딥러닝 기반 악성 트래픽 탐지 웹사이트 개발

최종 발표

NET-SERT (Network-Security Emergency Response Team)

팀원: 김지원 · 한유민 · 이상원 · 전나현

We are gonna talk about...

프로젝트 목표

기존의 목표와
수정된 목표

데이터셋

사용한 데이터셋과
데이터셋 정보

웹사이트 구현

스트림릿을 사용한
웹사이트 제작

느낀점과 확장성

기존 목표

1. ET-BERT 모델을 사용해서
2. 직접 수집한 데이터셋으로
3. 악성 트래픽 탐지 웹사이트 개발

기존 목표

1 ET-BERT 모델을 사용해서

2 다양한 데이터셋을 통합하여

3 악성 트래픽 탐지 웹사이트 개발

직접 수집한 데이터셋으로

악성 트래픽을 직접 수집하는데 따르는 **부담, 위험**

다양한 악성/양성 트래픽 데이터셋을 통합하고 **악성 트래픽 종류**를 분석

악성 트래픽 데이터란?

서비스 거부 공격 (DoS/DDoS)

SYN / UDP / ICMP Flooding

특정 서버 or 네트워크에 **과도한 트래픽** 발생 → 정상 사용자 접근 방해·시스템 마비

- 총 패킷 수↑ / 초당 패킷 수↑
- 연결 시도 횟수 대비 성공 횟수↓

악성코드 공격

C&C 통신 / Worm·Virus 전파 / Botnet

감염된 시스템이 외부 서버와 통신 or 다른 시스템으로 악성코드 확산

- 일정한 패킷 도착 간격
- 길고 주기적인 플로우 지속 시간
- 페이로드 엔트로피↑
- 비표준 HTTP 헤더 / 비정상적인 TLS Cipher Suite 사용

침입/탐색 공격 (Probing/Intrusion)

Port / Vulnerability Scanning

시스템 취약점·정보 파악 위한 사전 활동

- 짧은 플로우 지속 시간
- 목적지 포트 고유성↑

사용자·정보 탈취 공격

Web Attack / Data Exfiltration / Backdoor

시스템 취약점 이용해 **권한 획득** or 내부 기밀 **정보 유출**

- 총 바이트 수↑
- 특정 키워드 패턴

통합 대상 데이터셋

NSL-KDD

- DoS / Probe / R2L / U2R
- 약 14만 개 샘플
- 41개 feature (프로토콜 유형, 서비스, 플래그, 연결 시간, 전송 바이트 수 등)
- 다중 분류 라벨 (Benign + 상세 공격 클래스)

CSE-CIC-IDS2018

- Web attacks / PortScan / DoS·DDoS / Botnet
- 수백만 개 샘플 → 30% 사용
- 80개 feature (플로우 지속 시간, 패킷 길이 통계, 바이트 비율 등)
- 다중 분류 라벨

CTU-13

- Botnet (13가지 시나리오 : Neris, Rbot, Virut, Conficker 등)
- 약 8100만 개 샘플
- 15개 feature
- 2진 분류 라벨 (Normal / Attack)

USTC-TFC2016

- 10가지 악성 트래픽 + 10가지 정상 트래픽
- 수십만~수백만 플로우 (원시 트래픽 데이터 기반)
- 다중 분류 라벨

데이터셋 통합 과정

데이터셋 로드

↓

전처리

- 라벨 통일 : 정상 → 0 / 악성 → 1
- 데이터셋 출처 컬럼 추가
- NaN / 빈 라벨 삭제
- Text 변환

↓

최종 컬럼 구조

text | label_bin | label_raw | norm_label | source | feature1 | feature2 | ...

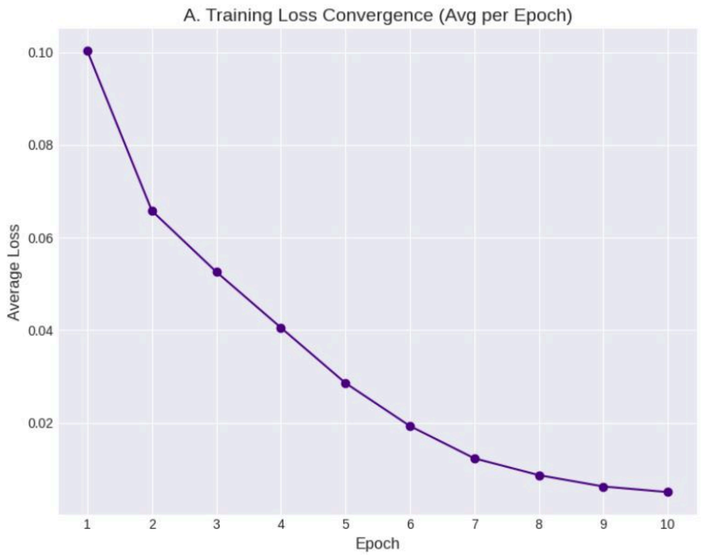
↓

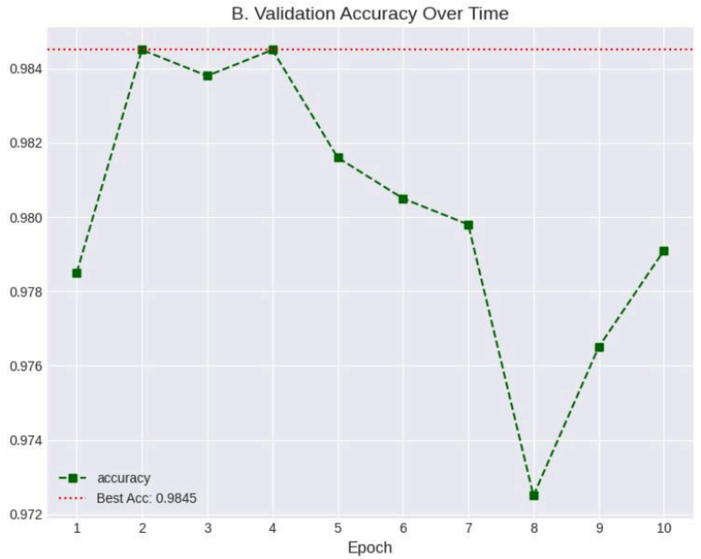
.tsv 변환

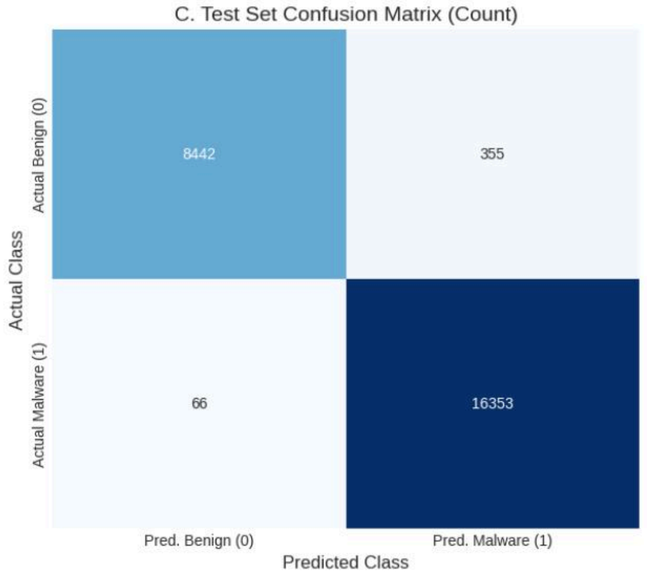
↓

train / val / test 분할

파인튜닝 결과







Accuracy

98.33%

Streamlit으로 구현

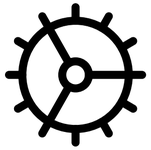
System Architecture

Input

PCAP 파일 업로드

(User Input)

→

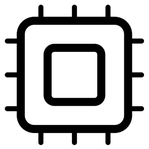


Process

TShark 전처리
& 벡터 변환

(Preprocessing)

→



Model

ET-BERT 모델 추론

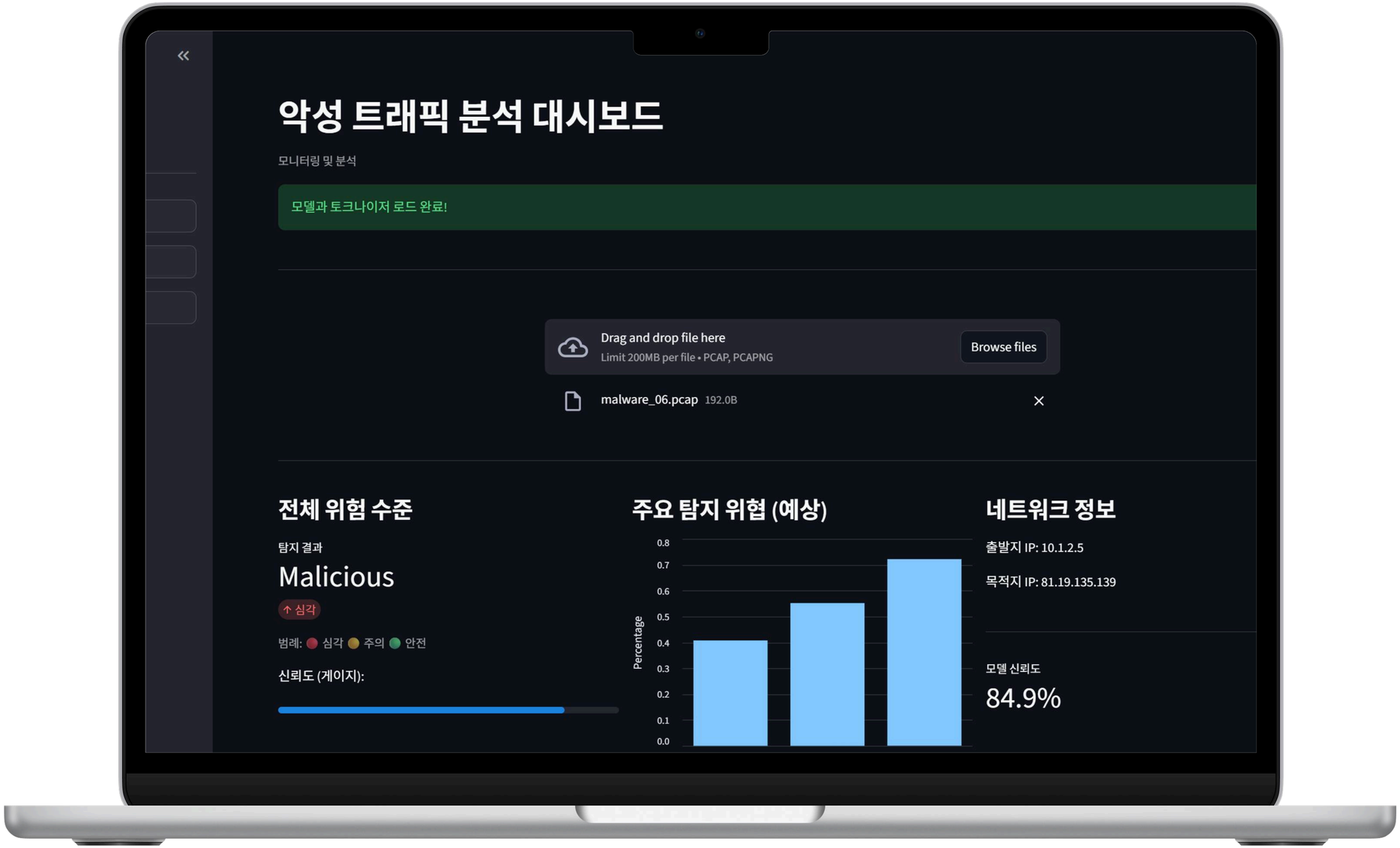
(AI Inference)

→

Output

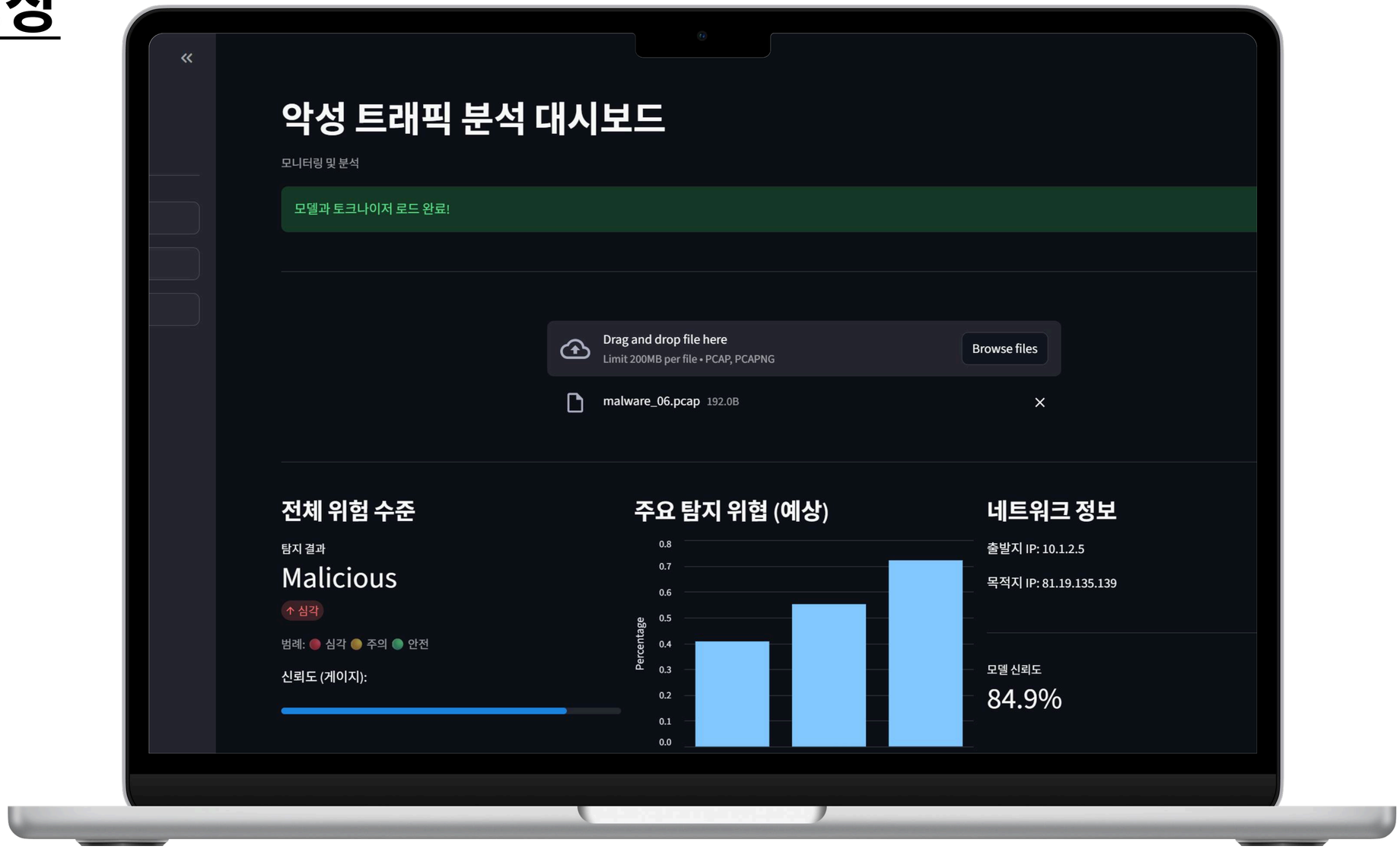
Streamlit 실시간
대시보드

(Visualization)


악성 트래픽 분석 대시보드
모니터링 및 분석
모델과 토크나이저 로드 완료!
Drag and drop file here
Limit 200MB per file • PCAP, PCAPNG
Browse files
malware_06.pcap 192.0B
전체 위험 수준
탐지 결과
Malicious
↑ 심각
범례: 심각 주의 안전
신뢰도 (게이지):
주요 탐지 위협 (예상)
Percentage
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
네트워크 정보
출발지 IP: 10.1.2.5
목적지 IP: 81.19.135.139
모델 신뢰도
84.9%

입력한 패킷이 정상인지, 악성인지 판별하는 웹사이트

시연 영상



프로젝트 확장성

메타데이터 파싱

- 현재 - UI에서 첨부한 패킷 파일의 정상/악성 여부, 모델의 신뢰도, 패킷의 출발지IP-목적지IP만 출력
- 모델이 패킷의 다른 메타데이터-패킷 전송 계층 프로토콜의 유형, 패킷의 바이트, 암호화 여부-도 함께 출력하도록 UI 개선

데이터셋 개선

- 모델이 사전학습한 데이터셋과 직접 미세조정한 데이터셋 모두 QUIC과 같은 비교적 최신의 프로토콜 패킷은 많이 포함하지 않음
  - 최신 악성 트래픽에 대해 불리할 가능성 높음
- 모델을 미세조정할 데이터셋의 최신의 패킷 데이터 비중을 높여 극복

악성 트래픽 탐지 자동화

- 현재 - 사용자가 첨부한 패킷에 대해서만 검사
- 사용자가 URL을 입력
  - 해당 URL에 접근하는 패킷에 대해 정상/악성 트래픽 여부 검사

프로젝트 의의 (느낀 점, 배운 점)

네트워크 보안 및 딥러닝 학습

- 다양한 네트워크 트래픽 분석 및 보안 관련 논문 학습
- ET-BERT 모델을 논문대로 재현
  - 다른 다양한 데이터셋으로 미세조정하고 결과를 확인
- ET-BERT 모델을 미세조정하는 과정에서 pytorch 딥러닝 학습
- 딥러닝 모델 구조 학습 → 실무에서 사용할 수 있는 기술과 툴 학습

데이터셋 구축 및 통합

- 정상 트래픽과 악성 트래픽을 직접 수집해보며 와이어샤크와 같은 패킷 캡처 기술 경험
- 수집한 트래픽을 분석
  - TCP, UDP, HTTP, QUIC 등 다양한 네트워크 패킷 학습
- 여러 패킷 데이터셋을 통합하여 재구축

배운 점

- NLP 과목에서 BERT에 대해 학습 중인데 보안 분야 파생 모델인 ET-BERT를 공부하고 실제로 재현하며 더 많은 이해와 흥미를 갖게 됨

모델 완성 및 UI 제작

- 통합하여 재구축한 데이터셋으로 (사전학습된) 모델 파인튜닝
  - 정확도가 98%까지 오르는 등 성능 향상 확인
- 도출한 모델로 사용자가 패킷을 첨부하면 그 패킷이 정상인지 악성인지 판별하는 UI 제작
  - 초기 목표 달성

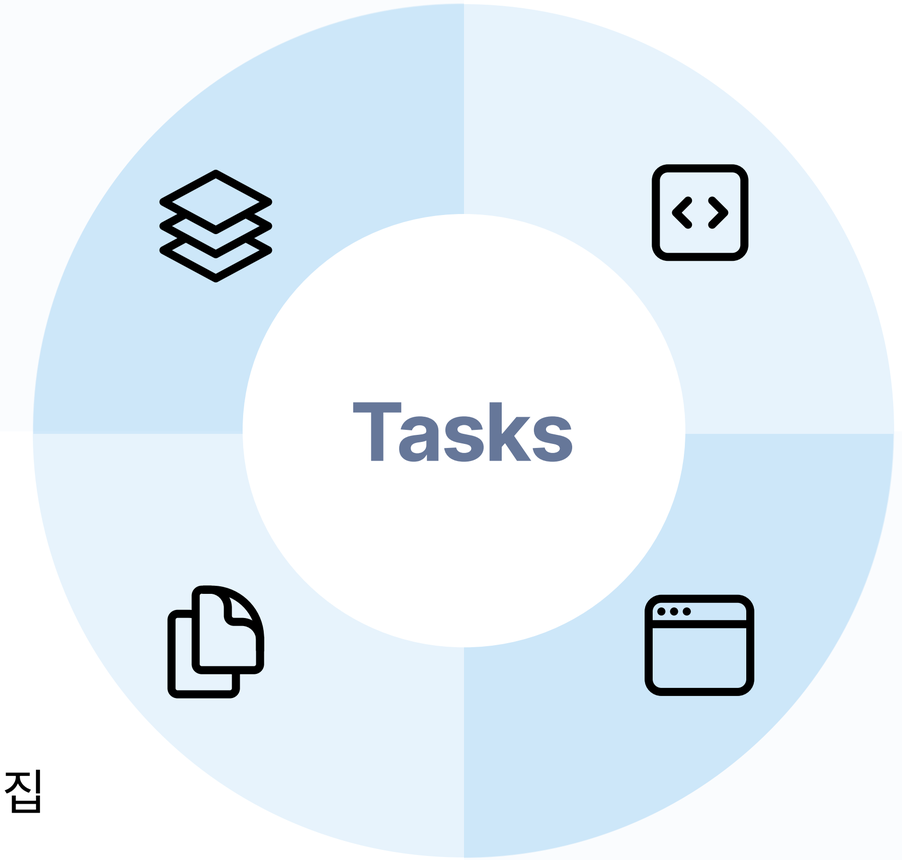
느낀점과 역할

김지원

데이터셋 통합

이상원

ET-BERT 모델 변형



전나현

정상 트래픽 수집

한유민

UI 구현

공통 - 발표자료 제작, 발표

Thanks.

질문 있으신가요?