017s
Blending	Logistic Regression	96%	96%	0.016s
Blending	Random Forest	97%	97%	0.016s

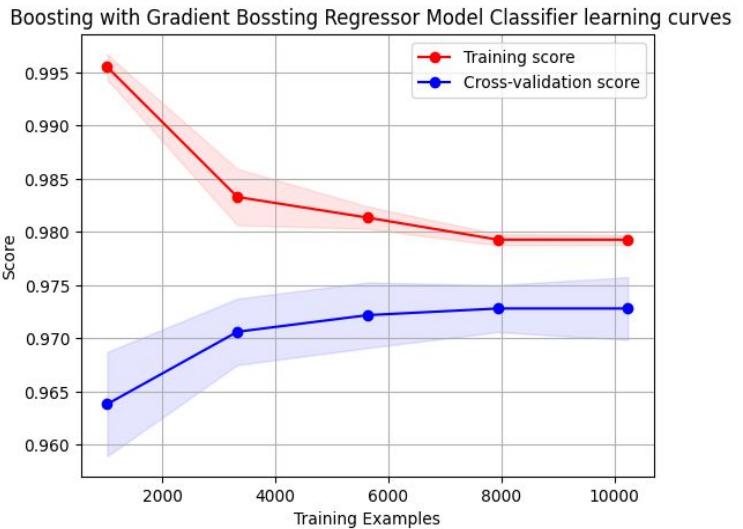
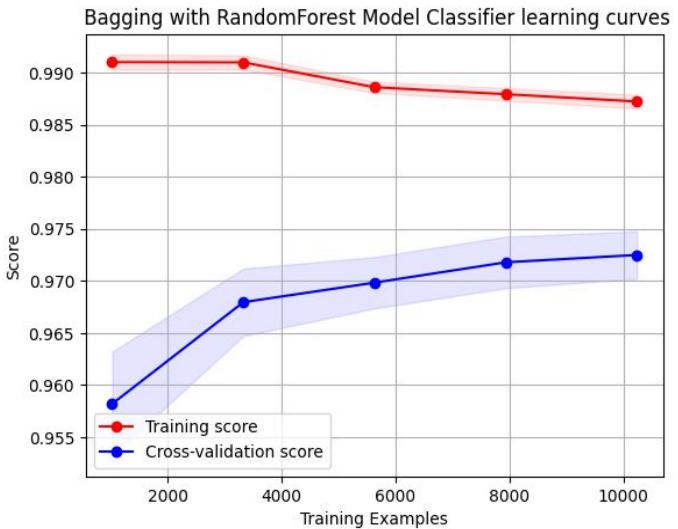
ENSEMBLE LEARNING

Model	Estimator-model	Macro F1	Accuracy	Time
Bagging Classifier	Random Forest	97%	97%	0.005s
Gradient Boosting Classifier		97%	97%	0.007s
XGBoosting		97%	97%	0.008s
LightGBM Classifier		98%	98%	0.005s
Random Forest		97%	97%	0.003s

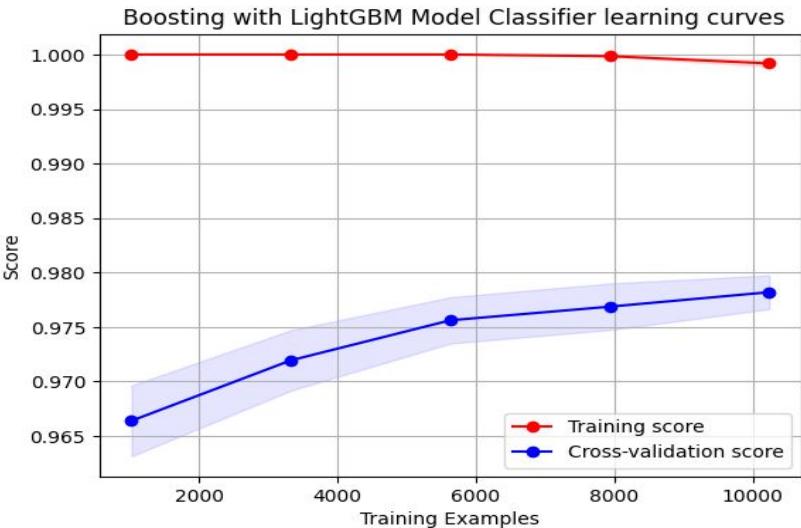
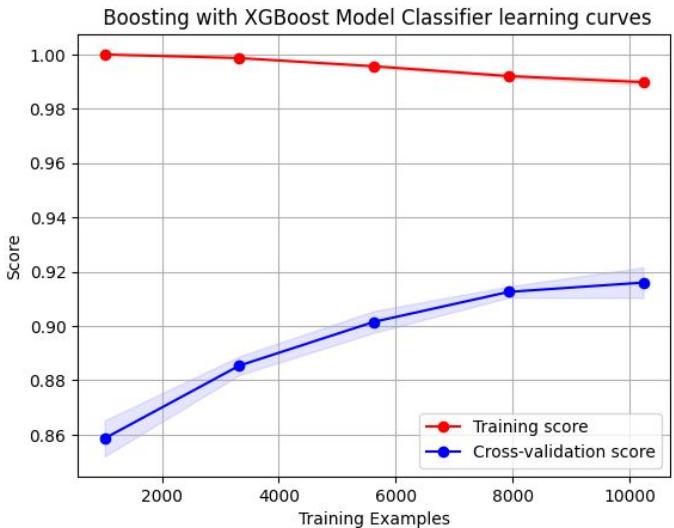
ĐỒ THỊ HỌC TẬP



ĐỒ THỊ HỌC TẬP



ĐỒ THỊ HỌC TẬP

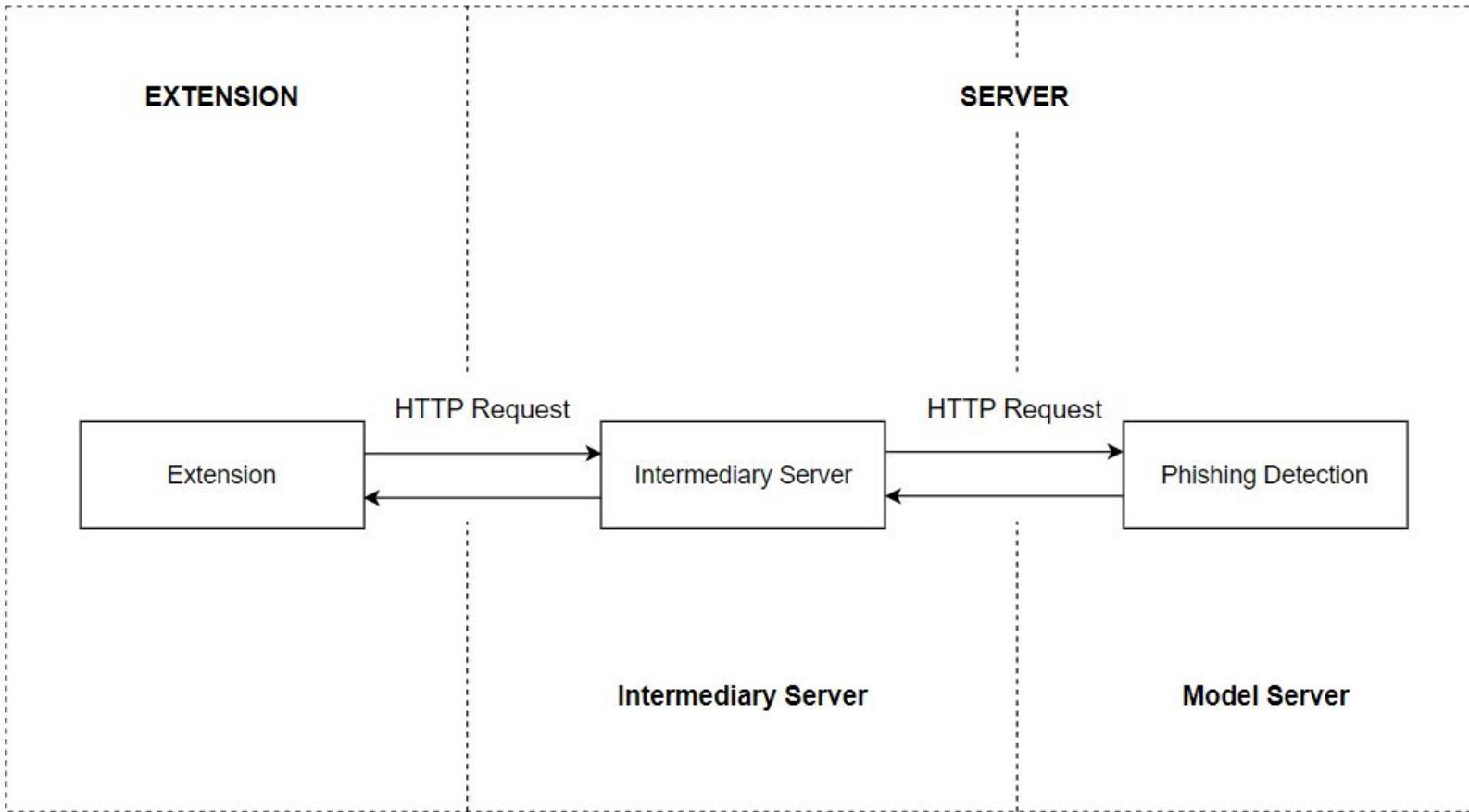


IV

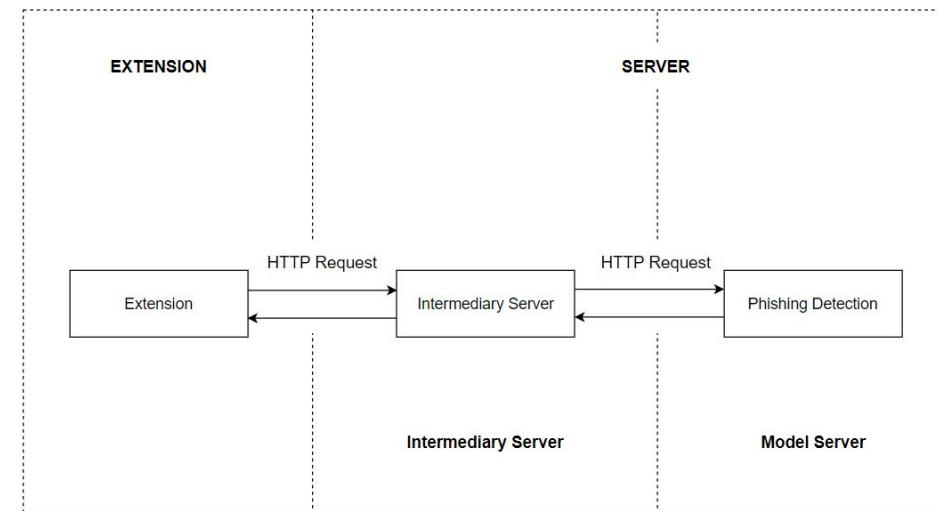
THIẾT KẾ ỨNG DỤNG HỆ THỐNG

1. Tổng quan kiến trúc hệ thống

IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG



- Tiện ích mở rộng
- Hệ thống phát hiện giả mạo
 - Hệ thống trung gian
 - Hệ thống dự đoán bằng mô hình học máy



- Tiện ích mở rộng:

Mỗi khi người dùng truy cập vào địa chỉ URL, tiện ích sẽ gửi HTTP Request đến phía Server để xử lý.

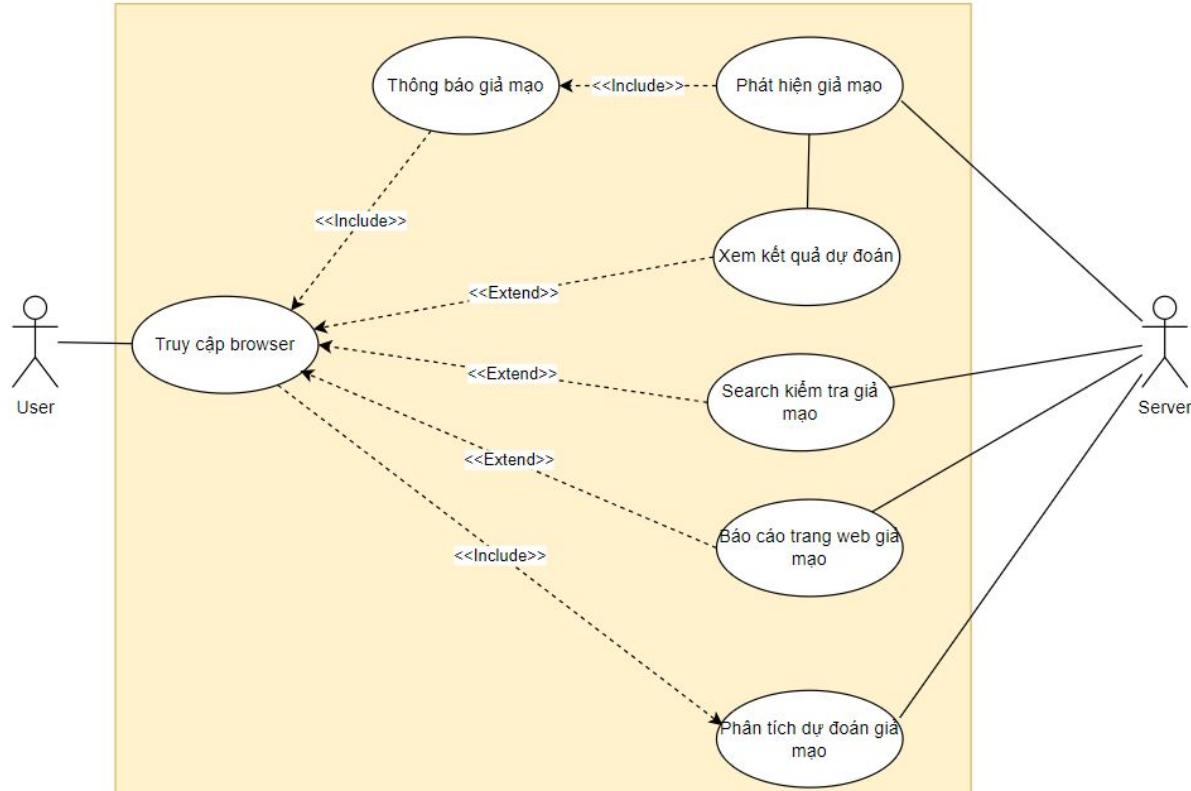
Sau khi nhận kết quả, lưu kết quả đã nhận được vào bộ nhớ cục bộ.

- Hệ thống trung gian:
 - Lấy địa chỉ URL từ trang truy cập của người dùng
 - Phân tích đặc trưng
 - Gửi và nhận yêu cầu dự đoán từ mô hình học máy
 - Lưu trữ các website được báo cáo gửi lên server

- Hệ thống dự đoán bằng mô hình học máy:
 - Nhận yêu cầu dự đoán từ mô hình trung gian
 - Dự đoán kết quả
 - Trả kết quả dự đoán về hệ thống trung gian

2. Hiện thực hệ thống

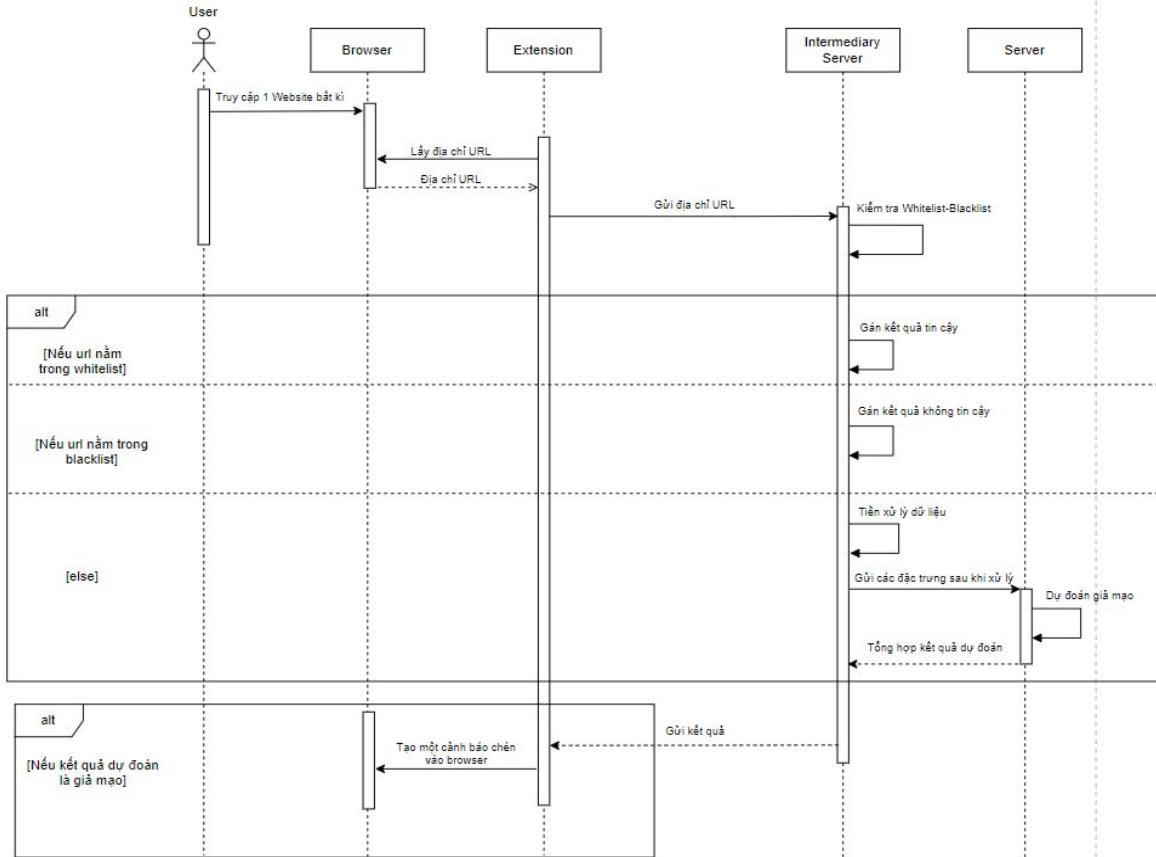
IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG



Use-case diagram

2. Hiện thực hệ thống

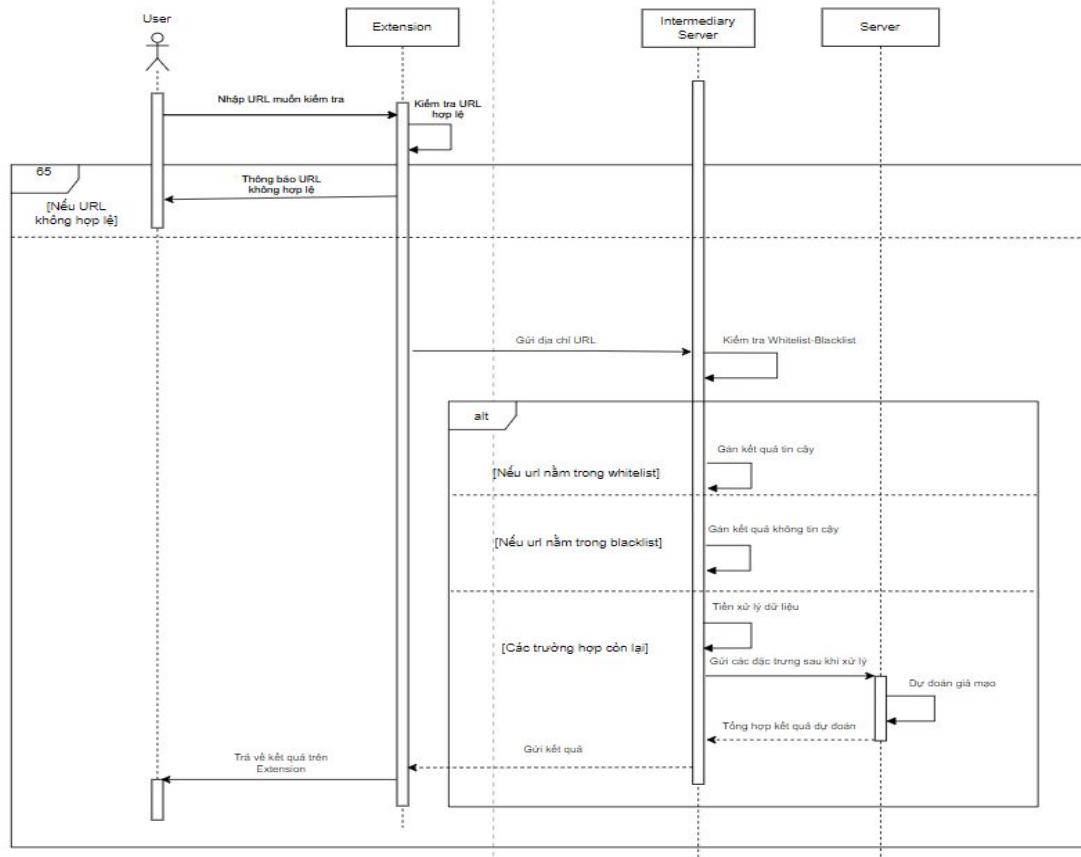
IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG



Sequence diagram chức năng cảnh báo người dùng

2. Hiển thực hệ thống

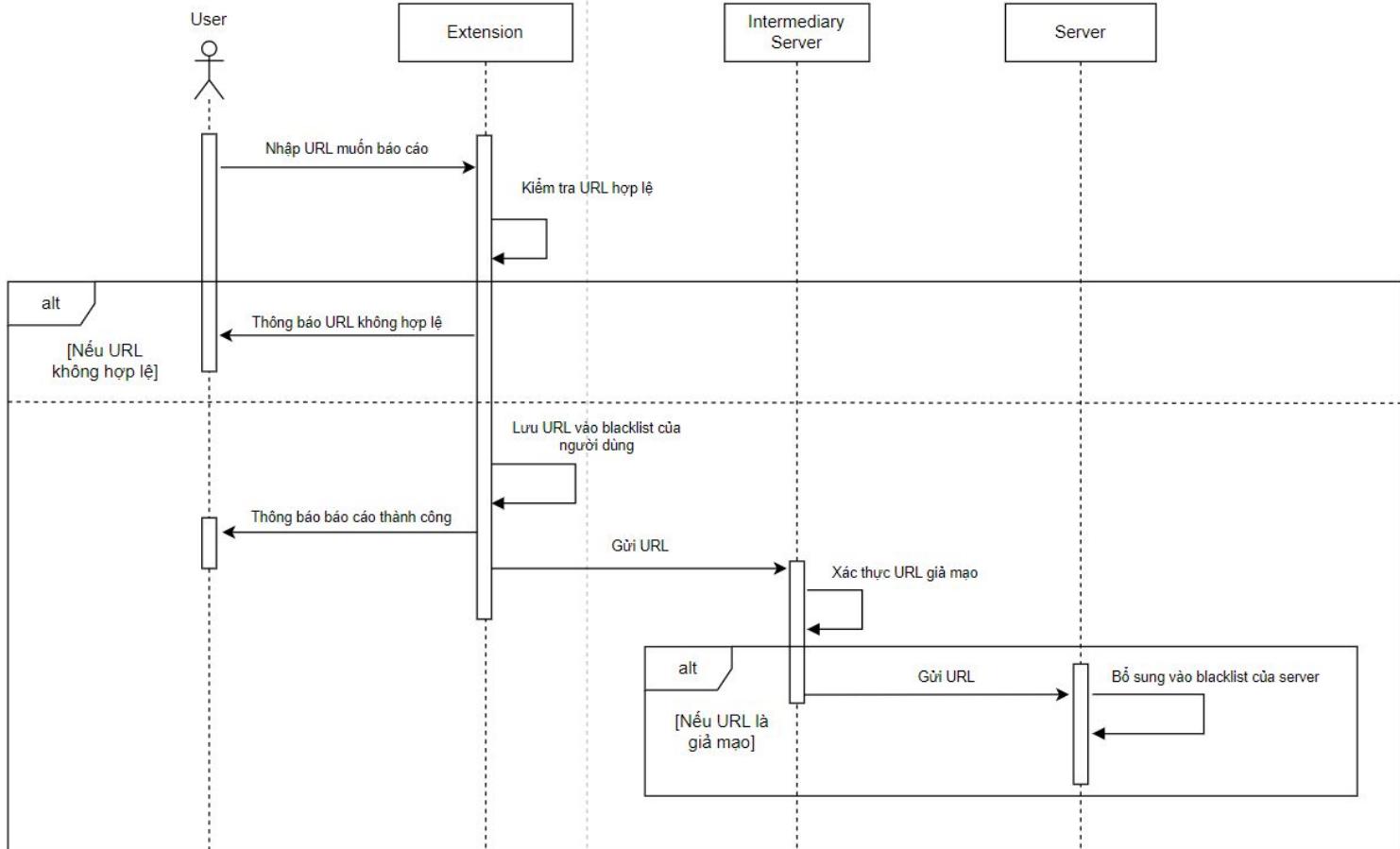
IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG



Sequence diagram chức năng kiểm tra website

IV. THIẾT KẾ ỨNG DỤNG HỆ THỐNG

2. Hiện thực hệ thống



Sequence diagram chức năng báo cáo lừa đảo

Tiện ích mở rộng cơ bản được viết bằng **HTML, CSS và Javascript**.

Gửi URL mà người dùng truy cập đến hệ thống.

Hiển thị kết quả được trả về và lưu kết quả vào bộ nhớ cục bộ.

Phishing Detector

docs.google.com



Trang web này an toàn



Hệ thống được triển khai bằng **Azure Application Services**, sử dụng framework Flask. (<https://phishingdetector.scm.azurewebsites.net/>)

Các HTTP Request được sử dụng

STT	Method	Endpoint	Chức năng
1	GET	/extractFeatures	Lấy các đặc trưng của một URL
2	GET	/getPrediction	Lấy kết quả dự đoán của URL dựa vào các đặc trưng
3	POST	/postReportUrl	Gửi URL mà người dùng báo cáo lên lưu trữ để sau này xử lý

Lưu các tên miền uy tín và lừa đảo vào 2 danh sách riêng biệt, và được kiểm tra với URL trước khi đem đi phân tích.

Sử dụng **Azure Blob Storage** để lưu trữ 1 file text chứa các URL mà người dùng báo cáo về cho hệ thống.

(https://phishingdetectorstorage.blob.core.windows.net/list/report_url.txt)



V

ĐÁNH GIÁ VÀ KẾT LUẬN

ĐÁNH GIÁ HỆ THỐNG

Tính năng	Đánh giá
Hệ thống cảnh báo người dùng khi truy cập một website giả mạo	<input checked="" type="checkbox"/>
Hiển thị kết quả giả mạo hoặc hợp lệ lên bảng popup extension	<input checked="" type="checkbox"/>
Quay trở lại trang trước thông qua phím tắt trên bảng thông báo	<input checked="" type="checkbox"/>
Tiếp tục truy cập website giả mạo sau khi tắt bảng cảnh báo	<input checked="" type="checkbox"/>
Truy cập một website bất kỳ nằm trong whitelist mà không báo giả mạo	<input checked="" type="checkbox"/>

ĐÁNH GIÁ HỆ THỐNG

Tính năng	Đánh giá
Kiểm tra một website đúng cú pháp vào thanh kiểm tra và trả về kết quả	<input checked="" type="checkbox"/>
Kiểm tra một website không đúng cú pháp vào thanh kiểm tra và không trả về kết quả	<input checked="" type="checkbox"/>
Hệ thống cập nhật website lên blacklist local của người dùng	<input checked="" type="checkbox"/>
Hệ thống cảnh báo khi truy cập vào website vừa báo cáo	<input checked="" type="checkbox"/>
URL vừa báo cáo được đẩy lên danh sách cần xử lý ở Server	<input checked="" type="checkbox"/>

THỜI GIAN XỬ LÝ

Trường hợp	Thời gian (s)
Người dùng truy cập một website nằm trong white-black list	0.32
Người dùng truy cập một website giả mạo nằm ngoài white-black list	4.12
Người dùng truy cập một website hợp lệ nằm ngoài white-black list	5.84
Kiểm tra một website không đúng cú pháp	0.14

THỜI GIAN XỬ LÝ

Trường hợp	Thời gian (s)
Kiểm tra một website đúng cú pháp nằm trong white-black list	0.24
Kiểm tra một website giả mạo không nằm trong white-black list	7.55
Kiểm tra một website hợp lệ không nằm trong white-black list	9.03
Báo cáo một website bất kỳ không đúng cú pháp	0.56
Báo cáo một website bất kỳ đúng cú pháp	0.37

THỜI GIAN XỬ LÝ CÙNG LÚC

Số lượng	Thời gian phản hồi lâu nhất (s)
1	5.34
2	5.69
3	10.43
4	15.39

Số lượng	Thời gian phản hồi lâu nhất (s)
5	18.73
6	21.34
7	25.62
8	29.83

KẾT QUẢ ĐẠT ĐƯỢC

- Tổng hợp nội dung và cung cấp kiến thức về thực trạng của việc giả mạo trên mạng hiện nay
- Tổng hợp, trích xuất các đặc trưng từ URL và so sánh tìm ra các đặc trưng quan trọng hỗ trợ cho việc phân loại.
- Tổ chức tổng hợp và đánh nhãn giả mạo từ nhiều nguồn website khác nhau
- Áp dụng Ensemble Learning trong việc xây dựng mô hình phát hiện giả mạo
- Xây dựng nhiều mô hình phân loại và so sánh, đánh giá các mô hình
- Xây dựng tiện ích mở rộng và đã đưa lên Cửa hàng tiện ích giúp cảnh báo người dùng khi truy cập website giả mạo



HẠN CHẾ TỒN TẠI

- Do còn phụ thuộc vào chi phí chi trả cho Azure App Service nên chưa tối ưu được thời gian phản hồi cho nhiều người dùng một lúc.
- Tiện ích mở rộng mới chỉ hỗ trợ được trên các trình duyệt có nhân Chromium, do đó một bộ phận người dùng sẽ không tiếp cận được với công cụ này



HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI

- Mở rộng thêm nhiều lựa chọn, giúp người dùng thoải mái tùy chỉnh extension.
- Cải thiện thêm tốc độ xử lý và khả năng chịu tải của hệ thống.
- Bổ sung tập dữ liệu mới nhất từ các trang web cung cấp website giả mạo để gia tăng độ hiệu quả khi phát hiện lừa đảo.
- Xem xét những góp ý, đánh giá của người dùng trên Cửa hàng Tiện ích mở rộng để hoàn thiện hệ thống hơn.
- Mở rộng tiện ích trên nhiều nền tảng khác nhau để đa dạng hóa và thu hút thêm người sử dụng.

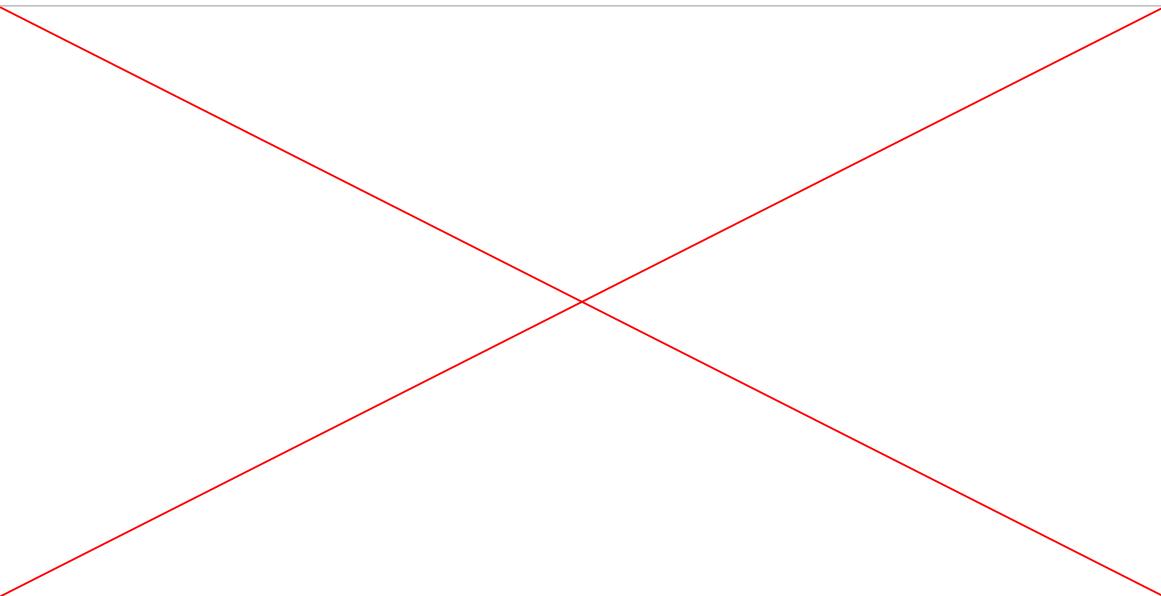
TÀI LIỆU THAM KHẢO

1. Abdul Basit, Maham Zafar, Xuan Liu, Abdul Rehman Javed, Zunera Jalil, Kashif Kifayat, "A comprehensive survey of AI-enabled phishing attacks detection techniques", 2020.
2. Anti-Phishing Working Group (APWG), "Phishing Activity Trends Report, 1st Quarter 2022", 2022.
3. Kang Leng Chiewa , Choon Lin Tan, KokSheik Wong , Kelvin S.C. Yong, Wei King Tiong, "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system", 2019.
4. Lizhen Tang , Qusay H. Mahmoud, "A Survey of Machine Learning-Based Solutions for Phishing Website Detection", 2021.
5. VISHAKHA PRASHANT RATNAPARKHI, SAHIL SIDDHARTH JAMBULKAR, "FRAMEWORK FOR DETECTION AND PREVENTION OF PHISHING WEBSITE USING MACHINE LEARNING APPROACH", 2020

6. Shafaizal Shabuddin, Nor S Sani, Mohd Aliff, "Feature Selection for Phishing Website Classification", 2020.
7. Ammara Zamir, Hikmat Ullah Khan and Tassawar Iqbal, Nazish Yousaf, Farah Aslam, Almas Anjum, Maryam Hamdani, "Phishing web site detection using diverse machine learning algorithms", 2020.
8. Jian Mao, Jingdong Bian, Wenqian Tian, Shishi Zhu, Tao Wei, Aili Li, Zhenkai Liang, "Phishing page detection via learning classifiers from page layout feature", 2019.
9. Ariyadasa, Subhash; Fernando, Shantha; Fernando, Subha, "Phishing Websites Dataset", Mendeley Data, V1, doi: [10.17632/n96ncsr5g4.1](https://doi.org/10.17632/n96ncsr5g4.1), 2021.
10. Abdelhakim Hannousse, Salima Yahiouche, "TOWARDS BENCHMARK DATASETS FOR MACHINE LEARNING BASED WEBSITE PHISHING DETECTION: AN EXPERIMENTAL STUDY", 2020.
11. Rami M. Mohammad, Fadi Thabtah, Lee McCluskey, "Phishing Websites Features", 2015.

12. Sai Nikhilesh Kasturi, "XGBOOST vs LightGBM: Which algorithm wins the race !!!", 2019.
13. Joseph Rocca, "Ensemble methods: bagging, boosting and stacking", retrieve from:
<https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205> , 2019
14. V. H. Tiệp, "Machine learning cơ bản," 2019.
15. Pham Minh Hoang , "Ensemble learning và các biến thể (P1)", Retrieve from:
<https://viblo.asia/p/ensemble-learning-va-cac-bien-the-p1-WAyK8oAkKxX> ,2020
16. Ting, K.M. & Witten, I.H, "Stacked generalization: when does it work?", 1997.

DEMO ỨNG DỤNG





THANKS FOR WATCHING

Do you have any questions?