AI기반 사기거래 탐지 중고거래 서비스

팀명 :	PIO
지도교수 :	이병천
팀장 :	김평안
팀원 :	송명석
	신준혁
	최중건

2025. 11.

중부대학교 정보보호학전공

목 차

1. 서론

1.1 연구 배경 및 필요성

1.2 연구 목적 및 주제 선정

2. 관련 연구

2.1 언어, API

2.1.1 fast API

2.1.2 AI서버 API구성

2.1.3 시스템 아키텍처

2.2 사기 탐지 및 이미지 비교

2.2.1 채팅 사기 탐지 모듈

2.2.1.1 Ollama 선택이유

2.2.1.2 채팅 사기 탐지 과정

2.2.2 이미지 유사도 검사 모듈

2.2.3 성능 검증

2.2.4 이미지 검증 과정

3. 본론

3.1 시스템 구성도

3.2 앱 제작

3.3 이미지 검증 과정 개발

3.4 채팅 내 사기 탐지 기능 개발

3.5 Spring Security, JWT

3.6 배포 환경 구성 및 연결

4. 서비스 안내

5. 결론

5.1 결론 및 기대효과

5.2 향후 계획

6. 별첨

6.1 소개 자료

6.2 프로젝트 소개

6.3 팀 소개

6.4 발표 자료

1. 서론

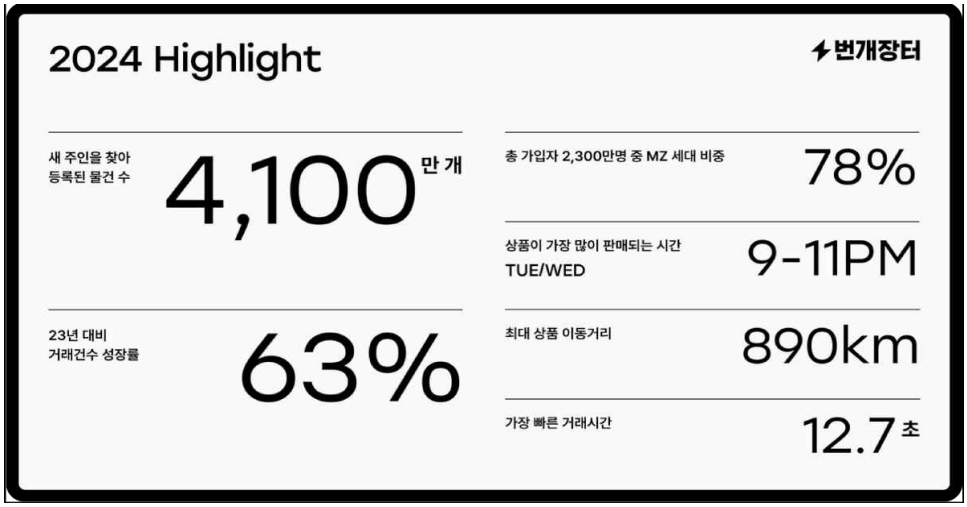
1.1 연구 배경 및 필요성

- 최근 국내 중고거래 시장은 급속한 성장세를 보이며 새로운 소비문화로 자리 잡고 있다. 2008년 약 4조 원 규모에 불과했던 시장은 2020년 약 20조 원 규모로 확대되었으며, 이는 단순한 개인 간 거래를 넘어 사회 전반의 경제 구조와 소비 패턴에 영향을 미치는 수준으로 발전했음을 보여준다. 대표적인 중고거래 플랫폼으로는 중고나라, 당근마켓, 번개장터가 있으며, 이 중 당근마켓은 2023년 12월 기준 누적 가입자 수 3,600만 명, 월간 이용자 수(MAU) 1,900만 명에 달하며 중고거래 시장 내 가장 높은 성장세를 기록하고 있다.

  이러한 변화는 단순한 거래의 확대를 넘어 순환경제의 활성화, 친환경 소비문화 확산, 합리적 소비 가치의 강화와 같은 사회적 변화를 이끌고 있다. 더불어 대기업 유통업계 또한 이 시장의 성장 가능성에 주목하여 적극적인 참여를 보이고 있다. 현대백화점그룹, 신세계, 롯데그룹 등 주요 유통사는 각각 리셀 매장 개설 또는 중고거래 플랫폼에 대한 전략적 투자를 통해 새로운 유통 생태계로의 전환을 시도하고 있다. 아울러 중고자동차(엔카), 중고 도서(알라딘), 중고 명품(트렌비), 패션 전문 플랫폼(세컨웨어) 등 세분화된 중고거래 플랫폼의 등장으로 시장은 다각화되고 있다.

  그러나 시장의 양적 성장이 지속되는 한편, 거래의 신뢰성과 안전성 확보는 여전히 해결되지 않은 핵심 과제로 남아 있다. 2023년 세이프타임즈의 조사 결과에 따르면, 중고거래 이용자들이 가장 불편함을 느끼는 요인은 '거래 물품의 품질·상태 확인(47%)'과 '사기 거래에 대한 불안감(46%)'으로 나타났다. 특히 사기 거래는 플랫폼 전체의 신뢰도를 저하시킬 뿐만 아니라, 신규 이용자의 진입을 가로막고 장기적으로 시장의 성장 가능성을 제한하는 요인으로 작용하고 있다.

  따라서 중고거래 시장의 지속적 성장을 위해서는 이용자 신뢰를 확보할 수 있는 기술적·제도적 보완책이 필수적이다. 특히 인공지능(AI) 기술을 활용한 자동화된 사기 탐지, 이상 거래 패턴 분석, 판매자 신뢰도 검증 등의 기능은 기존 수동적 대응 체계의 한계를 극복하고 거래 안정성을 강화하는 유력한 대안으로 주목받고 있다. 이러한 배경에서 본 연구는 AI 통한 사기 거래 예방 및 신뢰 기반의 거래 환경 조성의 필요성을 중심으로 문제를 탐색하고, 그 구체적 적용 방안을 제시하고자 한다.



1.2 연구 목적 및 주제 선정

- 최근 중고거래 시장은 플랫폼 중심의 디지털 생태계로 빠르게 확장되고 있으며, 개인 간 거래(C2C)의 활성화로 다양한 상품이 손쉽게 유통되고 있다. 그러나 이러한 성장의 이면에는 사기 거래 및 허위 매물로 인한 신뢰 저하라는 구조적 문제가 여전히 잔존하고 있다. 사용자들은 거래 안전성에 대한 불안감으로 인해 직접 거래를 주저하거나, 불필요한 검증 절차에 과도한 시간을 소모하는 등 이용 경험의 질적 저하를 겪고 있다. 이는 곧 플랫폼 이탈률 증가와 시장 위축으로 이어질 가능성을 내포한다.

  이러한 상황에서 거래 신뢰성을 체계적으로 확보하는 기술적·서비스적 접근은 중고거래 시장의 지속 가능성 및 경쟁력 확보를 위한 핵심 요건이라 할 수 있다. 기존 플랫폼들이 주로 신고 접수나 사후 보상과 같은 사기 발생 이후의 대응 방식에 머무른 반면, 본 연구는 사기 거래를 사전에 예측하고 차단하는 선제적(preventive) 대응 체계 구축에 초점을 둔다.

  본 연구의 주요 목적은 인공지능(AI) 기반 사기 거래 감지 및 이미지 검증 시스템을 설계하고, 이를 통해 중고거래 환경의 신뢰성과 안정성을 향상시키는 것이다. 구체적으로는 다음과 같은 세 가지 방향에서 연구를 진행한다.

1. 거래 데이터 분석 및 사용자 행동 패턴 감지: 채팅 데이터, 거래 빈도, 결제 방식 등의 정형·비정형 데이터를 분석하여 사기 가능성을 실시간으로 탐지한다.

2. 이미지 검증(Image Verification) 및 콘텐츠 신뢰성 확보: 판매자가 업로드한 상품 이미지를 기존 인터넷 이미지 데이터와 비교하여 도용, 중복, AI 합성 여부 등을 탐지함으로써 허위 매물을 선제적으로 필터링한다.

3. 플랫폼 내 신뢰 메커니즘 설계: AI 기반 판단 결과를 이용자에게 명확하고 투명하게 제시함으로써, 플랫폼 전반의 거래 안정성·공정성·브랜드 신뢰도를 강화한다.

   특히, 본 연구는 기존의 CLIP 기반 이미지 분석 모델 대비 표현 학습 능력이 고도화된 DINOv2 모델을 적용하여, 사기성 이미지의 미세한 시각적 유사성까지 정밀하게 검출할 수 있는 기술적 개선을 실현하였다. 이를 통해 사기 거래 감소뿐만 아니라, 선량한 판매자 보호 및 거래 효율성 향상이라는 부가적 효과도 기대할 수 있다.

   나아가, 본 연구는 기술적 측면에 국한되지 않고, AI 신뢰도에 대한 사용자 인식, 경고 알림의 수용성, 거래 경험 만족도 등 사용자 행태적 요인도 함께 분석함으로써, 기술과 사회적 신뢰 메커니즘이 상호 보완적으로 작동하는 지능형 중고거래 안전관리 모델(Intelligent Trust Management Model)을 제시하고자 한다. 이를 통해 중고거래 시장의 건전한 성장과 더불어, AI 기반 플랫폼 안전 기술의 새로운 표준화 방향을 제안하는 데 연구의 의의가 있다.

2. 관련연구

2.1 언어, API

2.1.1 fast API

- FastAPI는 ASGI 기반 파이썬 웹 프레임워크로, 타입 힌트와 Pydantic을 활용해 입력 검증과 직렬화를 자동화하고 Starlette 위에서 비동기 처리 성능을 제공합니다. OpenAPI 문서를 자동 생성하며 Swagger/ReDoc UI를 즉시 제공해 클라이언트와의 계약을 명확히 유지합니다.

  핵심 특징

- 타입 기반 요청·응답 검증, 의존성 주입, CORS·OAuth2·JWT·WebSocket 연동을 간결하게 지원합니다.
- async/await, Uvicorn으로 저지연·고동시성을 확보합니다.
- 미들웨어·백그라운드 태스크·스트리밍 응답 등 운영 필수 기능을 내장합니다.

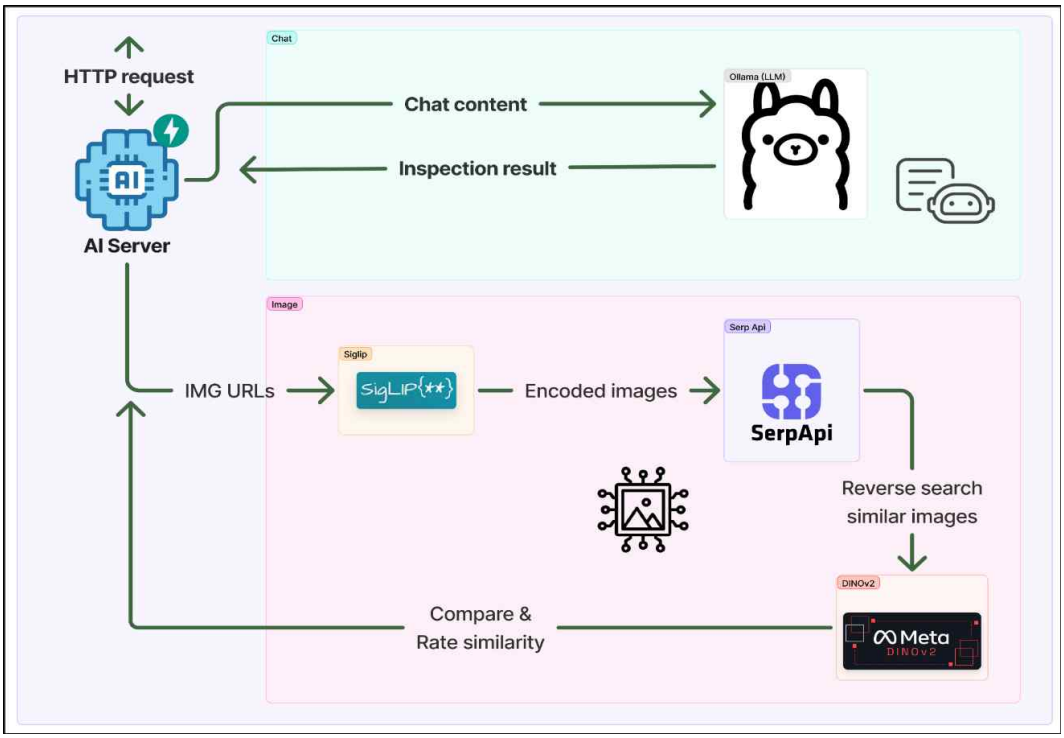
프로젝트 적용 맥락

- 이미지 유사도·대화 사기 탐지 엔드포인트를 라우터로 분리하고 Pydantic 모델로 스키마를 강제합니다.
- 로컬부터 ARM 컨테이너까지 동일한 실행 모델을 유지해 빠른 배포와 안정적 운영을 지원합니다.

2.1.2 AI서버 API구성

- 본 프로젝트는 두 가지 AI 기능을 웹 API 형태로 제공합니다. 서버는 FastAPI를 기반으로 하며, 사기 시도 탐지는 로컬 LLM 서빙 엔진인 Ollama를 통해 추론하고, 이미지 유사도 검사는 SigLIP과 DINOv2 임베딩을 결합해 수행한다. 모든 요청·응답은 JSON 형식을 따르며 Pydantic으로 스키마를 검증한다. 로컬 추론 구조를 채택해 개인정보의 외부 전송을 최소화하고 외부 API 의존으로 인한 지연과 장애 리스크를 줄였다.

2.1.3 시스템 아키텍처



- 클라이언트(앱/웹) → FastAPI 서버 → 사기 탐지 모듈(Ollama) 및 이미지 모듈(SigLIP+DINOv2) →역이미지검색(SerpAPI) →결과 저장·조회(DB)의 흐름으로 구성한다. 서버는 /check-fraud와 /check-fraud-images 엔드포인트를 제공하며, 요청 스키마 검증과 예외 처리를 통합했습니다. 헬스 체크와 문서화(/, /docs)를 통해 운영 모니터링을 지원한다.

2.2 사기 탐지 및 이미지 비교

- 이 절에서는 채팅 사기 탐지와 이미지 유사도 검사의 구현과 검증을 다룬다. 채팅 사기 탐지는 FastAPI와 Ollama 기반 LLM으로 대화 내 위험 징후를 분류하며, 이미지 유사도 검사는 SigLIP으로 카테고리를 추정한 후 DINOv2 임베딩을 사용해 이미지 간 유사도를 정밀 산출한다. 역이미지검색은 SerpAPI를 활용하여 유사 후보를 수집하고, 최종 결과는 SAFE, WARNING, DANGER 세 등급으로 제공한다.

2.2.1 채팅 사기 탐지 모듈

- 사기 탐지 모듈은 최근 대화 N건을 수집해 프롬프트로 구성하고, Ollama의 LLaMA 계열 모델로부터 JSON만 출력하도록 지시합니다. 모델 응답은 우선 json.loads로 파싱하며, 실패 시 첫 번째 중괄호 블록 추출, 그래도 실패 시 키워드 폴백(SAFE/WARNING/DANGER)으로 일관된 결과를 보장합니다. 분류 기준은 선결제 요구, 외부 채널 유도, 송금 링크 등 강한 위험 신호는 DANGER, 모호한 설명·비정상 가격 등은 WARNING, 구체적 거래 정보와 안전결제 선호 표현은 SAFE로 설정합니다. 운영 로그는 민감정보를 마스킹한 뒤 보관하며, 정밀도·재현율·평균 지연 등 품질 지표를 주기적으로 점검한다.

구분	테스트 시나리오	기대 결과
정상 거래	일반적인 가격 제시 및 거래 의사 표현	"정상"으로 분류, 경고 없음
선결제 유도	"입금 먼저 해주시면 바로 보내드릴게요" 등	높은 사기 점수 부여, 경고 문구 출력
안전결제 사칭	"○○안전결제 링크로 결제하세요" 등	사칭 문구 인식, 정책 위반 판정
외부 사이트 유도	"이 링크로 들어가서 결제해주세요" 등	외부 URL 탐지, 경고 처리
고가 거래 급매	"시세보다 저렴하게 급처합니다" 등	위험 패턴 인식, 주의 알림 표시

2.2.1.1 Ollama 선택 사유

- 개인정보 보호, 비용 예측 가능성, 낮은 지연, HTTP 연동의 단순성, 모델 교체와 버전 고정의 용이성, 양자화 모델로 인한 저사양 호환성, 배포의 단순성, 로깅·튜닝의 통제 용이성 등의 이유로 로컬 서빙 엔진인 Ollama를 채택했다. 외부 API 장애나 레이트리밋으로부터 독립적이며, 사내망·오프라인 환경에서도 동일한 품질을 유지한다.

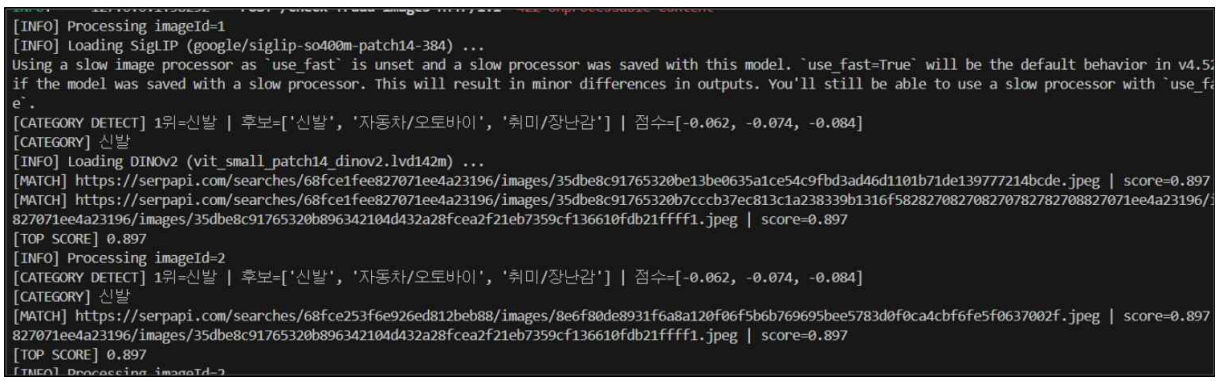
2.2.1.2 채팅 사기 탐지 과정

- 채팅 사기 탐지는 프롬프트 설계와 파서 안정화가 핵심이다. 모델에는 역할·출력 형식·판정 기준을 명확히 지시하고, 출력은 반드시 JSON만 생성하도록 제약합니다. 응답 파서는 JSON 파싱 실패 시에도 결과를 일관되게 반환하도록 폴백 계층을 둔다. 위험 시나리오(선결제, 안전결제 사칭, 외부 링크, 급매 등)에 대한 케이스 기반 테스트를 수행해 분류 기준을 점검하고, 임계치·프롬프트 파라미터(temperature, top_p)를 주기적으로 튜닝한다. 모든 대화 데이터는 익명화하여 저장하고, 민감정보는 마스킹 후 보관한다.

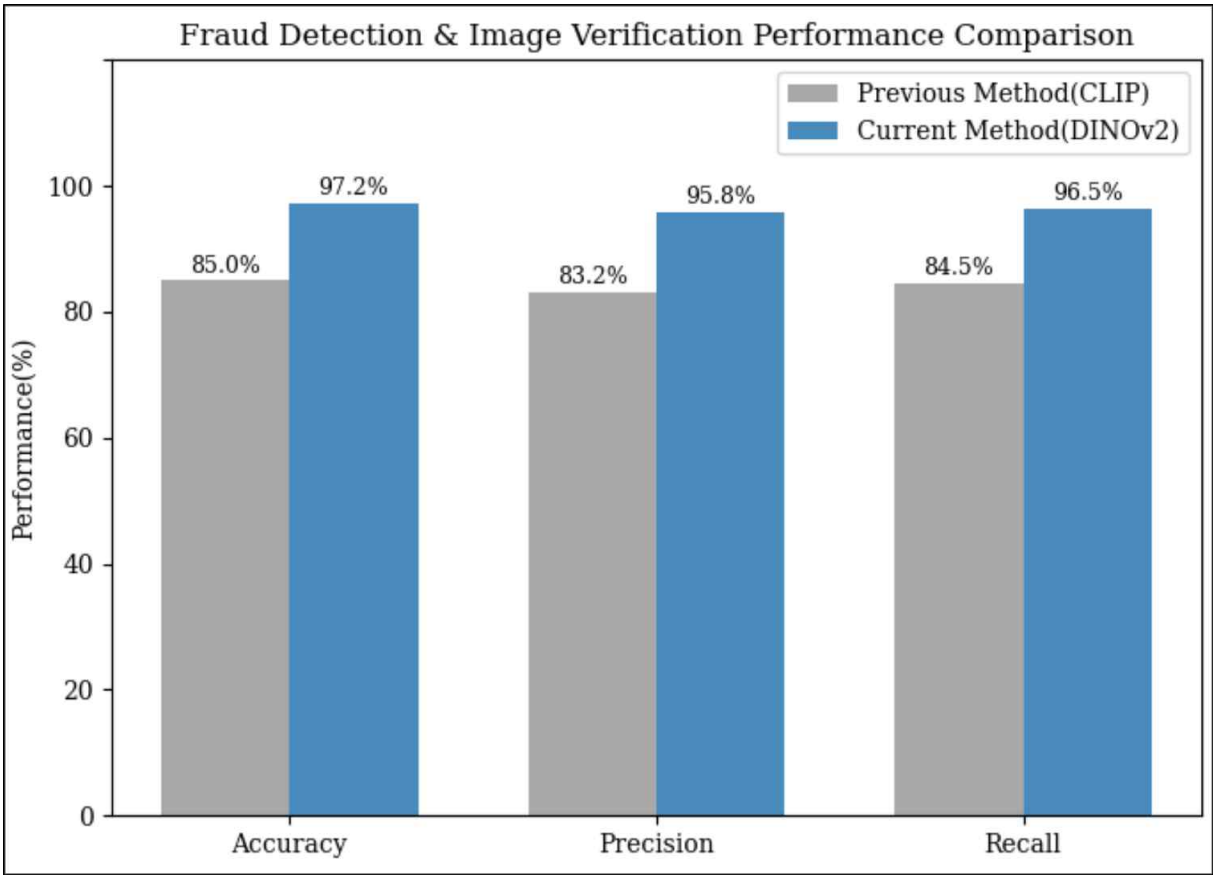
2.2.2 이미지 유사도 검사 모듈

- 이미지 검사는 5단계 파이프라인으로 동작한다. ① 입력 이미지 수신 ② SigLIP 텍스트–이미지 임베딩으로 카테고리 추정 ③ SerpAPI로 역이미지검색 후보 수집 ④ SigLIP 기반 카테고리 일치 여부 1차 필터 ⑤ DINOv2 이미지–이미지 임베딩 코사인 유사도로 최종 점수 산출입니다. 최고 유사도 점수에 따라 결과를 등급화하며, 초기 임계치는 DANGER(>0.75), WARNING(0.45~0.75), SAFE(≤0.45)로 두고 운영 데이터에 맞춰 주기적으로 재튜닝한다. 응답에는 상위 3개 유사 이미지의 URL과 점수를 함께 제공한다.

구분	테스트 시나리오	기대 결과
동일 이미지	완전히 동일한 이미지 업로드	1.0에 가까운 유사도, 중복 판정
리사이즈/크롭	크기 조정 또는 일부 영역 자르기	높은 유사도 유지 (0.85 이상)
밝기·색온도 변화	필터 적용, 색감 조절	일정 유사도 유지 (0.7 이상)
다른 각도	동일 상품의 다른 각도 촬영	중간 유사도, 참고 경고 표시
다른 상품	완전히 다른 물품 이미지	낮은 유사도 (<0.3), 정상 등록



2.2.3 성능 검증



- 본 연구에서는 제안된 AI 기반 사기 거래 감지 및 이미지 검증 시스템의 성능을 검증하기 위해 약 5,000개의 거래 데이터(채팅 로그 및 이미지 샘플)를 대상으로 실험을 수행하였다. 평가에는 학습 데이터와는 독립적인 검증 전용 데이터셋을 사용하였으며, 각 테스트 케이스에 대해 시스템이 사기 여부를 정확히 분류하는지를 측정하였다.

- 실험 결과, DINOv2 기반 이미지 검증 모델을 적용한 본 시스템은 정확도(Accuracy) 97.2%, 정밀도(Precision) 95.8%, 재현율(Recall) 96.5%를 달성하였다. 이는 기존의 CLIP(Contrastive Language–Image Pretraining) 모델을 활용한 규칙 기반 필터링 방식 대비 평균 약 12%p 성능이 향상된 수치로, DINOv2의 고차원 표현 학습 능력과 시각적 특징 인식 정밀도의 개선이 성능 향상에 주된 요인으로 분석된다.

- 특히, DINOv2 모델은 사기성 이미지의 미세한 시각적 유사성(예: 배경 패턴, 조명 조건, 부분적 크롭 등)을 더 효과적으로 구별하여, 기존 CLIP 모델이 간과하던 도용 이미지 검출 민감도(sensitivity)를 크게 개선하였다. 이러한 결과를 통해, DINOv2 기반의 이미지 검증 모듈이 실제 중고 거래 플랫폼에서도 높은 일반화 성능과 실용적 신뢰도를 확보할 수 있음을 확인하였다.

2.2.4 이미지 검증 과정

- 본 연구에서 제안하는 이미지 검증 모듈은 SigLIP(Sigmoid-enhanced Language–Image Pretraining)과 DINOv2(Self-Distillation with No Labels v2)의 상호 보완적 특성을 결합하여 구현하였다. 두 모델의 결합을 통해 대규모 웹 이미지로 학습된 범주 인식 능력(semantic alignment)과 고해상도 시각 특징 표현(visual representation)을 동시에 활용함으로써, 기존 단일 모델 기반 접근 대비 높은 신뢰도와 세밀한 검출 성능을 확보하였다.

  먼저, 1단계 전처리(카테고리 기반 필터링)에서는 SigLIP을 이용하여 거래 상품의 텍스트 기반 카테고리 임베딩과 이미지 임베딩을 정렬(alignment)시켜, 해당 이미지의 상위 의미적 범주를 결정한다. 이를 통해 입력 이미지와 카테고리 간 의미적 일치를 판단하고, 비정상적·불일치한 범주의 이미지를 1차적으로 제거하여 후속 검증 단계에서의 잡음(noise)을 최소화한다.

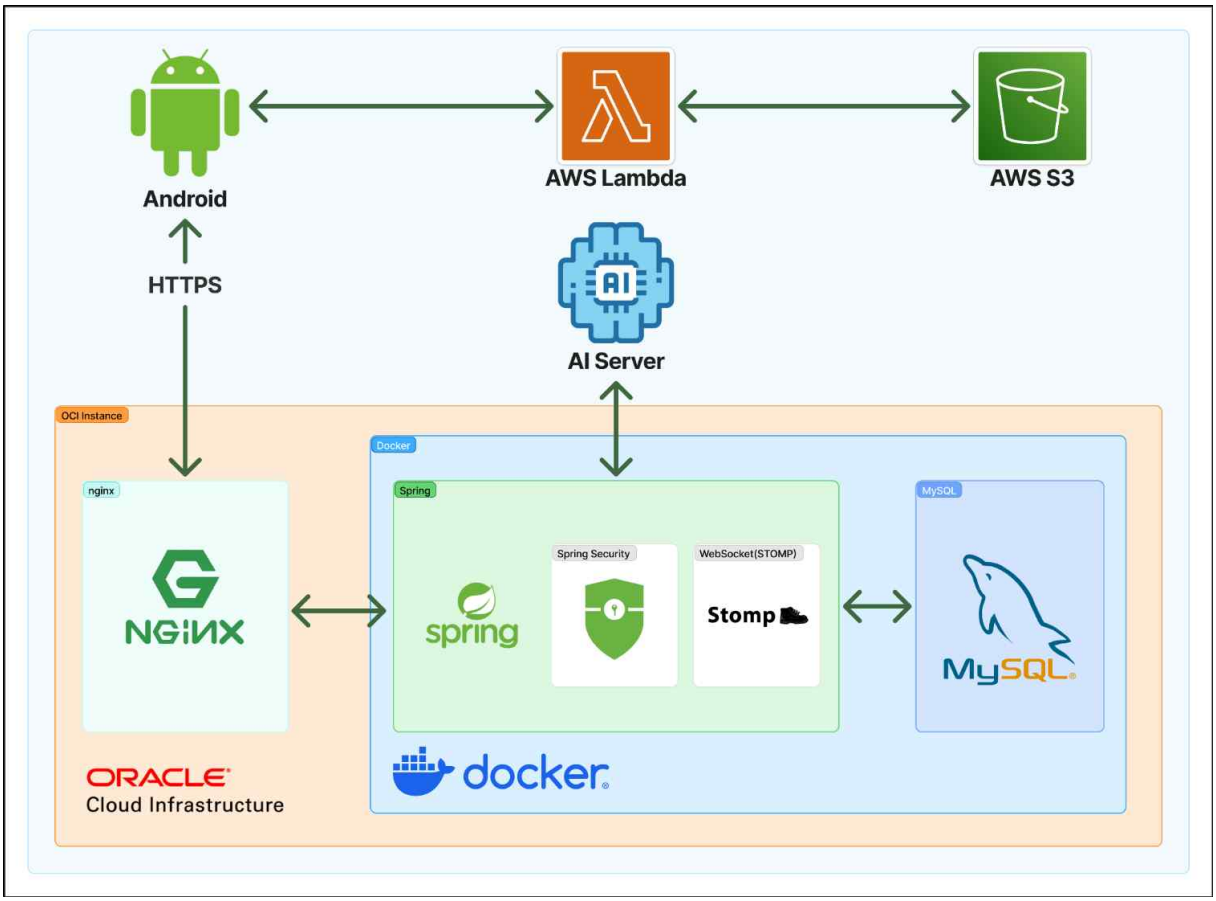
  다음으로, 2단계 정밀 검증(시각 유사도 비교)에서는 DINOv2를 활용하여 잔존한 후보 이미지와 원본 간의 임베딩 벡터 간 코사인 유사도(Cosine Similarity)를 계산한다. DINOv2는 자가 지도(self-supervised) 학습을 통해 이미지의 구조적·국소적 특징을 보존하므로, 크롭(crop), 밝기 및 색온도 변화, 회전(angle shift) 등 시각적 변형(visual augmentation)이 가해진 경우에도 안정적으로 동일 개체로 판별할 수 있다.

  최종적으로 유사도 점수가 사전에 정의된 임계값(threshold)을 초과할 경우, 해당 이미지를 도용·중복 가능성이 있는 항목으로 분류하며, 시스템은 자동으로 검수(review) 또는 경고(alert) 프로세스를 트리거한다.
  동시에, 사용자 및 운영자에게는 상위 매칭 이미지의 URL과 유사도 점수(score)를 근거로 함께 제공하여 판별의 투명성과 신뢰성을 확보하였다.

단계	적용 모델	핵심 기능	주요 효과
1단계: 카테고리 정렬	SigLIP	텍스트–이미지 임베딩 정렬을 통한 의미 기반 필터링	불일치 범주 제거 및 잡음 감소
2단계: 시각 유사도 분석	DINOv2	임베딩 간 코사인 유사도 계산, 변형 이미지 안정성 확보	정밀 유사도 판별 및 도용 탐지
결과 제공	SigLIP+DINOv2	상위 매칭 URL 및 유사도 점수 제공	검증 근거 확보 및 신뢰도 강화

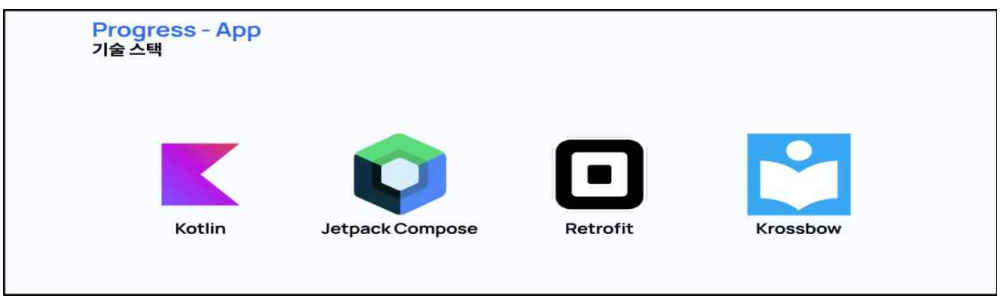
3. 프로젝트 내용

3.1 시스템 구성도



3.2 앱 제작

- 본 연구에서 개발된 애플리케이션은 안드로이드(Android) 운영체제를 기반으로 제작되었다. 안드로이드 플랫폼은 오픈소스 구조와 폭넓은 사용자층을 갖추고 있어, 다양한 디바이스 환경에서도 유연하게 동작할 수 있다는 장점을 지닌다. 이를 통해 사용자는 모바일 환경에서 손쉽게 사기 탐지 및 이미지 비교 기능을 이용할 수 있도록 설계되었다.



개발 언어로는 Kotlin을 사용하였다. Kotlin은 JetBrains에서 개발한 현대적인 프로그래밍 언어로, 자바(Java)와의 높은 호환성을 유지하면서도 간결한 문법과 안정성을 제공한다. 특히 널 안정성(null safety) 기능을 통해 런타임 오류 발생 가능성을 최소화하고, 비동기 처리를 위한 코루틴(Coroutines)을 활용하여 네트워크 요청 시의 효율적인 자원 관리가 가능하도록 구현하였다.

네트워크 통신을 위해 Retrofit 라이브러리를 사용하였다. Retrofit은 RESTful API 통신을 간소화하고, JSON 형태의 데이터를 자동으로 직렬화·역직렬화하는 기능을 제공한다. 이를 통해 Oracle Cloud Database와의 연동 과정에서 데이터 요청 및 응답 처리가 안정적이고 효율적으로 이루어지도록 하였다.

이미지 로딩과 표시를 위해서는 Glide 라이브러리를 채택하였다. Glide는 안드로이드 환경에서 이미지 캐싱 및 리사이징을 효율적으로 수행할 수 있는 라이브러리로, 사기 탐지 과정 중 사용자로부터 업로드되는 이미지의 미리보기 및 비교 시 높은 성능을 보장한다. 또한 Glide는 메모리 관리 측면에서도 우수하여, 이미지 비교 테스트에서 발생할 수 있는 메모리 누수 문제를 방지하는 데 기여하였다.

3.3 이미지 검증 과정 개발

- 본 시스템의 이미지 검증(Image Verification) 모듈은 판매자가 업로드한 상품 이미지를 대상으로, 인터넷 상의 기존 이미지 데이터와의 시각적 유사도를 정량적으로 산출하여 중복 등록 및 이미지 도용 가능성을 자동 탐지하도록 설계되었다. 구매자는 이를 선택적으로 확인함으로써, 게시 이미지의 진위성과 신뢰성을 직접 검증할 수 있다.

  이미지 분석 모델은 ResNet-50(Residual Network-50) 구조를 기반으로 구축되었다. ResNet-50은 합성곱 신경망(CNN) 계열의 모델로, 잔차 학습(residual learning) 구조를 통해 깊은 네트워크에서도 학습 효율과 정확도를 유지할 수 있는 것이 특징이다. 이를 통해 입력된 이미지로부터 고차원 특징 벡터(feature vector)를 추출하였으며, 이 벡터는 이후 유사도 계산 단계에서 핵심적으로 활용되었다.

유사도 계산에는 Cosine Similarity 알고리즘이 적용되었다. Cosine Similarity는 두 벡터 간의 방향적 유사성을 측정하는 방법으로, 결과값이 1에 가까울수록 두 이미지가 높은 유사도를 갖는 것으로 해석된다. 본 방식은 이미지의 크기, 해상도, 명암비 등 외형적 요소에 영향을 덜 받으므로, 다양한 환경에서 안정적인 유사도 판별이 가능하였다. 또한 Euclidean Distance 기반의 비교 알고리즘도 함께 적용하여, 이미지 간 거리 계산에 따른 정량적 차이를 분석하고 Cosine Similarity 결과와 교차 검증하였다. 이를 통해 모델의 일관성과 신뢰도를 높였다.

데이터 관리와 연산 환경은 Oracle Cloud Database를 기반으로 구성되었다. Oracle Cloud의 고성능 연산 자원을 활용하여 이미지 메타데이터 및 특징 벡터를 효율적으로 저장·조회할 수 있었으며, 대규모 이미지 처리 과정에서도 데이터 입출력 병목 현상을 최소화하였다.

모델의 검증은 다음과 같은 시나리오를 중심으로 수행되었다.

동일 이미지 비교: 완전히 동일한 이미지 간 유사도 계산의 정확성 검증

리사이즈 및 크롭 변형: 해상도 및 비율이 변경된 이미지의 인식 일관성 확인

밝기·색온도 변화: 조명 조건 변화에 따른 유사도 안정성 평가

다른 각도 이미지: 동일 피사체의 촬영 각도 차이에 따른 유사도 편차 측정

다른 상품 이미지: 상이한 피사체 간 명확한 구분 여부 확인

- 검증 결과, 시스템은 다양한 변형 조건에서도 일관된 유사도 판별 성능을 보였으며, 동일 이미지 및 경미한 변형 이미지에 대해서는 높은 정확도의 유사도 산출이 가능하였다. 이를 통해 본 이미지 검증 모듈은 사기 거래 탐지 시스템의 신뢰성을 강화하는 핵심 구성 요소로서 기능하였다.

3.4 채팅 내 사기 탐지 기능 개발

- 본 시스템의 채팅 내 사기 탐지(Fraud Detection in Chat) 기능은 판매자와 구매자 간 거래 과정에서 발생할 수 있는 사기 시도를 실시간으로 감지하도록 설계되었다. 이 기능은 사용자의 채팅 메시지를 분석하여, 거래 과정에서 과도한 개인정보 요구, 선결제 유도, 외부 사이트 링크 안내, 고가 거래 급매 등 사기 거래 의심 패턴을 식별하고, 해당 사용자에게 즉시 경고 메시지를 제공한다.

  사기 탐지 모델은 공공 데이터 포털 및 다양한 거래 사례에서 수집한 데이터와, 블랙리스트 기반 특정 단어 리스트를 활용하여 학습되었다. 이를 통해 정상적인 거래 문구와 사기 의심 문구를 구분할 수 있으며, 채팅 내 메시지의 의미적 패턴 및 문맥을 분석하여 거래 위험 점수(Fraud Risk Score)를 산출한다.

  모델 구현에는 Ollama 기반 LLM(Local Large Language Model)이 사용되었으며, FastAPI 서버를 통해 채팅 데이터와 실시간으로 연동된다. 사용자가 메시지를 전송하면, 서버는 해당 메시지를 모델에 전달하여 분석하고, 위험 점수에 따라 경고 문구를 생성하여 양측 사용자에게 반환한다. 이러한 구조는 사기 거래 발생 가능성을 사전 예방적으로 차단할 수 있도록 설계되었다.

  사기 탐지 기능의 검증은 다음과 같은 시나리오를 중심으로 수행되었다.

  정상 문구: 일반적인 가격 제시 및 거래 의사 표현

  선결제 유도: "입금 먼저 해주시면 바로 보내드립니다" 등

  안전결제 사칭: "○○안전결제 링크를 통해 결제하세요"

  외부 사이트 유도: "해당 링크에서 결제해주세요"

  고가 거래 급매: 시세보다 저렴한 급매 안내

- 검증 결과, 모델은 다양한 사기 거래 패턴을 안정적으로 감지하였으며, 위험 점수와 경고 메시지를 정확하게 산출하였다. 또한, 모든 채팅 내역과 탐지 결과는 Oracle Cloud Database에 저장되어, 관리자 모니터링과 모델 성능 분석에 활용되었다.

3.5 Spring Security

- 본 시스템에서는 Spring Security를 활용하여 사용자 인증(Authentication)과 권한 관리(Authorization)를 구현하였다. Spring Security는 스프링 프레임워크 환경에서 강력한 보안 기능을 제공하는 표준 라이브러리로, 사용자 인증과 접근 제어를 체계적으로 관리할 수 있다.

  인증 방식으로는 JWT(Json Web Token)를 사용하였다. JWT는 JSON 형식의 컴팩트하고 독립적인 데이터 구조를 통해, 사용자 정보를 안전하게 전송할 수 있는 개방형 표준(RFC 7519)이다. 토큰은 서버에서 발급 시 HMAC 알고리즘을 활용한 시크릿 키 또는 RSA/ECDSA 공개/비공개 키를 통해 서명되어 변조를 방지한다.

  시스템에서는 사용자가 로그인하면 서버가 JWT를 발급하고, 클라이언트는 이후 모든 API 요청에 해당 토큰을 포함시킨다. 서버는 전달받은 JWT를 검증하여 요청자의 신원을 확인하고, 권한에 따라 접근을 허용 또는 제한한다. 이를 통해 비인가 사용자의 데이터 접근을 방지하고, 사기 거래 감지 및 이미지 검증 기능과 같은 핵심 서비스의 보안을 확보하였다.

- Spring Security와 JWT 기반 인증 체계는 확장성과 보안성을 동시에 확보하며, 다양한 사용자 환경에서도 안전한 서비스 운영을 지원한다.

3.6 배포 환경 구성 및 연결

- 본 시스템은 안정적 운영과 실시간 서비스 제공을 위해 클라우드 기반 인프라와 모바일/웹 환경을 연동하여 배포되었다. 전체 아키텍처는 다음과 같은 구성 요소로 이루어졌다.

- 백엔드 서버

  FastAPI 서버를 중심으로 Ollama 기반 사기 탐지 모델과 CLIP 기반 이미지 검증 모델을 호스팅하였다.

  서버는 RESTful API를 제공하며, Oracle Cloud Database와 연동되어 이미지 메타데이터, 채팅 내역, 위험 점수, 로그 데이터를 관리하였다.

  JWT 기반 인증을 적용하여 Spring Security와 연동된 사용자 인증 및 권한 관리를 수행함으로써 API 접근을 보호하였다.

- 모바일 클라이언트

  안드로이드 앱은 Kotlin으로 개발되었으며, Retrofit을 활용한 API 호출과 Glide를 활용한 이미지 처리 기능을 제공하였다. 사용자는 앱을 통해 상품 등록, 이미지 업로드, 채팅, 사기 거래 알림 등 시스템 기능을 실시간으로 이용할 수 있다.

- 데이터베이스 및 인증 연동

  Oracle Cloud Database는 모든 거래 및 이미지 데이터를 중앙에서 관리하며, FastAPI 서버와 연동되어 AI 모델과 앱 간 데이터 흐름을 지원하였다.

  JWT를 기반으로 사용자 인증 정보를 안전하게 전달하고 검증함으로써, 클라이언트와 서버 간 데이터 무결성과 보안을 확보하였다.

  연결 및 배포 구조

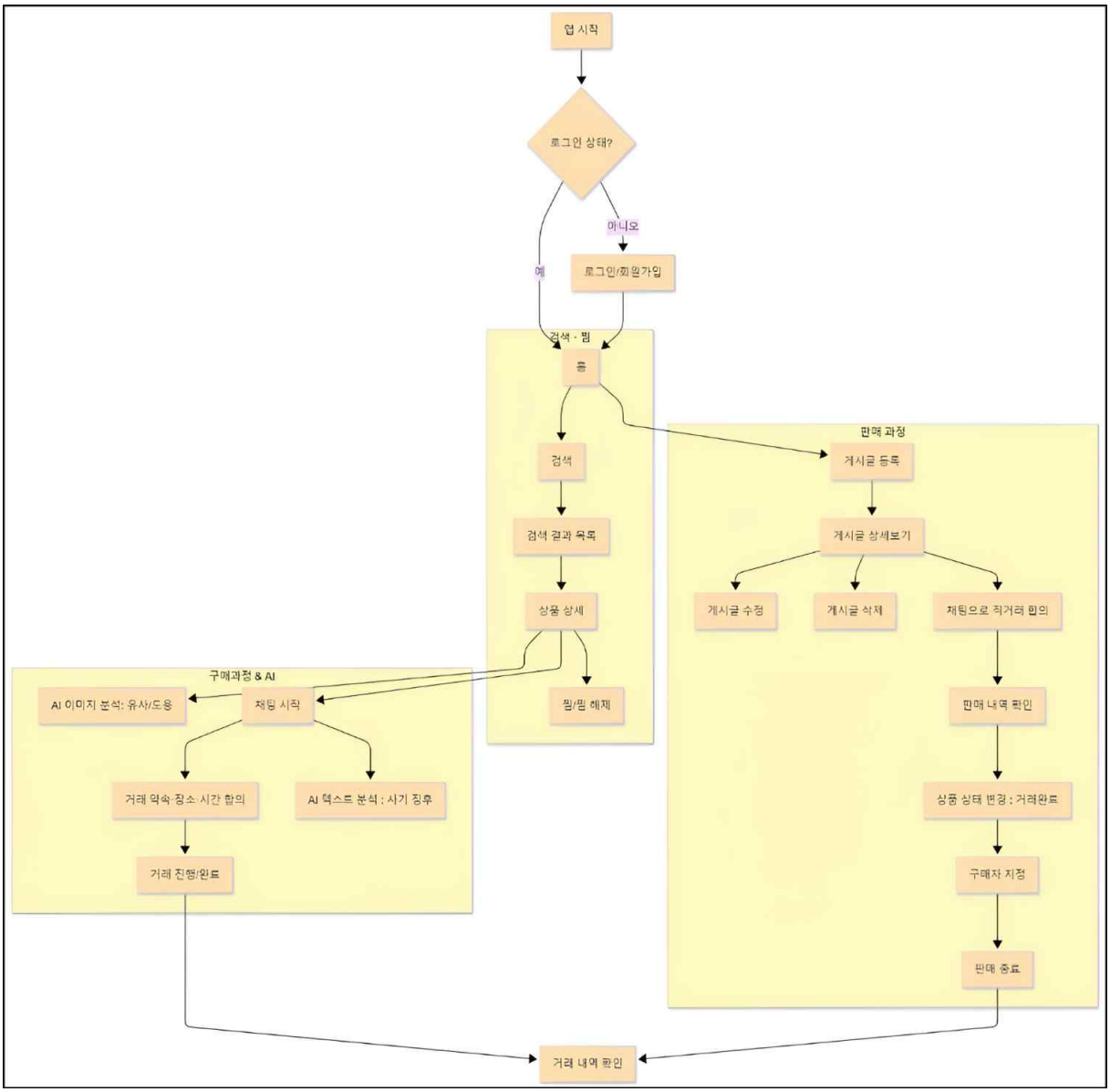
  클라이언트 ↔FastAPI 서버 ↔AI 모델(Ollama/CLIP) ↔Oracle Cloud DB

  이 구조를 통해 채팅 메시지와 이미지 데이터가 실시간으로 처리되며, 사기 탐지 및 이미지 검증 결과가 즉시 사용자에게 전달된다.

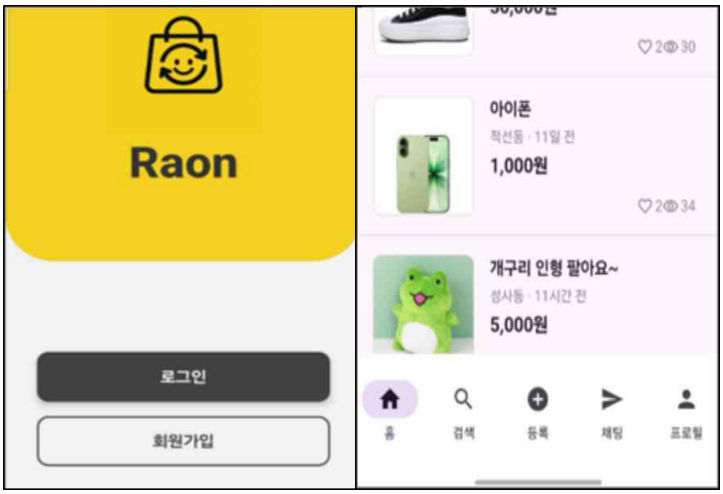
  모든 서버와 클라이언트 간 통신은 HTTPS를 통해 암호화되어 보안성을 강화하였다.

4. 서비스 안내

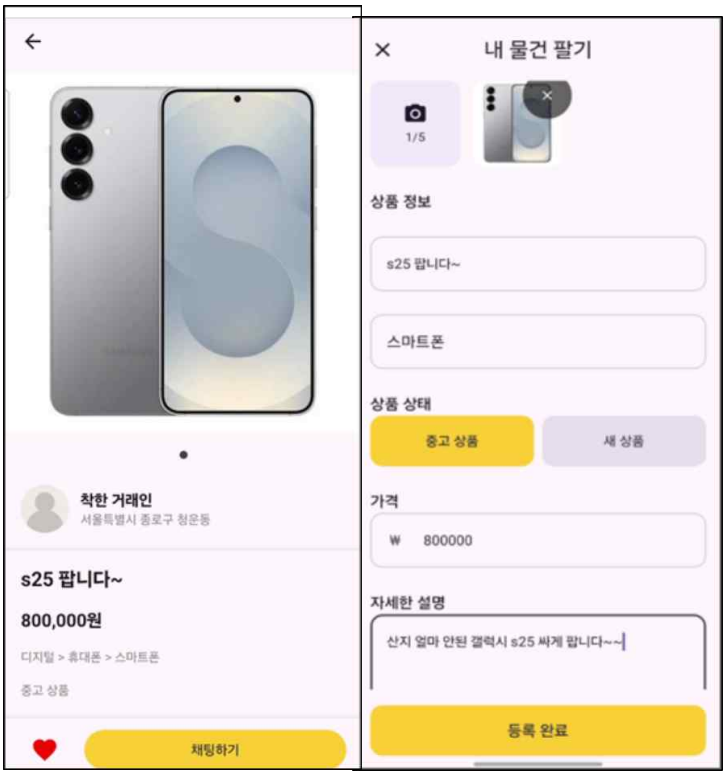
- UX flow



- 인 앱 스크린샷

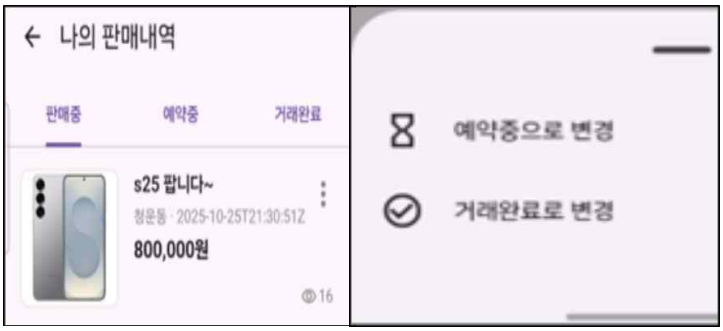


- 사용자는 "라온" 계정의 존재 유무에 따라 회원가입 또는 로그인 후, 홈 화면으로 이동해 네비게이션 바의 UI를 통해 자유롭게 탐색할 수 있다.



- 구매를 원하는 사용자는 자신이 원하는 상품의 판매 게시글을 올린 판매자에게 채팅을 보내 자신의 구매를 확정할 수 있다. 이 과정에서 AI 이미지 검증 및 텍스트 분석을 통해 사기 거래 방지가 가능하다.

- 반대로 판매를 원하는 경우에는 네비게이션 바의 "등록"을 통해 판매 게시글을 등록할 수 있고, 상세보기 기능을 통해 자신이 등록한 게시글을 자유롭게 수정/삭제 할 수 있다.



- 구매자와 판매자 간의 거래가 정상적으로 종료되었다고 판단될 경우, 네비게이션 바의 "프로필"에 존재하는 "나의 판매 내역"에서 상품의 상태를 "거래완료로 변경" 후 구매자를 선택해 확정지음으로써 마무리 할 수 있다.

5. 결론

5.1 결론 및 기대효과

- 본 시스템은 인공지능(AI)을 활용한 사기 거래 감지 및 이미지 검증 기술을 통합하여, 중고 거래 과정에서 발생하는 금전적·심리적 피해를 사전에 차단하고 거래 신뢰를 구조적으로 강화하는 것을 목표로 한다.

  텍스트 기반의 채팅 분석 모델은 거래 대화 중 사기성 패턴(예: 선입금 유도, 외부 결제 링크 유도, 개인정보 요구 등)을 실시간으로 탐지하여 사용자에게 경고하며, 이미지 검증 모델은 등록 상품의 이미지가 온라인상 기존 데이터와 중복되거나 허위 매물로 의심되는 경우 이를 자동 식별한다.

  이와 같은 다층적 AI 검증 절차를 통해 본 시스템은 기존 중고 거래 플랫폼에서 반복되어 온 사기 위험, 정보 비대칭, 신뢰 부족 문제를 근본적으로 개선하며, 사용자에게 안전하고 투명한 거래 경험을 제공한다.

- 구매자 측면: '거래 신뢰' 확보

  AI가 채팅 및 이미지 데이터를 실시간으로 분석하여, '선입금 유도', '타 사이트 결제 유도', '개인정보 요구' 등 사기 징후 패턴을 자동 감지하고 사용자에게 경고함으로써 금전적 피해를 사전에 차단한다. 또한 인터넷에서 도용된 이미지나 실물이 존재하지 않는 허위 매물을 선별하여 허위 게시물로 인한 피해와 시간 낭비를 최소화한다.

  이로 인해 구매자는 사기 위험을 신경 쓰지 않고 거래에만 집중할 수 있으며, 플랫폼에 대한 신뢰와 만족도가 높아져 거래 피로도가 감소한다.

- 판매자 측면: '공정하고 신뢰받는 판매 환경' 조성

  AI가 허위 매물과 비정상 판매자를 자동 필터링함으로써, 정상적으로 활동하는 판매자들이 더 많이 노출될 기회를 얻는다. 구매자들은 플랫폼 자체를 신뢰하기 때문에 판매자에 대한 불필요한 의심이 줄어들고, 결과적으로 판매 완료까지의 시간(Lead Time)이 단축된다.

  또한, 만일 AI 모델이 오탐을 일으키는 경우에도 간단한 검증 절차를 통해 빠르게 정상

등록이 가능하므로, 판매자의 신뢰도와 편의성이 함께 확보된다.

- 플랫폼(서비스) 측면: '경쟁력' 및 지속 가능 성장 강화

  본 시스템은 플랫폼에 '안전한 거래 환경'이라는 강력한 브랜드 이미지를 부여한다. "사기당하기 싫으면 라온을 사용해야 한다"는 인식은 차별화된 경쟁력을 창출한다. 안전한 경험에 만족한 사용자는 플랫폼에 충성도를 갖고 재방문하며, 이는 자연스러운 Lock-in 효과로 이어진다. 또한 사기 거래를 사전에 걸러냄으로써, 고객센터(CS)의 신고·분쟁·보상 처리 비용 등 운영 리소스를 절감할 수 있다.

  신뢰도가 높아진 플랫폼에서는 구매 결정 속도가 빨라지고, 결과적으로 거래 성사율 및 전체 거래액(GMV)이 증가하는 선순환을 기대할 수 있다.

- 종합 효과

  결과적으로 본 시스템은 모든 유형의 중고 거래 이용자에게 복잡한 추가 절차 없이 사기 위험을 효과적으로 감소시키는 환경을 제공한다. AI 기반의 자동 검증과 실시간 대응 기능을 통해 거래의 안전성·편의성·효율성을 동시에 향상시키며, 더 나아가 중고 거래 플랫폼 전반의 신뢰 확보와 거래 활성화에 크게 기여할 수 있다.

5.2 향후 계획

- 모델 고도화

  사기 탐지 및 이미지 비교 모델의 정확도를 지속적으로 향상시키기 위해, 다양한 거래 패턴과 이미지 변형 데이터를 추가 학습한다.

  Cosine Similarity, Euclidean Distance 외에도 다양한 유사도 알고리즘을 적용하여 최적의 결과를 도출하고, 오탐률을 최소화한다.

- 실시간 처리 성능 향상

  동시에 증가하는 사용자 요청을 효율적으로 처리하기 위해 서버 부하 분산 및 캐싱 전략을 강화한다.

  클라우드 기반 인프라를 확장하고, FastAPI 서버와 AI 모델 간 통신 최적화를 통해 응답 지연 시간을 감소시킨다.

- 서비스 확장 및 사용자 편의 개선

  모바일 앱과 웹 클라이언트에서 제공되는 UI/UX를 지속적으로 개선하여, 사용자가 보다 직관적으로 사기 위험 정보를 확인하고 대응할 수 있도록 한다.

  채팅 내 경고 기능과 이미지 검증 결과를 통합하여, 판매자와 구매자 모두에게 편리하고 신뢰성 높은 거래 환경을 제공한다.

- 보안 및 데이터 관리 강화

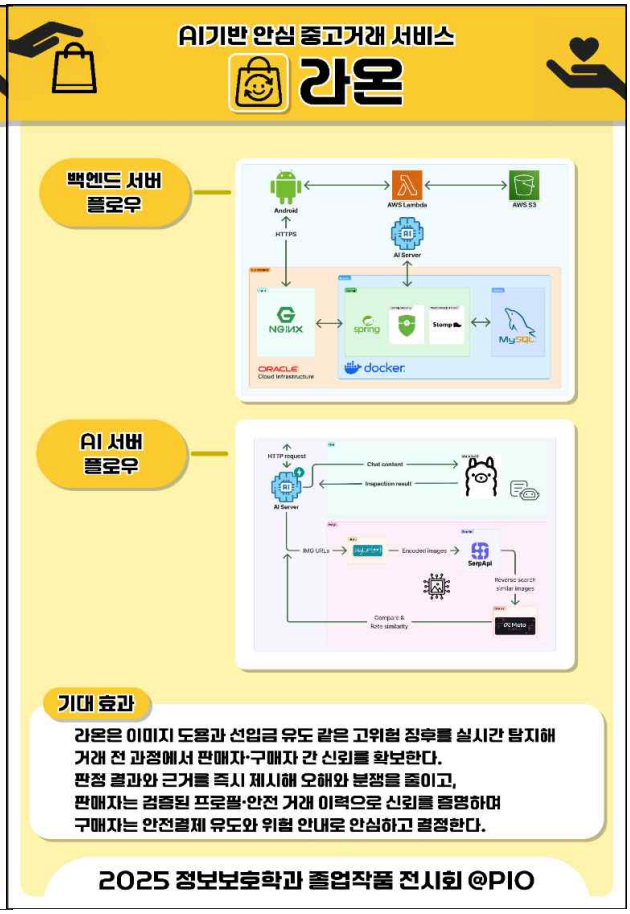
  JWT 기반 인증 및 Spring Security 체계를 지속적으로 검토하여, 사용자 데이터 보호와 API 접근 제어를 강화한다.

  거래 데이터와 이미지 정보를 안전하게 저장·관리하며, 데이터 분석과 모델 개선에 활용한다.

6. 별첨

6.1 소개자료


AI기반 안심 중고거래 서비스
라온
팀 구성
지도교수: 이병천 교수님
팀장: 김명안
팀원: 송명석, 신준혁, 최중건
개요
라온은 중고 거래에서 발생하는 사기 시도를 사전에 탐지하고
차단하는 AI 기반 보안 서비스이다.
이미지 도용·허위 매물과 선입금 유도 등 대화형 사기 패턴을
AI 기능을 통해 식별해 사용자에게 경고하고
이를 통해 사용자 간 거래의 신뢰도를 확보한다.
UX 플로우
Raon
Landing screen
Search product
Product detail
Chatting screen
AI Warning
2025 정보보호학과 졸업작품 전시회 @PIO

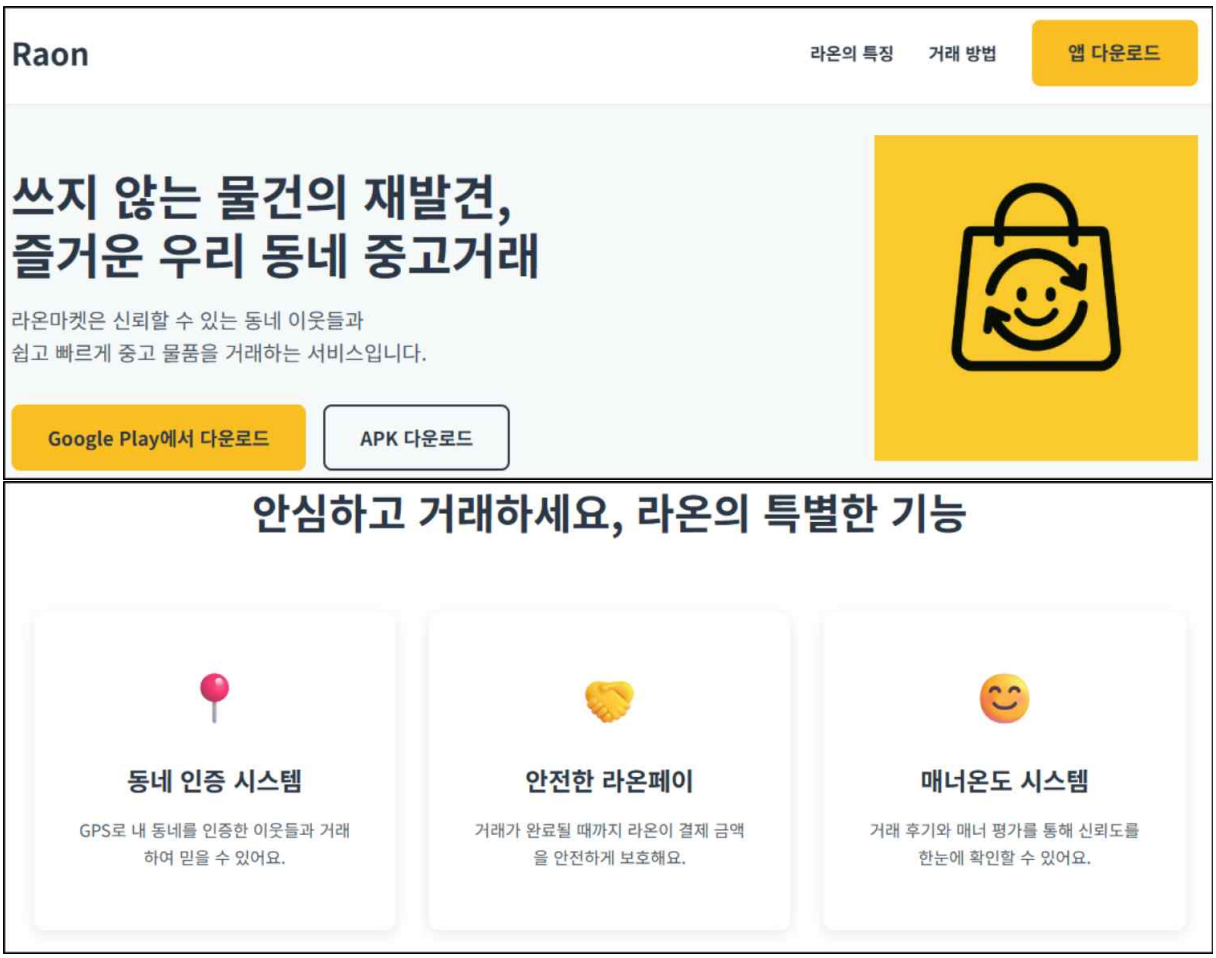

AI기반 안심 중고거래 서비스
라온
백엔드 서버
플로우
Android
HTTPS
AWS Lambda
AWS S3
AI Server
NGINX
spring
Stomp
MySQL
ORACLE
Cloud Infrastructure
docker
AI 서버
플로우
HTTP request
Chat content
Inspection result
AI Server
IMG URLs
Encoded Images
SerpApi
Reverse search similar images
Compare & Rate similarity
기대 효과
라온은 이미지 도용과 선입금 유도 같은 고위험 징후를 실시간 탐지해
거래 전 과정에서 판매자·구매자 간 신뢰를 확보한다.
판정 결과와 근거를 즉시 제시해 오해와 분쟁을 줄이고,
판매자는 검증된 프로필·안전 거래 이력으로 신뢰를 증명하며
구매자는 안전결제 유도와 위험 안내로 안심하고 결정한다.
2025 정보보호학과 졸업작품 전시회 @PIO

6.2 프로젝트 소개

- 깃허브 주소 : https://github.com/JB-PIO

- 시연 영상 : https://www.youtube.com/watch?v=266t0DIgQus
- 시연 영상(긴 영상) : https://youtu.be/Gg0A2mZady8

- 앱 배포 주소 : https://jb-pio.github.io/raon-homepage/



6.3 팀 소개

이름	GitHub 주소	수행 파트
김평안	https://github.com/crew852	PM, AI, 프론트엔드
송명석	https://github.com/MyungSeokSong	앱 개발
신준혁	https://github.com/JunHyeokShin	백엔드, 서버
최중건	https://github.com/Ariera226	AI, 프론트엔드

6.4 발표 자료

(별도 첨부)