**논문 요약**

**날짜: 2020.10.14**

**이름: 전우진**

|  |  |
| --- | --- |
| 논문제목 | Movie Recommendation Algorithm Using Social Network Analysis to Alleviate Cold-Start Problem |
| 논문 요약 | 최근 몇 년 동안, 인터넷에서 이용할 수 있는 정보는 빠르게 성장하고 있음. 정보의 과부화를 관리하기 위해 추천 시스템이 개발됨. 추천 시스템은 많은 고객이 관심을 갖는 제품의 정보 분석과 고객 및 과거의 구매 활동을 바탕으로 고객의 취향에 맞는 제품이나 아이템을 추천하는 소프트웨어 툴임. 추천 시스템의 목적은 사용자에게 흥미가 있을 수 있는 것을 파악하는 것임. 우리는 협업 필터링(CF) 기법, 콘텐츠 기반 필터링 기법, 하이브리드 기법을 포함한 다양한 기법을 사용하는 추천 시스템을 활용할 수 있음. CF 방법은 가장 성공적이고 추천 시스템에서 널리 사용되는 방법이지만 콜드 스타트 문제와 같은 문제도 가지고 있음. 콜드 스타트 문제는 시스템이 사용자에게 아이템을 추천하지 못하는 문제임. 모든 추천 시스템에선 사용자의 선호도 및 취향을 고려하여 사용자의 프로필을 구축하는데 새로운 사용자나 새로운 항목이 시스템에 진입할 때, 시스템이 결정을 내릴 충분한 정보를 가지고 있지 않은 사용자나 항목에 대해 부정확한 추천을 할 수 있음. 본 논문은 콜드 스타트 문제를 완화하기 위해 SNA를 이용한 영화 추천 알고리즘을 제시함. 본 논문의 주요 목표는 개인 특성을 바탕으로 타겟 유저에게 가장 적합한 영화를 추천할 수 있는 기법을 개발하는 것임.  추천 시스템은 평가와 사용자 만족도를 예측하기 위해 한 가지 유형의 필터링을 사용하며, 이를 통해 사용자는 자신의 관심사나 필요에 따라 제품을 구입할 수 있음. 추천 시스템은 검색 알고리즘의 유용한 대안임. 왜냐하면 추천 시스템은 사용자들이 찾지 못할 수도 있는 것을 발견하는데 도움을 줄 수 있기 때문임. 추천 시스템의 다양한 접근법으로는 Content-based ﬁltering, Collaborative ﬁltering, Hybrid recommender system 등이 있음. 그 중 Collaborative ﬁltering(CF)이 가장 자주 사용되는 방법임.  CF 방식은 사용자 행동, 활동 또는 사용자 선호도에 대한 대량의 데이터를 수집, 분석하는 데서 도출되며, 어떤 사용자가 다른 사용자와 유사할 것인지 예측함. 본 논문은 사용자 기반 CF 알고리즘을 적용함.  Social Network Analysis는 소셜 네트워크 속성을 분석하는 데 사용되는 방법임. 중심성은 SNA에서 사용되는 척도임. 사이 중심성은 가장 짧은 경로를 기준으로 그래프의 중심을 측정하는 것임.  Edge betweenness로 community detection도 가능함.  추천시스템의 데이터 수집과 workflow는 다음과 같음. 1. 시스템은 유저데이터와 영화 목록을 수집해 테스트 데이터로 사용할 데이터베이스에 저장함. 2. 신규 유저의 유입 3. 신규 유저의 데이터를 수집하고 영화 목록 또한 수집함. 4. 유저간의 정보를 이용해 인접행렬을 만듦. 5. 유저테이블과 인접행렬을 이용해 community detection의 결과와 비교하고 평가함. 6. 평가를 위한 그룹이 설정되면 신규 유저와 그룹의 코사인 유사도를 측정해 수정함. 7. 가장 높은 유사도를 가진 그룹으로 신규 유저가 배치됨. 8. 해당 그룹 안에서 정렬된 영화 목록에서 신규유저와 적합한 영화를 추천해줌. |
| 논문 의의 | 본 연구는 추천시스템의 cold start 문제에 대한 해결방안을 제시함. 실제 실험을 통해 다른 분류기와 추천시스템보다 나은 성능을 보여줌. |
| 논문의 한계점 및 기타 | 본 연구에선 제안된 프레임워크로 R프로그래밍을 이용했는데, 이때 꽤 오랜 시간이 걸려서 향후 연구로 다양한 타입의 데이터셋을 이용해 정확도와 소요시간을 향상시킬 필요가 있음. |