

성분 및 리뷰 기반 맞춤형 스킨케어 하이브리드 AI 추천 시스템

정진희, 김서영, 최선호, 김현희
동덕여자대학교 정보통계학과

20201077@dongduk.ac.kr, 20201043@dongduk.ac.kr, 20200274@dongduk.ac.kr, heekim@dongduk.ac.kr

Hybrid AI Recommendation for Personalized Skincare Based on Ingredients and Reviews

Jin-Hui Jeong, Seo-Young Kim, Seon-Ho Choi, Hyun-Hee Kim
Dept. of Information and Statistics, Dongduk Women's University

요약

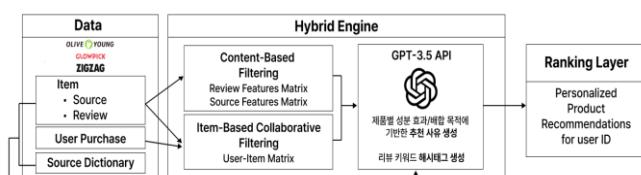
본 연구는 스킨케어 제품 구매시 제품의 내재적 특성을 중시하는 소비자의 스마트 소비 경향을 반영하여 구매 이력, 리뷰, 함유 성분 및 효능 정보를 통합한 하이브리드 추천 시스템을 제안한다. 리뷰 임베딩 및 성분 기재 원칙을 활용해 제품 간 유사도를 정교하게 측정하고, 생성형 AI를 활용해 추천 사유와 키워드를 생성하여 설명가능성을 강화하였다. 이를 통해 소비자에게 신뢰도 높은 맞춤형 추천 경험을 제공하는 데 기여할 수 있다.

1. 서론

H&B 스토어 올리브영은 소비자 구매 이력에만 의존하여 제품을 추천하고 있다. 하지만 최근 소비자들은 특히 스킨케어 제품 구매 시 성분, 효능, 리뷰 등 제품의 내재적 특성에 기반한 스마트 소비 경향을 보이고 있으며, 특히 성분을 직접 확인하고 구매하는 경향이 증가하고 있다. 이처럼 성분에 대한 관심은 높아지고 있으나, 전성분을 비교하는 데에는 여전히 어려움이 따른다[1]. 이에 본 연구에서는 구매 이력, 리뷰, 성분 및 효능 정보를 통합한 하이브리드 추천 시스템을 제안한다[2].

2. 하이브리드 AI 추천 시스템 개요

소비 트렌드 반영과 소비자 요구 충족을 위하여 스킨케어 제품군을 대상으로 하이브리드 추천 시스템을 구축하였다[3]. 또한, 생성형 AI를 활용해 추천 사유와 해시태그를 함께 생성함으로써 소비자 구매 의사결정을 돋도록 하였다. 최종 시스템 구조도는 (그림 1)과 같다.



(그림 1) 하이브리드 추천 시스템 구조도.

3. 하이브리드 추천 알고리즘

올리브영에서 제품 정보와 리뷰를 포함한 1,474 개의 스킨케어 제품 데이터를 수집하였다. 또한, 올리브영과 글로우픽에서 수집 가능한 리뷰어 2,000 명의 구매기록 및 평점 데이터를 수집하였다. 지그재그에서는 제품별 성분의 효과와 배합목적 데이터를 수집해 2,639 개의 성분 사전을 구축하였다. 본 연구에서는 아이템 기반 협업 필터링과 콘텐츠 기반 필터링을 가중 평균하여 하이브리드 추천 알고리즘을 구축하였다. 또한, GPT-3.5 모델을 활용하여 성분 사전 데이터를 기반으로 제품별 성분 효과 및 배합 목적에 기반한 추천 사유를 생성하고, 리뷰 데이터를 기반으로 제품별 주요 키워드를 해시태그 형태로 생성하도록 구현하였다. 한편, 딥러닝 모델은 추천 근거를 소비자에게 직관적으로 제공하기 어렵고 많은 학습 데이터가 요구되는 한계가 있어, 본 연구에서는 제한된 데이터 환경에서도 구현 가능하고 소비자에게 성분 및 리뷰 기반 유사도 수치를 명시적으로 제공할 수 있는 하이브리드 모델을 택하였다. 딥러닝 모델과의 성능 비교는 선행연구를 통해 이론적으로 검토하였다.

3.1 아이템 기반 협업 필터링

사용자 평점 데이터를 기반으로 사용자-아이템 평점 행렬을 생성한 후, 아이템 기반 협업 필터링(Item-Based Collaborative Filtering)을 적용하였다. 데이터 희소성 문제 보완과 미평가 항목의 평점 예측을 위해

확률적 경사하강법(Stochastic Gradient Descent, SGD)을 활용한 행렬 분해 기법을 사용하였다. 이후 아이템 간 코사인 유사도(Cosine Similarity)를 계산하여 유사도 행렬을 구축하였다. 최종적으로 사용자가 평가한 제품과 유사한 제품에 대해 유사도와 예측 평점의 가중 평균을 활용하여 추천 점수를 산정하였다.

3.2 콘텐츠 기반 필터링

제품별 리뷰 및 성분을 기반으로 제품 특성을 반영한 추천을 위해 콘텐츠 기반 필터링(Content-Based Filtering)기법을 적용하였다.

리뷰 데이터를 기반으로 BERT(Bidirectional Encoder Representations from Transformers)모델을 사용하여 문장 단위 임베딩 벡터를 생성하였다. 이후 K-Means 알고리즘을 적용하여 의미적으로 유사한 리뷰들을 군집화하고, 제품별 클러스터 중심 벡터를 계산하였다. 이후 제품 간 코사인 유사도를 계산하여 리뷰 기반 유사도를 계산하였다. 이 방식은 단순 평균 임베딩보다 의미 기반 특징을 더 효과적으로 반영해 더 정밀한 유사도 비교가 가능하다.

화장품에 사용된 모든 성분의 정확한 함량은 알 수 없으나, 함량이 많은 순서로 표기된다는 화장품 성분 기재 원칙을 고려하여 제품별 성분 데이터를 기반으로 성분 유사도 계산식을 도출하였다. 자카드 유사도(Jaccard Similarity)와 켄달 타우 상관계수(Kendall's Tau Correlation Coefficient)를 결합한 방식으로 유사도를 산출하였다. 자카드 유사도로 제품 간 공통으로 포함된 성분을 기준으로 유사도를 계산하였고, 켄달 타우 상관계수로 제품 간 공통 성분의 상대적 순서를 비교하여 성분 배열 간 순위 상관관계를 반영하여 유사도를 계산하였다. 이러한 리뷰 기반 유사도와 성분 기반 유사도는 아이템 잠재요인으로서, 콘텐츠 기반 필터링의 최종 유사도 계산식을 <식 1>과 같이 도출하였다.

$$\alpha \cdot S_{\text{review}} + (1 - \alpha) \cdot [\beta \cdot S_{\text{Jaccard}} + (1 - \beta) \cdot S_{\text{Kendall}}]$$

<식 1> 콘텐츠 기반 필터링 최종 유사도 계산식.

각 가중치는 MAP(Mean Average Precision)값이 최적이 되는 지점을 찾아 결정하였으며, 값은 <표 1>과 같다.

Review (α)	Source ($1-\alpha$)	Jaccard (β)	Kendall ($1-\beta$)
0.32	0.68	0.64	0.36

<표 1> 콘텐츠 기반 필터링 가중치 구성.

4. 최종 추천 결과 및 성능 평가

최종 추천 결과는 사용자별 구매 기록을 활용하여 리뷰와 성분이 유사한 새로운 제품을 맞춤형으로 추천하고, 제품 유사도 수치, 성분 정보에 기반한

추천 사유, 리뷰에 기반한 해시태그를 함께 제공한다. (그림 2)는 ‘플라워날’ 사용자의 실제 추천 결과이다.



(그림 2) ‘플라워날’ 사용자의 추천 결과.

제안한 알고리즘의 효과를 검증하기 위해 구매기록이 20 개 이상인 리뷰 데이터 1,751 개를 대상으로 모델별 성능평가를 진행하였다. Grid Search를 통해 각 모델의 가중치를 최적화한 결과, 콘텐츠 기반 필터링 0.49, 협업 필터링 0.51 으로 설정되었다. MAP@8에서 하이브리드 모델이 약 19% 향상된 성능을 보였으며, 결과는 <표 2>와 같다. 이는 사용자 취향과 제품 특성을 동시에 반영할 때 정밀한 추천이 가능함을 시사한다.

	Collaborative	Hybrid
MAP@8	0.3213	0.5199

<표 2> 성능평가 결과.

5. 결론

본 연구에서는 협업 필터링과 콘텐츠 기반 필터링의 장점을 결합한 하이브리드 추천 시스템을 제안하고 구현하였다. 또한, 생성형 AI를 활용해 추천 결과에 대한 설명 가능성을 확보하였다.

최종적으로 협업 필터링으로 예측된 사용자 선호도와 콘텐츠 기반 필터링으로 도출된 제품 간 유사도를 함께 고려함으로써, 사용자 취향과 제품 특성을 동시에 반영한 정교한 추천 시스템을 구현하였다. 특히 리뷰와 성분 정보를 활용한 유사도는 단순 평점 패턴보다 개인화와 정확도 측면에서 우수한 성능을 보였다. 이러한 접근은 평점 데이터가 희소한 환경에서도 안정적인 추천 품질을 유지할 수 있다. 또한, 구매기록 등 플랫폼 내부 데이터 확보에 제약이 있는 상황으로 제한된 데이터를 활용하였지만, 향후 플랫폼 연계를 통해 안정적인 내부 데이터 확보가 가능할 경우, 시스템의 실용성과 확장 가능성 향상에 기여할 수 있을 것이다.

참고문헌

- [1] 정상훈. (2024). 화장품 성분 인식과 구매행동에 관한 연구 (석사학위논문). 성신여자대학교 뷰티융합대학원.
- [2] 한국경영정보학회. (2021). 고객의 특성 정보를 활용한 화장품 추천시스템 개발. 경영정보학연구, 23(4), 69-86.
- [3] Cano, E., & Morisio, M. (2019). Hybrid recommender systems: A systematic literature review. Intelligent Data Analysis, 23(1), 148-164.