

AI 기반 웹 피싱 탐지 사이트 구축

Development of an AI-based Web Phishing Detection Website

팀 명	딱걸렸조
지도교수	유승재
팀 원	오 준 석 (팀장) 권 기 범 (팀원) 김 응 찬 (팀원) 김 현 진 (팀원)

2025. 11. 4

중부대학교 미래융합공학부 정보보호학전공

목 차

1. 서론
 - 1.1 연구 배경
 - 1.2 연구 필요성
 - 1.3 연구 목적 및 주제선정
2. 관련 연구
 - 2.1 피쳐 설계 연구 분석
 - 2.2 XGBoost
 - 2.3 Flask
3. 본론
 - 3.1 시스템 구성
 - 3.2 피쳐 선정 기준 및 설계 방향
 - 3.3 데이터 처리 및 피쳐 추출 과정
 - 3.4 머신러닝 모델 설계 및 학습
 - 3.5 세계 지도 기반 피싱 탐지 시각화
4. 결론
 - 4.1 결론 및 기대효과
 - 4.2 향후 계획
5. 참고 자료
6. 별첨
 - 6.1 발표 PPT
 - 6.2 소스 코드

1. 서론

1.1 연구 배경

최근 디지털 금융 서비스와 전자상거래의 급격한 확산으로 인해 사용자의 온라인 활동이 폭발적으로 증가하고 있다. 이에 따라 피싱(Phishing) 공격, 스팸 문자, 악성 링크 등을 통한 개인정보 유출 및 금전적 피해 사례도 지속적으로 발생하고 있다.

특히 공격자들은 정교한 URL 조작과 위장된 도메인을 활용하여 사용자가 정상 사이트로 착각하도록 유도하고 있으며, 이러한 피싱 사이트는 기존의 블랙리스트 기반 탐지 기법만으로는 신속한 대응이 어렵다는 한계를 가지고 있다.

이에 본 연구에서는 머신러닝 기반 인공지능(AI) 기법을 활용하여 웹 피싱 URL을 분석 및 분류하는 시스템을 구축하고자 한다. 본 시스템은 URL 문자열 및 구조적 특징을 기반으로 주요 피처(Feature)를 추출하고, 학습된 XGBoost 모델을 통해 피싱 가능성을 예측함으로써 사용자에게 신뢰성 높은 탐지 결과를 제공하는 것을 목표로 한다.

1.2 연구 필요성

기존의 피싱 탐지 기법은 블랙리스트 데이터베이스를 사용하기 때문에, 새로운 패턴의 피싱 사이트나 신생 도메인에 대한 탐지율이 낮다는 한계가 존재한다. 또한, 실시간 대응이 불가능하거나 탐지 정확도가 낮은 문제가 존재한다.

따라서 다양한 URL 기반의 피처를 추출하고, 머신러닝 모델을 통해 피싱 가능성을 학습, 예측하는 시스템이 필요하다. 이러한 시스템은 단순한 문자열 비교를 넘어, 도메인의 구조적 패턴과 통계적 특성을 기반으로 피싱 행위를 식별할 수 있어, 기존 블랙리스트 탐지 방식보다 높은 정확도와 효율성을 기대할 수 있다.

본 연구는 피처 엔지니어링을 중심으로 한 AI 학습 모델 설계를 통해, 피싱 URL의 패턴을 정량적으로 분석하고, 웹 기반 서비스 형태로 사용자에게 간편하게 제공될 수 있는 실용적 탐지 시스템 구축에 중점을 둔다.

1.3 연구 목적 및 주제선정 이유

본 연구는 AI 기술을 활용하여 피싱 사이트를 탐지하는 실용적 보안 시스템을 구현함으로써, 급증하는 사이버 보안 위협에 대응하고 학문적·산업적 가치 모두를 달성하는 것을 목표로 한다. 본 연구 주제인 「인공지능 기반 웹 피싱 탐지 사이트 구축」은 다음과 같은 이유로 선정되었다.

1) 사회적 필요성

피싱 공격으로 인한 개인정보 유출 및 금융 피해가 매년 증가하면서 단순 차단이 아닌 AI 기반의 능동적 탐지 기술이 필요하다는 점에서 사회적 가치가 높다고 판단하였다.

2) 기술적 도전성

단순한 웹 개발 프로젝트가 아니라, 머신러닝 모델 설계와 피처 엔지니어링을 결합한 보안 기술 연구라는 점에서 기술적 깊이가 있다.

3) 확장 가능성

웹 서비스에 국한되지 않고 실제 보안 솔루션으로 발전할 수 있는 잠재력을 지닌다.

2. 관련 연구

2.1 피쳐 설계 연구 분석

기존의 피싱 URL 탐지 연구들은 대부분 URL의 단순한 구조적 특성에 기반하여 탐지 모델을 구축해왔다. 그러나 이러한 단순 피쳐는 피싱 사이트의 다양하고 교묘한 형태를 잡기에는 어려운 한계가 있다.

이를 개선하기 위해 김영준·이재우(2022)의 연구에서는 「URL 주요특징을 고려한 악성 URL 머신러닝 탐지모델 개발」에서는 문자열 특성 외에 **URL의 생성 시기, 도메인 등록정보, DNS 비정상 패턴** 등 심화된 피쳐 그룹을 새롭게 제안하였다. 해당 연구에서는 총 43개의 주요 피쳐를 3가지 범주로 구분하여 제시하였으며, 각각의 특징은 다음과 같다.

1) URL Days

도메인의 생성일·갱신일·만료일을 기반으로, 사이트의 생명 주기 및 지속성을 분석하여 악성 여부를 판단하는 피쳐이다.

2) URL Words

Netname과 ASN(자율 시스템 번호) 정보를 이용해 등록 기관과 IP 네트워크 연관성을 분석하는 피쳐이다.

3) URL Abnormal

DNS 질의를 통해 TTL(Time To Live) 값과 네임 서버 일치 여부 등을 측정하여 비정상적인 도메인 설정 여부를 탐지하는 피쳐이다.

이 연구는 Random Forest, Gradient Boosting 등 트리 기반 알고리즘을 활용하였으며 특히 Random Forest 모델에서 *정확도 **98.5%**, F1-score **98.5%***로 높은 성능을 기록하였다. 또한 피쳐 중요도 분석 결과, Domain Create Day, Domain Expiration Day, DNS TTL과 같은 도메인 기반 속성이 탐지 정확도 향상에 큰 기여를 하였다.

이 논문은 기존의 단순 문자열 피쳐만으로는 한계가 있었던 피싱 탐지 문제에 ‘시간적·네트워크적 맥락 정보’를 결합하여 정확도를 향상시킨 사례로, 본 연구의 피쳐 설계 방향에도 직접적인 참고가 되었다.

다만, 실제 시스템 구현 과정에서는 WHOIS 및 DNS 질의 오류, 응답 지연 등으로 인해 일부 도메인 기반 피쳐의 활용에 기술적 제약이 존재하였다. 이에 따라 본 연구에서는 해당 요소들을 보조 피쳐로 분류하고, URL 구조적·통계적 특성 중심의 피쳐 엔지니어링에 초점을 맞추어 모델을 설계하였다.

2.2 XGBoost 모델 설명

본 절에서는 본 연구의 탐지 모델로 사용된 XGBoost 알고리즘의 구조와 작동 원리를 설명한다. XGBoost는 Gradient(경사) Boosting 계열의 머신러닝 알고리즘으로 여러 개의 의사결정나무를 순차적으로 학습시켜 예측 오차를 단계적으로 줄이는 **앙상블** 형태의 학습 모델이다. 경사는 오차의 방향을 의미하며 오차를 줄이는 방향으로 조금씩 가중치를 업데이트 하는 방식이다. XGBoost는 이 효과를 더 빠르고 효율적으로 만든 모델로 그 구조를 최적화해서 더 빠르고 안정적으로 학습하게 만든 업그레이드 버전 모델이다.

XGBoost의 주요 특징은 다음과 같다.

1) 고속 학습(High Speed)

내부적으로 병렬 연산과 캐시 최적화 기법을 사용하여 학습 속도가 매우 빠르다. 대용량 데이터셋에서도 효율적으로 모델을 학습시킬 수 있다.

2) 과적합 방지(Regularization)

모델의 복잡도를 제어하는 정규화항(regularization term)을 포함하고 있어 훈련 데이터에 과도하게 맞추는 현상을 효과적으로 방지한다.

3) 결측치 자동 처리(Missing Value Handling)

일부 피처가 누락된 경우에도 자동으로 최적의 경로를 찾아 학습을 진행하므로 데이터의 불완전성에도 비교적 안정적인 성능을 유지한다.

4) 피처 중요도 계산(Feature Importance)

학습된 모델은 각 피처가 탐지 결과에 미친 기여도를 수치적으로 계산할 수 있어 결과 해석이 용이하다.

모델 선택(XGBoost)의 이유

기존 연구에서는 Random Forest가 가장 높은 정확도를 보였으나, 본 연구에서는 학습 효율성과 피처 중요도 시각화 측면에서 우수한 XGBoost(eXtreme Gradient Boosting) 알고리즘을 채택하였다.

XGBoost는 Gradient Boosting 계열 중에서도 과적합 방지, 학습 속도 향상, 피처 중요도 해석 측면에서 우수한 성능을 보여 보안 서비스와 같이 신뢰성이 중요한 분야에서 적합한 모델로 평가된다.

또한 트리 기반 알고리즘이므로 각 피처의 기여도(Feature Importance)를 정량적으로 분석할 수 있어 '왜 해당 URL이 피싱으로 판단되었는가'를 시각적으로 설명할 수 있다.

2.3 Flask 설명

본 연구에서는 학습된 피싱 탐지 모델(XGBoost)을 실시간 웹 서비스 형태로 제공하기 위해 Python 기반의 경량 웹 프레임워크인 Flask를 사용하였다.

Flask는 WSGI(Web Server Gateway Interface) 기반으로 동작하며, 가볍고 유연한 구조를 갖추고 있어 머신러닝 모델과의 연동에 적합하다. 특히 복잡한 설정이 필요하지 않으며, 간단한 라우팅과 REST API 구성을 통해 모델을 바로 웹 서비스로 배포할 수 있다는 장점이 있다. 본 연구에서는 Flask를 사용하여 다음과 같은 기능을 구현하였다.

1) URL 입력 인터페이스 제공

사용자가 웹 페이지에서 URL을 입력하면 Flask 서버가 이를 수신한다.

2) 피처 추출 및 모델 호출

입력된 URL은 feature_builder 모듈을 통해 피처가 추출되고, 학습된 XGBoost 모델을 호출하여 예측을 수행한다.

3) 결과 반환 및 시각화

예측 결과(정상/피싱)는 Flask의 템플릿 엔진을 통해 웹 화면에 즉시 표시된다.

3. 프로젝트 내용

3.1 시스템 구성 개요

본 연구에서 구축한 인공지능 기반 웹 피싱 탐지 시스템은

Feature Extraction - Machine Learning Model - Database - Web Interface로 구성된 4단계의 통합 구조를 가진다.

사용자가 웹 페이지를 통해 URL을 입력하면 서버 측에서는 입력된 주소를 Feature Builder 모듈을 통해 분석하여 각 URL의 통계적·구조적 속성을 수치화한다. 이 피쳐들은 학습된 XGBoost 모델의 입력값으로 전달되어 피싱 여부(정상: 0, 피싱: 1)를 예측하며, 예측결과를 MySQL 데이터베이스에 자동으로 저장된다.

전체 시스템은 Flask 웹 프레임워크를 중심으로 구성되어 있으며, 피쳐 추출 → 모델 예측 → 결과 반환까지의 과정이 하나의 API 파이프라인으로 통합되어 있다. 이러한 구조를 통해 사용자는 브라우저에서 URL을 입력하기만 해도 즉시 피싱 가능성을 분석할 수 있으며, 관리자는 DB를 통해 분석 이력과 피쳐별 로그를 추적·관리할 수 있다.

3.2 피쳐 선정 기준 및 설계 방향

기존 연구에서 제안된 피쳐 중, URL 구조 및 도메인 분석만으로 추출 가능한 주요 항목들을 중심으로 선별하였다.

1) URL 구조 기반

URL 길이, 서브도메인 개수, 하이픈 포함 여부

2) 보안 속성 기반

HTTPS 사용 여부, 파일 확장자, 단축 URL(Tiny URL)

3) 통계적 특성 기반

특수문자 비율, 숫자 비율 등

이들은 모두 '접근성이 높고 탐지 근거로 해석 가능한 속성'을 기준으로 선정되었으며, 모델의 설명 가능성을 확보하기 위해 수치화 가능한 형태로 변환하였다.

3.3 데이터 처리 및 피쳐 추출 과정

본 시스템의 데이터 전처리와 피쳐 추출은 Python 기반 모듈 feature_builder.py에서 수행된다. 해당 모듈은 URL 구조를 분석하여 문자열·구조·리스크 기반 정보를 중심으로 약 10개 이상의 피쳐를 생성한다. 아래는 이를 카테고리별로 구분한 피쳐 목록이다.

1) 문자열 기반 피쳐

URL·도메인 길이, 서브도메인 개수, 특수문자·숫자 비율, 하이픈·언더스코어 포함 여부 등을 기반으로 URL의 복잡도를 수치화하고, 피싱 URL의 난독화 정도를 분석한다.

2) 도메인 및 보안 피쳐

HTTPS 존재 여부, 파일 확장자, 단축 URL(Tiny URL) 사용 여부, DNS TTL, 도메인 갱신·만료일 등을 분석하여 도메인과 연결된 보안 요소를 평가하고, URL의 신뢰도와 정상 사이트 여부를 판단한다.

각 피처는 수치형 형태로 변환되어 Pandas DataFrame 구조로 저장되며, 이후 CSV 파일(phish.csv) 형태로 관리되어 모델 학습 단계에서 활용된다.

피처 추출 과정에서는 불필요한 파라미터와 공백을 제거하는 정규화(Normalization) 및 로그 스케일링(Log-Scaling) 기법을 적용하여 극단값(Outlier)으로 인한 학습 편향을 최소화하였다.

※ WHOIS 기반 피처(도메인 생성일, 만료일 등)는 네트워크 질의 오류가 빈번하게 발생하는 점을 고려하여 현재는 보조 피처(secondary feature)로 분류하였으며, 모델의 핵심 입력값은 URL 구조적·통계적 피처로 구성하였다.

3.4 머신러닝 모델 설계 및 학습

본 시스템의 탐지 모델은 XGBoost 알고리즘을 기반으로 설계하였다. 이는 수치화된 피처(feature) 값을 입력으로 받아 학습하는 트리 기반 부스팅(Boosting) 기법으로, 각 피처가 탐지 결과에 미치는 영향을 수치적으로 계산하여 예측 성능을 점진적으로 향상시키는 방식이다.

XGBoost는 학습 속도가 빠르고 과적합 방지에 강하며, 각 피처의 기여도(Feature Importance)를 정량적으로 분석하고 시각화할 수 있다는 장점이 있다.

학습 데이터 구성

본 연구의 학습 데이터는 Phishtank, PhishStorm, KISA의 정보들을 직접 정제하여 사용하였다. 데이터는 URL과 Label 두 개의 컬럼으로 구성되어 있으며, Label 값은 정상(0)과 피싱(1)으로 수동 라벨링하였다.

원본 데이터는 중복 URL 제거, 결측치 보정, 문자열 정규화 등의 전처리 과정을 거쳐 최종적으로 총 약 10만개(정상 약 50,000개 / 피싱 약 50,000개)의 데이터셋(1:1 비율)으로 구성하였다. 데이터셋은 학습:테스트 비율 8:2로 분할하여, 학습 단계에서는 모델의 일반화 성능을 확보하고, 테스트 단계에서는 실제 예측 성능을 평가하였다.

모델 학습 및 결과 분석

모델 학습은 Python 환경에서 수행했고 다음과 같은 하이퍼파라미터를 적용하였다.

1) Learning Rate (0.1)

모델의 학습 속도를 조절하는 하이퍼파라미터로, 최적값에 수렴하는 단계를 제어한다.

2) Max Depth (5)

트리의 최대 깊이를 의미하며, 깊이가 클수록 과적합이 발생하므로 5로 제한하였다.

3) Estimators (100)

학습에 사용되는 트리의 개수로, 정확도·연산 효율의 균형을 위해 100개로 설정하였다.

4) Evaluation Metric (Accuracy, F1-score)

모델의 예측 정확도와 정밀도·재현율의 균형을 평가하기 위한 지표로 사용하였다.

Classification Report

Overall Accuracy: 0.9964 (99.64%)

class	precision	recall	f1-score	support
0 (Normal)	1.00	0.99	1.00	9,193
1 (Phishing)	1.00	1.00	1.00	10,807
accuracy			1.00	20,000
macro avg	1.00	1.00	1.00	20,000
weighted avg	1.00	1.00	1.00	20,000

Feature Importance

feature	importance
subdomain_count	0.8019
hostname_length	0.1010
has_hyphen	0.0473
has_equals	0.0231
special_ratio	0.0091
digit_ratio	0.0065
has_question	0.0062
path_length	0.0016
url_length	0.0026
alpha_ratio	0.0005
has_https	0.0002
has_at	0.0000
has_ip	0.0000

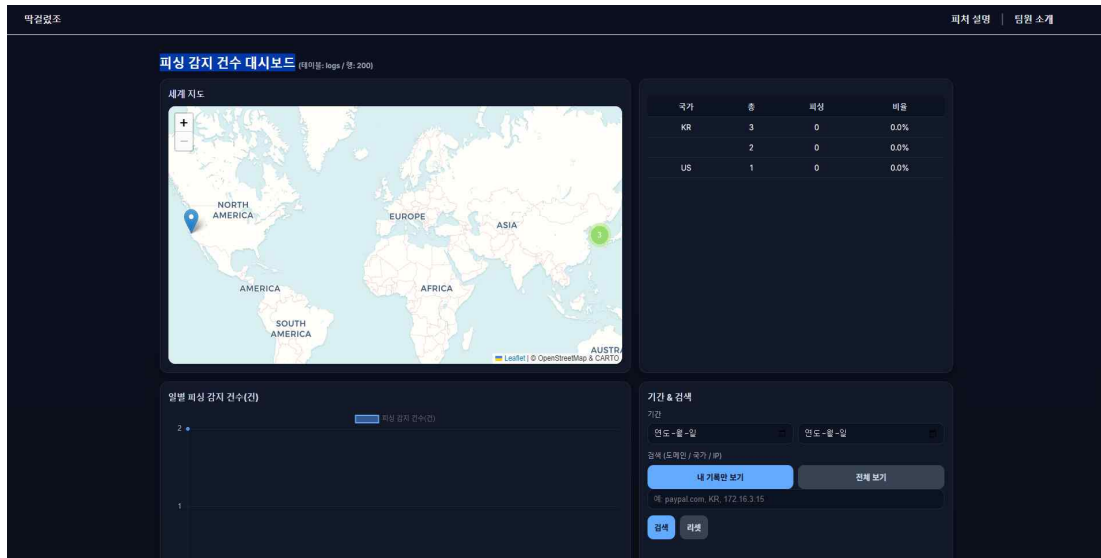
[그림] 모델 학습 결과표

3.5 세계 지도 기반 피싱 탐지 시각화 기능

본 시스템은 URL의 피싱 여부를 판별하는 데 그치지 않고, 탐지된 결과를 **세계 지도 형태로 시각화**하여 국가별 피싱 발생 현황과 분포를 한눈에 파악할 수 있도록 설계하였다.

서버 측에서는 Flask 프레임워크를 기반으로 탐지 결과를 MySQL 데이터베이스에 저장하며, /api/map/points 및 /api/map/country-stats 엔드포인트를 통해 도메인, IP 주소, 위·경도(latitude, longitude), 국가 코드(country code), 탐지 확률 등의 정보를 JSON 형식으로 제공한다. 클라이언트 측에서는 JavaScript 기반 지도 시각화 라이브러리(Leaflet.js, Google Maps API 등)을 활용하여 수집된 데이터를 지도에 마커(marker) 형태로 표시하고, 국가별 평균 피싱 확률과 탐지 비율을 색상 그라데이션으로 표현하였다. 이를 통해 사용자는 대시보드 상에서 실시간으로 수집되는 피싱 URL의 지리적 분포와 활동 밀집 지역을 직관적으로 확인할 수 있다.

또한 관리자 페이지에서는 탐지 로그를 기반으로 국가별 위험 점수(Risk Score)를 자동 계산하여, 피싱 공격의 확산 경향과 위험도를 시각적으로 비교·분석할 수 있도록 구성하였다.



[그림] 피싱 감지 건수 대시보드

4. 결론

4.1 결론 및 기대효과

본 연구에서는 인공지능(AI)과 머신러닝(ML) 기법을 활용하여 웹 피싱 URL을 자동으로 탐지하는 AI 기반 피싱 탐지 웹 시스템을 구현하였다.

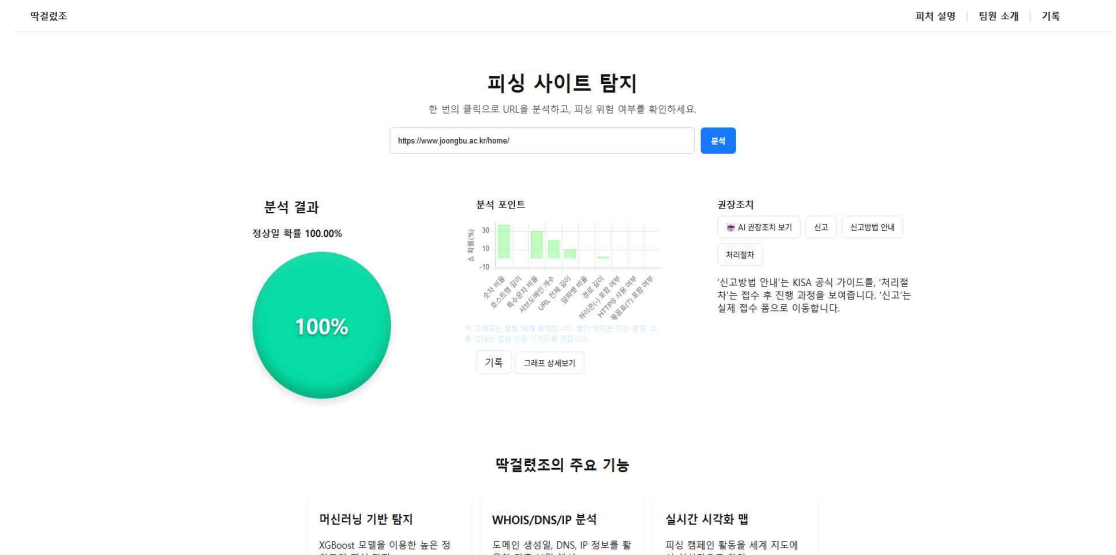
학습 과정에서는 XGBoost 알고리즘을 적용하여 입력된 URL이 정상(0)인지 피싱(1)인지를 정밀하게 예측하도록 설계하였으며, Flask 프레임워크를 이용하여 웹 환경에서 실시간으로 탐지 결과를 확인할 수 있도록 구성하였다. 분석 결과는 MySQL 데이터베이스(DB)에 자동 저장되어 로그 관리 및 재학습에 활용될 수 있도록 설계하였다.

전체 데이터에 대한 탐지 정확도(Accuracy)는 **0.9964(99.64%)**로 나타났으며, 피싱(1)과 정상(0) 모두에 대해 정밀도(Precision), 재현율(Recall), F1-score가 **1.00**에 근접한 높은 성능을 보였다. 피쳐 중요도 분석 결과, subdomain_count(0.8019)가 가장 높은 영향력을 보였으며, 이는 피싱 URL이 일반적으로 서브도메인을 과도하게 포함하거나 비정상적으로 긴 도메인 구조를 갖는 경향이 있음을 보여준다.

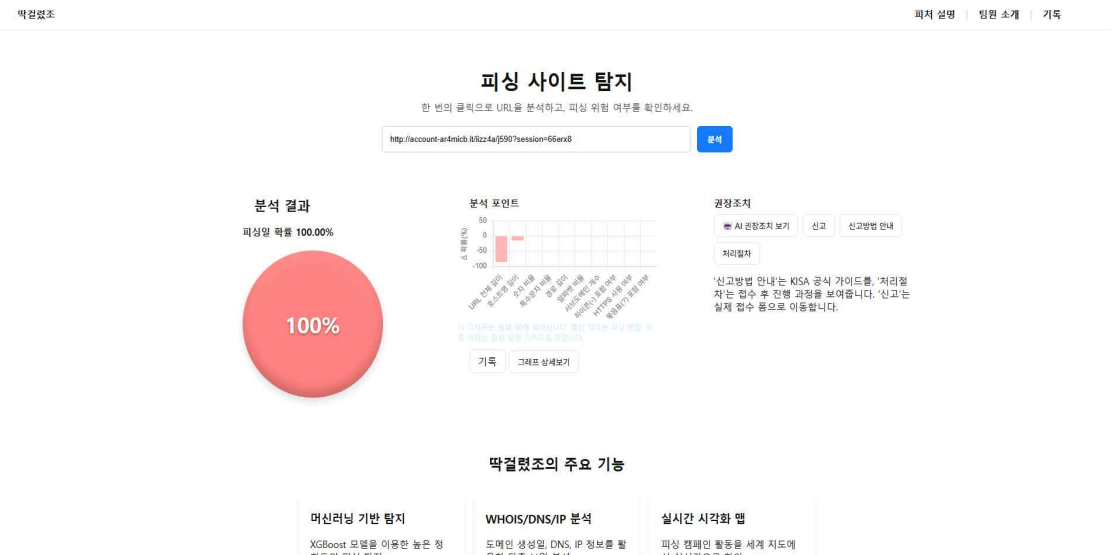
Flask-MySQL-XGBoost 통합 구조를 구축함으로써, 단일 모델 학습을 넘어 사용자가 직접 탐지 결과를 확인할 수 있는 웹 기반 실시간 분석 시스템을 완성하였다. 사용자는 URL 입력만으로 피싱 여부 및 예측 확률을 확인할 수 있으며, 탐지 이력은 DB에 자동 저장되어 로그 분석과 향후 모델 개선에 재활용할 수 있다.

결과적으로 본 시스템은 기존의 블랙리스트 탐지 방식보다 높은 정확도, 유연성, 그리고 설명 가능성(Explainability)을 제공한다. 또한 보안 비전문가도 손쉽게 활용할 수 있는 실용적 피싱 탐지 솔루션으로서의 가능성을 입증하였으며, 피쳐 중심의 머신러닝 접근 방식을 기반으로 로그 분석을 통한 보안 자동화 및 위협 인텔리전스(Threat Intelligence) 확장 가능성도 확인하였다. 더불어, 탐지 결과를 세계 지도 기반으로 시각화하여 국가별 피싱 발생 현황과 위험도를 한눈에 파악할 수 있도록 함으로써, 시스템의 가시성(Visibility)과 분석 효율성(Analytical Efficiency)을 크게 향상시켰다.

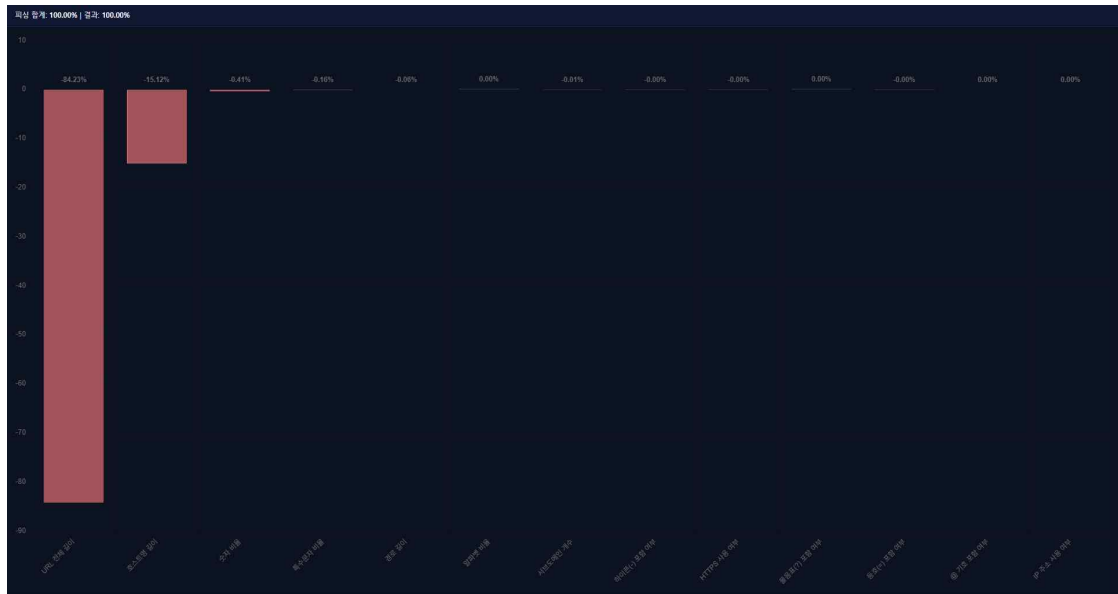
딱결렸조 Web 이미지



[그림 1] 정상 URL 분석 결과 창



[그림 2] 피싱 URL 분석 결과 창



[그림 3] Feature Bars

[그림 3] Feature Bars는 각 피처의 중요도(Feature Importance)를 막대그래프로 시각화한 결과이다.

그래프를 통해 모델이 어떤 피처를 기반으로 URL을 분류하는지 한눈에 확인할 수 있으며, 모든 막대그래프 값의 합은 왼쪽 원 그래프의 전체 비율(정확도 지표)과 일치한다.

4.2 향후 계획

본 연구는 약 10만개의 URL 데이터를 기반으로 학습된 모델을 통해 높은 정확도를 달성하였으나, 향후에는 모델의 데이터 다양성, 학습 자동화, 사용자 배포 서비스를 강화하는 방향으로 발전시킬 계획이다.

1) 데이터셋 확장 및 다양화

국내외 피싱 신고 데이터와 실시간 크롤링을 통해 실제 환경에서 수집 가능한 다양한 URL을 확보하여 모델의 일반화 성능(Generalization)을 강화할 예정이다. 또한 최신 피싱 트렌드를 반영하기 위해 데이터 수집 및 전처리 과정을 자동화(Auto Pipeline)하는 기능을 추가할 계획이다.

2) WHOIS.DNS 기반 피처 복원 및 개선

현재 제외된 도메인 기반 피처를 API 질의 방식으로 안정적으로 복원하고, 도메인 수명(Domain Lifetime), 등록 기관, 네임서버 일치 여부 등 보안 속성을 강화하여 피싱 사이트의 신뢰도 분석(Trustworthiness Analysis) 기능을 향상시킬 예정이다.

3) 웹 서비스 고도화 및 사용자 편의성 강화 (배포)

관리자용 대시보드를 추가하여, 탐지 결과 통계, 주요 피처 기여도, 피싱 패턴 변화 등을 확인할 수 있는 위협 인텔리전스(Threat Intelligence) 시각화 기능을 제공한다.

4) 학문적 및 산업적 확장성

본 연구에서 구축한 AI 피싱 탐지 모델은 향후 이메일·메신저 피싱 탐지 시스템으로 확장 가능하며, 보안 기관이나 금융권의 사이버 위협 인텔리전스(CTI) 플랫폼과 연계하여 실제 산업 현장에서 활용될 수 있는 가능성을 지닌다.

5. 참고 자료

- (1) 김영준, 이재우, 「URL 주요특징을 고려한 악성URL 머신러닝 탐지모델 개발」, 한국정보통신학회논문지, 제26권 제12호
- (2) Kaggle, Phishing Website Dataset (URL classification data)
- (3) T. Chen, C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
- (4) Flask Documentation, Flask Web Framework (Python).
- (5) MySQL Official Documentation, MySQL Reference Manual 8.0.
- (6) PhishTank, Phishing Website Verification Dataset.
- (7) URLHaus, Malicious URL Repository.

6. 별첨

6.1 소개 자료 (폼보드)








AI를 활용한 URL 탐지 및 분석 웹 서비스

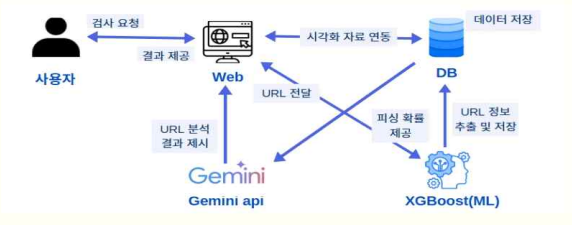
개요

최근 이메일 및 URL을 이용한 피싱 공격으로 인한 보안 위협이 커지고 있음에 따라, 실시간으로 URL의 피싱 여부를 판별해 주는 웹 서비스를 제공한다.

사이트의 URL 구조, 도메인 정보 등 다양한 요소를 종합적으로 분석하여 사용자에게 피싱 여부를 직관적으로 보여준다.

별도의 설치 없이 웹을 통해 누구나 쉽게 이용할 수 있도록 설계되어, 보안 지식이 없는 일반 사용자도 손쉽게 피싱 위험도를 확인할 수 있다.

구상도



2025 정보보호학전공 졸업작품 전시회 딱걸렸조

사이트 소개




기대 효과

사용자가 스스로 의심스러운 링크를 확인하고 안전하게 인터넷을 이용할 수 있는 생활형 보안 환경을 제공한다. 자동 분석과 시각화를 통해 피싱 공격의 특성을 한눈에 파악할 수 있으며, 축적된 데이터는 향후 AI 기반 보안 연구와 교육 자료로 활용될 수 있다.

이를 통해 개인의 보안 의식 향상과 사회 전반의 피싱 피해 예방에 기여할 수 있다.

팀원 소개



팀장 오준석
ML
피쳐코드
코드통합



팀원 권기범
UI/UX
API



팀원 김웅찬
ML
피쳐코드



팀원 김현진
크롤링
DB
API

2025 정보보호학전공 졸업작품 전시회 딱걸렸조

6.2 프로젝트 소개

깃허브 주소

https://github.com/JunSeok5093/graduation_project/tree/main/%EB%94%B1%EA%B1%B8%EB%A0%B8%EC%A1%B0/%EC%A1%B8%EC%9E%91%EC%88%98%EC%A0%95

시연 영상

https://youtu.be/Hle3BTV7t_U

6.2 팀 소개

1. 팀장: 오준석(92003228) - 백엔드(총괄), ML 및 피쳐코드 모델 개발
2. 팀원: 권기범(91913153) - 프론트엔드, UI / UX 개발, 실시간 API 연동
3. 팀원: 김응찬(91913244) - 백엔드, ML 및 피쳐코드 모델 개발, 데이터 전처리
4. 팀원: 김현진(92002339) - 백엔드, 데이터셋 수집, API 연동, MySQL 구축