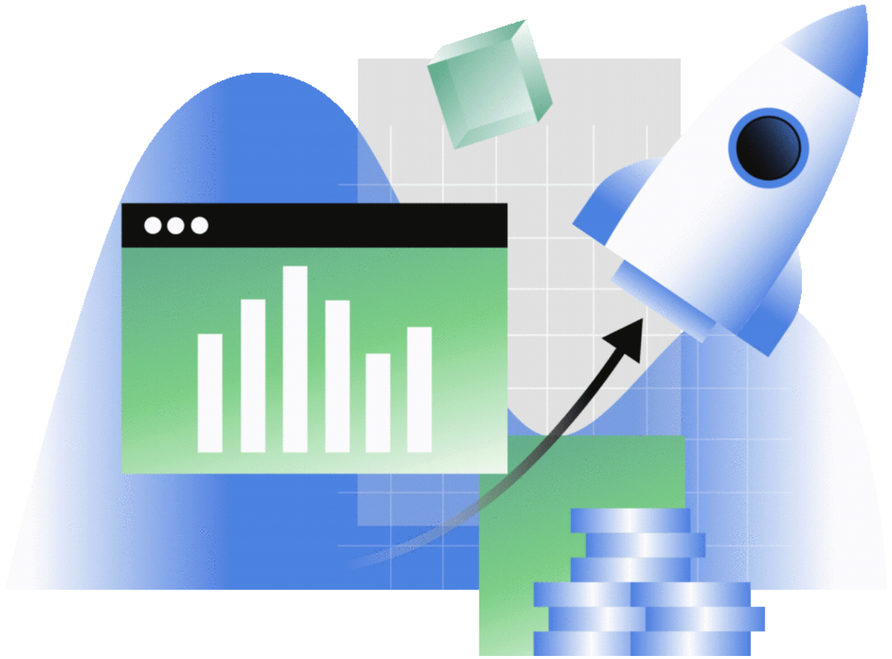
Suricata기반 ML
웹 공격 탐지 시스템

잠은죽어서자조
발표자 윤현식

발표 목차



팀원 소개

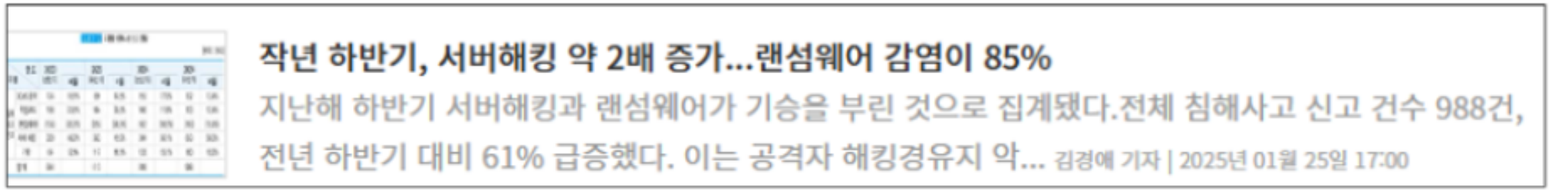
PM 윤현식
데이터 수집·전처리(정)
모델 개발(부)
데이터 분석

김동현
시스템 배포(정)
시각화·결과 분석(부)
데이터 분석

최경규
모델 개발(정)
데이터 수집·전처리(부)
데이터 분석

윤지현
시각화·결과 분석(정)
시스템 배포(부)
데이터 분석

문제 정의 및 주제 선정 이유



작년 하반기, 서버해킹 약 2배 증가...랜섬웨어 감염이 85%

지난해 하반기 서버해킹과 랜섬웨어가 기승을 부린 것으로 집계됐다.전체 침해사고 신고 건수 988건, 전년 하반기 대비 61% 급증했다. 이는 공격자 해킹경유지 악... 김경애 기자 | 2025년 01월 25일 17:00

아카마이 "전체 웹 트래픽 중 봇이 42% 차지‥65%가 악성"

A 정종길 기자 | ⏱ 입력 2024.06.26 17:19 | 댓글 0

국내 사이트 30여 곳 해킹 피해 확인...11만건 이상 개인정보 유출

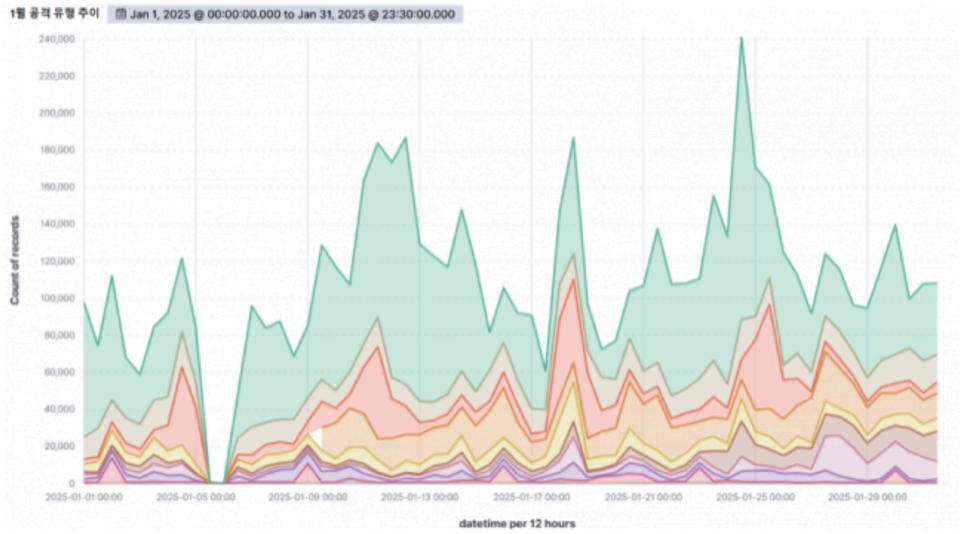
오아시스 시큐리티, CTI 플랫폼 '아가사' 통해 대규모 해킹 정황 포착...교육·의료·쇼핑·언론·금융 등 광범위한 피해 발생

공격자, 국내 약 100여 개 사이트 침투 시도...이 가운데 최소 30개 이상의 사이트 실제 해킹에 성공

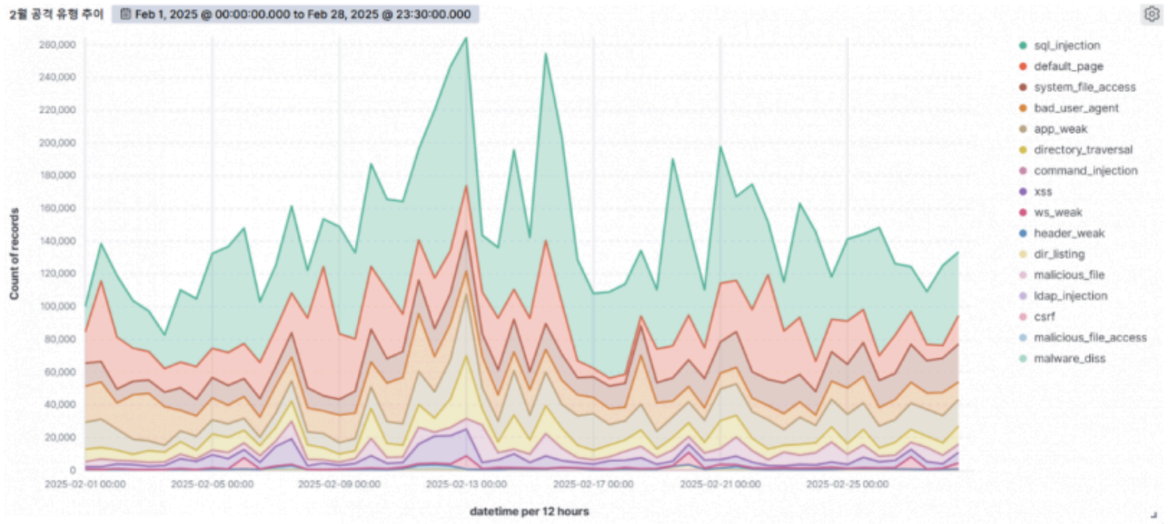
한국어 기반 웹사이트 구조와 보안 취약점을 정밀하게 파악...한국 환경에 익숙한 공격자로 추정

웹 취약점 공격 동향

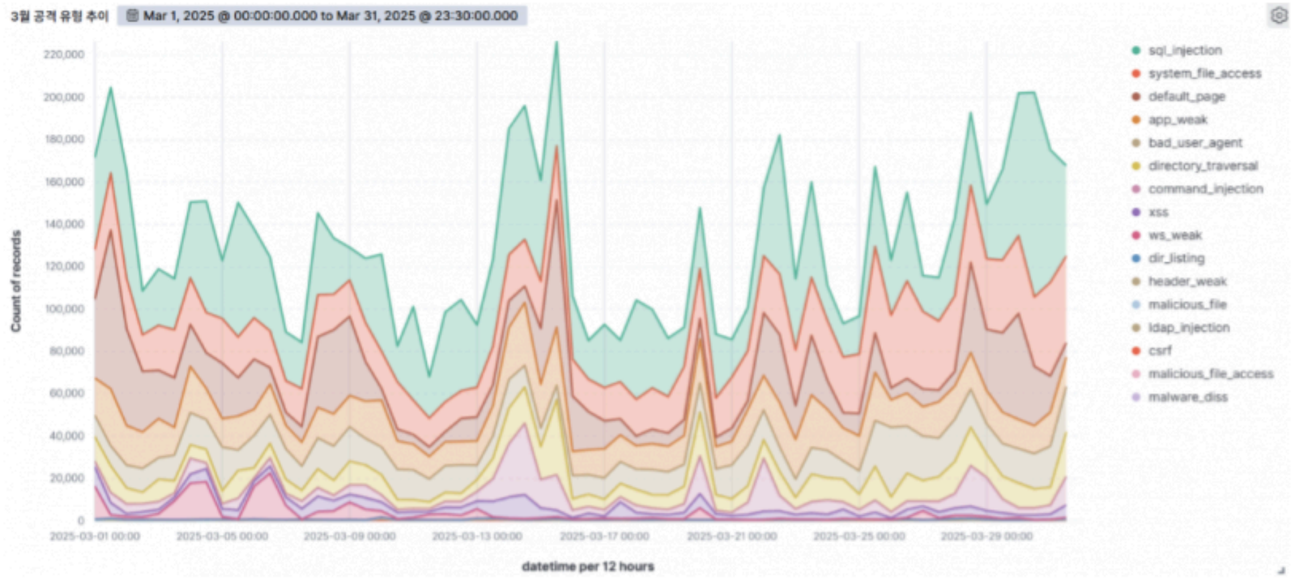
1월



2월



3월



- 모니터랩의 웹취약점 공격 보고서에서 발췌한 그래프

- 꾸준히 다양한 유형의 사이버 공격 증가

문제 정의 및 주제 선정 이유

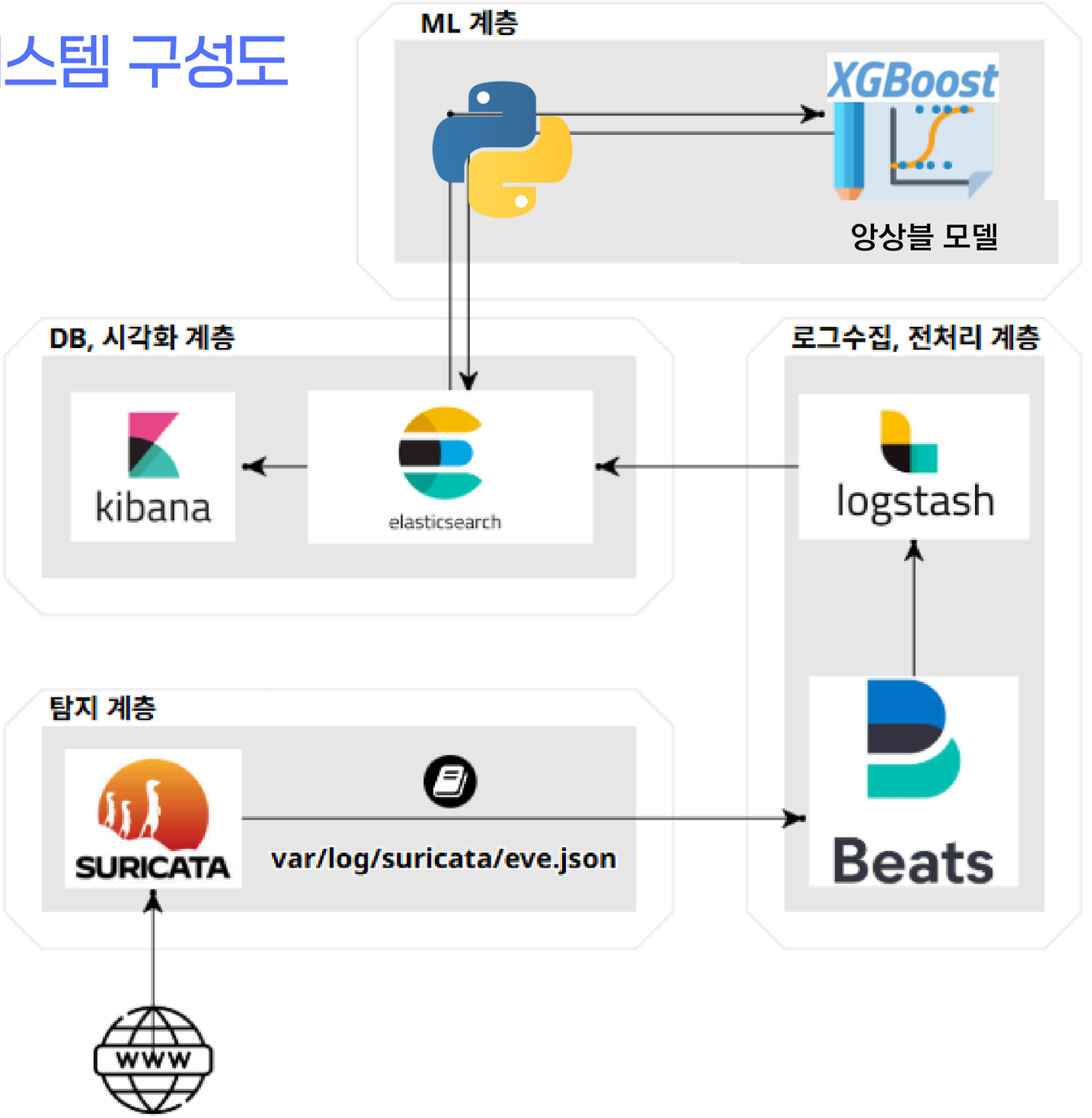
기존 SURICATA IDS의 한계

- 룰 기반 탐지
- 오탐 문제
- 수동 룰 관리의 부담
- 트래픽 양 증가 시
  관리자 부담 증가

프로젝트 목표

- 정적 룰 기반 탐지의 한계 극복
- 정탐률 향상 및 관리자 부담 감소
- SURICATA + ML 탐지 체계
- 웹 기반 공격에 특화된, 유연하고 확장 가능한 머신러닝 모델 설계

시스템 구성도



[웹서버에서 공격 발생]

↓

[Suricata (탐지)] → 로그 생성

↓

[Beats (수집)]

↓

[Logstash (전처리)]

↓

[Elasticsearch (저장)]

↓

[ML 모델 (분석)]

↓

[Elasticsearch에 예측 결과 기록]

↓

[Kibana (시각화)]

프로젝트 진행

공격 유형

Automatically_Searching_Infor	자동 정보 수집 시도
Cross_Site_Scripting	크로스 사이트 스크립팅 (XSS)
Directory_Indexing	디렉터리 인덱싱 노출
HOST_Scan	호스트 스캔
Leakage_Through_NW	네트워크를 통한 정보 유출
Path_Disclosure	경로 정보 노출
SQL_Injection	SQL 인젝션
System_Cmd_Execution	시스템 명령 실행
Vulnerability_Scan	취약점 스캔 탐지

공통 피처

payload_length	Payload 전체 길이
entropy_level	데이터 복잡도
digit_ratio	숫자 비율
url_encoding_ratio	URL 인코딩 비율
symbol_density	특수문자 밀도
repetition_ratio	반복 문자 비율
uri_depth	URL 경로 깊이
numeric_token_ratio	숫자 토큰 비율
token_count	전체 토큰 개수

프로젝트 진행

특화 피처

port_like_token_count	:80, :443 등 포트 번호 패턴 개수
scan_keyword_count	nmap, sqlmap, nikto 등 스캔 키워드 개수
word_diversity_feature	단어 다양도 (정보 탐색 유형 탐지)
url_encoding_ratio	URL 인코딩 비율
sql_keyword_count	SELECT, UNION, SLEEP 등 SQL 키워드 개수
param_key_entropy	파라미터 key 복잡도 (ex: ?id=...)
uri_depth	URL 경로 깊이
param_value_entropy	파라미터 value 복잡도

script_obfuscation_count	eval, fromCharCode 등 스크립트 난독화 함수 개수
event_handler_count	onclick, onload 등 DOM 이벤트 핸들러 개수
html_tag_depth	<태그> 중첩 깊이
path_depth	URL 경로 깊이
basename_entropy	마지막 경로 요소의 복잡도 (파일명 무작위성)
error_pattern_match	fopen, include(), Fatal error 등 에러 메 시지 포함 여부
ip_like_pattern	IP 주소 형식 탐지 (ex: 192.168.x.x)
error_code_count	4xx, 5xx 등 에러 응답 코드 수
is_backup_file	.bak, .old, ~ 등 백업 파일 여부

프로젝트 진행

피처 맵핑

```
common_features = [
    "payload_length", "entropy_level", "digit_ratio", "url_encoding_ratio",
    "symbol_density", "repetition_ratio", "uri_depth", "numeric_token_ratio", "token_count"
]

attack_specific_features = {
    "Cross_Site_Scripting": ["script_obfuscation_count", "event_handler_count", "html_tag_depth", "symbol_density"],
    "Path_Disclosure": ["path_depth", "basename_entropy", "error_pattern_match"],
    "HOST_Scan": ["ip_like_pattern", "port_like_token_count", "scan_keyword_count"],
    "System_Cmd_Execution": ["symbol_density", "payload_length", "word_diversity_feature"],
    "SQL_Injection": ["sql_keyword_count", "param_key_entropy", "param_value_entropy", "symbol_density"],
    "Vulnerability_Scan": ["scan_keyword_count"],
    "Leakage_Through_NW": ["error_code_count", "repetition_ratio", "path_depth"],
    "Directory_Indexing": ["path_depth", "is_backup_file", "basename_entropy"],
    "Automatically_Searching_Infor": ["word_diversity_feature", "payload_length"]
}
```

프로젝트 진행

모델 튜닝

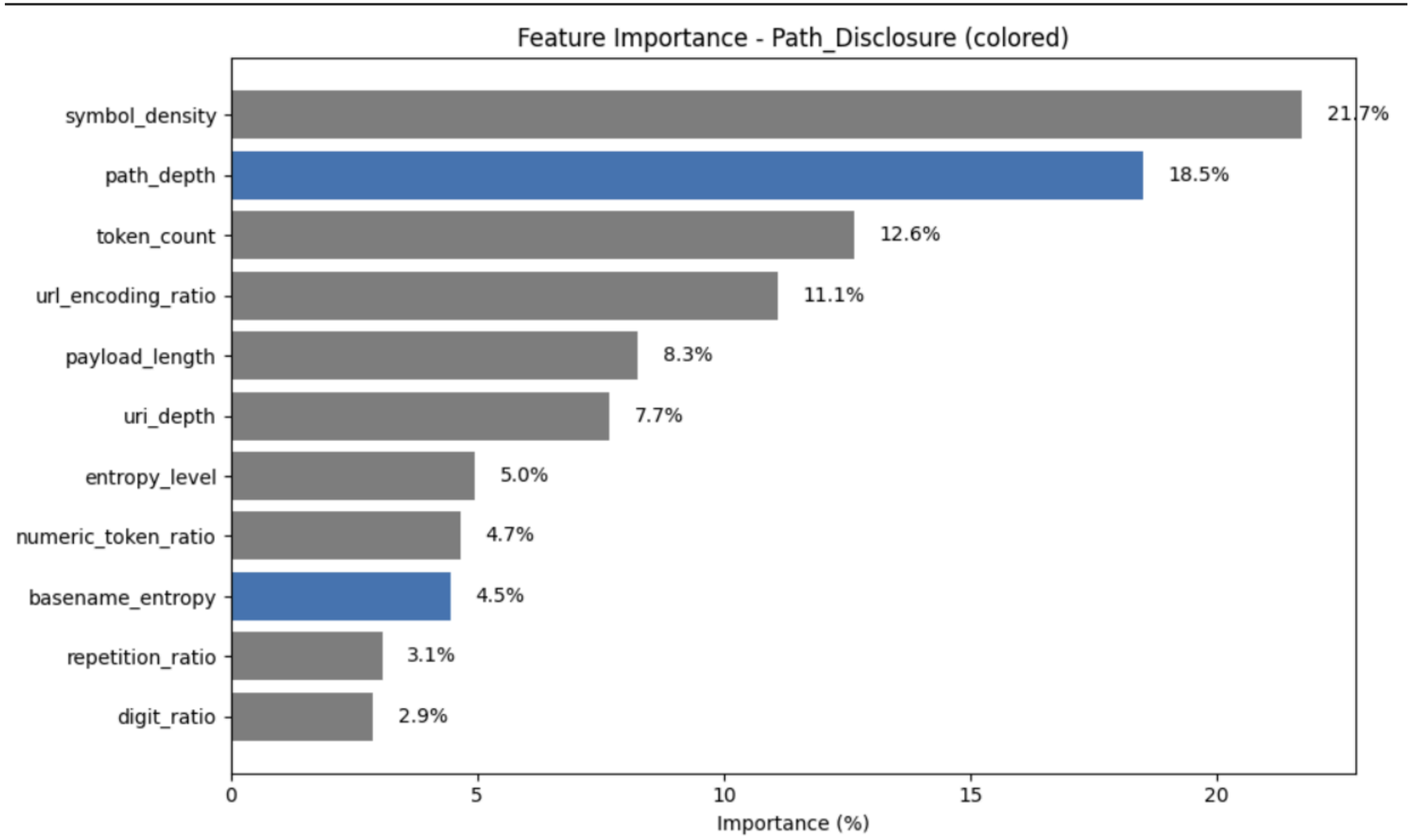
공격 유형	max_depth	learning_rate	min_child_weight	gamma	subsample	colsample_bytree
Automatically_Searching_Infor	7	0.1446854424	5	0.65280916	0.92058466	0.52911690
Cross_Site_Scripting	5	0.1598924361	3	0.57395099	0.94132574	0.56313157
Directory_Indexing	4	0.2315272505	5	0.36012924	0.87182668	0.59909577
HOST_Scan	5	0.2908860087	9	0.11259135	0.74811974	0.77853308
Leakage_Through_NW	3	0.1451088344	3	0.45976749	0.83437732	0.56924001
Path_Disclosure	5	0.2150271184	1	0.65325139	0.81170142	0.66048466
SQL_Injection	3	0.0745425576	2	0.64396210	0.95692935	0.81835750
System_Cmd_Execution	6	0.2552005438	1	0.43409407	0.86284998	0.50717594
Vulnerability_Scan	5	0.2699156238	1	0.60229054	0.96981412	0.66161460

optuna 라이브러리를 통해
하이퍼파라미터 튜닝 진행

```
param = {
    'verbosity': 0,
    'objective': 'binary:logistic',
    'eval_metric': 'logloss',
    'use_label_encoder': False,
    'tree_method': 'hist',
    'max_depth': trial.suggest_int('max_depth', 3, 10),
    'learning_rate': trial.suggest_float('learning_rate', 0.01, 0.3),
    'min_child_weight': trial.suggest_int('min_child_weight', 1, 10),
    'gamma': trial.suggest_float('gamma', 0.0, 1.0),
    'subsample': trial.suggest_float('subsample', 0.5, 1.0),
    'colsample_bytree': trial.suggest_float('colsample_bytree', 0.5, 1.0),
    'n_estimators': 300
}
```

프로젝트 진행

피처 중요도



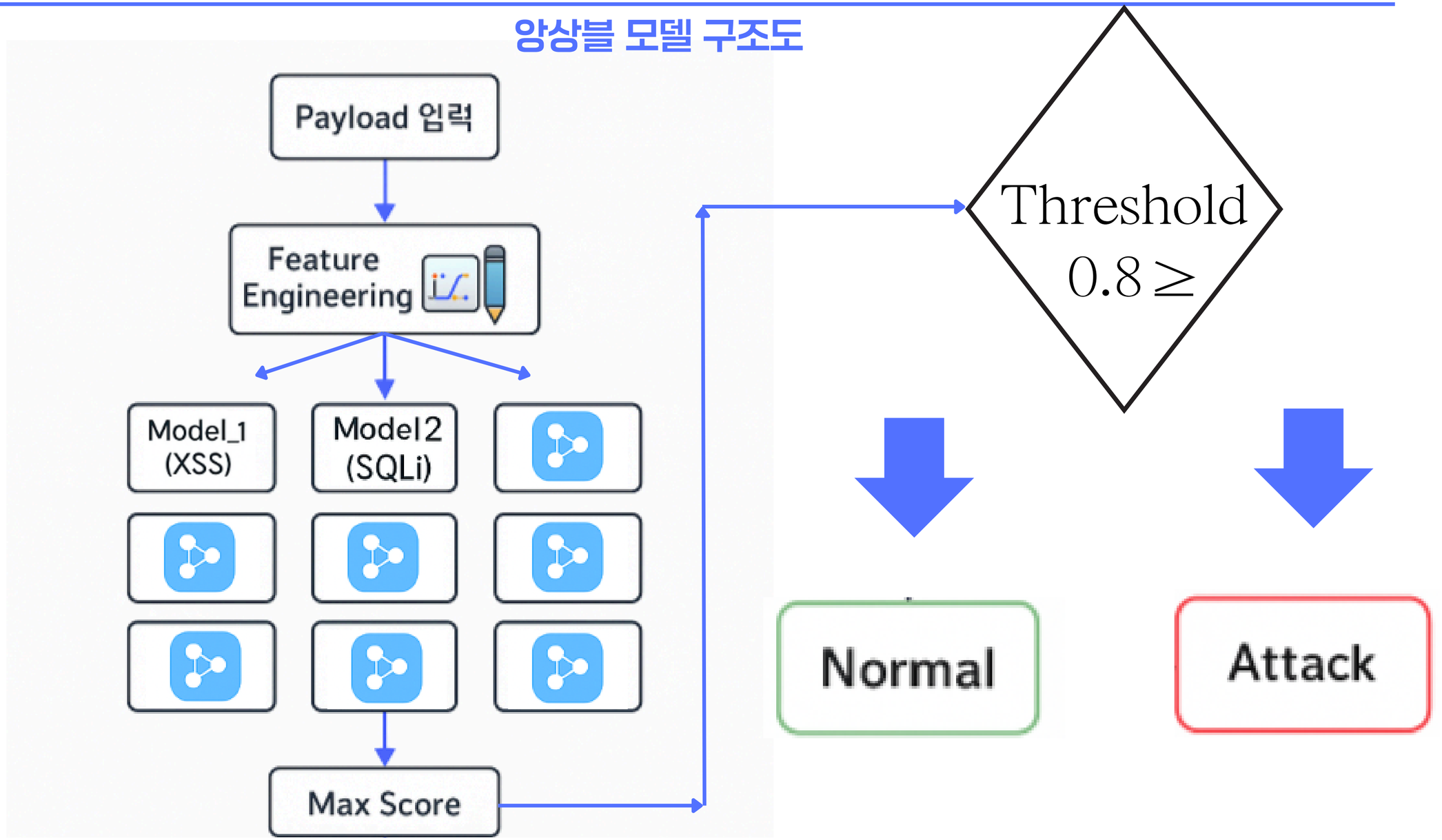
프로젝트 진행

공격 유형	Accuracy (정확도)	precision (정밀도)	Recall (재현율)	F1-Score
Automatically_Searching_Infor	100%	100%	100%	100%
Cross_Site_Scripting	99.8%	99.95%	99.64%	99.80%
Directory_Indexing	98.81%	98.81%	98.81%	98.81%
HOST_Scan	99.66%	100%	99.32%	99.66%
Leakage_Through_NW	99.10%	98.21%	100%	99.10%
Path_Disclosure	99.84%	99.9%	99.79%	99.84%
SQL_Injection	99.93%	100%	99.85%	99.93%
System_Cmd_Execution	99.92%	100%	99.85%	99.92%
Vulnerability_Scan	99.95%	100%	99.89%	99.95%

모든 유형에서
98%이상의 준수한
탐지 성능

프로젝트 진행

앙상블 모델 구조도



프로젝트 결과 - 성능평가

```
🎯 Binary F1 Score (Attack class): 0.9474
✅ Accuracy: 0.9130
📊 Macro F1 Score: 0.8474
📉 Weighted F1 Score: 0.9047

📋 Classification Report:
              precision    recall  f1-score        support
0              0.982766  0.602961  0.747379    3783.000000
1              0.902480  0.997131  0.947447   13940.000000
accuracy       0.912994  0.912994  0.912994       0.912994
macro avg      0.942623  0.800046  0.847413   17723.000000
weighted avg   0.919617  0.912994  0.904742   17723.000000

🧩 Confusion Matrix:
             Pred_Normal  Pred_Attack
True_Normal         2281         1502
True_Attack           40        13900

🟤 ROC AUC: 0.9372
🟢 PR AUC: 0.9744
```

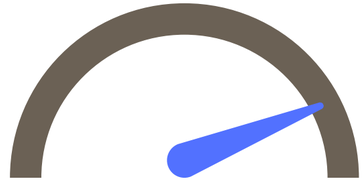
프로젝트 결과 - 성능평가

```
🎯 Binary F1 Score (Attack class): 0.9781
✅ Accuracy: 0.9780
📊 Macro F1 Score: 0.9780
📉 Weighted F1 Score: 0.9780

📋 Classification Report:
              precision    recall  f1-score       support
0              0.978385  0.97770  0.978042  10000.00000
1              0.977716  0.97840  0.978058  10000.00000
accuracy       0.978050  0.97805  0.978050      0.97805
macro avg      0.978050  0.97805  0.978050  20000.00000
weighted avg   0.978050  0.97805  0.978050  20000.00000

🧩 Confusion Matrix:
             Pred_Normal  Pred_Attack
True_Normal         9777          223
True_Attack          216         9784

🟠 ROC AUC: 0.9978
🟢 PR AUC: 0.9976
```



성능 평가

정상 트래픽 오탐 2.23%
공격 탐지 실패 2.16%

임계치

공격 탐지의 신뢰도를 높이기 위해 임계치를 0.8로 조정

처리 속도

초당 300~500건 처리

프로젝트 결과 - 성능평가

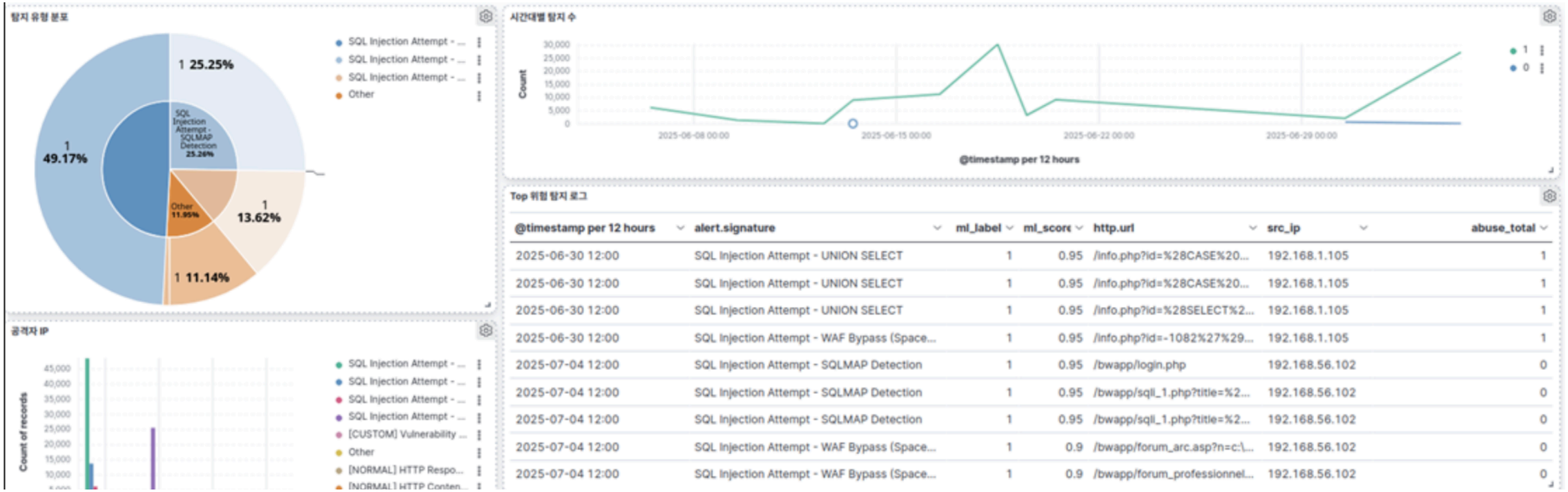
```
2025-07-04 18:34:04 INFO Fetch ▶📥 Fetch 500건 , 큐=500, 0.046s
2025-07-04 18:34:05 INFO Fetch ▶📥 Fetch 500건 , 큐=500, 0.119s
2025-07-04 18:34:05 INFO Fetch ▶📥 Fetch 500건 , 큐=700, 0.153s
2025-07-04 18:34:05 INFO Fetch ▶📥 Fetch 500건 , 큐=1200, 0.096s
2025-07-04 18:34:05 INFO Fetch ▶📥 Fetch 500건 , 큐=1700, 0.087s
2025-07-04 18:34:06 INFO Worker-2 ▶⏱Worker 100건  ML 0.604s  AbuseIPDB 0.564s
 총 1.167s
2025-07-04 18:34:06 INFO Worker-7 ▶⏱Worker 100건  ML 0.614s  AbuseIPDB 0.574s
 총 1.188s
2025-07-04 18:34:06 INFO Worker-6 ▶⏱Worker 100건  ML 0.740s  AbuseIPDB 0.597s
 총 1.337s
2025-07-04 18:34:06 INFO Fetch ▶📥 Fetch 480건 , 큐=1880, 0.855s
2025-07-04 18:34:06 INFO Worker-3 ▶⏱Worker 100건  ML 0.709s  AbuseIPDB 0.629s
 총 1.338s
2025-07-04 18:34:06 INFO Worker-1 ▶⏱Worker 100건  ML 0.718s  AbuseIPDB 0.636s
 총 1.353s
2025-07-04 18:34:06 INFO Worker-5 ▶⏱Worker 100건  ML 0.588s  AbuseIPDB 0.635s
 총 1.223s
2025-07-04 18:34:06 INFO Worker-8 ▶⏱Worker 100건  ML 0.711s  AbuseIPDB 0.669s
 총 1.381s
2025-07-04 18:34:06 INFO Bulk ▶💾 Bulk 100건 저장
2025-07-04 18:34:06 INFO Fetch ▶📥 Fetch 0건 , 큐=1480, 0.101s
2025-07-04 18:34:06 INFO Bulk ▶💾 Bulk 100건 저장
2025-07-04 18:34:06 INFO Worker-4 ▶⏱Worker 100건  ML 0.743s  AbuseIPDB 1.047s
```

```
2025-07-04 18:33:55 INFO Worker-2 ▶⏱Worker 100건  ML 0.151s  AbuseIPDB 0.525s
 총 0.676s
2025-07-04 18:33:55 INFO Bulk ▶💾 Bulk 100건 저장
2025-07-04 18:33:55 INFO Worker-3 ▶⏱Worker 33건  ML 0.079s  AbuseIPDB 0.645s
총 0.724s
2025-07-04 18:33:56 INFO Worker-6 ▶⏱Worker 100건  ML 0.634s  AbuseIPDB 1.781s
 총 2.415s
2025-07-04 18:33:56 INFO Bulk ▶💾 Bulk 100건 저장
2025-07-04 18:33:56 INFO Bulk ▶💾 Bulk 33건 저장
2025-07-04 18:33:59 INFO Fetch ▶📥 Fetch 500건 , 큐=500, 0.042s
2025-07-04 18:33:59 INFO Fetch ▶📥 Fetch 500건 , 큐=500, 0.194s
2025-07-04 18:33:59 INFO Fetch ▶📥 Fetch 500건 , 큐=700, 0.115s
2025-07-04 18:33:59 INFO Fetch ▶📥 Fetch 307건 , 큐=1007, 0.072s
2025-07-04 18:33:59 INFO Fetch ▶📥 Fetch 0건 , 큐=1007, 0.059s
2025-07-04 18:34:00 INFO Worker-7 ▶⏱Worker 100건  ML 0.702s  AbuseIPDB 0.500s
 총 1.201s
2025-07-04 18:34:00 INFO Worker-6 ▶⏱Worker 100건  ML 0.881s  AbuseIPDB 0.553s
 총 1.433s
2025-07-04 18:34:00 INFO Worker-2 ▶⏱Worker 100건  ML 0.738s  AbuseIPDB 0.508s
 총 1.246s
2025-07-04 18:34:00 INFO Worker-5 ▶⏱Worker 100건  ML 0.881s  AbuseIPDB 0.561s
 총 1.441s
2025-07-04 18:34:00 INFO Worker-4 ▶⏱Worker 100건  ML 0.912s  AbuseIPDB 0.524s
 총 1.437s
```

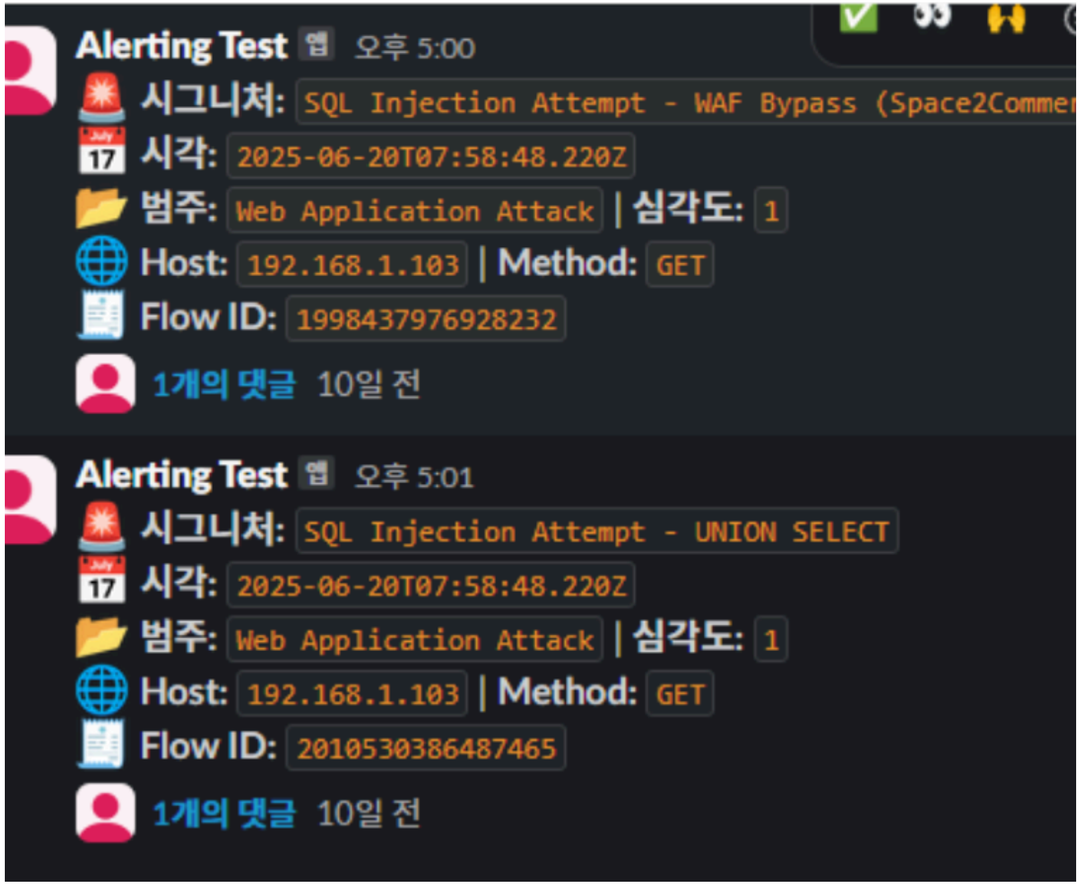
처리 속도

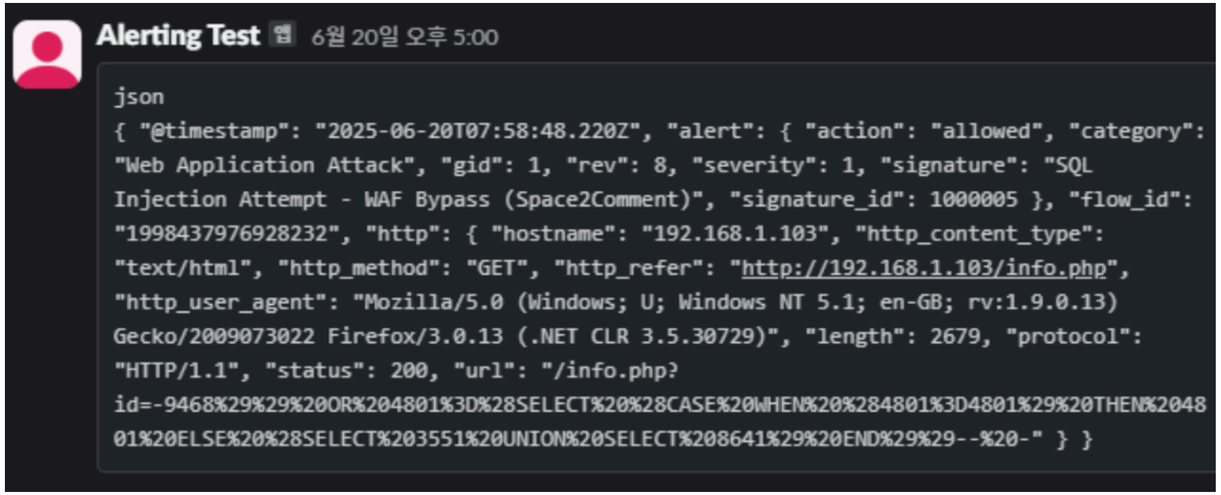
초당 300~500건 처리

프로젝트 결과 - 대시보드

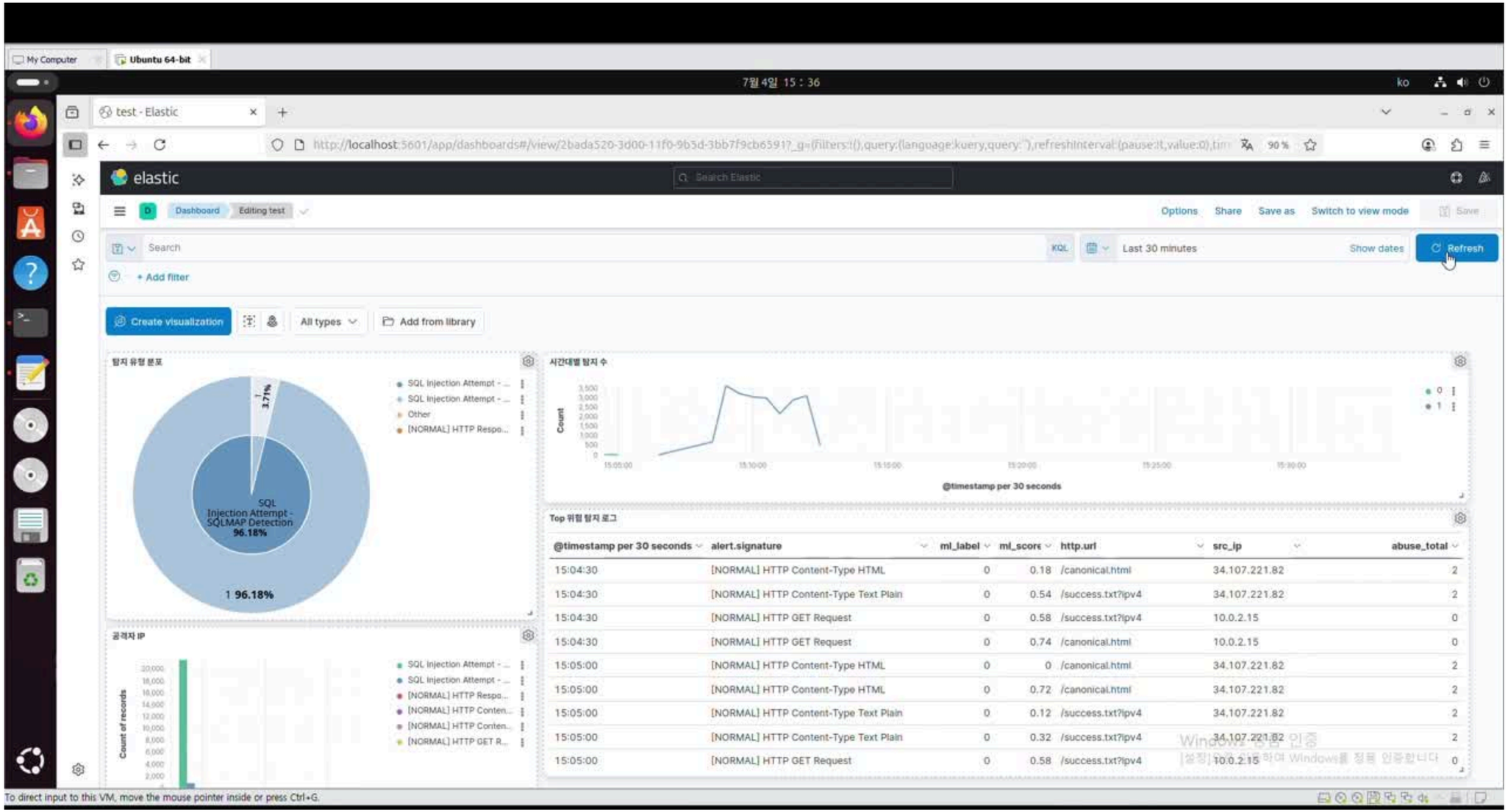

탐지 유형 분포
1 25.25%
1 49.17%
1 13.62%
1 11.14%
SQL Injection Attempt - SQLMAP Detection 25.26%
Other 11.95%
SQL Injection Attempt - ...
SQL Injection Attempt - ...
SQL Injection Attempt - ...
Other
시간대별 탐지 수
Count
30,000
25,000
20,000
15,000
10,000
5,000
0
2025-06-08 00:00
2025-06-15 00:00
2025-06-22 00:00
2025-06-29 00:00
@timestamp per 12 hours
1
0
공격자 IP
Count of records
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
SQL Injection Attempt - ...
SQL Injection Attempt - ...
SQL Injection Attempt - ...
SQL Injection Attempt - ...
[CUSTOM] Vulnerability ...
Other
[NORMAL] HTTP Respo...
Top 위협 탐지 로그
@timestamp per 12 hours
alert.signature
ml_label
ml_score
http.url
src_ip
abuse_total
2025-06-30 12:00 SQL Injection Attempt - UNION SELECT 1 0.95 /info.php?id=%28CASE%20... 192.168.1.105 1
2025-06-30 12:00 SQL Injection Attempt - UNION SELECT 1 0.95 /info.php?id=%28CASE%20... 192.168.1.105 1
2025-06-30 12:00 SQL Injection Attempt - UNION SELECT 1 0.95 /info.php?id=%28SELECT%2... 192.168.1.105 1
2025-06-30 12:00 SQL Injection Attempt - WAF Bypass (Space... 1 0.95 /info.php?id=-1082%27%29... 192.168.1.105 1
2025-07-04 12:00 SQL Injection Attempt - SQLMAP Detection 1 0.95 /bwapp/login.php 192.168.56.102 0
2025-07-04 12:00 SQL Injection Attempt - SQLMAP Detection 1 0.95 /bwapp/sqli_1.php?title=%2... 192.168.56.102 0
2025-07-04 12:00 SQL Injection Attempt - SQLMAP Detection 1 0.95 /bwapp/sqli_1.php?title=%2... 192.168.56.102 0
2025-07-04 12:00 SQL Injection Attempt - WAF Bypass (Space... 1 0.9 /bwapp/forum_arc.asp?n=c:\... 192.168.56.102 0
2025-07-04 12:00 SQL Injection Attempt - WAF Bypass (Space... 1 0.9 /bwapp/forum_professionnel... 192.168.56.102 0

프로젝트 결과 - 슬랙 알림


Alerting Test 웹 오후 5:00
시그니처: SQL Injection Attempt - WAF Bypass (Space2Commer
시각: 2025-06-20T07:58:48.220Z
범주: Web Application Attack | 심각도: 1
Host: 192.168.1.103 | Method: GET
Flow ID: 1998437976928232
1개의 댓글 10일 전
Alerting Test 웹 오후 5:01
시그니처: SQL Injection Attempt - UNION SELECT
시각: 2025-06-20T07:58:48.220Z
범주: Web Application Attack | 심각도: 1
Host: 192.168.1.103 | Method: GET
Flow ID: 2010530386487465
1개의 댓글 10일 전


Alerting Test 웹 6월 20일 오후 5:00
json
{ "@timestamp": "2025-06-20T07:58:48.220Z", "alert": { "action": "allowed", "category": "Web Application Attack", "gid": 1, "rev": 8, "severity": 1, "signature": "SQL Injection Attempt - WAF Bypass (Space2Comment)", "signature_id": 1000005 }, "flow_id": "1998437976928232", "http": { "hostname": "192.168.1.103", "http_content_type": "text/html", "http_method": "GET", "http_refer": "http://192.168.1.103/info.php", "http_user_agent": "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-GB; rv:1.9.0.13) Gecko/2009073022 Firefox/3.0.13 (.NET CLR 3.5.30729)", "length": 2679, "protocol": "HTTP/1.1", "status": 200, "url": "/info.php?id=-9468%29%29%20OR%204801%3D%28SELECT%20%28CASE%20WHEN%20%284801%3D4801%29%20THEN%204801%20ELSE%20%28SELECT%203551%20UNION%20SELECT%208641%29%20END%29%29--%20-" } }

시연영상


My Computer
Ubuntu 64-bit
7월 4일 15 : 36
ko
test - Elastic
elastic
Search Elastic
Dashboard
Editing test
Options
Share
Save as
Switch to view mode
Save
Search
KQL
Last 30 minutes
Show dates
Refresh
+ Add filter
Create visualization
All types
Add from library
탐지 유형 분포
SQL Injection Attempt - ...
SQL Injection Attempt - ...
Other
[NORMAL] HTTP Respo...
3.71%
SQL Injection Attempt - SQLMAP Detection 96.18%
1 96.18%
시간대별 탐지 수
Count
@timestamp per 30 seconds
Top 위협 탐지 로그
@timestamp per 30 seconds
alert.signature
ml_label
ml_score
http.url
src_ip
abuse_total
15:04:30 [NORMAL] HTTP Content-Type HTML 0 0.18 /canonical.html 34.107.221.82 2
15:04:30 [NORMAL] HTTP Content-Type Text Plain 0 0.54 /success.txt?ipv4 34.107.221.82 2
15:04:30 [NORMAL] HTTP GET Request 0 0.58 /success.txt?ipv4 10.0.2.15 0
15:04:30 [NORMAL] HTTP GET Request 0 0.74 /canonical.html 10.0.2.15 0
15:05:00 [NORMAL] HTTP Content-Type HTML 0 0 /canonical.html 34.107.221.82 2
15:05:00 [NORMAL] HTTP Content-Type HTML 0 0.72 /canonical.html 34.107.221.82 2
15:05:00 [NORMAL] HTTP Content-Type Text Plain 0 0.12 /success.txt?ipv4 34.107.221.82 2
15:05:00 [NORMAL] HTTP Content-Type Text Plain 0 0.32 /success.txt?ipv4 34.107.221.82 2
15:05:00 [NORMAL] HTTP GET Request 0 0.58 /success.txt?ipv4 10.0.2.15 0
공격자 IP
Count of records
To direct input to this VM, move the mouse pointer inside or press Ctrl+G.

프로젝트 결과

프로젝트 목표 달성

- SURICATA의 한계를 극복하고, 머신러닝 기반의 하이브리드 탐지 체계 구축
- 정탐률 향상 및 관리자 운영 부담 감소에 기여

주요 기술적 성과

- 공격 유형별 앙상블 모델 학습 및 튜닝
- ML 예측 + TI 기반 분류
- ELASTICSEARCH + KIBANA 대시보드 시각화 완료

시사점 및 확장 가능성

- 정적 룰 기반 탐지 시스템의 한계를 머신러닝으로 보완 가능
- 향후 IP 기반 TI 외에 도메인, URL, 행동 기반 탐지 확장 예정
- 실시간 대응 시스템과 연계 시 SOC 운영 최적화 가능성 기대

Q&A

궁금하신 점이 있다면 자유롭게 질문해주세요

감사합니다