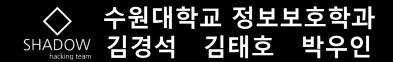
머신러닝을 활용한 피싱 사이트 탐지 방안

발표자 김경석



1. 소개

- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

Contents

Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

소개 – 팀원





* Mentor 이호진 *

- K-Shield Jr. 강사, 멘토
- 사이버보안빅데이터센터 강사

Mentee 김경석 Mentee 박우인 - 관심분야 -- 관심분야 -악성코드 분석 모의해킹 Mentee 김태호 - 관심분야 -모의해킹

- 수원대학교 정보보호학과 4학년 재학 중
- 수원대학교 보안 동아리 SHADOW 소속
- K-Shield Jr. 2기 보안사고 분석대응 과정 TOP20

소개 – 프로젝트



K-Shield Jr. 보안사고 분석대응 프로젝트

- * 수행 기간
 - 4월 ~ 7월
- * 사용 기술
 - Python
 - Machine Learning
 - HTML/CSS
 - Javascript
 - PHP
 - Database
- * 사이버보안빅데이터센터
 - 활용 과정에서 본 프로젝트를 교육 자료로 사용 중

Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

피싱 사이트란? – 피싱





Private data(개인정보)와 Fishing(낚시)에서 유래한 피싱은 이메일 등을 통해 신뢰할 수 있는 사람 또는 기관으로 위장하여 개인정보, 금융정보를 부정하게 얻으려는 사회공학적 기법

피싱 사이트란? – 피싱 종류



보이스 피싱, 스미싱, 스피어 피싱, 피싱 사이트 . . .

피싱 사이트란? – 피싱 종류



보이스 피싱, 스미싱, 스피어 피싱, 피싱 사이트 ...

피싱 사이트란? – 피싱 사이트



피싱 사이트는 정부기관, 금융기관 등의 웹사이트로 위장하여 사용자의 개인정보, 금융정보 등을 탈취하는 악의적인 웹사이트를 의미

피싱 사이트란? – 국내 피싱 사이트 현황



o 2006~2012년 발생건수 (단위:건)

구분	2006년	2007년	2008년	2009년	2010년	2011년	2012년	합계
발생건수	4	2	2	4	8	1,849	1,218	3,087

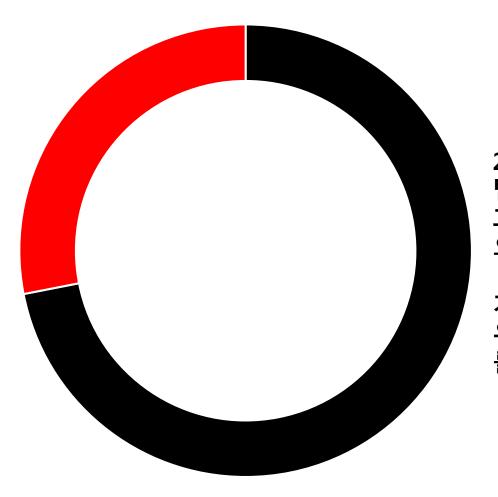
* 출처 : KISA

구 분	2013년	2014년	2015년	2016년	2017년	2018년 7월	합계
피싱사이트 차단	5,019	5,114	6,503	4,286	10,469	6,185	37,576

* 출처 : 과학기술정보통신부

피싱 사이트란? – 해외 피싱 사이트 현황





2018년 TrendMicro社 보안 인프라에서 약 410억 건의 이메일 공격을 탐지했다고 밝혔으며 이메일을 통한 피싱, 스팸 공격 등은 높은 파급력과 해킹 성공률로 인해 폭발적으로 증가하고 있어 이에 대한 대비가 필요함

가짜 office 365 로그인 양식을 이메일로 보내 피싱 사이트로의 연결을 유도하는 등의 크리덴셜 피싱은 2018년도 고위험 이메일 위협 중 40% 를 차지함

* 출처: KISA 2019년 1분기 사이버 위협 동향 보고서 중

■ Credential phishing - 40% (3,530,495 detections)

피싱 사이트란? - 국내 사례 - KISA



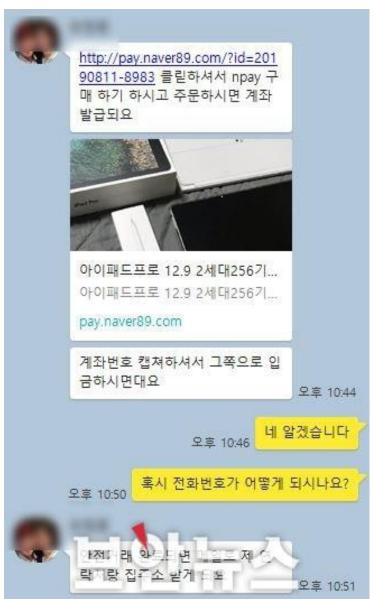


* 출처 : KISA

주민등록번호 이용내역 확인 서비스를 위장하여 주민등록번호 입력 유도 / 2012

피싱 사이트란? - 국내 사례 - NAVER



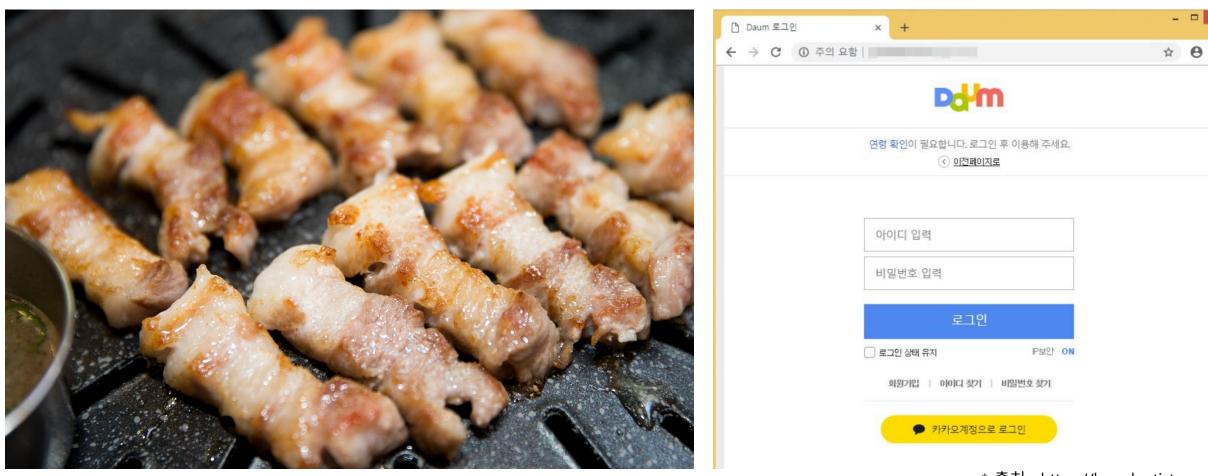




중고나라를 통해 거래를 제안하고 네이버페이를 사용하도록 하여 계정 정보 입력 유도 / 2018

피싱 사이트란? – 국내 사례 - Daum





* 출처 : https://la-nube.tistory.com/

자극적인 동영상처럼 보이는 사진을 클릭하면 계정 정보 입력 유도 / 2019

피싱 사이트란? - 해외 사례 - eBay



To: johndoe@fraudtest.com

From: "verify@ebay.com" <verify@ebay.com>

Subject: Verify Your eBay Account Date: Wed, 07 Jul 2004 04:37:52 +0200



home | pay | services | site map | help @



Dear eBay user,

You are urged to establish your proof of identity with ID Verify - an easy way to help others trust you as their trading partner as our service quality is changing. The process takes about 5 minutes to complete and involves updating your information over a secure connection. When you're successfully verified, you will receive an ID Verify icon in your feedback profile. This feature applies to residents of the United States and U.S. territories (Puerto Rico, US Virgin Islands and Guam), Europe (EU), Japan and Australia.

Fee: Verification is free of charge and is valid until your name, home address or phone number change.

Click here and ID Verify Now, help us improve the service we are providing to our customers.

Note: After successful verification, you will not be able to modify your contact information for 3D days. Also, the ID Verify icon is used for 3 different verification processes. Verisign provides individual verification on eBay.com. GeoTrust, Inc. is used for the verification of certain businesses selling on eBay.com. Deutsche Post is used for the verification of German members on eBay.de.

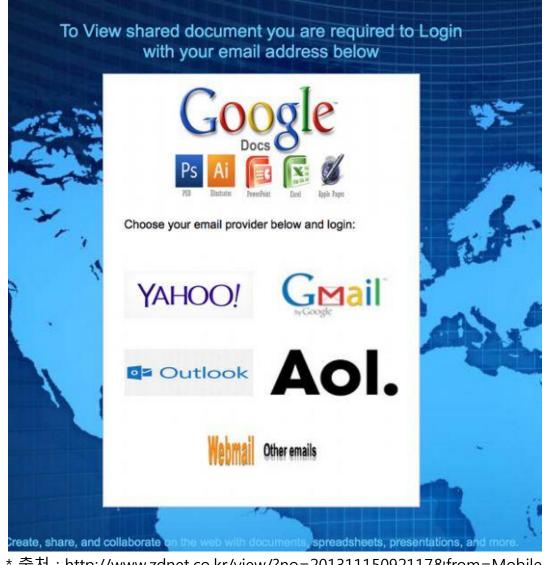
Warning: Failure to Verify your ID may result in Account Suspension.

* 출처 : http://inews24.com/view/125650

"보안상의 위험으로 계정이 일시 차단됐으니 재등록해야..." 신용카드번호, 개인인증번호, 사회보장번호 등 입력 유도 / 2003

피싱 사이트란? - 해외 사례 - Google



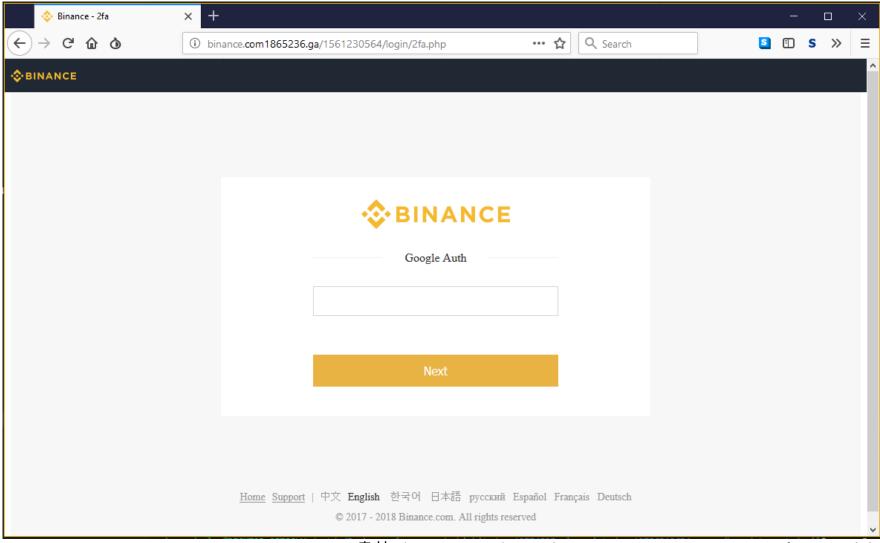


* 출처: http://www.zdnet.co.kr/view/?no=20131115092117&from=Mobile

구글 독스를 위장하여 계정 정보 입력 유도 / 2013

피싱 사이트란? – 해외 사례 – BINANCE



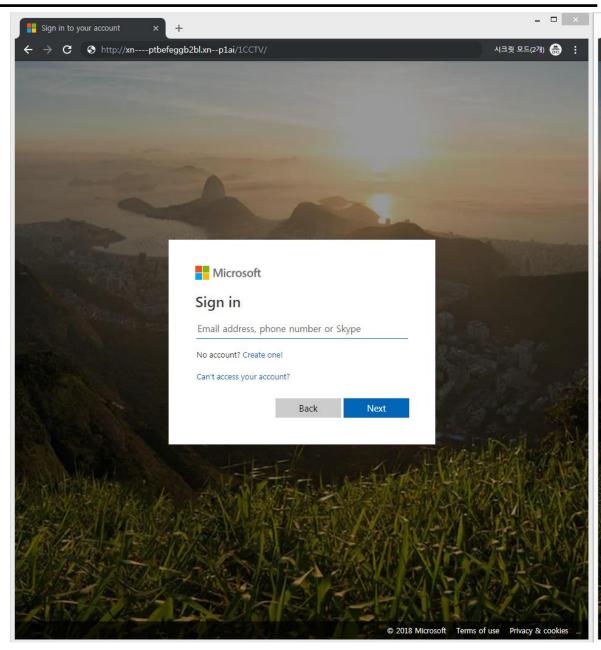


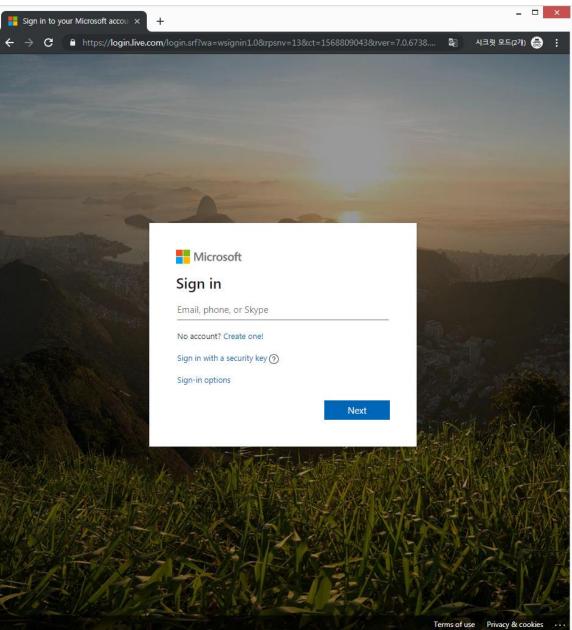
* 출처: http://www.zdnet.co.kr/view/?no=20131115092117&from=Mobile

가상화폐 거래소를 위장하여 계정 정보 입력 유도 / 2019

피싱 사이트란? – Outlook 최신 피싱 사이트

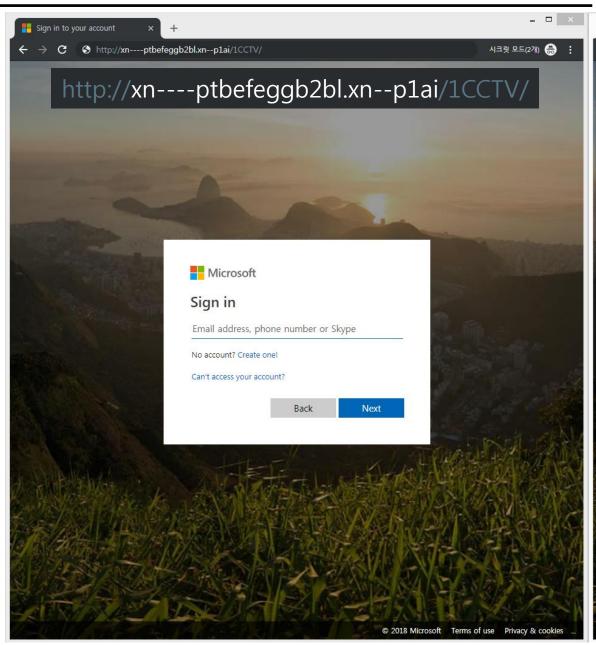


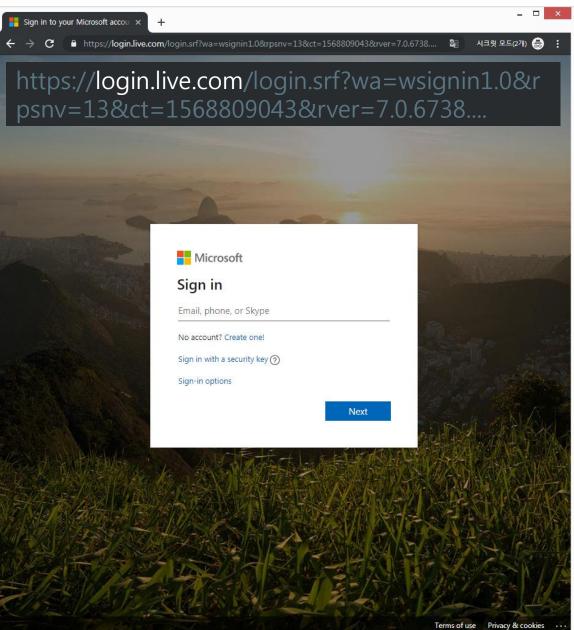




피싱 사이트란? - Outlook 최신 피싱 사이트







Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

데이터 수집 및 전처리 – Alexa

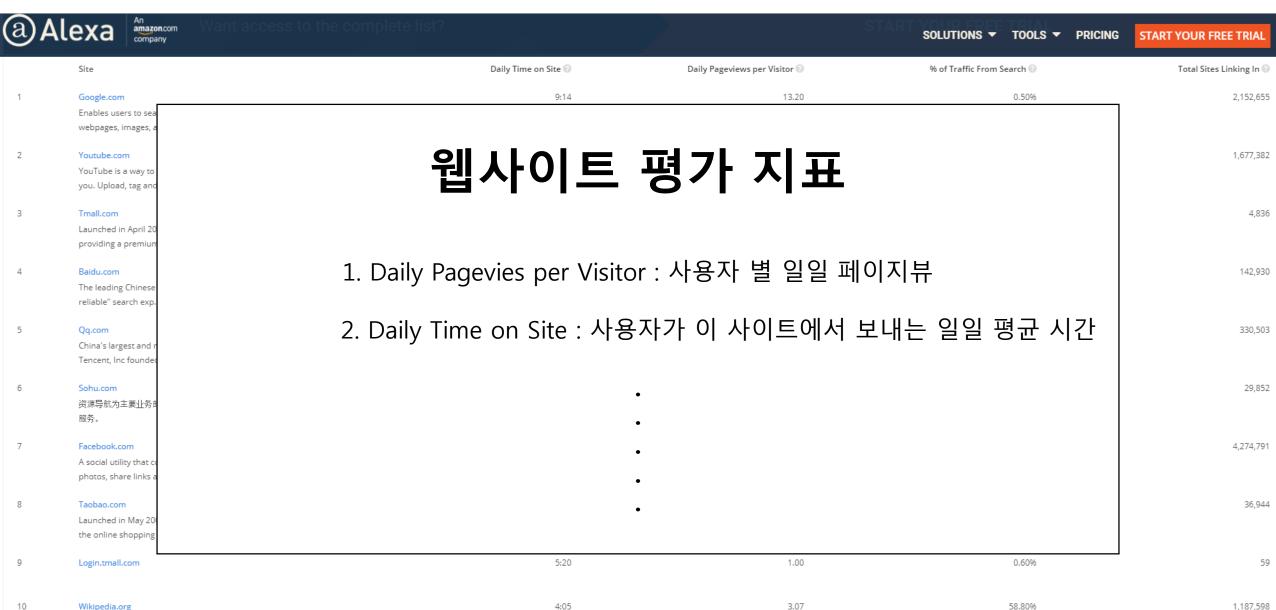


(a) Al	.exa An amazon.com Want access to the complete	te list?		SOLUTIONS ▼ TOOLS	▼ PRICING START YOUR FREE TRIAL
	Site	Daily Time on Site 🕢	Daily Pageviews per Visitor 💿	% of Traffic From Search 💿	Total Sites Linking In 🕢
1	Google.com Enables users to search the world's information, including webpages, images, and videos. OffersMore	9:14	13.20	0.50%	2,152,655
2	YouTube is a way to get your videos to the people who matter to you. Upload, tag and share yourMore	9:04	5.70	12.20%	1,677,382
3	Tmall.com Launched in April 2008, Tmall.com (www.tmall.com) is dedicated to providing a premium shoppingMore	7:03	2.90	1.10%	4,836
4	Baidu.com The leading Chinese language search engine, provides "simple and reliable" search expMore	6:48	4.90	6.80%	142,930
5	Qq.com China's largest and most used Internet service portal owned by Tencent, Inc founded in NovMore	3:20	3.80	6.50%	330,503
6	Sohu.com 资源导航为主要业务的门户网站,经营综合性业务,社区,无线等增值 服务。	3:46	4.34	2.60%	29,852
7	Facebook.com A social utility that connects people, to keep up with friends, upload photos, share links andMore	12:49	6.05	7.80%	4,274,791
8	Taobao.com Launched in May 2003, Taobao Marketplace (www.taobao.com) is the online shopping destination ofMore	5:58	3.64	4.00%	36,944
9	Login.tmall.com	5:20	1.00	0.60%	59
10	Wikipedia.org A free encyclopedia built collaboratively using wiki software.	4:05	3.07	58.80%	1,187,598

데이터 수집 및 전처리 – Alexa

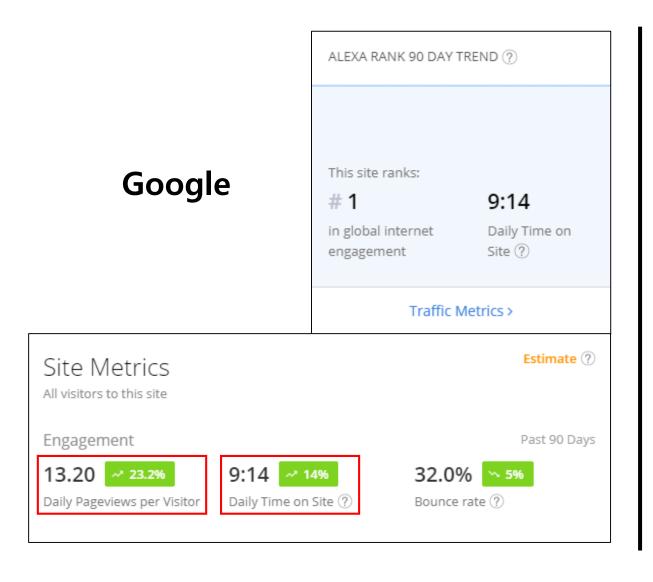
A free encyclopedia built collaboratively using wiki software.





데이터 수집 및 전처리 – Alexa – Google vs KISA









Drive More Website Traffic with Competitive Analysis

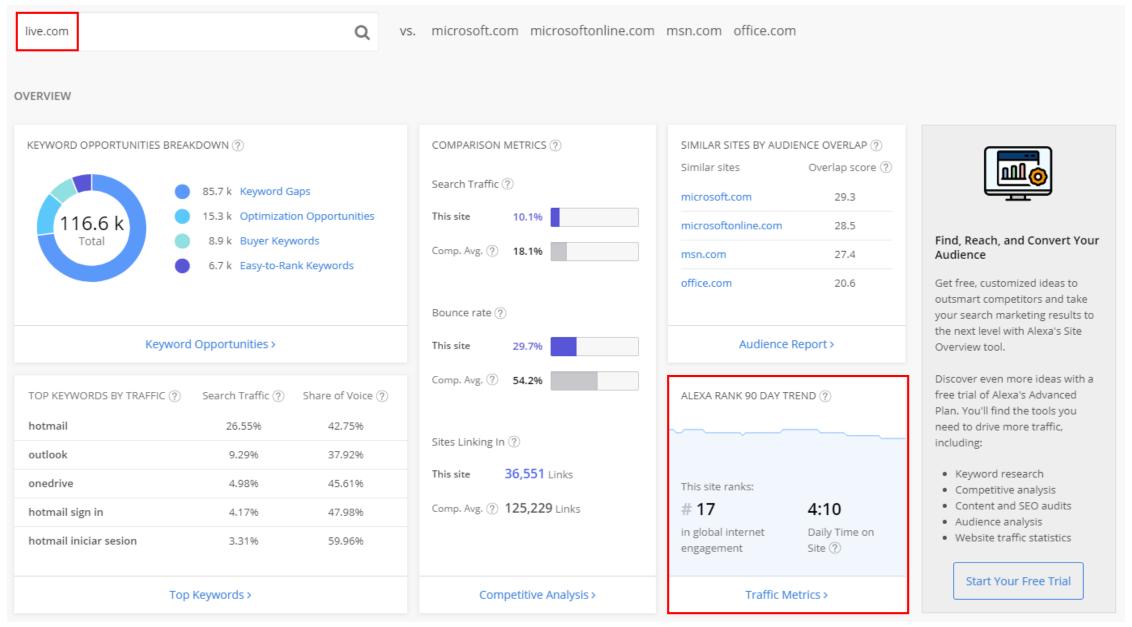
Get a free, automated website analysis to quickly gain insights into a site's strengths, weaknesses, and digital marketing opportunities.

Enter a site. For example: mysite.com

Run Analysis

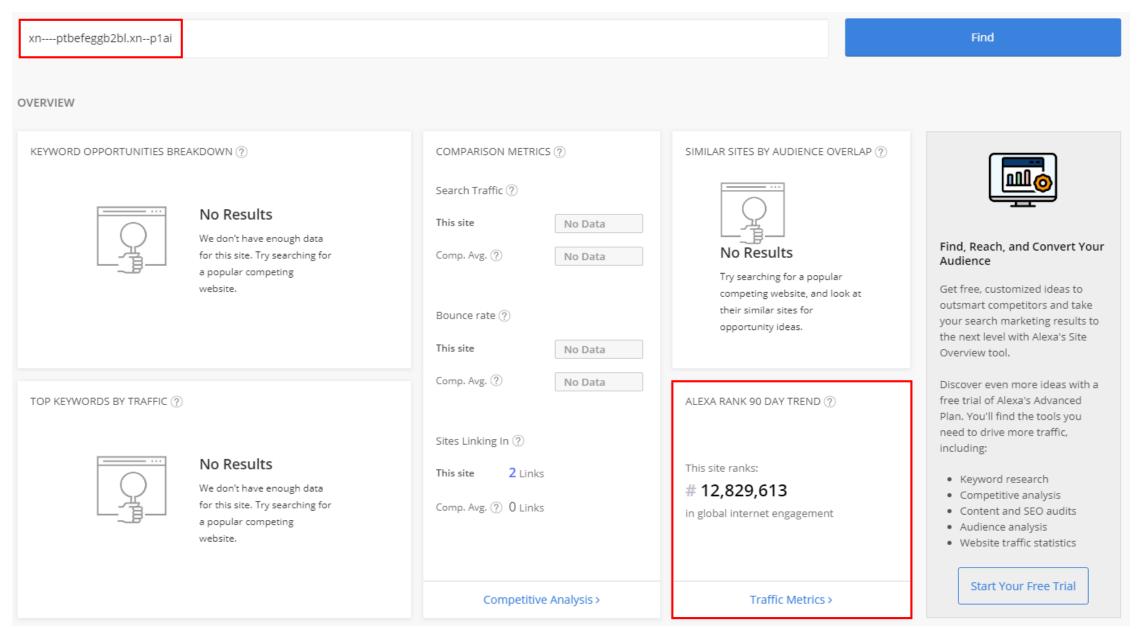
데이터 수집 및 전처리 – Alexa 웹사이트 분석 서비스 – Outlook 정상 사이트





데이터 수집 및 전처리 – Alexa 웹사이트 분석 서비스 – Outlook 피싱 사이트





데이터 수집 및 전처리 – Alexa Topsites 수집



	Α	В
1	1	google.com
2	2	youtube.com
3	3	tmall.com
4	4	baidu.com
5	5	sohu.com
6	6	qq.com
7	7	taobao.com
8	8	facebook.com
9	9	login.tmall.com
10	10	360.cn
11	11	wikipedia.org
12	12	yahoo.com
13	13	jd.com
14	14	amazon.com
15	15	sina.com.cn
16	16	pages.tmall.com
17	17	reddit.com
18	18	weibo.com
19	19	live.com
20	20	netflix.com

다운로드 url: http://s3.amazonaws.com/alexa-static/top-1m.csv.zip



최상위 순위 약 9,000개의 웹사이트를 정상 사이트의 데이터로 활용

데이터 수집 및 전처리 – PhishTank





Phish Search

Valid? Valid p	hishes ▼ Online? All ▼ Search			
ID	Phish URL	Submitted	Valid?	Online?
6198726	http://atacadaodeofertasnet.xyz/PljsahVVaddbh4420000sa/index.php?adeir added on Sep 15th 2019 1:13 PM	by morganamarchand	VALID PHISH	ONLINE
6198720	https://www.apple-id.map-login.in/fmi added on Sep 15th 2019 1:01 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198719	https://www.apple.com.5fa.co.za/ added on Sep 15th 2019 1:01 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198718	http://sourcegen.in/wp-includes/ppid0125/myaccount/signin/ added on Sep 15th 2019 1:01 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198717	http://sourcegen.in/wp-includes/ppid0125/myaccount/signin added on Sep 15th 2019 1:01 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198716	http://sourcegen.in/wp-includes/ppid0125 added on Sep 15th 2019 1:00 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198715	http://undistinguishing-sp.000webhostapp.com/Yahoo/Yahoo/Mail/index4_f added on Sep 15th 2019 1:00 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198714	https://stojanovic019.rs/wp-content/plugins/social-media-gallery/langu added on Sep 15th 2019 1:00 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198713	http://rasorial-carloads.000webhostapp.com/SunTrustOnline-Zelle/SunTru added on Sep 15th 2019 1:00 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198712	http://nostalgic-conjectur.000webhostapp.com/ppaayy/PayPallast/Blockch added on Sep 15th 2019 1:00 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198696	https://sardegnanatural.com/wp-includes/images/Cache/Continue/continue added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198695	https://sardegnanatural.com/wp-includes/images/Cache/Continue/continue added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198694	https://paypal.co.uk.syyg.icu/ added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198693	http://sourcegen.in/wp-includes/ppid0125/myaccount/signin/?country.x=U added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198692	http://sourcegen.in/wp-includes/ppid0125/myaccount added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198687	https://icloud.com.find-support.info/mobivs.php?ID=shuzanne.ong@icloud added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198686	http://bounteous-lane.000webhostapp.com/Yahoo/Yahoo/Mail/index4_files/ added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198685	http://facturas.apple.com-eup.top/sparkasse added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE
6198684	https://micrologic.co.za/wp/wp-includes/SimplePie/HTTP/access/info/aud added on Sep 15th 2019 12:02 PM	by <u>cleanmx</u>	VALID PHISH	ONLINE



Join the fight against phishing

Submit suspected phishes. Track the status of your submissions.

Verify other users' submissions. Develop software with our free API.

Found a phishing site? Get started now — see if it's in the Tank:

http://

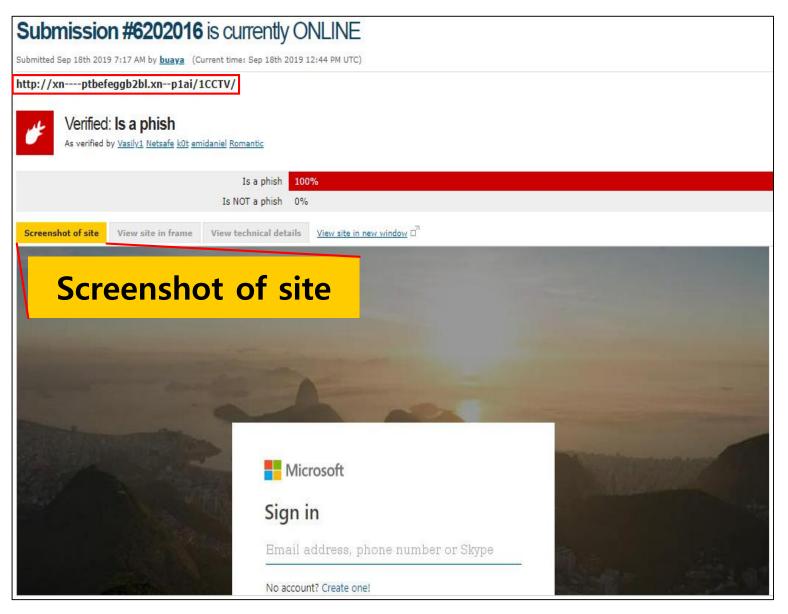
Is it a phish?

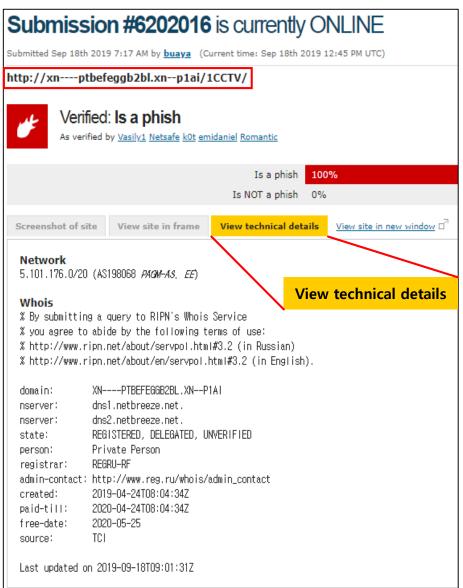
PhishTank에 보고된 피싱 사이트만 조회가 가능

■ 알려지지 않은 최신 피싱 사이트는 조회 불가

데이터 수집 및 전처리 – PhishTank 웹사이트 조회 서비스 – Outlook 피싱 사이트







데이터 수집 및 전처리 – PhishTank Valid Phishingsites 수집



Developer Information

Get the Database

If you'll be doing lots of lookups, the best option is to take advantage of our downloadable databases. Available in multiple formats and updated

The data is available in a variety of formats to make it as easy as possible for you to implement. We're always open to suggestions for additional

If you do intend to fetch these files automatically, please register for an application key and see below for instructions on how to use it to request

Format Options

XML
http://data.phishtank.com/data/online-valid.xml
http://data.phishtank.com/data/online-valid.xml.gz
http://data.phishtank.com/data/online-valid.xml.bz2

CSV	
http://data.phishtank.com/data/online-valid.csv	
http://data.phishtank.com/data/online-valid.csv.gz	
http://data.phishtank.com/data/online-valid.csv.bz2	

Serialized PHP
http://data.phishtank.com/data/online-valid.php_serialized
http://data.phishtank.com/data/online-valid.php_serialized.gz
http://data.phishtank.com/data/online-valid.php_serialized.bz2

JSON
http://data.phishtank.com/data/online-valid.json
http://data.phishtank.com/data/online-valid.json.gz
http://data.phishtank.com/data/online-valid.json.bz2

데이터 수집 및 전처리 – PhishTank Valid Phishingsites 수집



	Α	В	С	D	E	F	G	Н
1	phish_id	url	phish_detail_url	submission_time	verified	verification_time	online	target
2	6198728	http://meubussines2virtual.com/promocao.php	http://www.phishtank.com/phish_detail.php?phish_id=6198728	2019-09-15T13:16:40+00:00	yes	2019-09-15T13:18:11+00:00	yes	Other
3	6198726	http://atacadaodeofertasnet.xyz/PljsahVVaddbh4	http://www.phishtank.com/phish_detail.php?phish_id=6198726	2019-09-15T13:13:18+00:00	yes	2019-09-15T13:14:22+00:00	yes	Other
4	6198720	https://www.apple-id.map-login.in/fmi	http://www.phishtank.com/phish_detail.php?phish_id=6198720	2019-09-15T13:01:11+00:00	yes	2019-09-15T13:03:45+00:00	yes	Other
5	6198719	https://www.apple.com.5fa.co.za/	http://www.phishtank.com/phish_detail.php?phish_id=6198719	2019-09-15T13:01:06+00:00	yes	2019-09-15T13:03:45+00:00	yes	Other
6	6198718	http://sourcegen.in/wp-includes/ppid0125/myacc	http://www.phishtank.com/phish_detail.php?phish_id=6198718	2019-09-15T13:01:02+00:00	yes	2019-09-15T13:02:59+00:00	yes	Other
7	6198717	http://sourcegen.in/wp-includes/ppid0125/myacc	http://www.phishtank.com/phish_detail.php?phish_id=6198717	2019-09-15T13:01:01+00:00	yes	2019-09-15T13:02:59+00:00	yes	Other
8	6198716	http://sourcegen.in/wp-includes/ppid0125	http://www.phishtank.com/phish_detail.php?phish_id=6198716	2019-09-15T13:00:56+00:00	yes	2019-09-15T13:02:59+00:00	yes	Other
9	6198715	http://undistinguishing-sp.000webhostapp.com/Y	http://www.phishtank.com/phish_detail.php?phish_id=6198715	2019-09-15T13:00:52+00:00	yes	2019-09-15T13:02:13+00:00	yes	Other
10	6198714	https://stojanovic019.rs/wp-content/plugins/socia	http://www.phishtank.com/phish_detail.php?phish_id=6198714	2019-09-15T13:00:48+00:00	yes	2019-09-15T13:02:13+00:00	yes	Other
11	6198713	http://rasorial-carloads.000webhostapp.com/Sun	http://www.phishtank.com/phish_detail.php?phish_id=6198713	2019-09-15T13:00:44+00:00	yes	2019-09-15T13:02:13+00:00	yes	Other
12	6198712	http://nostalgic-conjectur.000webhostapp.com/p	http://www.phishtank.com/phish_detail.php?phish_id=6198712	2019-09-15T13:00:33+00:00	yes	2019-09-15T13:02:13+00:00	yes	Other
13	6198696	https://sardegnanatural.com/wp-includes/images	http://www.phishtank.com/phish_detail.php?phish_id=6198696	2019-09-15T12:02:55+00:00	yes	2019-09-15T12:09:14+00:00	yes	Other
14	6198695	https://sardegnanatural.com/wp-includes/images	http://www.phishtank.com/phish_detail.php?phish_id=6198695	2019-09-15T12:02:51+00:00	yes	2019-09-15T12:08:28+00:00	yes	Other
15	6198694	https://paypal.co.uk.syyg.icu/	http://www.phishtank.com/phish_detail.php?phish_id=6198694	2019-09-15T12:02:46+00:00	yes	2019-09-15T12:08:28+00:00	yes	Other
16	6198693	http://sourcegen.in/wp-includes/ppid0125/myacc	http://www.phishtank.com/phish_detail.php?phish_id=6198693	2019-09-15T12:02:42+00:00	yes	2019-09-15T12:08:28+00:00	yes	Other
17	6198692	http://sourcegen.in/wp-includes/ppid0125/myacc	http://www.phishtank.com/phish_detail.php?phish_id=6198692	2019-09-15T12:02:41+00:00	yes	2019-09-15T12:07:42+00:00	yes	Other
18	6198687	https://icloud.com.find-support.info/mobivs.php?	http://www.phishtank.com/phish_detail.php?phish_id=6198687	2019-09-15T12:02:20+00:00	yes	2019-09-15T12:09:59+00:00	yes	Other
19	6198686	http://bounteous-lane.000webhostapp.com/Yaho	http://www.phishtank.com/phish_detail.php?phish_id=6198686	2019-09-15T12:02:16+00:00	yes	2019-09-15T12:06:56+00:00	yes	Other
20	6198685	http://facturas.apple.com-eup.top/sparkasse	http://www.phishtank.com/phish_detail.php?phish_id=6198685	2019-09-15T12:02:11+00:00	yes	2019-09-15T12:06:56+00:00	yes	Other

다운로드 url: http://data.phishtank.com/data/online-valid.csv



피싱 사이트로 검증된 약 8500개의 웹사이트를 피싱 사이트 데이터로 활용

데이터 수집 및 전처리 – 18개 특징 선별 및 분류 기준 조정 과정



피싱 사이트 관련 논문 및 서적 수집 및 연구



웹사이트 특징 30~50개 테스트



최신 피싱 사이트를 탐지할 수 있도록 분류 기준 조정



정상 사이트와 피싱 사이트를 구분할 수 있는 18개 특징 선별

데이터 수집 및 전처리 – 정상 사이트와 피싱 사이트를 구분할 수 있는 18개 특징



URL 기반 특징

- 01. URL 길이
- 02. URL에 '-' 문자 포함 여부
- 03. 서브 도메인 개수
- 04. HTTPS 사용
- 05. HTTPS 위장

도메인 기반 특징

06. 도메인 등록 기간

HTML 및 Javascript 기반 특징

- 07. HTML 소스코드 내 도메인 존재 여부
- 08. HTML 소스코드의 길이
- 09. <head>, <body> 태그 중복 사용
- 10. Favicon 로드 경로
- 11. <form> 태그의 action 속성 값
- 12. <meta>, <script>, <link> 태그의 비율
- 13. <a> 태그의 외부 링크 비율
- 14. <script> 태그의 외부 링크 비율
- 15. 외부 요청 URL 비율

비정상적인 특징

- 16. 피싱 도메인 및 IP 사용
- 17. 리다이렉션 횟수
- 18. Alexa 순위

데이터 수집 및 전처리 – Dataset 구축



0	1																	
0	*	1	1	-1	1	1	1	1	-1	-1	-1	1	0	-1	0	1	-1	
•	-1	1	1	1	1	1	1	1	1	1	-1	1	1	-1	1	1	-1	
1	1	0	0	1	1	1	1	1	1	1	-1	1	1	1	0	1	-1	
1	1	1	1	-1	1	1	1	1	-1	-1	-1	1	-1	-1	0	1	-1	
1	-1	1	0	1	1	1	1	1	1	1	-1	1	1	-1	0	1	-1	
1	-1	1	0	1	1	1	1	1	1	1	-1	1	1	-1	0	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	0	-1	0	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	0	-1	0	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	0	-1	0	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	0	-1	0	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	1	-1	1	1	-1	
-1	1	1	1	1	1	1	1	1	1	1	-1	1	1	-1	1	1	-1	
1	-1	1	0	-		-	-			-	-	-	-	-1	0	1	-1	
1	-1	1	0											-1	0	1	-1	
1	-1	1	-1											-1	0	1	-1	
1	1	0	0											-1	-1	-1	-1	
1	-1	0	1											.1	-1	1	-1	
1	-1	1	-1							_				-1	0	1	-1	
-1	1	1	1										_	-1	1	1	-1	
1	-1	1	1						\		~	7 P	•	-1	0	1	-1	
													-	-	-	-	_	
	1	1	0						v					.1	0	1	-1	_
-1	1	1	0					V	y	L		J I		1	0	1	-1 -1	
-1	1	0	0					Μ.	y	L	h	J I		1	0	1	-1	
-1 -1	1	0	0			•		Υ.	y	L		JI		1	0	1	-1 -1	
-1 -1 -1	1 1 1	0 0 1	0			•		Υ.	y	LI		JI	•	1 1 -1	0	1 1 1	-1 -1 -1	
-1 -1 -1	1 1 1 1	0 0 1 1 1	0 0			•		Υ.	y	LI		<i>)</i>		1 1 ·1	0 0 0	1 1 1 1	-1 -1 -1 -1	
-1 -1 -1 -1 -1	1 1 1 1	0 0 1 1 1 1	0 0 0 0			•		Υ.	y	LI		/ 1		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0	1 1 1 1	-1 -1 -1 -1 -1	,
-1 -1 -1 -1 -1	1 1 1 1 1 1 1 1	0 0 1 1 1	0 0 0 0 0											1 1 ·1	0 0 0 0 0 0	1 1 1 1 1	-1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1 1	0 0 1 1 1 1	0 0 0 0 0	1	1	1	1	1	-1	1	-1	1	1	1 1 1 1 1 1 1 1 1	0 0 0 0 0	1 1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0	1	1	1 1	1	1 1	-1	1	-1 -1	1 1	1	1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0	1 1 1 1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 1 1	1 1 1 1 1 1 1 1 1 0	0 0 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0	1 -1	1	1 1 1	1 1 1	1 1 1	-1 -1 1	1 1 1	-1 -1 -1	1 1 1 1	1 1 -1	1 1 1 1 1 1 1 1	0 0 0 0 0 0 0	1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 1 1 -1	1 1 1 1 1 1 1 1 1 0	0 0 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1	1 -1 1	1 1 1	1 1 1	1 1 1	1 1 1	-1 -1 1	1 1 1	-1 -1 -1 -1	1 1 1	1 1 -1	1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1	1 1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 1 1 -1 -1 -1	1 1 1 1 1 1 1 1 1 0	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1	1 -1 1 1	1 1 1	1 1 1 1	1 1 1 1	1 1 1 1	-1 -1 1 1	1 1 1 1	-1 -1 -1 -1 -1	1 1 1 1	1 1 -1 1	1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 0 1 1 1 0	1 1 1 1 1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	
-1 -1 -1 -1 -1 -1 1 1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1 0 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 1 0 0	1 -1 1 1 -1	1 1 1 1	1 1 1 1 1	1 1 1 1 1	1 1 1 1 1	-1 -1 1 1	1 1 1 1 1	-1 -1 -1 -1 -1 -1	1 1 1 1 1	1 1 -1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 0 1 1 1 0	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	1 1 1 1 1 1 1 1 1 0 1 1 0	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0	1 -1 1 1 -1	1 1 1 1 1	1 1 1 1 1 1	1 1 1 1 1 1	1 1 1 1 1 1	-1 -1 1 1 1	1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1	1 1 -1 1 1 -1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 0 1 1 1 0 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0	1 -1 1 1 -1 1	1 1 1 1 1 1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	-1 -1 1 1 1 1	1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1	1 1 -1 1 1 -1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 0 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0 0	1 -1 1 1 -1 1 1	1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	-1 -1 1 1 1 1 1	1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1	1 1 -1 1 1 -1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 0 1 1 0	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 1 -1 -1 -1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0 0 1	1 -1 1 1 -1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	-1 -1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1	1 1 -1 1 1 -1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 0 1 1 0	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 1 -1 -1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0 0 1 0 0	1 -1 1 -1 1 -1 1 1 1	1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1	-1 -1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	1 1 1 1 1 1 1 1 1 1	1 1 -1 1 1 -1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 0 1 1 0 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	
-1 -1 -1 1 -1 -1 -1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 1	0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 1 1 1 1 0 0 0 1	1 -1 1 -1 1 -1 1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	-1 -1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1	1 1 1 1 1 1 1 1 1	1 1 -1 1 1 -1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 1 1 1 0 1 1 0	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	

Alexa/PhishTank 웹사이트 17,636개 특징 추출(정상 1 / 의심 0 / 피싱 -1) 및 결과 값 (정상 1 / 피싱 -1) 추가

Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

머신러닝 모델 구축 및 탐지 결과 – 테스트 알고리즘



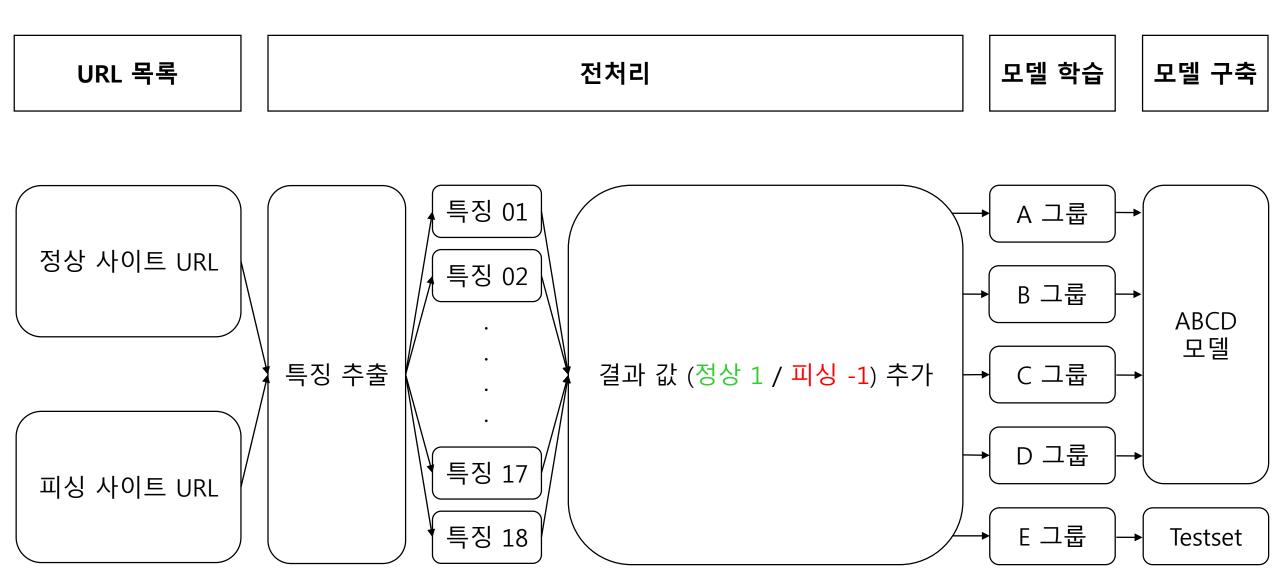
Logistic regreesion

Decision tree

Random forest

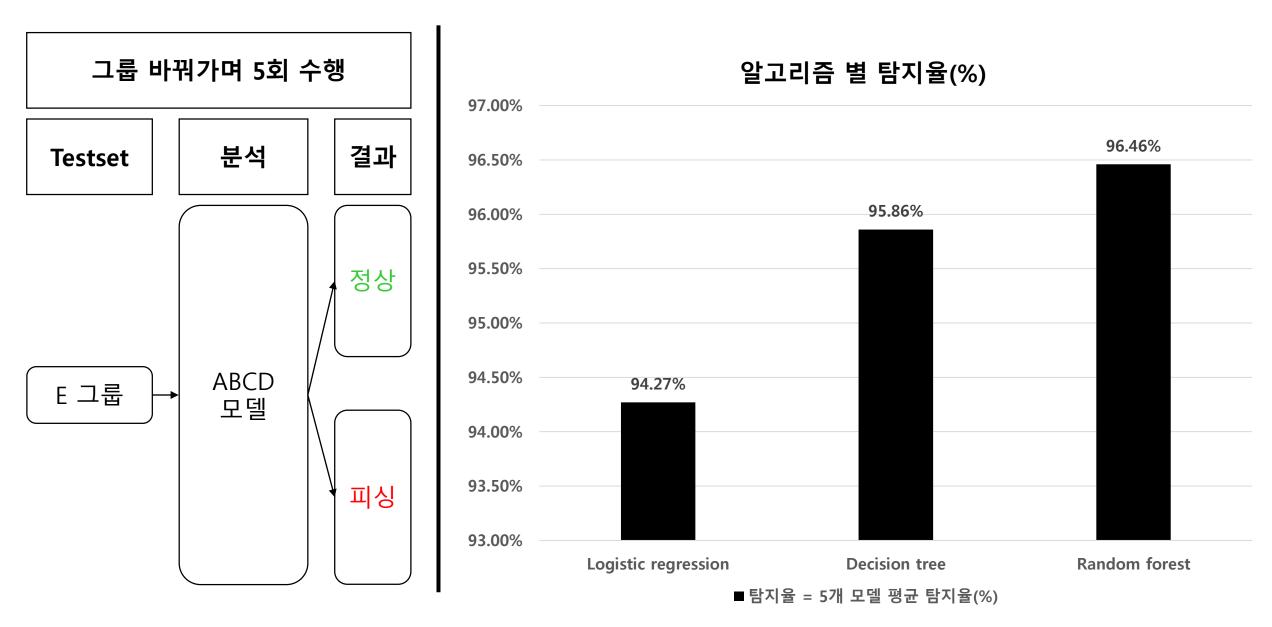
머신러닝 모델 구축 및 탐지 결과 – 알고리즘 별 모델 구축





머신러닝 모델 구축 및 탐지 결과 – 탐지 결과





Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

피싱 사이트 특징 - 01. URL 길이



Description

일부 피싱 사이트들은 긴 URL을 이용하여 웹브라 우저의 주소 표시줄에서 의심스러운 부분을 숨길 수 있습니다.

Example

- 피싱 사이트 URL http://paypal.com.webmpp-informationsteprEview eb.u4561925id.ha003.t.justns.ru/
- URL 길이 78자

Target webite: Paypal

Classification criterion

if URL 길이 > 75:
return 피싱
elif URL 길이 >= 54:
return 의심
else:
return 정상

Classification result 10000 9033 9000 8000 7000 6000 5000 4170 4000 2492 3000 1831 2000 1000 55 55 정상 사이트 피싱 사이트 ■정상 ■의심 ■피싱

피싱 사이트 특징 - 02. URL에 '-' 문자 포함 여부



Description

일부 피싱 사이트들은 '-' 문자를 사용하여 사용자로 하여금 정상적인 웹사이트라고 착각하게 할 수 있습니다.

Example

- 피싱 사이트 URL http://paypal-bestellungen.com/de/signin

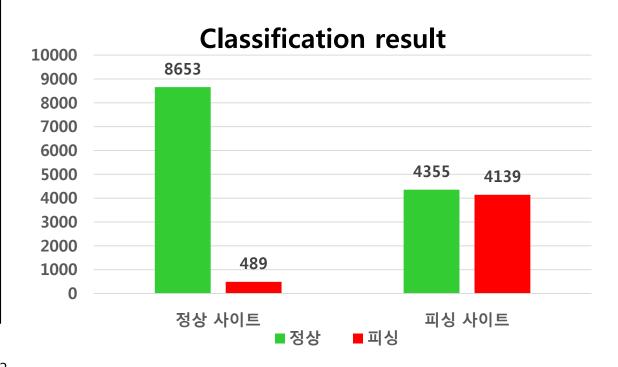
Target website: Paypal

Classification criterion

if '-' in URL:

return 피싱

else:



피싱 사이트 특징 – 03. 서브 도메인 개수



Description

일부 피싱 사이트들은 2개 이상의 서브 도메인을 갖습니다.

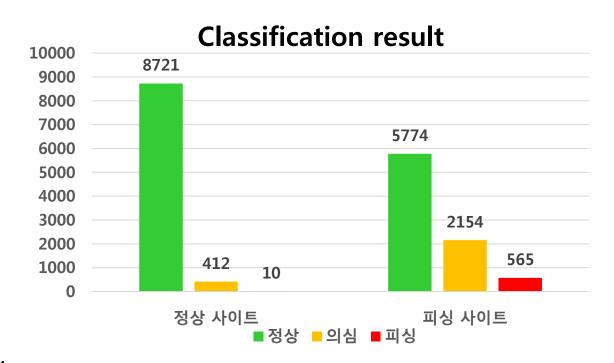
Example

- 피싱 사이트 URL https://account-update.amazon.co.jp.ama888.com /
- 서브 도메인 개수 <mark>4개</mark>

Target website: Amazon

Classification criterion

if 서브 도메인 개수 > 1:
return 피싱
elif 서브 도메인 개수 == 1:
return 의심
else:



피싱 사이트 특징 - 04. HTTPS 사용 여부



Description

대부분의 피싱 사이트들은 HTTPS를 사용하지 않으며, 사용하더라도 신뢰할 수 없습니다.

Example

- 피싱 사이트 URL http://twowheelcool.com/twitter.com/logine034.ht ml

Target website : Twitter

Classification criterion

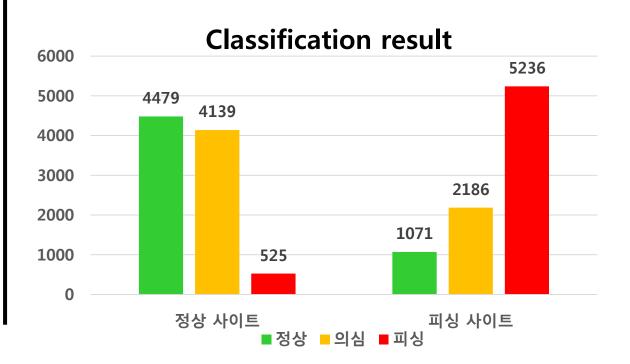
if HTTPS 미사용:

return 피싱

elif 신뢰할 수 없는 HTTPS 사용 or 사용 기간 < 365:

return 의심

elif 신뢰할 수 있는 HTTPS 사용 and 사용 기간 >= 365:



피싱 사이트 특징 – 05. HTTPS 위장



Description

어떤 피싱 사이트들은 사용자를 속이기 위해 도메 인에 "HTTPS" 토큰을 추가할 수 있습니다.

Example

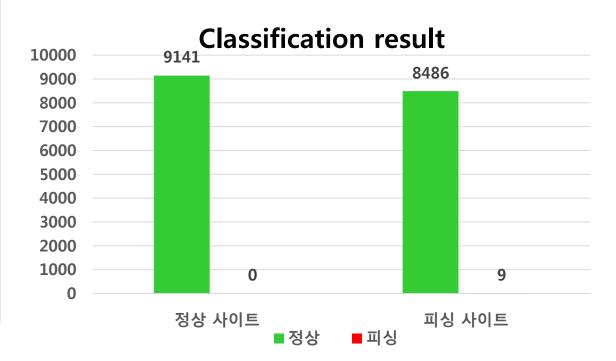
- 피싱 사이트 URL http://https.mail.office.40tbr2dzkekhd8qzjti8t8eoe spdf.hfkfv8ck1gbpkdckyjehk.mg67lq2fafsrsb1c6hk.dxu67egggsdh73hi88zd8ujdk.4q5qtc6quhbwk9hh 6.7svpk1q49iyi0nhevvad9h4kim632.qyhxy2bm583 zzhozuoypluytb8.kkzzkkzkza.duckdns.org/365New OfficeG15/yukari.kato@jti.com/paul/

Target website: Microsoft

Classification criterion

if "https" in 도메인 : return 피싱

else:



피싱 사이트 특징 – 06. 도메인 등록 기간



Description

수명이 짧은 특성을 가지는 대부분의 피싱 사이트 들은 도메인 등록 기간이 짧습니다.

Example

- 피싱 사이트 URL http://xn--b3c0ack6ahcca0kxa4t.com/cgi-bin/vp2
- 도메인 등록 기간 349일

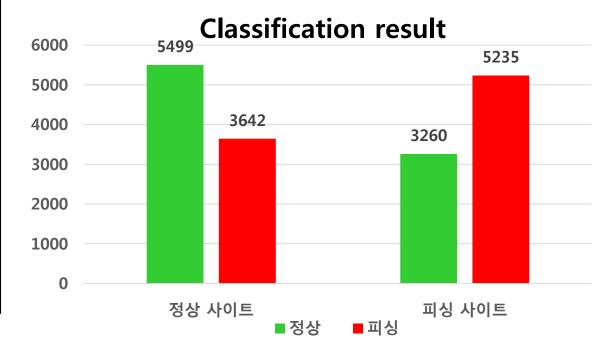
Target website: IRS

Classification criterion

if 도메인 등록 기간 < 365일 :

return 피싱

else:



피싱 사이트 특징 - 07. HTML 소스코드 내 도메인 존재 여부



Description

대부분의 정상 사이트들의 HTML 소스코드에는 도메인이 포함되어 사용됩니다.

Example

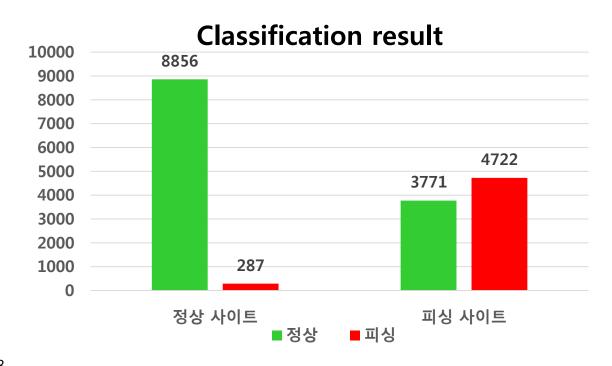
- 피싱 사이트 URL http://amazon-co-jpehsh.com/
- 주 도메인 amazon-co-jpehsh

Target website : Amazon

Classification criterion

if 도메인 not in 소스코드 : return 피싱

else:



피싱 사이트 특징 - 08. HTML 소스코드의 길이



Description

대부분의 피싱 사이트들은 정상 사이트에 비해 상대적으로 소스코드의 길이가 짧습니다.

Example

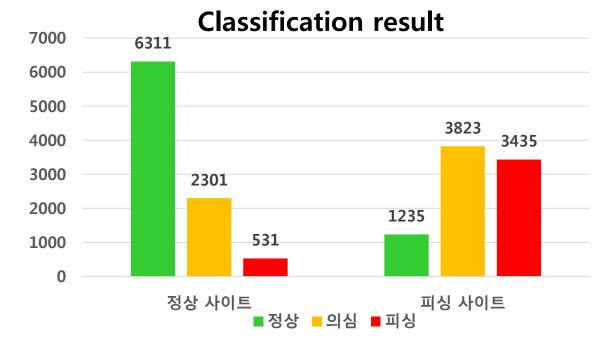
- 피싱 사이트 URL https://mansouramobileacademy.com/cgi-bn/pay pal.com/8830a520994762b1c791e3d2939a785aN DI2NmZmYWE4ZmIyZTM4M2ZmZGRkYmFiNWM1 NDE2Yzc=

- HTML 소스코드 길이 3571

Target website : Paypal 정상 Paypal 로그인 페이지 길이 : 150224

Classification criterion

if 소스코드 길이 < 5000:
return 피싱
elif 소스코드 길이 < 50000:
return 의심
else:
return 정상



피싱 사이트 특징 - 09. <head>, <body> 태그 중복 사용



Description

일부 피싱 사이트들은 HTML 소스코드 내에 <head>, <body> 태그가 중복되거나 없습니다.

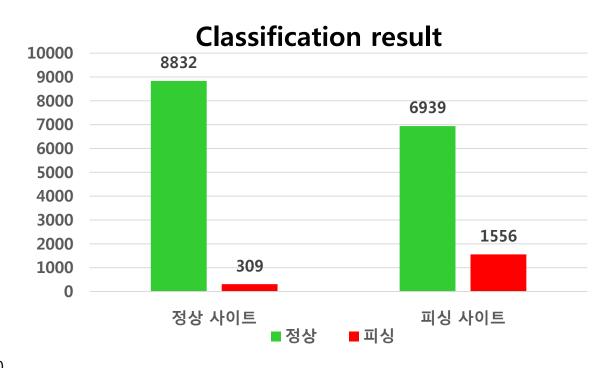
Example

- 피싱 사이트 URL http://espresses.gq/Swbt
- 중복 태그 <body> 태그 <mark>2회</mark> 사용

Target website: BT

Classification criterion

if <head>, <body> 중복 사용 : return 피싱 else : return 정상



피싱 사이트 특징 – 10. Favicon 로드 경로



Description

Favicon은 웹페이지를 대표하는 아이콘으로, 대부분의 피싱 사이트들은 Favicon 이미지를 사용하지 않거나 외부에서 로드합니다.

Example

- 피싱 사이트 URL http://twowheelcool.com/twitter.com/logine034.ht ml
- Favicon을 로드하는 코드 k href=http://abs.twimg.com/favicons/favicon.ico>

Target website : Twitter

Classification criterion

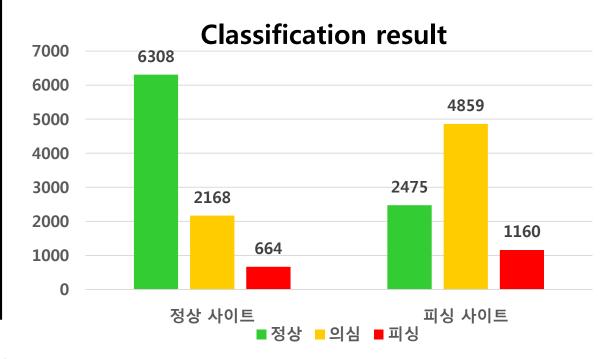
if 외부에서 로드된 Favicon:

return 피싱

elif Favicon not in 소스코드 :

return 의심

else:



피싱 사이트 특징 - 11. <form> 태그의 action 속성 값



Description

<form> 태그의 action 속성을 통해 정보를 전송할수 있습니다. 대부분의 피싱 사이트들은 action 속성 값이 외부 도메인이거나 공백입니다.

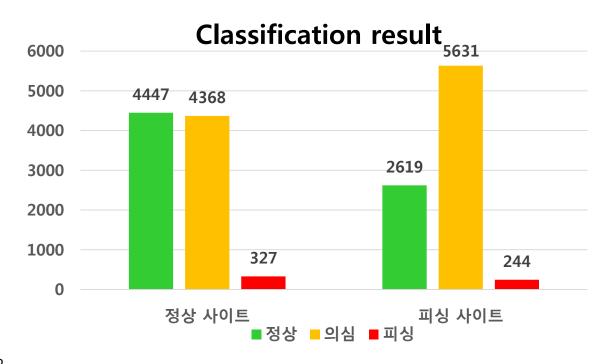
Example

- 피싱 사이트 URL http://paypal.infinitydeals.ml/customer_center/cust omer-IDPP00C255/myaccount/signin
- action 속성 값 <form for="9945-xMARVELxDCxCOMIC18x-11888" action="" method="post" ... 생략>

Target website: Paypal

Classification criterion

```
if action 속성 값 == "":
return 피싱
elif action 속성 값 == 외부 도메인:
return 의심
else:
return 정상
```



피싱 사이트 특징 – 12. <meta>, <script>, <link> 태그의 비율



Description

일부 피싱 사이트들은 전체 태그 중 <meta>, <script>, , , 니다.

Example

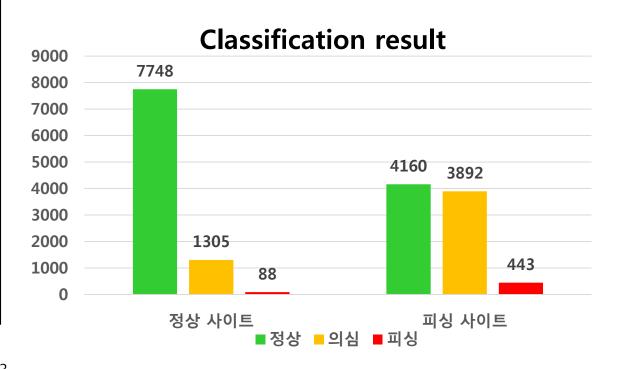
- 피싱 사이트 URL http://100-customersatisfaction-000.com/Am azon/278e6
- <meta>, <script>, <link> 태그 비율 31%

Target website : Amazon - 정상 사이트의 위 태그 비율 KISA(4.2%), NAVER(3.4%), Google(3.0%)

Classification criterion

if 전체 태그 중 왼쪽 태그들의 비율 > 81% :
return 피싱
elif 전체 태그 중 왼쪽 태그들의 비율 >= 17% :
return 의심

else:



피싱 사이트 특징 - 13. <a> 태그의 외부 링크 비율



Description

일부 피싱 사이트들은 <a> 태그의 링크들이 URL 내 도메인과 다르거나 웹페이지에 연결되어 있지 않습니다.

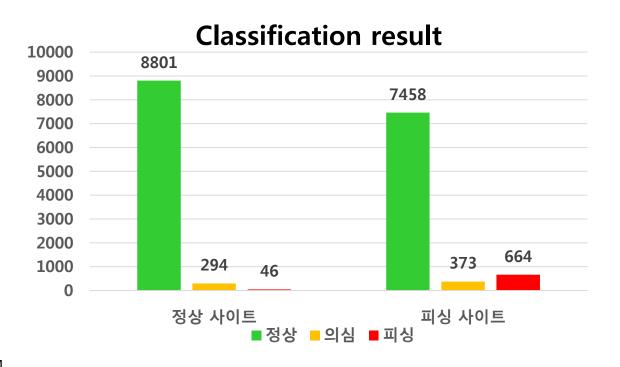
Example

- 피싱 사이트 URL http://serwer1904063.home.pl/appl/bas/base/
- 메뉴 선택 시 정상 사이트(다른 도메인)로 연결하는 <a> 태그 코드
-

Target website: Coinbase

Classification criterion

if 전체 <a> 태그 중 왼쪽 조건들의 비율 > 67%:
return 피싱
elif 전체 <a> 태그 중 왼쪽 조건들의 비율 >= 31%:
return 의심
else:



피싱 사이트 특징 - 14. <script> 태그의 외부 링크 비율



Description

대부분의 피싱 사이트들은 자바스크립트를 가져오기 위해 사용하는 <script> 태그를 100% 외부에서로드하거나 아예 로드하지 않습니다.

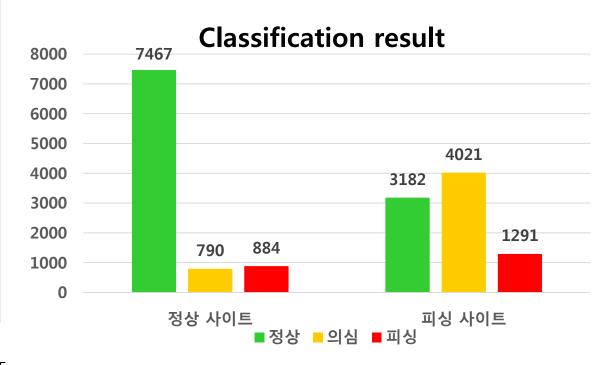
Example

- 피싱 사이트 URL https://palatial-diagnoses.000webhostappp.com/c ox.full/cox.full/Zo/stepCoxMail/log/sycho/index.ht ml
- 사용 중인 <script> 코드 <script src="https://webcdn2.cox.com/ui/5_0/tsw/j s/lib.js.jgz"> </script>

Target website : Cox

Classification criterion

if 외부에서 로드된 script 비율 == 100%:
return 피싱
elif 로드된 script 비율 == 0%:
return 의심
else:



피싱 사이트 특징 - 15. 외부 요청 URL 비율



Description

일부 피싱 사이트들은 웹페이지 내에 포함된 이미지, 비디오 등을 다른 도메인에서 로드합니다.

Example

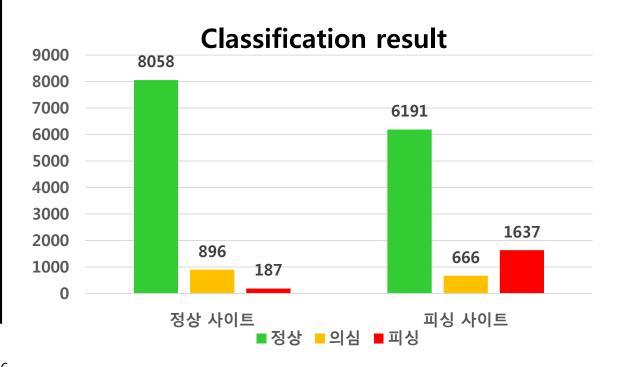
- 피싱 사이트 URL http://serwer1904063.home.pl/appl/bas/base/
- 아이콘을 로드하는 코드 k rel="icon" href="https://www.coinbase.com/ Img/favicon/favicon-196x196.png" sizes="196x196">

Target website : Coinbase

Classification criterion

if 외부에서 로드된 개체 비율 >= 61%:
return 피싱
elif 외부에서 로드된 개체 비율 >= 22%:
return 의심

else:



피싱 사이트 특징 – 16. 리다이렉션 횟수



Description

일부 피싱 사이트들은 두 번 이상 리다이렉션합니다.

Example

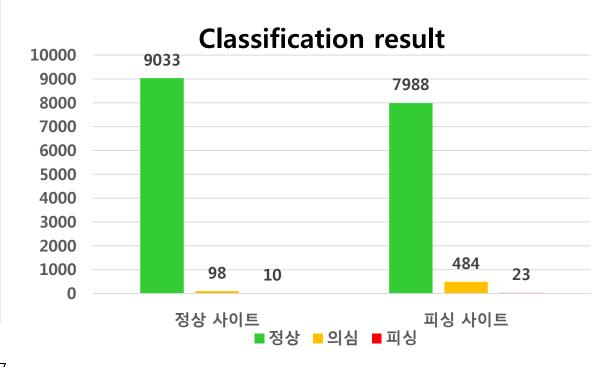
- 피싱 사이트 URL http://tutkuav.com/wp-content/themes/twentysixt een/-/bin-de/0fccb
- 리다이렉션 횟수 3<mark>회</mark>

Target website: Paypal

Classification criterion

if 리다이렉션 횟수 >= 4 : return <mark>피싱</mark> elif 리다이렉션 횟수 > 1 : return 의심

else:



피싱 사이트 특징 - 17. 피싱 도메인 및 IP 사용 여부



Description

PhishTank, StopBadware는 피싱 사이트의 도메인과 IP에 대한 통계 보고서를 작성합니다. 해당 보고서에서 많이 사용되는 피싱 사이트 도메인과 IP를 참조합니다.

Example

- 피싱 사이트 URL http://pruebadepaginate.webcindario.com/of-ame rica-signOnV2Screen1im2y1njaindexphpcmd_upd ate_account&sessionNjU1MDQ5NjkxNDQyNDA0 MA.html

Target website: Bank of America

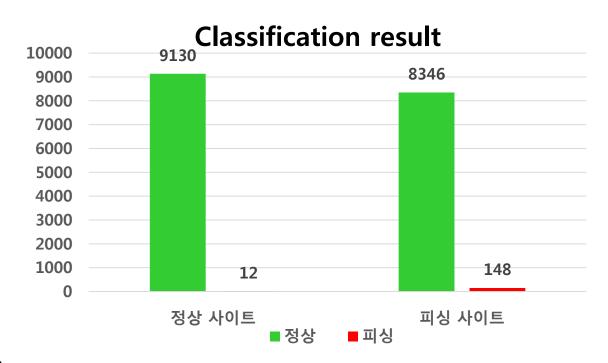
피싱 IP: 5.57.226.202

Classification criterion

if phishtank, stopbadware 통계에 존재:

return 피싱

else:



피싱 사이트 특징 - 18. Alexa 순위



Description

Alexa 데이터베이스는 웹사이트 트래픽을 수집하여 순위를 알려줍니다. 대부분의 피싱 사이트들은 짧은 기간 동안 활동하기 때문에 웹사이트 트래픽이 상대적으로 낮습니다. 따라서 Alexa 순위에 랭크되지 않을 가능성이 높습니다.

Example

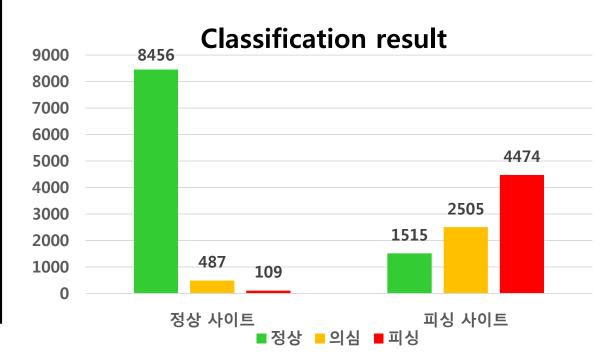
- 피싱 사이트 URL https://mypostfinance.net/woa/5a6eaf

Target website : PostFinance

Alexa 순위 : No Rank

Classification criterion

if URL not in Alexa 순위:
return 피싱
elif Alexa 순위 > 100,000:
return 의심
else:



Contents

- 1. 소개
- 2. 피싱 사이트란?
- 3. 데이터 수집 및 전처리
- 4. 머신러닝 모델 학습 및 탐지 결과
- 5. 피싱 사이트 특징
- 6. 활용

활용 – 크롬 확장 프로그램을 활용한 피싱 사이트 탐지 서비스

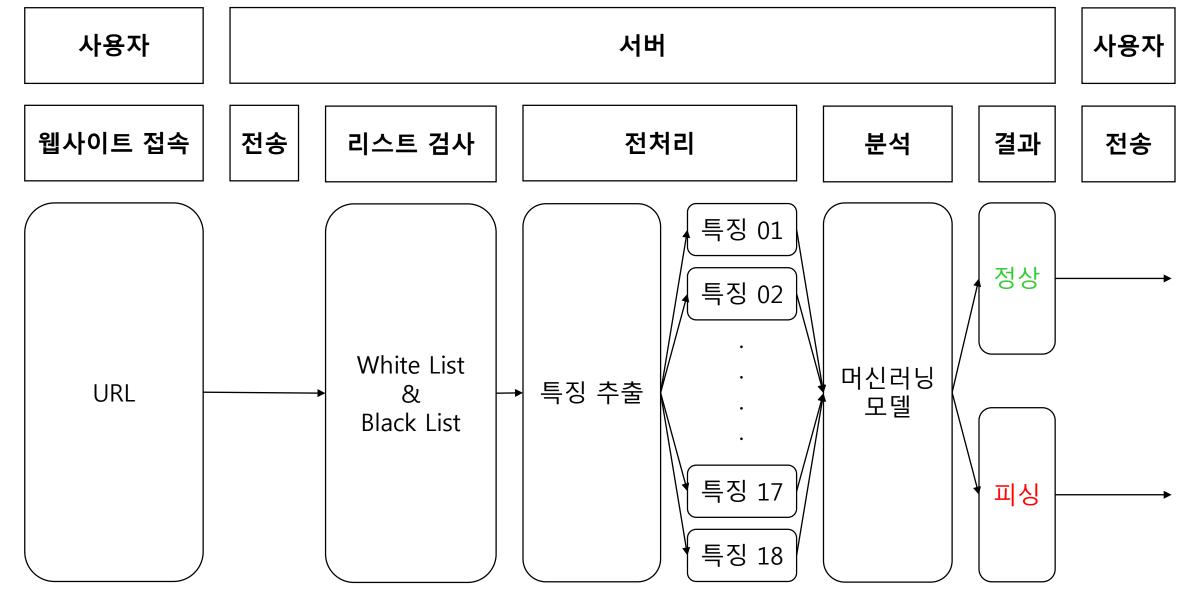


기능

- 1. 사용자가 접속한 웹사이트 자동 검사 기능
- 2. 별도 입력 창을 통해 웹사이트 검사 기능
- 3. 피싱 사이트일 경우 알림 기능
- 4. White List & Black List 활용
- 5. 관리자 페이지 / 머신러닝 모델에 의해 분석된 웹사이트 재검토용

활용 – 서비스 처리 과정





소요시간 : 1 ~ 2초





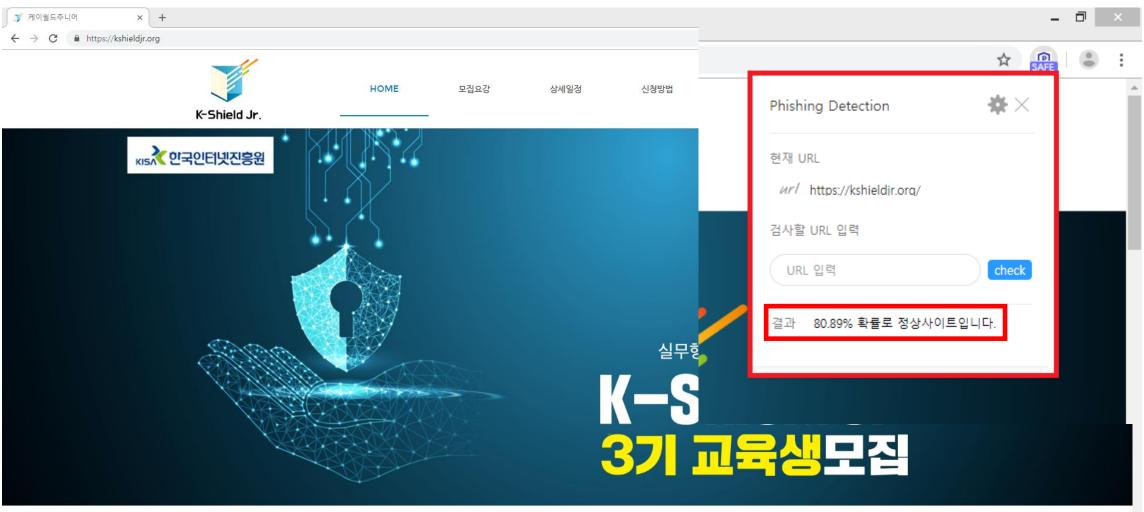














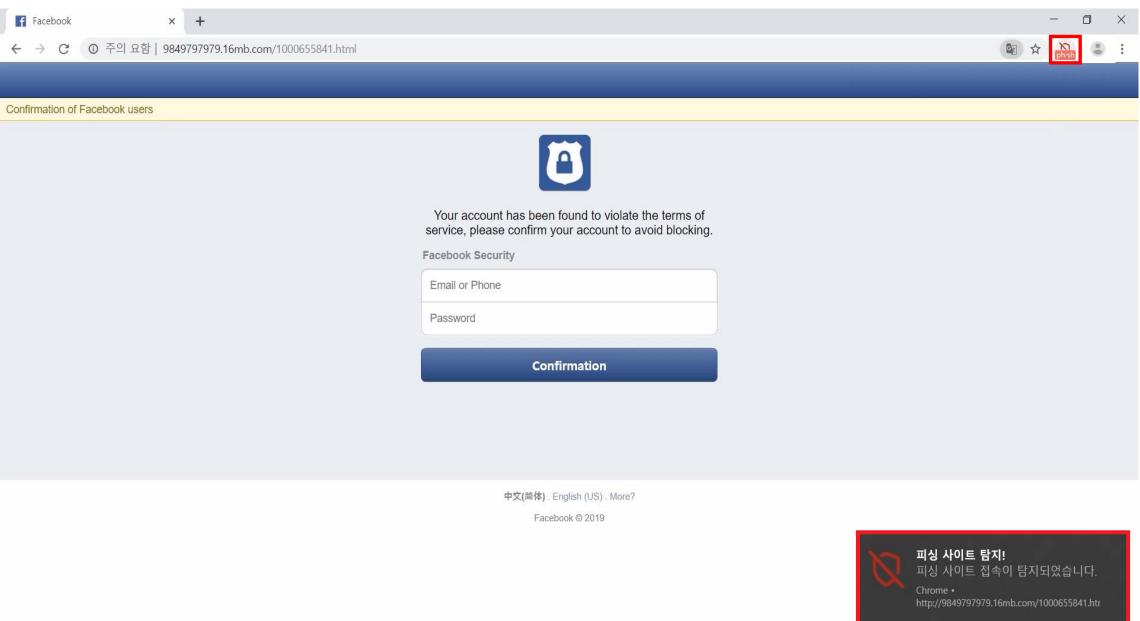






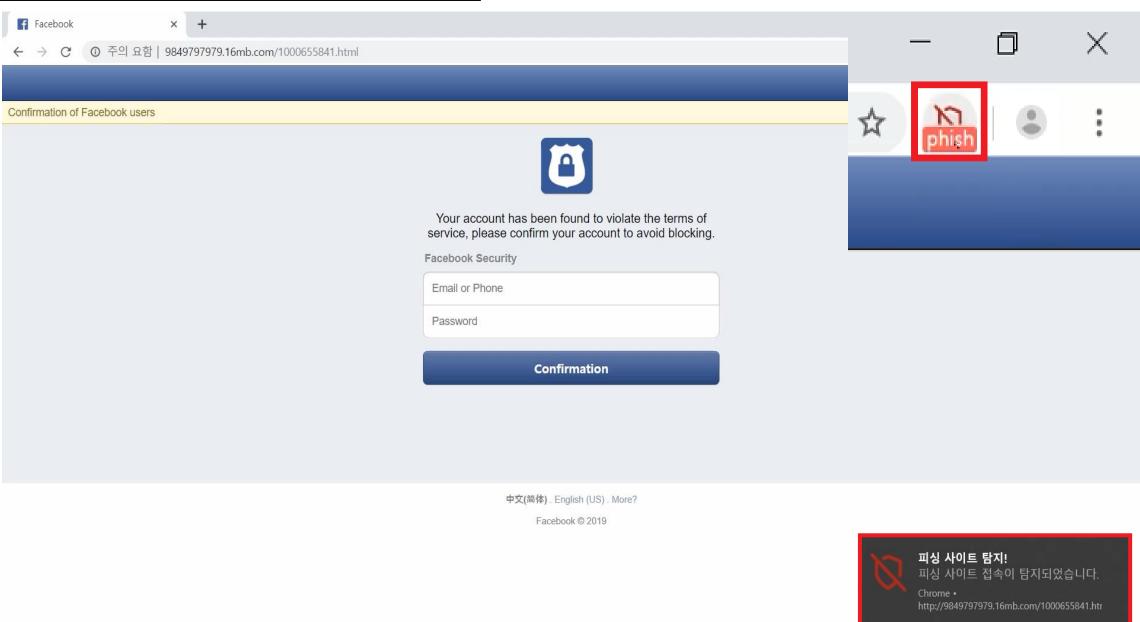
활용 – Facebook 피싱 사이트 테스트





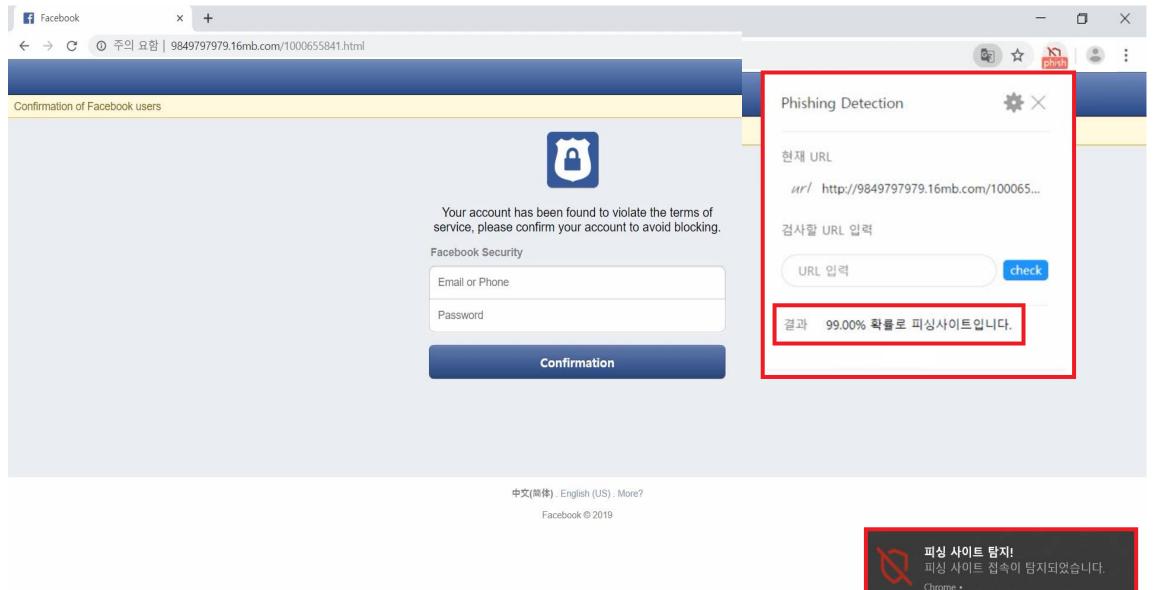
활용 – Facebook 피싱 사이트 테스트





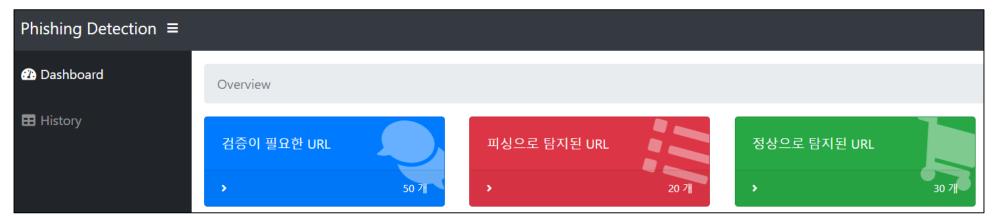
활용 – Facebook 피싱 사이트 테스트

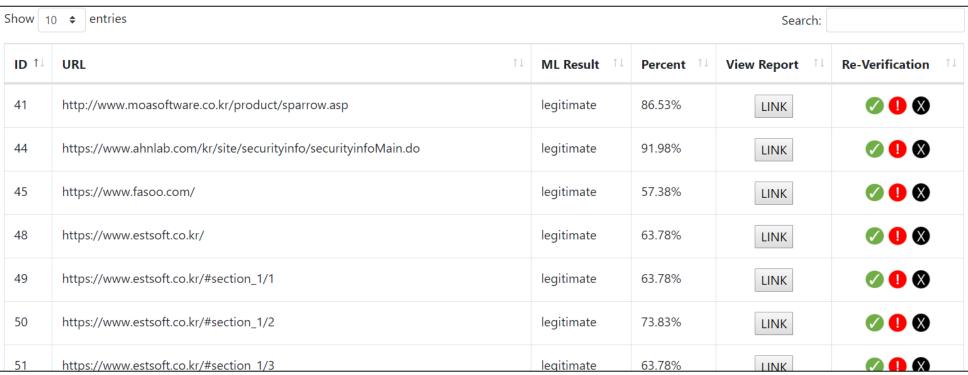




활용 – 관리자 페이지







Reference

논문

- 01. Adaboost를 이용한 효과적인 정적 분석 기반 피싱 탐지 기법 연구 / 김동현 외 3명
- 02. Detection of Phishing URLs Using Machine Learning Techniques / Joby James 외 2명
- 03. HTML 태그 순서를 이용한 불법 사이트 탐지 자동화 기술 / 이기룡 외 1명
- 04. KoNLPy와 Word2Vec을 활용한 한국어 자연어 처리 및 분석
- 05. Machine Learning based Phishing Website Detection System / Pradeep Tiwari 외 1명
- 06. Phishing Websites Detection Using Machine Learning Based Clssification Techniques / Mazharul Islam 외 1명
- 07. Phishing Website Detection based on Supervised Machine Learning with Wrapper Features Selection / Waleed Ali
- 08. 내부 네트워크에서 알려지지 않은 피싱사이트 탐지방안 / 박정욱 외 1명
- 09. 딥 러닝을 이용한 피싱 웹 페이지 분류 제안 / 이창명, 이현우, 김현구, 김혜정
- 10. 문자열 기반의 피싱 URL 탐지 방법 비교 / 김태리 외 3명
- 11. 웹사이트 특징을 이용한 휴리스틱 피싱 탐지 방안 연구 / 이진이 외 2명
- 12. 피싱사이트 실시간 탐지 기법 / 사준호 외 1명
- 13. 피싱/파밍 사례 및 대응방안 분석 / 강지윤 외 2명

서적

01. Mastering Machine Learning for Presentation Testing / Chebbi, Chiheb

Thank you

