
MMGCN: Multi-modal Graph Convolution Network for Personalized Recommendation of Micro-video

Introduction

Introduction1

- 최근 Tiktok이나 Kwai와 같은 micro-video 공유 플랫폼이 큰 성공을 이루면서 micro-video 추천시스템이 관심 받고 있음
- micro-video는 3가지 모달리티(visual, acoustic, textual)를 포함하고 있어 풍부한 정보를 담고 있음
- 멀티모달 정보를 사용하면 micro-video에 대한 유저의 선호도를 더 자세히 파악할 수 있음

Introduction2

- 유저는 모달리티별 다른 취향을 가질 수 있음(예를 들어, 포스터가 마음에 들었으나 사운드 트랙에 실망할 수 있음)
- Figure1에서 i_1 과 i_2 는 Visual Space에서 가깝게 표현되지만, 다른 topic word로 인해 Textual Space에서는 멀게 표현됨
- 만약 u_1 이 시각적인 요소를 중요시 한다면 i_2 를 클릭할 것이고, 설명을 중요시 한다면 i_3 를 선택할 것임
- 따라서 모달리티별 유저의 선호도를 파악하는 것이 매우 중요함



Figure 1: An illustration of modal-specific user preferences.

Introduction

- 본 연구에서는 다양한 모달리티의 정보를 활용하여 더 자세한 유저의 선호도를 파악하고, 모달리티에 대한 추천시스템 성능에 영향을 미치는지 연구함
- 그래프를 활용해 각 모달리티의 정보가 반영된 유저의 representation을 얻음
- Tiktok, Kwai, MovieLense 데이터셋을 활용해 모델의 효과를 입증함

Framework

- 저가 시청 이력이 있는 아이템의 modality를 Visual, Acoustic, Textual로 나눔
- 각 modality 별로 이분형 그래프(bipartite graph)로 나타내고 Aggregation Layer와 Combination Layer를 통과하여 중심 노드로부터 연결된 modality의 정보를 표현할 수 있는 최종 노드를 생성
- 이러한 과정을 모든 modality에 대해 아이템 노드와 유저 노드를 기준으로 수행합니다.
- 최종적으로 유저 노드를 기준으로 합쳐진 modality와, 아이템 노드를 기준으로 합쳐진 modality가 하나의 벡터를 만들어 행렬 연산을 통해 유저에게 새로운 아이템을 추천해줍니다.
- 이때 각 modality마다 features vector로 만들어 주기 위해 Visual modality는 pre-trained ResNet50을 적용하고, Acoustic Modality는 VGGish, Textual Modality는 Sentence2Vector를 적용했다고 합니다.

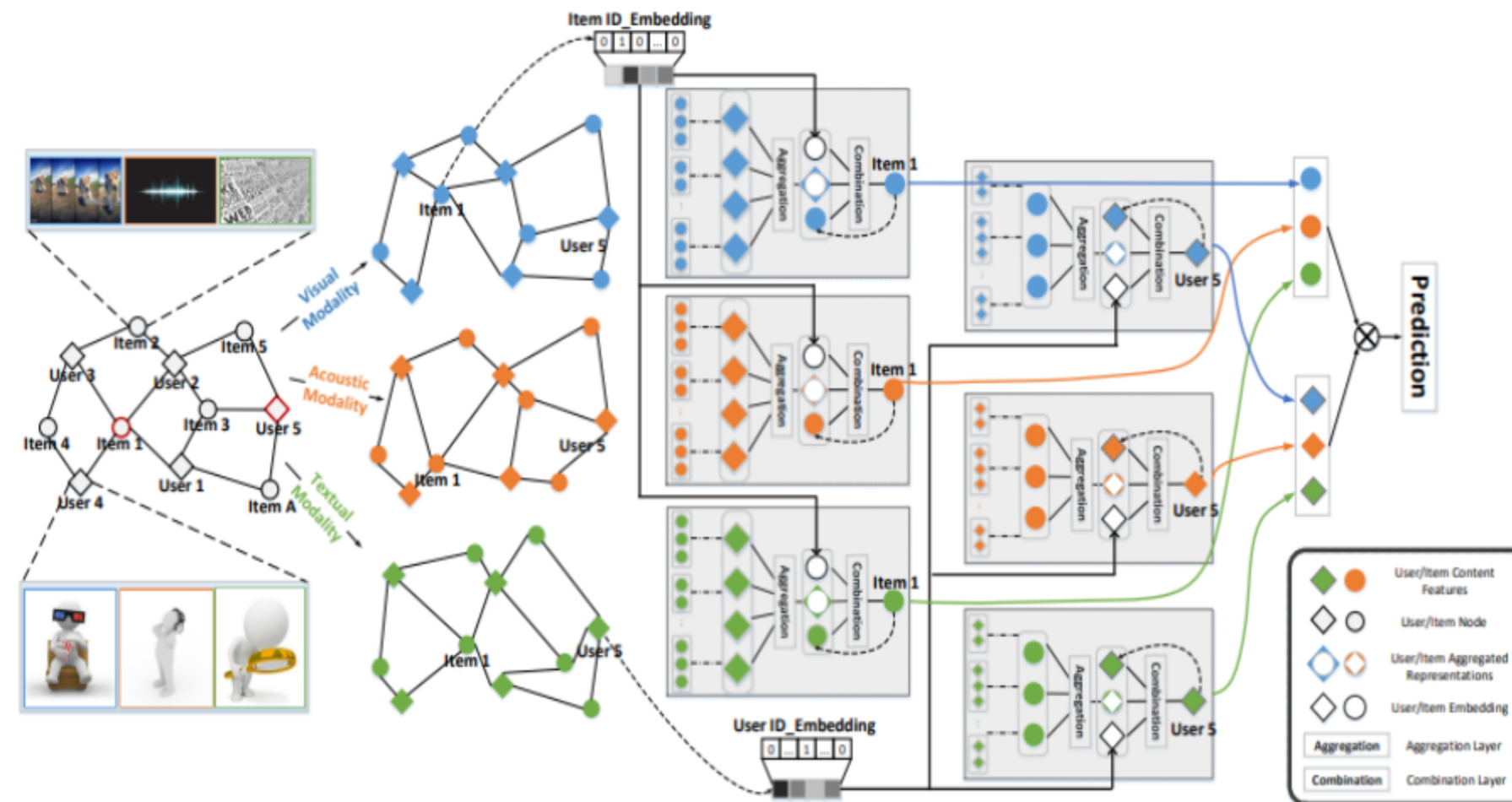


Figure 2: Schematic illustration of our proposed MMGCN model. It constructs the user-microvideo bipartite graph for each modality to capture the modal-specific user preference for the personalized recommendation of micro-video.

The user preferences on different modalities



- 유저 5명을 랜덤 샘플링해서 다른 modality에 대해 선호도의 차이가 있는지 t-SNE를 통해 시각화
- 노란색으로 표현된 유저1의 경우 Visual Modality에서 뚜렷한 차이를 보이지 않고 분산되어 나타나 있지만, Textual Modality에서는 전쟁과 로맨스의 특성들로 나누어진 것을 알 수 있음
- 마찬가지로, 파란색으로 표현된 유저2의 경우 Visual Modality에서 고전주의 포스터와 애니메이션 포스터로 특성들이 나누어졌지만, Textual Modality에서는 구분된 특징을 찾을 수 없음
- 이를 통해서 유저들이 서로 다른 modality에서 각각 다른 선호도를 갖고 있다는 것을 알 수 있습니다.

Discussions

- 유저와 아이템의 연결성을 고려하는 것이 중요할 뿐만 아니라 아이템 콘텐츠의 다양한 modality들을 고려하고자 함
- micro-video의 멀티 모달리티와 유저 사이의 관계를 통해 micro-video의 선호도를 파악하여 게임을 추천하고자 함

감사합니다