

SK네트웍스 Family AI과정 3기

데이터 전처리 학습된 인공지능 모델

□ 개요

- 산출물 단계 : 데이터 전처리
- 평가 산출물 : 학습된 인공지능 모델
- 제출 일자 : 2025.03.07
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN06-FINAL-2Team>
- 작성 팀원 : 이세화, 박미현, 안형진, 전하연, 정유진

데이터

1) 사용 데이터

- 웹툰 데이터
 - 플랫폼: 네이버, 카카오페이지
 - 데이터 양: 23,585개
- 웹소설 데이터
 - 플랫폼: 네이버시리즈, 카카오페이지
 - 데이터 양: 156,973개
- 참조 문서
 - 네이버 지식인에서 "웹툰 추천", "웹툰 ㄷㄷ" 관련 질문과 답변 데이터 활용

2) Vector Store 및 RAG 적용

- 벡터화 과정
 - "BAAI/bge-m3" 모델(HuggingFace)을 활용하여 데이터 벡터화
 - 웹툰 및 웹소설 데이터를 장르별로 벡터화하여 Vector Store 구성
- RAG 기반 추천 시스템
 - 검색 과정에서 Vector Store를 기반으로 관련 작품을 Retrieval
 - Retrieval된 데이터를 LLM(Large Language Model)과 결합하여 자연스러운 추천 문장 생성
 - 키워드 기반 검색과 벡터 검색을 함께 활용하여 정확도 향상

기술

1) 자체 score 시스템

- 조회수, 별점, 댓글 수, 첫 연재일, 최근 댓글 수 등을 조합하여 자체 점수(Score) 계산
- 조회수, 댓글 수는 log scaling 적용하여 급격한 값 차이를 완화
- 각 요소의 가중치를 장르별로 다르게 설정하여 점수 계산
- 최종적으로 MinMaxScaling을 적용하여 0~1 범위로 정규화
- 높은 점수를 가진 작품이 추천될 확률이 증가하도록 반영

2) RAG 기반 추천 모델 설계

(1) 페르소나 기반 LLM 말투 조정

각 모델에게 특정한 ****페르소나(Persona)****를 부여하여 답변의 말투를 다르게 생성하도록 설정함.

- 모델별 페르소나
 - **romance fantasy** - 로판의 남자주인공인 북부대공 캐릭터:
"영애의 취향을 고려하여 몇 가지 작품을 골라 보았습니다.
혹여 탐탁지 않으시다면, 직접 고르셔도 좋습니다."
 - **romance** - 집착광공 캐릭터: "ㅇㅇ씨는 참 재주가 많아요. 그중
제일 탁월한 건 사람 미치게 만드는 거?"
 - **historical** - 단호하고 절도있는 돌같은 캐릭터: "강호에서
살아남는 자, 결국 검을 진 자뿐이요. 당신을 위해 최고 인기
무협 웹툰 3개를 소개한다."
 - **Fantasy** - 우아하지만 말 많은 캐릭터: "오호, 사랑스러운
존재여. 마법을 배우고 싶다고요? 음... 하지만 마법이란
단순한 기술이 아니라, 조화와 흐름을 이해하는 예술이랍니다.
예를 들면, 달빛 아래에서 춤을 추는 물결을 보았나요? 바람에
흔들리는 꽃잎의 속삭임을 들었나요? 그 모든 것이 마법의
일부이지요. 후훗, 선택은 신중히 하세요. 마법은 달콤하지만,
때때로 치명적인 향기를 품고 있습니다."

(2) Agent 활용 및 Context 조정

- LLM을 직접 호출하는 것이 아니라, 에이전트가 **Query**를 해석하고
적절한 실행 방식을 결정하는 방식.
- **Multi-Agent** 구성
 - **의도분석 Agent**: 사용자의 쿼리가 무슨 의도인지 목록화 하여
선택하도록 수행.
 - **추천 Agent**: 웹툰/웹소설 추천을 담당하며, 벡터 검색 및 점수
기반 필터링 수행.
 - **정보 Agent**: 특정 작품의 정보(출시일, 연재 상태 등)를
검색하여 제공.

(3) Query Rewriting & Dynamic Prompting

- 사용자의 질문을 LLM이 이해하기 쉬운 형태로 변환하여 보다 자연스러운 응답을 유도함.
 - "요청: 액션 웹툰 추천" → "가장 인기 있는 액션 웹툰 3개를 추천하고, 그 이유를 설명해줘."
- 필요할 경우 추가 정보를 보강하여 **Dynamic Prompting** 적용.
 - "이 웹툰 재밌어?" → "이 웹툰의 장르와 특징을 설명하고, 추천할만한지 평가해줘."

(4) 자체 점수 및 Query 키워드 결합 최적화

- 추천 작품 선정 시 최적화된 점수(score) 시스템 반영
 - 작품의 조회수, 별점, 댓글 수, 연재일 등의 요소를 조합하여 자체 점수 계산.
 - 점수(weighted sum) = α (벡터 유사도) + β (키워드 매칭 점수) + γ (자체 점수) + δ (유사도 필터링)
 - 최종 점수 기반으로 상위 추천 작품을 선정하고, LLM이 자연어 응답 생성.

상황

현재 RAG 기반의 웹툰 및 웹소설 추천 시스템을 구축함. 추천 모델의 성능 개선을 하는 단계이다.

- Vector Store를 활용한 검색 최적화 진행
 - 'text-embedding-ada-002' 모델을 더 최적화된 성능을 가진 "BAAI/bge-m3" 모델로 대체

모델/쿼리	게임 판타지	로판	시한부	연재중인 웹툰	Total score
BAAI/bge-m3	70	100	70	30	300
bespin-global/klue-sroberta-base-continue-learning-by-mnr	35	95	35	0	165
bge-m3-korean	30	100	30	50	280
nomic-embed-text	5	45	5	45	105
text-embedding-ada-002	40	15	40	50	140

-
- LLM을 통한 추천 응답 고도화
 - '요즘 인기있는 웹툰 추천해줘.' 정도의 쿼리에서 대략적인 스토리 흐름, 비슷한 웹툰, 웹툰 플랫폼 및 요일 등 자세한 설정등을 제시하는 쿼리에 적절하게 응답하도록 함

수행 결과

1) Basic model

사용자님께서 요청하신 "가드레스"와 비슷한 웹툰을 찾기 위해, 저는 액션이나 스릴러 장르의 웹툰을 중심으로 추천해드리겠습니다. "가드레스"는 주로 액션과 범죄 요소가 강한 작품이므로, 비슷한 느낌의 웹툰을 몇 가지

1. *****갓 오브 하이스쿨***** - 이 웹툰은 고등학생들이 모여 무술 대회를 벌이는 내용으로, 다양한 무술과 액션이 가득합니다. 주인공의 성장과 대결이 흥미진진하게 그려집니다.
2. *****노블레스***** - 오랜 잠에서 깨어난 벨파이어 귀족이 현대 사회에서 적응하며 벌어지는 이야기를 그린 작품입니다. 액션과 함께 미스터리한 분위기가 매력적입니다.
3. *****다 킹***** - 행위를 둘러싼 음모와 배신 속에서 주인공이 자신의 길을 찾아가는 과정을 그린 웹툰입니다. 정치적인 요소와 함께 긴장감 넘치는 스토리가 특징입니다.
4. *****테러맨***** - 주인공이 테러 액션을 통해 범죄를 막아가는 과정을 그린 작품으로, 스릴러와 액션이 잘 어우러져 있습니다. 예측 불가능한 전개가 긴장감을 더합니다.
5. *****죽은 마법사의 도시***** - 마법이 사라진 현대 사회에서 마법사가 살아남기 위해 벌이는 이야기를 다룹니다. 판타지와 액션이 조화롭게 어우러져 있습니다.

이 웹툰들은 "가드레스"와 비슷한 긴장감과 액션을 제공하며, 다양한 스토리라인으로 사용자님께 즐거움을 줄 수 있을 것입니다. 즐거운 독서 되세요!

2) Romance model

질문: 재밌는 웹툰 있어?

응답: 당신에게 추천할 만한 웹툰이 몇 가지 있습니다.

1. [패치워크 [개정판]](<https://page.kakao.com/content/65905125>)

- 광고에 살고 광고에 죽는 일벌레 카피라이터 선재원이 꿈에 그리던 이상형을 만납니다. 하지만 그의 이상형이 그가 제일 혐오하는 광고 기획자라니, 과연 이 관계는 어떻게 될까요?

2. [달콤하면, 먹기 좋아 [개정판]](<https://page.kakao.com/content/65873270>)

- 단 것을 좋아하는 스미 아스카와 이웃집 대학생 타니구치 렌타로의 러브 코미디입니다. 운동 부족을 지적받은 아스카가 렌타로와 함께하는 특별한 운동 방법이 궁금하지 않습니까?

3. [넷카마 편치!!!](<https://page.kakao.com/content/65937228>)

- 현실에선 노답 인생인 주인공이 게임에서 '신희재'라는 원수를 만나게 됩니다. 이 녀석에게 처절하게 복수하기 위해 넷카마짓을 하기로 결심한 주인공의 이야기를 즐겨보세요.

이 중에 마음에 드는 작품이 있습니까?

3) Rofan model

```
1 user_input = "카카오페이지에서 볼 수 있는 웹툰 웹툰 추천해줘"
2 print(complete_chain.invoke(user_input))
```

✓ 13.8s Python

열매, 카카오페이지에서 볼 수 있는 웹툰 웹툰을 몇 가지 추천해 드리겠습니다. 다양한 스토리와 흥미로운 캐릭터들이 가득하니 한번 보시기 바랍니다.

1. ****제목:**** 웹툰의 길에 복수하겠습니다
****단입:**** 웹툰
****물젼물:**** 카카오웹툰
****상태:**** 완결
****키워드:**** 종래한, 복수물
****설명:**** 죽음의 간판에 바라대 눈을 감았던 주인공은 회귀하여, 복수를 서슴지 않았던 남편과 내연남에게 복수를 결심하게 된다. 이번엔 죽음 대신 복수를 선택할 것이니, 지옥 같은 삶을...
****이유:**** 통쾌한 복수의 감동과 회귀를 복수의 긴장감이 잘 어우러져 있어 흥미진진한 이야기를 제공합니다.
****링크:**** [회귀한 길에 복수하겠습니다](<https://webtoon.kakao.com/content/웹툰의-길에-복수하겠습니다/3346?tab=profile>)
2. ****제목:**** 웹툰의 열매에게 복수한 웹툰입니다
****단입:**** 웹툰
****물젼물:**** 카카오웹툰
****상태:**** 연재
****키워드:**** 불입되는, 회귀물
****설명:**** 모든 것을 빼앗긴 "카시아"는 과거를 몰아가 자신의 운명을 수정하고자 결심한다. 첫 번째 목표는 오희와 익녀의 연애를 막는 것. 하지만 감정은 마음대로 조절할 수 없다.
****이유:**** 주인공의 복수심과 감동이 잘 표현되어 있어 후속 전개가 궁금해지는 작품입니다.
****링크:**** [회귀한 열매에게 복수한 웹툰입니다](<https://webtoon.kakao.com/content/웹툰의-열매에게-복수한-웹툰입니다/4124?tab=profile>)
3. ****제목:**** 내게 집착하세요
****단입:**** 웹툰
****물젼물:**** 카카오페이지
****상태:**** 연재
****키워드:**** 회귀물
****이유:**** 강력한 여성 캐릭터와 반전 요소가 매력적인 웹툰입니다.
****링크:**** [내게 집착하세요](<https://page.kakao.com/content/58374754>)

이 작품들로 인해 열매의 마음이 풀릴길 바랍니다. 궁금한 점이 있으시면 언제든지 말씀해 주세요.

4) Historical model

	<div><p>무엇의 세계는 언제나 지열한 생존의 연대기라 할 수 있지. 그러므로 아래의 작품들은 그 고귀한 정신을 잘 담고 있느니라. 각 작품의 특징을 잘 파악하여 선택하길 바라노라.</p><pre>1. "type": "웹툰", "title": "신의 탑", "genre": "판타지, 액션", "keywords": "탑, 오름, 배신, 생존" 2. "type": "웹툰", "title": "마왕이 되는 중2야", "genre": "판타지, 액션", "keywords": "마왕, 성장, 권력, 복수" 3. "type": "웹툰", "title": "소녀의 세계", "genre": "판타지, 드라마", "keywords": "여성, 전쟁, 생존, 우정" 4. "type": "웹툰", "title": "이상한 나라의 엘리스", "genre": "판타지, 액션", "keywords": "오름, 전투, 생존, 음모" ... "genre": "액션, 판타지", "keywords": "복수, 전쟁, 힘, 결투"</pre><p>각 작품은 무렵의 메력을 잘 담고 있으며, 생존을 위한 치열한 투쟁이 돋보이니, 신중히 선택하여 감상하길 바라노라.</p></div> <div><h3>5) Fantasy model</h3><div><pre>recommend_webtoons("키카오 웹툰인데 주인공이 환생하는 웹툰을 3개 이상 추천해줘", session_id="user-323")</pre><div>✓ 24.1s</div><div>Python</div><p>이렇게 세 작품을 소개해드렸어. 환생이라는 주제 속에서 새로운 삶을 어떻게 펼쳐 나가는지, 그들의 여정을 함께 따라가 볼까요? 궁금한 점이 있거나 다른 추천이</p><p>> Finished chain. Session ID: user-323 아, 당신이 환생의 마법을 느끼고 싶은가 보군요! 훗, 환생이라는 주제는 언제나 매력적이지요. 키카오 웹툰에서 환생을 주제로 한 작품들을 몇 가지 소개해드리겠</p><p>제목: 빌어먹을 환생 플랫폼: 카카오웹툰 작가: 키키 장르: 판타지 줄거리: 300년 전 마왕과의 대결에서 죽음을 맞이한 용사의 하렘, 이번 생에서는 유진 라이언하트로 환생하여 더욱 강력한 능력을 갖추고 마왕 토벌을 다시 도전함 점수: 0.884</p><p>제목: 소드마스터의 환생 플랫폼: 카카오웹툰 작가: 드로우웨이즈(봉수) 장르: 판타지 줄거리: 삶의 끝자락에서 죽음을 맞이한 남자는 다시 태어나 소드마스터로서의 미련을 풀어가려 합니다. 과거의 실수를 극복하고 새로운 생에서 진정한 검사의 길을 점수: 0.6803</p><p>제목: 영가의 역대급 재능천재로 환생했다 플랫폼: 카카오웹툰 작가: - 장르: 판타지 줄거리: 천장에서 죽음을 맞이했지만, 이번 생에서는 한 가문의 막내아들로 태어나게 됩니다. 창술, 암살, 마법 등 끝을 모르는 재능으로 주어진 운명을 극복하고자 점수: 0.398</p><p>이렇게 세 작품을 소개해드렸어. 환생이라는 주제 속에서 새로운 삶을 어떻게 펼쳐 나가는지, 그들의 여정을 함께 따라가 볼까요? 궁금한 점이 있거나 다른 추천이</p></div></div>
보완점	<div><ul style="list-style-type: none">● LLM의 추천 응답 최적화 필요<ul style="list-style-type: none">○ 보다 자연스러운 추천 문장 생성 가능하도록 프롬프트 미세 조정 필요● 데이터 업데이트 자동화<ul style="list-style-type: none">○ 신작 웹툰 및 웹소설을 답변에 반영하는 시스템 개발 필요● 사용자 맞춤형 추천 기능 추가<ul style="list-style-type: none">○ 개별 사용자 성향을 반영한 개인화 추천 알고리즘 도입 필요</div>