

**SK네트웍스 Family AI과정 3기**

**프로젝트 기획서**



**□ 개요**

* 평가 산출물 : 프로젝트 기획서
* 제출 일자 : ~ 11/14
* 깃허브 경로 : https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN03-FINAL-2Team
* 작성 팀원 : 이준석, 박중헌

| **프로젝트 주제** | LLM을 이용한 맞춤형 문화 예술 컨텐츠 추천 플랫폼 개발 |
| --- | --- |
| **문제정의** | 현대 사회에서는 디지털 기술의 발전과 함께 문화 예술 콘텐츠의 접근성이 향상되었다. 하지만 인터넷과 모바일 기기를 통해 수많은 뮤지컬, 전시회, 공연 정보에 쉽게 접근할 수 있으나 그만큼 정보의 양이 방대해져서 개인이 자신의 취향에 맞는 콘텐츠를 선택하기 어려워고 검색형 추천시스템이 대부분이다.   * **선택 과부하:** 방대한 정보로 인해 사용자가 원하는 콘텐츠를 찾는 데 시간이 많이 소요됨. * **개인화 부족:** 현재 대부분의 플랫폼은 개인의 취향이나 이전 활동을 고려하지 않은 일반적인 추천을 제공 * **사용자 경험 저하:** 맞춤형 추천이 부족하여 사전 지식이 없으면 사용자가 만족할 만한 콘텐츠를 발견하기 어렵다.   따라서, 개인의 선호도와 취향을 고려하여 뮤지컬과 전시회를 추천해주는 지능형 대화형 추천 플랫폼의 필요성을 느낌. |
| **시장조사** | **뮤지컬 시장**   * 시장 성장: 국내 뮤지컬 시장은 2022년 티켓 판매액이 4,253억 원을 기록하며 역대 최고치를 달성했다. * 관람층 확대: 20대 및 청소년층까지 관람층이 확대되고 있다. * 디지털 리뷰의 영향력: 온라인 리뷰와 평점은 관객들의 선택에 큰 영향을 미치며, SNS를 통한 입소문 마케팅이 활발합니다.   **전시회 시장**   * 다양성 증가: 현대 미술, 사진전, 디지털 아트 등 다양한 형태의 전시회가 증가하고 있습니다. * 문화 소비 트렌드 변화: 경험 중심의 문화 소비가 증가하면서 전시회 방문이 여가 활동의 중요한 부분으로 자리매김하고 있습니다. * 이미지 기반 콘텐츠 소비: 인스타그램 등 이미지 중심의 SNS 플랫폼에서 전시회 관련 콘텐츠 소비가 활발합니다.   **경쟁 현황**   * 기존 플랫폼 한계:  1. 단순 정보 제공: 일정, 장소, 간단한 소개 등 기본 정보만 제공하는 경우가 많습니다. 2. 개인화 서비스 부족: 사용자 선호도를 반영한 맞춤형 추천 기능이 미흡합니다.  * 사례:  1. 인터파크티켓: 다양한 공연과 전시회의 예매를 지원하며, 사용자 리뷰와 평점을 제공하나 단순한 정보 제공에 그쳐 사용자 맞춤 컨텐츠 정보를 제공하는 것은 부족하고 관람이력이나 선호도 기반의 추천도 없음 2. 아트맵: 전시회 정보를 중심으로 한 플랫폼으로, 위치 및 관심사 기반의 전시회 검색과 정보 제공을 주요 기능으로 한다. 전시회 정보에 집중되어 있어, 뮤지컬이나 공연 등의 정보는 부족 3. 멜론티켓: 음악 관련 공연과 페스티벌 예매를 주로 다루며, 아티스트 정보와 공연 소식을 제공한다. 음악 공연에 집중되어 있어, 전시회나 연극 등의 정보는 제공하지 않는다. 4. 예스24: 도서, 공연, 영화 등 다양한 문화 콘텐츠의 예매와 구매를 지원하는 종합 플랫폼이다. 다양한 분야를 다루지만, 특정 분야에 대한 깊이 있는 추천보다는 일반적인 추천에 그친다.   이러한 기존 플랫폼들의 한계를 보완하기 위해, 사용자 개인의 선호도와 이전 활동을 기반으로 한 대화형 맞춤 추천 기능을 갖춘 새로운 문화 예술 콘텐츠 플랫폼의 개발이 필요로 한다. |
| **시스템 구성** |  |
| **모델링 방안** | * OCR   + OCR 모델 레퍼런스 조사   + OCR 모델 비교 및 선정   + OCR 모델 파인튜닝 * 추천시스템   + 추천시스템 모델 레퍼런스 조사   + 추천시스템 모델 비교 및 선정   + 추천시스템 모델 파인 튜닝 * RAG   + RAG Pipeline 레퍼런스 조사   + RAG Pipeline 설계 및 구축 * DeepFM   + 모델 파인튜닝 |
| **사용데이터** |  |
| **R&R** | **데이터 수집 레이어 구축:** 준석, 종명   * 웹 크롤링 * API 연동 * 데이터 업데이트 자동화   **ETL 파이프라인 구축:**   * 데이터 정제 스크립트 작성 * 데이터 변환 로직 구현: JSON, XML, CSV 등 다양한 형식의 데이터 처리. * 데이터 적재 및 관리: 데이터베이스 성능 최적화 및 백업 전략 수립.  1. **뮤지컬:** 허지원, 이주원, 박종명  * DeepFM 모델 사용 * 모델 훈련 및 튜닝 * 성능 평가 및 보고서 작성 * RAG 파이프라인 구축 * 벡터스토어 구축  1. **전시회:** 이준석, 중헌  * OCR 모델 파인튜닝 * RAG 파이프라인 구축 * 벡터스토어 구축   **웹 페이지 구현:** 이준석, 박중헌, 허지원, 이주원, 박종명   * UI/UX 디자인(Figma): 허지원 * 메인페이지 구현: 이주원, 박종명, 이준석 * 뮤지컬 페이지 구현: 이주원, 허지원 * 전시회 페이지 구현: 이준석, 박중헌   **데이터베이스 설계:** 박중헌   * 데이터베이스 관리: NoSQL 데이터베이스 설계 및 최적화. * 캐싱 전략: Redis 등을 활용한 데이터 캐싱으로 성능 향상.   **CI/CD 파이프라인 구축:** 이준석  **프로젝트 매니저:**   * 전체 일정 관리 및 조율: 이준석 * 요구사항 분석 및 문서화: 이준석 |