산업 · 사회 문제 해결을 위한 서울 지역사회 경험학습 공모전

사회문제- 사이버범죄 및 폭력

큐싱 범죄 차단을 위한 AI 기반 QR 코드 URL 분석 및 실시간 경고 시스템 제안

연세대학교 정보대학원

BK21 FOUR

바른 ICT연구소 Barun ICT Research Center

박세연 석사과정
비즈니스 AI 빅데이터 분석 전공

황성아 석사과정
비즈니스 AI 빅데이터 분석 전공

허수정 석사과정
비즈니스 AI 빅데이터 분석 전공

박창릉 KIST
연구원

목차

Part 1-1

제안 배경

큐싱(Qshing) 이란?

- **QR + Phishing**
- **악성 코드**나 **불법 웹사이트**로 유도하는 QR 코드를 통해 **개인정보를 탈취**하거나 **금전적 피해**를 입히는 범죄

'큐싱' 주의보…공유자전거 QR 찍었다가 통장 털려

진영화 기자 cinema@mk.co.kr

입력 : 2024-06-17 17:49:57 수정 : 2024-06-17 19:49:50

자영업자 박 모씨는 최근 소상공인을 상대로 낮은 이자로 대출해주겠다는 이메일을 받았다. 마침 대출을 알아보고 있던 그는 "자세한 내용을 확인하려면 QR코드 촬영 후 전자금융사기 예방 서비스 앱을 설치하라"는 문구를 보고 의심 없이 카메라 앱으로 QR코드를 비춰 애플리케이션(앱)을 다운받았다. 하지만 해당 앱은 개인정보를 빼내가는 해킹 앱이었고 박씨가 입력한 공인인증서 비밀번호 등 금융정보가 고스란히 해커의 손으로 넘어갔다. 며칠 뒤 박씨의 통장에서는 1000만원이 빠져나갔다.

출처: 매일경제(24.06.17)
https://www.mk.co.kr/news/society/11043900

(소비자유의사항) 출처가 불분명한 QR코드*는 스캔하지 않도록 유의

* 웹 주소와 달리 **악성 QR코드**를 **시각적**으로 **구별**하는 것은 거의 **불가능**

이미 QR코드를 **스캔**한 경우 접속되는 **웹 주소**를 **한번 더 확인**하고,
의심스러운 앱 설치 또는 **개인·금융정보**를 **입력**하지 **않도록** 유의

출처: 금융감독원, QR코드 이용 금융사기(큐싱) 관련 대응 (24.06.18)

큐싱 범죄의 증가

QR 코드가 간편 결제, 인증 등 다양한 분야에서 널리 활용되고 있어 더욱 확산

악성 QR 코드 식별의 어려움

일반 사용자가 악성 QR 코드를 사전에 식별하기가 매우 어려움

Part 1-2

필요성

피해 예방

- 일반 사용자의 악성 URL 사전 식별 불가능
- 금전적 손실과 함께 개인 정보 유출의 2차 피해 가능성

기존 보안 솔루션의 한계

- QR 코드 스캔 후 웹사이트 접속 시 실시간 URL 분석 기능이 없어 즉각적 경고 및 차단 불가능
- 실시간 URL 분석 및 악성 여부 판단 시스템 필요

AI 활용 필요성

- 범죄 패턴의 다양화로, 기존 정적 보안 기술로 탐지 불가능
- URL 이상 패턴 학습 및 새로운 악성 URL 실시간 탐지를 위한 AI 활용 필요

Part 1-3

프로젝트 개요

분석 프로세스

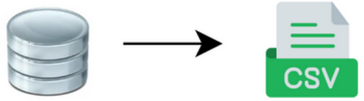
1. 데이터 수집

Mendeley Data

'Phishing Website Dataset'

Non-phishing(0): 49,737

Phishing(1): 29,964



* 라벨값 불균형 해결

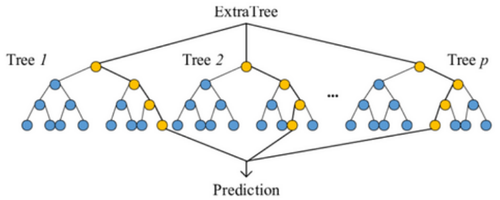
2. 데이터 전처리

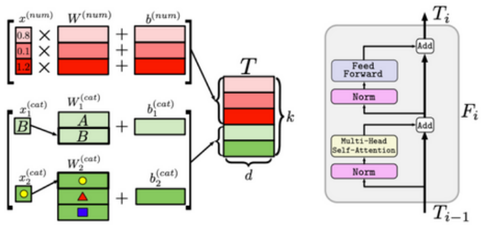
URL 기반 다양한 파생변수 생성

3. EDA

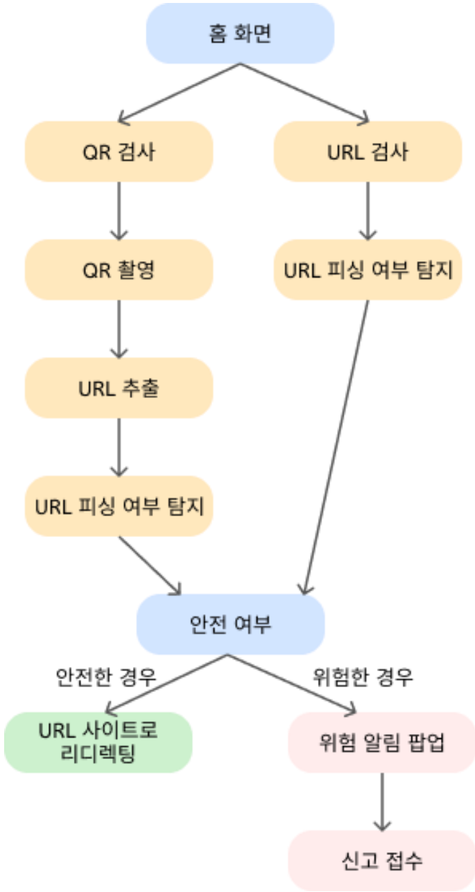
피싱 여부에 따른 독립변수 특징 분석

4. 모델링





시스템 프로세스



주요 기여점

- **QR 코드 URL 악성 탐지:** QR 코드로 직접 확인할 수 없는 악성 URL을 탐지할 수 있는 모델 개발
- **다양한 파생 변수 생성:** URL을 분석하여 악성 URL을 식별하기 위한 다양한 파생 변수 생성
- **URL 특징 분석:** 악성 URL과 정상 URL의 특징들을 시각화를 통해 확인
- **시스템 UI 제공:** 구축한 모델을 실질적으로 활용할 수 있는 시스템 UI 제공

Part 2-1

데이터 및 전처리

데이터 소개

Mendeley Data

Phishing Websites Dataset

Published: 17 November 2021 | Version 1 | DOI: 10.17632/n96ncsr5g4.1
Contributors: Subhash Ariyadasa, Shantha Fernando, Subha Fernando

피싱 웹사이트 탐지를 위한 URL 정보와 특징을 포함한 데이터셋

Ariyadasa, Subhash; Fernando, Shantha; Fernando, Subha (2021), "Phishing Websites Dataset", Mendeley Data, V1, doi: 10.17632/n96ncsr5g4.1

데이터 구조

red_id	레코드 번호
url	웹페이지의 URL
website	HTML 페이지 파일명
result	URL의 피싱 여부 (0:합법적, 1: 피싱
created_date	다운로드한 날짜

수집 기간

2020-2021년

수집 경로

합법 사이트: Google 검색, Ebbu2017 Phishing Dataset
불법 사이트: PhishTank, OpenPhish, PhishRepo

데이터 형식 변환

SQL형식 to CSV 파일

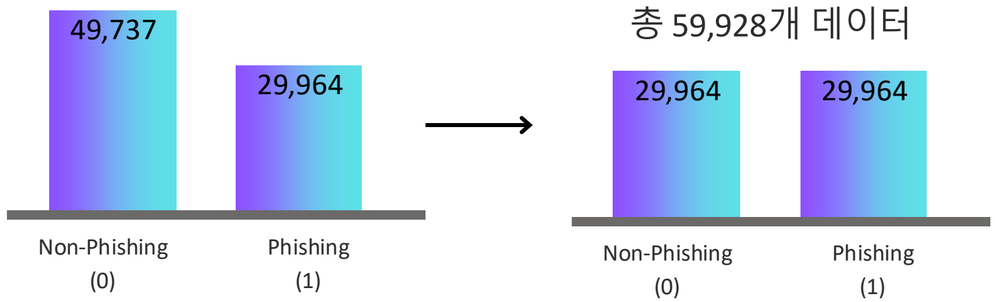
```
INSERT INTO `index` (`rec_id`, `url`, `website`, `result`, `created_date`
 VALUES (1, 'http://intego3.info/EXEL/index.php', '1613573972338075.html', 1, '2021-02-17 20:29:32'), ...
```

Python →

	A	B	C	D	E
1	rec_id	url	website	result	created_date
2	1	http://intego3.info/	1613573972	1	2021.2.17 20:29
3	2	https://www.matho	1635698138	0	2021.10.31 16:35
4	3	https://www.compu	1635699228	0	2021.10.31 16:53
5	4	https://www.investc	1635750062	0	2021.11.1 12:31

데이터 불균형 해결

Random Undersampling



Part 2-1

데이터 및 전처리

파생변수 생성

피싱 여부에 영향이 큰 URL 특징을 토대로 각 파생변수를 생성하고 값을 부여

**여부에 관한 변수인 경우 0이 아니오, 1이 예를 의미함

URL: https://sndc-card.com.od200z.cn/smbc/smbcupdatebill.php

common_term_sum 공통 키워드 개수	3	hostname_length 호스트 이름 길이	23	digit_ratio URL 내 숫자 비율 = 숫자길이 / URL길이	0.054	has_suspicious_text 의심확장자 포함 여부	1

*공통 키워드 : www, com, http, //

url_length URL 전체 길이	55	has_ip IP 주소 포함 여부	0	has_port 포트번호 포함 여부	0

*e.g. 192.168.1.1

*e.g. 8080

*의심확장자:
'.exe', '.bat', '.cmd', '.msi', '.apk', '.sh', # 실행 파일
'.js', '.vbs', '.cgi', '.php', '.asp', '.aspx', # 스크립트 파일
'.zip', '.rar', '.7z', # 압축 파일
'.doc', '.docx', '.xls', '.xlsx', '.pdf' # 문서 파일

Created_date: 2021/6/1-15

created_year 생성년도	2021	created_month 생성월	6		
created_day 생성일	1	created_hour 생성시간	15		

Part 2-1

데이터 및 전처리

파생변수 생성

피싱 여부에 영향이 큰 URL 특징을 토대로 각 파생변수를 생성하고 값을 부여

URL https://sndc-card.com.od200z.cn/smbc/smbcupdatebill.php

num_subdomains
서브 도메인 개수
2

has_prefix_suffix
하이픈 존재 여부
1

special_char_sum
특수문자 개수
10

*특수문자:
'.', '-', '@', '?', '&', '=', '_', '~', '%', '/', '*', ':', ',', ';', '$', '+', '#', '(', ')', '[', ']'

phish_word
피싱 단어 포함 개수
1

*특수문자:
'login', 'secure', 'bank', 'verify', 'free', 'account', 'update', 'validate', 'authenticate', 'password', 'signin', 'signup'

abnormal_subdomain
비정상 서브 도메인
0

*e.g.
https://v123.example.com

is_shortened
단축 서비스 여부
0

*e.g.
https://forbusiness.snapchat.com/

brand_in_domain
브랜드명 포함 개수
0

*브랜드명:
'google', 'paypal', 'amazon', 'facebook', 'microsoft', 'apple', 'netflix', 'linkedin', 'twitter', 'instagram', 'youtube'

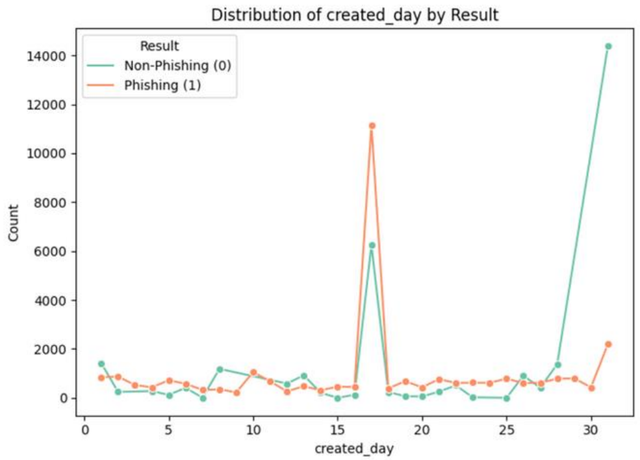
abnormal_tld_in_path
비정상 TLD 위치
0

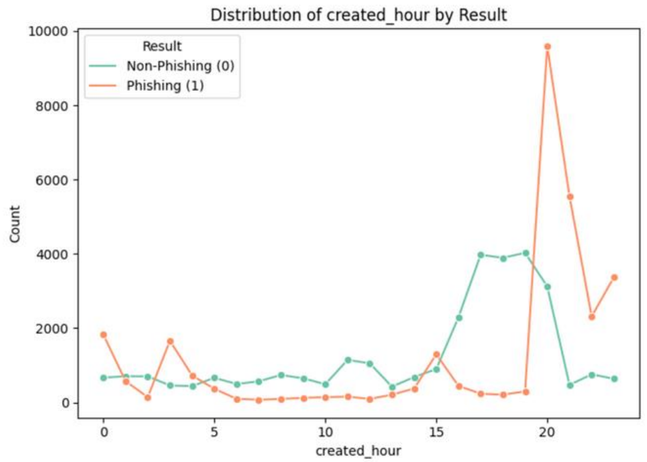
*TLD(Top-Level Domain):
도메인의 최상위 레벨을 나타내며, URL의 도메인 끝에 위치

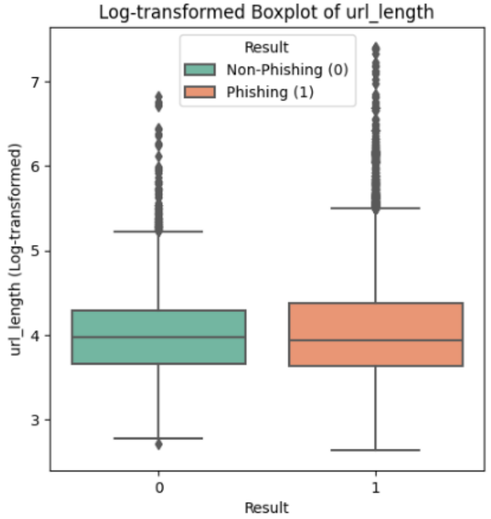
result
피싱여부
1

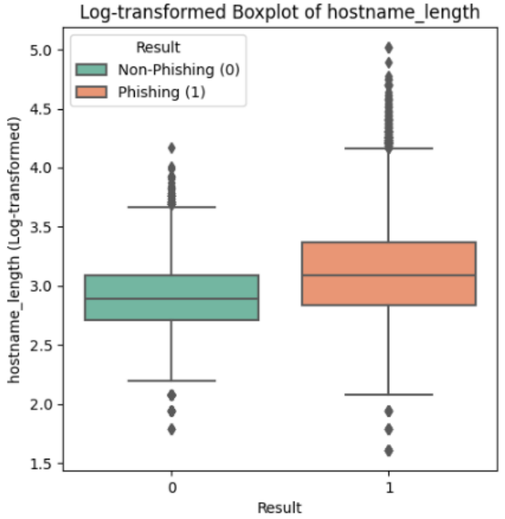
Part 2-2

EDA









피싱 URL 생성은 30일에 급증

30일은 월말 청구서나 급여와 같은 민감한 주제를 악용한 피싱 시도가 많을 것으로 예상

피싱 URL은 주로 20시 이후에 생성

사용자 활동이 감소하고 경계심이 낮아지는 야간 시간대를 타겟으로 삼는 경향이 높음

피싱 URL은 더 긴 길이와 호스트 이름을 가짐

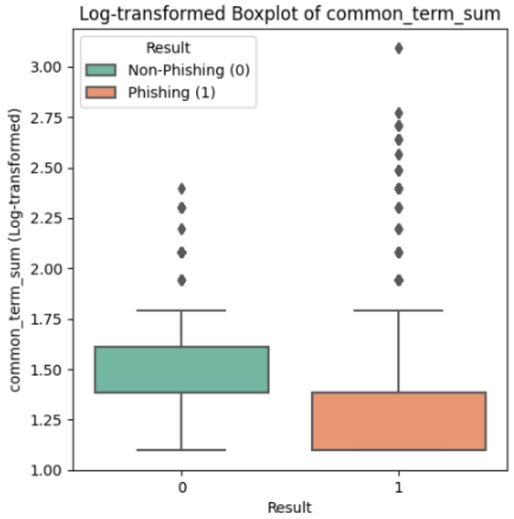
길이가 긴 URL로 악성 코드를 포함한 복잡한 경로를 감추려는 시도

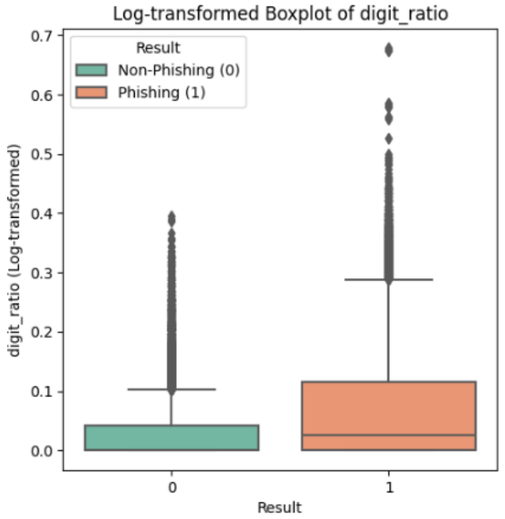
피싱 URL은 공통 용어 사용 빈도가 낮음

혼란을 주기 위해 의도적으로 비정상적이고 독특한 단어를 포함할 가능성이 높음

피싱 URL은 숫자 비율 높음

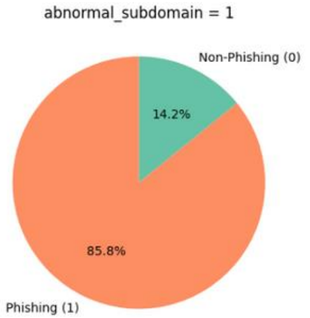
숫자는 직관적으로 이해하기 어렵기 때문에 숫자를 활용해 정교한 속임수 사용

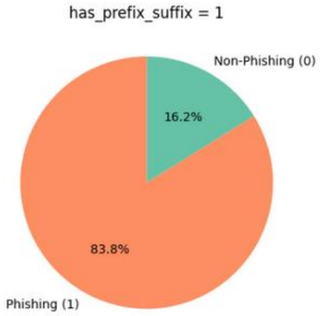


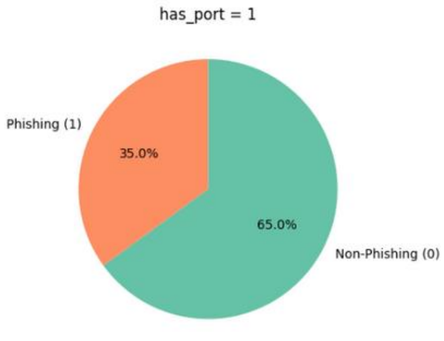


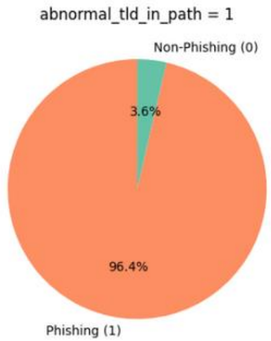
Part 2-2

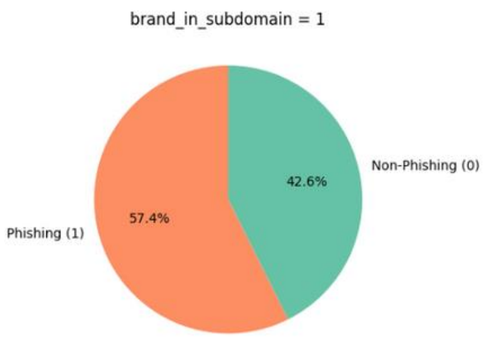
EDA

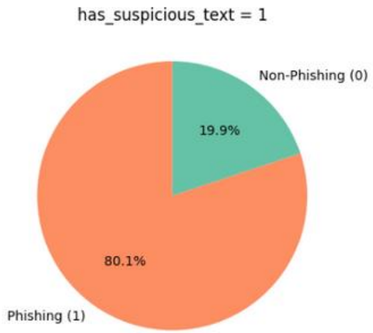


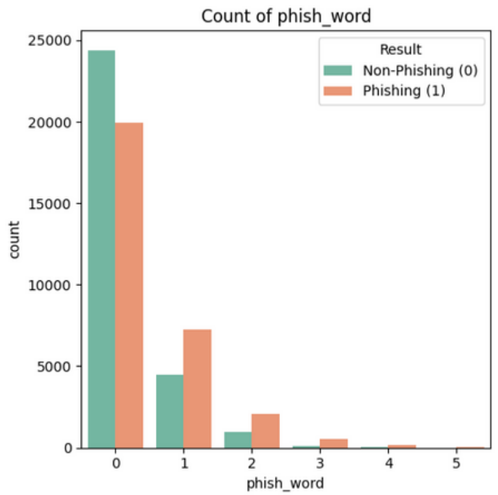


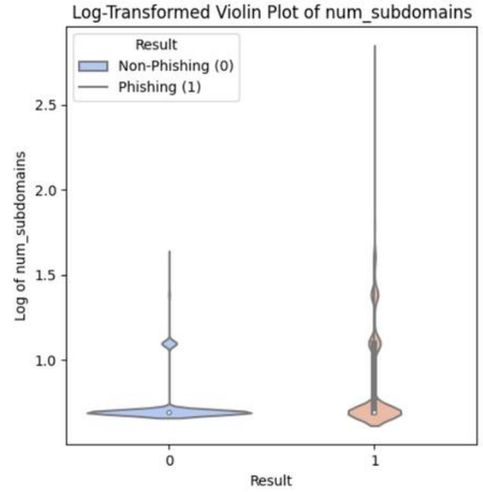












피싱 URL 중 약 85%가 비정상적인 서브도메인 및 prefix-suffix 구조 포함

피싱 URL은 비정상적인 서브도메인 혹은 '-' 같은 구분자를 활용하여 정상처럼 보이도록 위장

피싱 URL의 96.4%가 경로에 비정상적인 TLD(최상위 도메인)를 포함

경로 내 비정상 TLD를 포함하여 사용자 혼란을 유도

피싱 URL의 57.4%가 서브도메인에 브랜드명을 포함

유명 브랜드를 서브도메인에 포함하여 사용자가 신뢰하도록 유도

피싱 URL의 80.1%가 의심스러운 텍스트와 1개 이상의 피싱 관련 단어 포함

피싱 URL은 특정 키워드를 활용하여 사용자 행동을 유도

피싱 URL은 정상 URL에 비해 서브 도메인 수가 더 많음

피싱 URL은 서브도메인을 복잡하게 만들어 사용자를 혼란스럽게 함

Part 2-3

모델링

모델 소개

Machine Learning

Gradient Boosting

약한 학습기를 순차적으로 학습해 오류를 보완하며, 높은 예측 성능을 보이는 앙상블 기법

Extra Trees

결정 트리 분할을 무작위로 선택해 더 빠르고 다양성을 높인 앙상블 모델

Random Forest

여러 결정 트리의 결과를 앙상블하여 안정적이고 과적합에 강한 모델

Decision Tree

데이터를 조건에 따라 반복적으로 분리해 예측을 수행하는 단순한 트리 모델

Deep Learning

FT-Transformer

정형 데이터를 처리하기 위해 Transformer 구조를 활용한 딥러닝 모델

Part 2-3

모델링

모델링 결과

Model	Accuracy	Recall	Precision	F1-Score
Gradient Boosting	0.9017	0.9026	0.8948	0.9106
Random Forest	0.9845	0.9845	0.9832	**0.9858**
Decision Tree	0.9844	0.9844	**0.9849**	0.9838
Extra Trees	**0.9846**	**0.9846**	0.9840	0.9852
FT-Transformer	0.9774	0.9787	0.9760	0.9774

하이퍼파라미터 튜닝

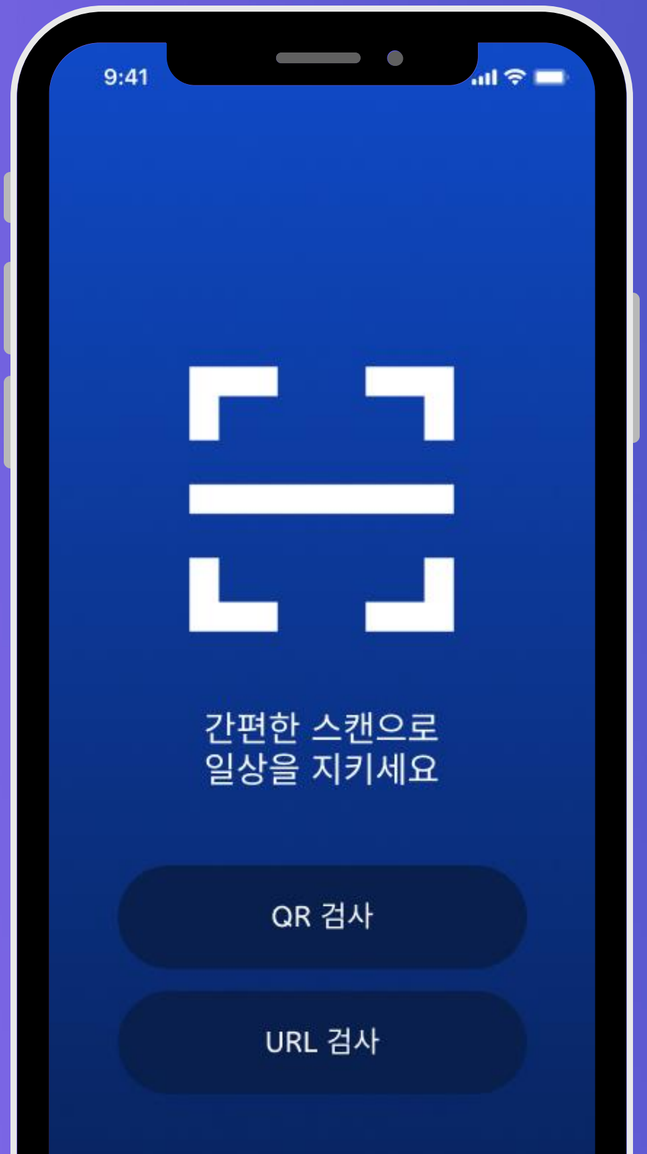
Accuracy	Recall	Precision	F1-Score
0.9908	0.9931	0.9885	0.9908

max_depth : 47
max_features : 0.8578
min_samples_leaf : 1
min_samples_split : 2
n_estimators : 300

max_depth: 트리의 최대 깊이
max_features: 각 분할에서 고려할 특성의 최대 수
min_samples_leaf: 리프 노드에 있어야 하는 최소 샘플 수
min_samples_split: 노드를 분할하기 위해 필요한 최소 샘플 수
n_estimators: 생성할 결정 트리의 수

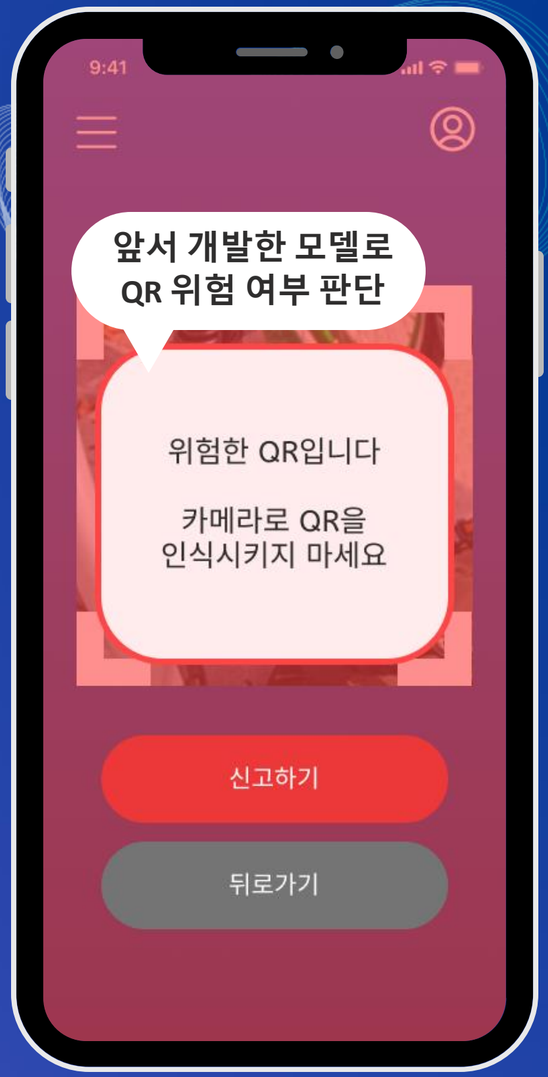
Part 3-1

시스템



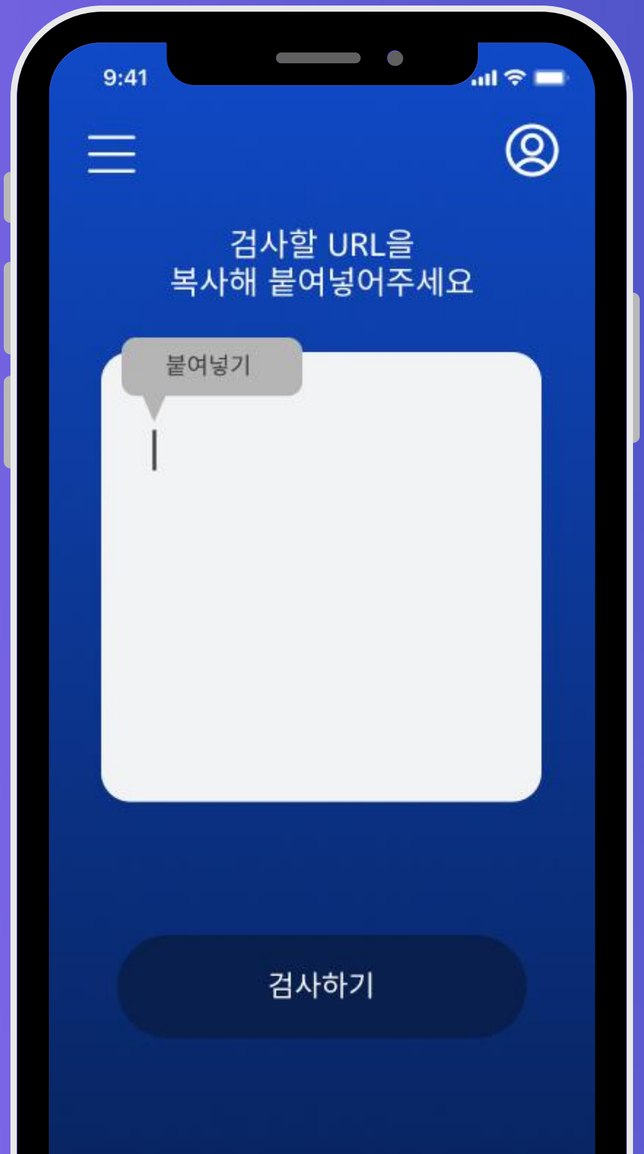
간단하게
QR 코드 이미지
업로드만으로
위험 QR을 식별

위험한 QR인 경우
알림이 뜨며
신고 가능



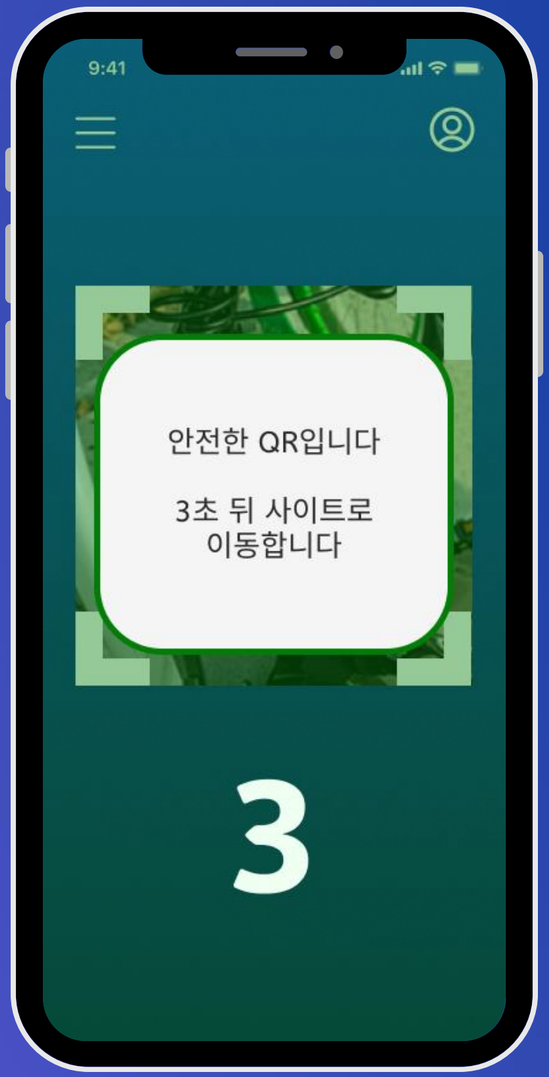
Part 3-1

시스템



URL 또한
CTRL+C,V만으로
안전 여부를 식별

안전한 URL이면
3초 뒤
사이트로 리디렉션



Part 3-2

기대효과

재산과 일상의 보호

간단한 스캐닝으로 악성 QR 코드로 인한 피싱 공격과 악성 소프트웨어 설치를 사전에 차단

불안감 해소 및 QR 제공자에 대한 신뢰도 향상

QR과 링크된 URL 주소를 눈으로 확인할 수 없던 불안감 해소

악성 여부를 판단함으로써 안전한 QR 제공자에 대한 신뢰도 향상

새로운 피싱 기법 탐지 및 공공 안전 강화

서비스 내 새로운 QR 코드 데이터를 지속적으로 분석해 새로운 피싱 기법 탐지

유관기관 및 보안업계와 공유함으로써 공공 안전을 강화