

Gemma 3를 활용한 사용자 프로필 기반 추천 시스템

문에스터¹, 송도연¹, 조준희², 문남미³

¹호서대학교 빅데이터AI학과 학부생

²호서대학교 컴퓨터공학부 학부생

³호서대학교 컴퓨터공학부 교수

esterstst@gmail.com, thdehdus0101@gmail.com,

jjunhuui@gmail.com, nammee.moon@gmail.com

User Profile-Based Recommendation System Using Gemma 3

Esther Moon¹, Doyeon Song¹, Junhui Cho², Nammee Moon²

¹Dept. of Big Data and AI, Hoseo University

²Dept. of Computer Engineering, Hoseo University

요 약

본 연구는 추천 시스템에서 신규 사용자 환경 중 cold-start 문제를 완화하기 위해, 초경량 언어 모델 Gemma 3 270M을 활용한 사용자 프로필 생성 프레임워크(GUPF)를 제안하였다. GUPF는 사용자 리뷰 텍스트로부터 선호·비선호 요인과 감정적 태도를 추출하고, 이를 구조화된 프로필로 변환하여 추천 알고리즘의 입력으로 활용한다. Amazon Beauty 데이터 셋에 대한 실험 결과, 리뷰 기반 프로필은 기존 평점 기반 협업 필터링 대비 신규 사용자 추천 성능을 Precision 약 4.0pp, Recall 약 20.2pp, NDCG 약 18.1pp 개선하였다. 또한 GUPF는 초경량 모델임에도 리뷰 텍스트 기반 사용자 프로필 생성을 효율적으로 수행해 실시간 서비스 적용 가능성을 보여준다.

1. 서론

최근 대형 언어 모델(LLM)은 추천 시스템에도 활용되며, 사용자 리뷰나 대화 기록을 분석해 선호를 정교하게 반영함으로써 추천 성능을 향상시키고 있다[1]. 그러나 이러한 모델들은 파라미터 규모가 크고 연산 비용이 높아, 실제 서비스 환경에서 적용하기 어렵다는 한계를 지닌다[2]. 또한 전통적인 협업 필터링과 마찬가지로, 사용자 정보가 부족한 cold-start 문제에서는 여전히 추천 성능이 저하되는 문제가 존재한다[3].

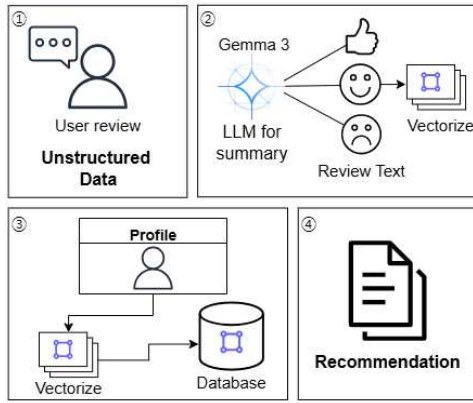
이에 본 논문은 초경량 언어 모델인 Gemma 3 270M을 활용한다[4]. 해당 모델을 활용하여 사용자 리뷰 텍스트를 수집하고, 이를 기반으로 정형화된 사용자 프로필을 구축하는 GUPF(Gemma-based User Profiling Framework)를 제안한다. 제안한 프레임워크는 기존 LLM 대비 비용과 지연을 크게 줄이면서도, 특히 신규 사용자 환경에서 소수의 리뷰만으로도 초기 프로필을 생성해 추천 품질을 개선할 수 있음을 보이는 것을 목표로 한다.

2. GUPF

본 논문에서 제안하는 GUPF는 전통적인 추천 시스템의 한계를 극복하고, 특히 신규 사용자 환경에서의 추천 성능을 향상시키기 위한 프레임워크이다. GUPF는 초경량 언어 모델인 Gemma 3 270M을 활용하여 사용자 리뷰 기반의 프로필을 구축하여 이를 기존 협업 필터링 기반 추천 모델과 결합하는 하이브리드 구조를 다음과 같이 제안한다.

- ① **텍스트 입력**: 사용자가 남긴 리뷰 텍스트를 수집하여, 이후 프로필 생성 단계에 활용한다.
- ② **LLM 요약·분석**: Gemma 3 270M 모델을 통해 리뷰 텍스트로부터 (1) 선호 카테고리, (2) 비선호 속성, (3) 감정적 태도(긍·부정)를 요약한다.
- ③ **프로필 생성**: 요약된 정보를 기반으로 사용자 프로필을 벡터화하여 데이터베이스에 저장한다.
- ④ **추천 시스템 연동**: (a) 기존 협업 필터링 기반 추천 점수와 (b) GUPF 프로필 기반 점수를 결합하여 최종 점수를 산출하고, 이를 바탕으로 아이템을 재정렬(reranking)하여 개인화된 추천 결과를 생성한다.

주요 단계에 대한 GUPF의 전체 구성도는 (그림 1)과 같다.



(그림 1) GUPF 전체 구성도

3. 실험

3.1 데이터 셋 전처리

실험에는 전자상거래 리뷰 데이터로 널리 활용되는 Amazon Beauty 데이터 셋을 사용한다[5].

본 논문에서는 각 사용자의 최근 10개 리뷰 텍스트를 대상으로 불용어 제거 및 토큰화 과정을 거쳐 입력 길이를 80 이상 100 이하의 토큰으로 제한하여 Gemma 3 270M 모델에 입력한다. 데이터 셋은 총 2000명의 사용자와 4,865개의 리뷰 텍스트로 구성되며, 리뷰가 3개 이하인 사용자는 1,862명이다.

이후 각 리뷰는 LLM이 세 가지 요소로 요약하여 사용자 프로필로 변환한다. 해당 프로필은 하이브리드 추천 단계에서 최종 순위 산출에 사용한다.

3.2 실험 설계

- Baseline: 사용자 - 아이템 평점 행렬만을 활용한 전통적 협업 필터링 방식(Matrix Factorization).
- GUPF: 리뷰 텍스트를 기반으로 Gemma 3 270M 모델이 생성한 사용자 프로필을 추천 모델 입력으로 활용.

이때 cold-start 환경을 재현하기 위해, 리뷰 수가 3개 이하인 사용자를 신규 사용자 그룹으로 설계한다[6]. 각 신규 사용자는 초기 리뷰를 GUPF에 입력해 프로필을 생성한다. 생성된 프로필을 기반으로 추천 성능을 평가함으로써, 제안한 프레임워크가 기존 협업 필터링 대비 cold-start 문제를 어느 정도 완화할 수 있는지 검증한다.

성능은 Precision, Recall, NDCG 등의 평가 지표를 활용하여 측정한다. 또한 신규 사용자 그룹과 기존 사용자 그룹을 나누어 성능을 비교함으로써, 제안된 프레임워크의 효과를 보다 명확히 검증한다.

3.4 실험 결과

<표 1>은 신규 사용자 및 기존 사용자 그룹에 대한 평가 결과를 보여준다. 제안한 GUPF는 신규 사용자 그룹에서 baseline 대비 모든 지표에서 유의미

한 향상을 나타냈다. Precision은 약 4.0pp, Recall은 약 20.2pp, NDCG는 약 18.1pp 개선되었다. 반면 기존 사용자 그룹에서는 Baseline 대비 성능이 다소 낮게 나타났는데, Precision은 9.6pp, Recall은 9.9pp, NDCG는 10.8pp 감소했다. 이는 협업 필터링만으로 충분한 정보를 활용할 수 있는 기존 사용자 환경과 달리, 신규 사용자 환경에서 GUPF가 cold-start 문제를 완화하는 데 강점을 보임을 의미한다.

<표 1> 성능 평가 결과

사용자 그룹	Model	Precision	Recall	NDCG
신규 사용자	Baseline	0.0004	0.0021	0.0009
	GUPF	0.0407	0.2037	0.1821
기존 사용자	Baseline	0.1200	0.1800	0.1500
	GUPF	0.0240	0.0810	0.0420

4. 결론

본 논문에서는 Gemma 3 270M을 활용하여 사용자 리뷰 텍스트로부터 구조화된 사용자 프로필로 변환하는 GUPF를 제안하였다. 제안한 프레임워크는 기존 LLM과 대비해 효율적이면서도, 신규 사용자 환경에서 발생하는 cold-start 문제를 완화할 수 있는 가능성을 보여준다.

ACKNOWLEDGEMENT

본 연구는 2025년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음(No. 2025-0-00040).

참고문헌

- [1] Lika, Blerina et al. "Facing the cold start problem in recommender systems." Expert systems with applications 41.4 2014: 2065-2073.
- [2] Hoffmann, Jordan et al. "raining compute-optimal large language models." arXiv preprint arXiv:2203.15556 2022.
- [3] Lika, Blerina et al. "Facing the cold start problem in recommender systems." Expert systems with applications 41.4 2014: 2065-2073.
- [4] Team, Gemma et al. "Gemma 3 technical report." arXiv preprint arXiv:2503.19786 2025.
- [5] Ni, Jianmo et al. "Justifying recommendations using distantly-labeled reviews and fine-grained aspects." Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP). 2019.
- [6] Lee, Hoyeop et al. "Melu: Meta-learned user preference estimator for cold-start recommendation." Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019.