

메타패스 기반 이종 그래프 모델을 활용한 패션 커뮤니티 사용자의 패션 착장 선호도 연구

(A Study on the Personal Fashion Preference in Social Media using Meta-path based Heterogeneous Graph Modeling)

김 은 지 [†]

(Eunji Kim)

여 해 인 [†]

(Haein Yeo)

한 경 식 ⁺⁺

(Kyungsik Han)

요약 어떠한 패션 착장이 멋스러운지에 대한 판단은 주관적이기 때문에, 의상 추천에는 사용자의 선호가 함께 반영될 필요가 있다. 본 연구에서는 메타패스 기반 이종 그래프 기반의 딥러닝 모델을 활용하여 소셜 미디어에 포스팅 된 의상 착장 이미지를 통해 사용자별 의상 선호를 확인하는 추천 모델을 제안한다. 대표적인 패션 커뮤니티 사이트 Lookbook.nu에서 2,497명의 사용자가 게재한 456,329장의 의상 착장 이미지를 수집하였고, 이미지를 바탕으로 328개의 의상의 속성 정보를 추출하였다. 이미지, 사용자, 의상 속성 정보를 노드로 둔 이종 그래프 모델을 구축하고 메타패스(meta-path) 기반 노드 학습을 거쳐 사용자 임베딩 방식에 따른 사용자 선호 의상 추천 성능을 평가하였다. 본 연구 결과는 패션 영역 뿐 아니라 여러 분야의 추천 시스템에서 소셜 미디어 포스팅 이미지 분석을 통한 개인의 선호 분석의 가능성을 제시한다.

키워드: 인공지능, 그래프 신경망, 패션, 추천 시스템

Abstract Given the subjective nature of fashion, it is essential to consider user preferences when recommending fashion items. In this study, we introduce a graph-based model for look attribute preference modeling—specifically, a meta-path based heterogeneous graph deep learning model. This model utilizes outfit posts from social media to better understand user preferences. We gathered data from a representative online fashion community, Lookbook.nu, which includes 456,329 outfit images from 2,497 users, along with 328 extracted fashion attributes. A heterogeneous graph was created, with nodes representing users, item images, and fashion attributes. Using meta-path-based graph learning, we updated the multi-modal node features to reflect the relationships among the different node types. We then evaluated the effectiveness of our model in recommending fashion items based on user preferences. Our findings highlight the potential for analyzing personal preferences through social media posts to enhance recommender systems across various domains, including fashion.

Keywords: artificial intelligence, graph neural network, fashion, recommender system

- 본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구입니다.(No.RS-2020-II201373, 인공지능대학원지원원(한양대학교))
- 이 논문은 2023 한국소프트웨어융합학술대회에서 ‘메타패스 기반 이종 그래프 모델을 활용한 패션 커뮤니티 사용자의 패션 착장 선호도 연구’의 제목으로 발표된 논문을 확장한 것임

[†] 비회원 : 한양대학교 인공지능학과 학생
angiekim@hanyang.ac.kr
haeinyeo@hanyang.ac.kr

⁺⁺ 비회원 : 한양대학교 인공지능학과 교수(Hanyang Univ.)
kyungsikhan@hanyang.ac.kr
(Corresponding author)

논문접수 : 2024년 3월 20일

(Received 20 March 2024)

논문수정 : 2024년 11월 6일

(Revised 6 November 2024)

심사완료 : 2024년 11월 8일

(Accepted 8 November 2024)

Copyright © 2025 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용 행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.
정보과학회 컴퓨터의 실제 논문지 제31권 제1호(2025. 1)

1. 서 론

개인의 개성과 자기 표현의 중요성이 커지면서 자신을 드러내는 방식으로서 패션의 역할이 확대되고 있다. 소셜미디어에서도 개인의 패션 취향 이미지를 올리고 타인과 공유하는 상호작용이 활발하다. 이를 배경으로 선행 연구에서는 소셜미디어 데이터를 통해 개인의 패션 취향을 분석하고자 하였다[1].

어떠한 패션 취향이 멋스러운지 판단하는 데에는 사용자의 취향과 주관이 크게 영향을 미친다. 또한 디자인, 패턴, 브랜드 등 다양한 의상 속성 요소는 어떻게 조합되는지에 따라 착장의 분위기에 큰 영향을 주며, 함께 착장을 이루는 상의와 하의의 의상 속성 요소 간 다양한 조합은 복잡한 관계망을 가지고 있다. 이러한 패션의 주관성과 복잡성으로 인해 개인의 취향에 맞춰 의상을 추천하는 것은 매우 어렵다. 이를 극복하기 위하여 연관된 요소 간의 관계를 학습하는 이종 그래프 모델을 사용하여 사용자와 패션 요소 간의 관계를 학습하기 위한 연구가 수행되어 왔다[2].

최근 연구에서는 이종 노드 그래프 내의 복잡한 다중 흐름 구조를 고려할 수 없다는 종전 연구들의 한계를 극복하기 위해 사용자와 의상 속성 간의 관계를 분석하여 이를 활용한 기반 모델을 제시하였다[3].

복하기 위하여, 사전 정의한 메타패스(Meta-path)를 통해 노드 간 관계를 탐색할 수 있게 해주는 MAGNN (Meta-path Aggregated Graph Neural Network)을 패션 추천 모델의 베이스라인으로 활용하여 높은 성능을 보였다[3].

우리는 개인의 개별 선호를 기반으로 하는 착장 추천 모델을 개발하기 위해 소셜미디어에서 수집된 착장 이미지를 활용해 메타패스 기반 이종 그래프 모델로 사용자와 의상 간 선호 관계를 학습하고자 하였다. 실생활에서 사용자가 직접 코디하여 착용한 데이터를 얻기 위해 온라인 패션 커뮤니티에서 사용자 정보와 포스팅 이미지를 수집하여 의상 속성 탐지 모델로 의상 속성을 추출하였다[4].

우리는 사용자와 의상 속성 간의 선호 관계를 분석하기 위해, 사용자(U) - 의상(I) - 의상 속성(A)으로 이종 노드 그래프를 구성한 뒤 딥러닝 그래프 모델 MAGNN을 사용하여 사전 정의된 6개의 메타패스 기반으로 학습시켰다[2,5]. 학습된 모델은 사용자의 선호 아이템을 추천할 수 있는 성능을 보였다.

본 연구에서는 온라인 쇼핑몰 기반의 벤치마크 데이터를 활용한 기존 연구와 차별화하여[2,3], 사용자가 개별 선호를 기반으로 하는 착장 추천 모델을 제시하였다.

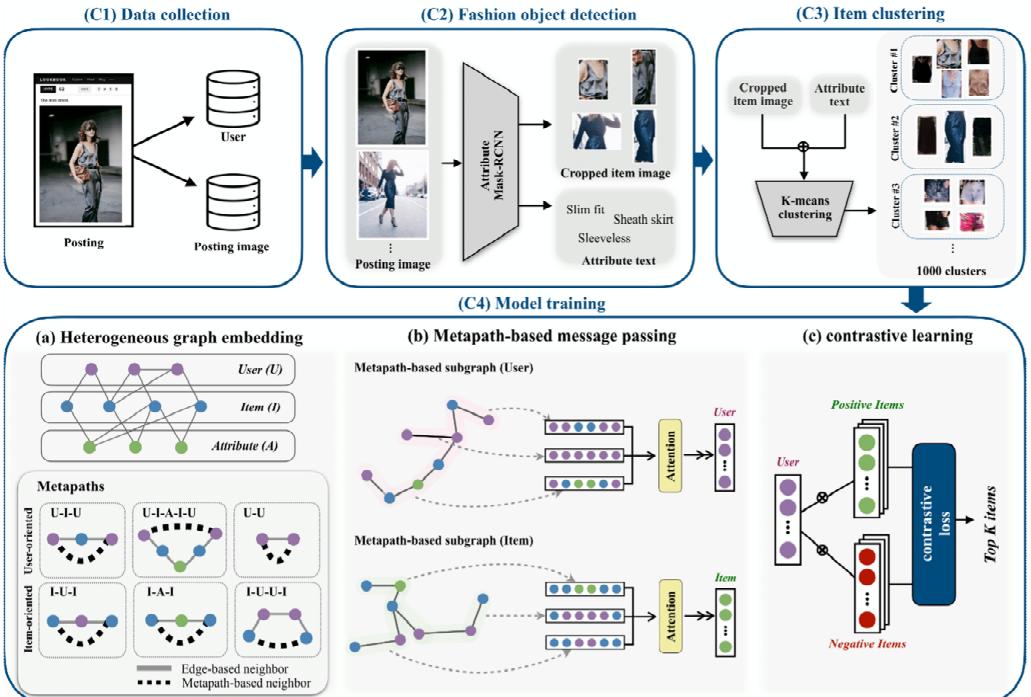


그림 1 패션 개인 선호 판별을 위한 메타패스 기반 이종 그래프 모델링 개요

Fig. 1 Overview of meta-path based heterogeneous graph modeling for fashion preference learning

재한 대량의 소셜미디어 데이터를 활용하여 군집화된 의상과 이미지에서 탐지된 패션 속성을 중심으로 노드를 표현하였다. 또한 다양한 멀티 모달 임베딩 중 사용자의 개별 패션 선호를 가장 잘 나타낼 수 있는 임베딩 방식을 제안하고 탐색한 데에서 기여점을 찾을 수 있다. 이를 통해 패션 뿐 아니라 다양한 영역의 추천시스템에서 사용자와 아이템 간의 관계를 분석하는 데에 통찰을 제시할 수 있을 것으로 기대한다.

2. 데이터 수집 및 전처리

대표적인 온라인 패션 커뮤니티 Lookbook.nu에서 2,497명의 사용자가 제작한 456,329개의 착장 이미지를 수집하였다. 이미지로부터 패션 지식에 기반한 의미 정보를 얻기 위하여 Fast R-CNN 기반 속성 탐지(Attribute detection) 모델을 사용하였다[4]. 48,825장의 패션 이미지로 사전 학습된 해당 모델을 통해 착장 이미지에서 객체를 탐지하여 상의, 하의 이미지를 얻은 후 각 아이템의 네크라인(neckline)이나 소매 길이, 패턴 등 속성 텍스트 정보를 추출하였다. 의상 이미지는 K-means 군집화로 1,000개의 아이템으로 분류하였다.

3. 메타패스 기반 이종 그래프 모델링

3.1 그래프 노드와 메타패스 정의

그래프 노드 간 복합적인 관계를 나타내는 정렬된 노드 순서를 의미하는 메타패스(Meta-path) 개념을 모델 학습에 활용하였다. 메타패스 기반 그래프에서는 사전 정의된 메타패스 관계의 이웃 노드끼리 서로 메세지를 주고 받으면서 노드 임베딩을 업데이트하고, 이 과정에서 노드 간의 관계에 대한 지식을 함축한다. 수집된 포스팅에서 3가지 요소인 1) 사용자(U), 2) 의상(I), 3) 아이템 의상 속성 정보(A)를 추출하고 이를 간의 관계를 반영할 수 있도록 세 개의 노드 타입을 가진 이종 그래프(Heterogeneous graph)를 구성하였다. 요소 간 가질 수 있는 연결 관계를 선행 연구에서 사전 정의한[5] 총 6개의 메타패스를 동일하게 구성하였다(표 1). 그 예시로, 메타패스 UIAIU은 같은 의상 속성(A)를 공유하는 아이템을 포스팅한 이력이 있는 사용자(U)들 간의 관계를 내포하고, UU는 서로를 팔로우한 사용자(U)들 간 관계

표 1 제안된 모델에서 구성된 그래프 정보

Table 1 Graph construction in the proposed model

Node	Edge	Meta-path
#User (U): 2,497	U-I	UIU, UIAIU,
#Item (I): 445,329	I-A	UU, IUI,
#Attribute (A): 328	U-U	IAI, IUUI

표 2 사용자 임베딩 사용 방식

Table 2 User embedding options

User embedding	Formular
$e_{u_i}^1$	One - hot vector
$e_{u_i}^2$	$\frac{1}{N^v k} \sum e_{v_k}, e_{v_k} = ViT(v_k)$
$e_{u_i}^3$	$\frac{1}{N^a k} \sum e_{a_k}, e_{a_k} = BERT(a_k)$
$e_{u_i}^4$	$concat [(e_{u_i}^2, e_{u_i}^3)]$

를 의미한다.

3.2 노드 임베딩

3.2.1 아이템과 의상 속성 임베딩

사용자 u_i 가 포스팅한 패션아이템 m_k 의 이미지 m_k^v 는 전체 착장 이미지에서 상의/하의 아이템으로 탐지된 바운딩 박스를 기준으로 잘린 이미지를 사전 학습된 Vision Transformer를 통해 추출된 벡터 e_{m_k} 로 임베딩하였다[6]. 의상 속성 m_k^a 로는 잘린 이미지에서 탐지된 카테고리와 속성 텍스트정보를 사전 학습된 BERT 모델을 활용하여 임베딩 벡터 e_{a_k} 로 환산하여 사용하였다[7].

$$e_{m_k} = ResNet(M_k^v), e_{a_k} = BERT(m_k^a)$$

3.2.2 사용자 임베딩

우리는 총 4가지 방식의 사용자 임베딩 e_{u_i} 을 설정하여 실험을 진행하였다(표 2). 임베딩 방식으로는, 1) 원 핫(One-hot) 벡터 임베딩, 2) 사용자가 포스팅한 이미지 v_k 를 Vision Transformer로 추출한 임베딩들의 평균, 3) 사용자의 포스팅에서 감지된 의상 속성 a_k 들을 BERT로 인코딩한 임베딩들의 평균, 4) 사용자 포스팅 이미지들의 평균 임베딩과 포스팅 룩의 의상 속성들의 평균 임베딩을 결합하는 방식을 사용하였다.

3.3 그래프 학습

3.3.1 메타패스 기반 이웃 노드 간 메세지 패싱

사전 정의된 메타패스 $P=\{p_1, p_2, \dots, p_6\}$ 는 사용자 중심 메타패스 $P_u=\{\text{UIU}, \text{UIAIU}, \text{UU}\}$ 와 아이템 중심 메타패스 $P_m=\{\text{IUI}, \text{IAI}, \text{IUUI}\}$ 로 분류된다. 각 메타패스 별로 너비 우선 탐색(Breadth first search)를 통해 구성된 서브 그래프 상에서 각 노드 v 는 각자의 임베딩 메세지를 이웃 $n \in N_v^p$ 에게 전달하고, 여러 개의 이웃 노드들에서 받은 임베딩 정보 e_{vn}^p 들은 어텐션(Attention) 기반 매커니즘을 통해 어텐션 벡터 a_p 와의 내적 값으로 환산되어 자기 자신의 임베딩 값으로 반영된다[2]. P_u 기반 이웃 노드 정보로 업데이트 된 사용자 임베딩 h_{u_i} 과 P_m 기반 이웃 정보로 업데이트 된 아이템 임베딩 h_{m_k} 의

내적으로 유사도 점수 S_{u,m_k} 를 산출하여 의상 착장 선호도 판별 과업을 수행하였다.

$$e_{vn}^b = \text{LeakyReLU} (a_v^\top \cdot \text{concat}[e_v, e_n])$$

$$h_v^b = \sigma (\sum_{n \in N_v^P} a_{vn}^P \cdot e_{vn}^b)$$

$$h_{u_i} = \sum_{P \in P_u} (\beta_P \cdot h_u^P), h_{m_k} = \sum_{P \in P_m} (\beta_P \cdot h_m^P)$$

$$a_{vn}^P = \frac{\exp(e_{vn}^P)}{\sum_{s \in N_v^P} \exp(e_{vs}^P)}, \beta_P = \frac{\exp(e_P)}{\sum_{P \in P_A} \exp(e_P)}$$

$$S_{u,m_k} = h_{u_i}^\top \cdot h_{m_k}$$

3.4 사용자별 패션 착장 선호도 판별

3.4.1 Triplets

실험에 사용된 Triplet은 (u_i, m_{k+}, m_{k-}) 로 구성되며, m_{k+} 는 해당 사용자 u_i 가 실제 포스팅한 이력이 있는 positive 아이템, m_{k-} 는 전체 아이템 중 랜덤 샘플링된 사용자와 관련 없는 negative 아이템을 의미한다. 손실 함수로는 사용자와 positive 아이템 간의 임베딩 거리는 좁히고 negative 아이템 간과의 거리는 멀어지도록 모델을 학습하는 contrastive loss를 사용하였다[9]. 주어진 아이템 중 사용자가 실제 포스팅 하였던 아이템을 추천할 수 있다면 모델이 사용자의 착장 선호를 파악한 것으로 판단하였다.

$$L = -\sum \log a(h_{u_i}^\top \cdot h_{m_{k+}}) - \sum -\log a(h_{u_i}^\top \cdot h_{m_{k-}})$$

4. 실험 및 실험 결과

4.1 실험

포스팅 이미지 456,329개의 중 299,319개(60%)가 모델 사전학습을 위한 훈련 세트로 사용되었고, 검증 세트, 테스트 세트로는 각각 99,773개(20%), 99,774개(20%)의 데이터를 사용하였다. Positive와 Negative 비율은 1:1로 설정하였으며, 학습 파라미터로는 학습률 0.005, 가중치 감쇠 0.001, 드롭아웃 0.5, 배치 사이즈 128, 에폭 5, 멀티 헤드 어텐션(Multi-head Attention) 개수 8개로 두어 실험을 수행하였다.

4.2 실험 결과

학습된 모델로 네 가지 사용자 임베딩을 비교 실험하여 추천시스템에서 주로 사용되는 F1-score와 NDCG 지표로 선호 아이템 판별 성능을 측정하였다(그림 2).

1000개의 아이템 클러스터 중 10개 아이템을 추천하여 그 중 사용자의 착장 이력이 있는지 확인한 결과,

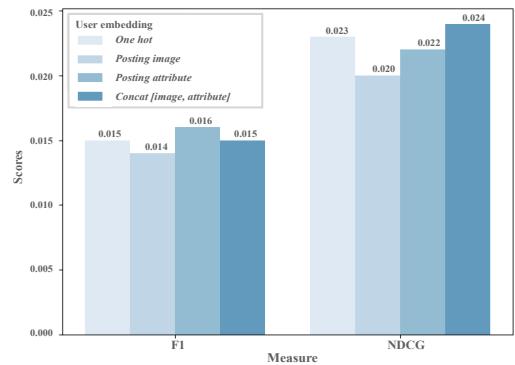


그림 2 사용자 임베딩에 따른 사용자 선호 제품 추천 성능

Fig. 2 Recommendation performance by user embedding

F1-score에서는 모든 임베딩에서 0.01를 상회하는 성능을 확인하였으며, 특히 의상 속성 텍스트가 가장 높은 성능을 보였다(F1-score: 0.016). 다음으로는 이미지와 의상 속성의 조합(F1-score: 0.015), 원핫벡터(F1-score: 0.015), 의상 이미지(F1-score: 0.014)가 순서대로 높은 성능을 보였다.

NDCG@10 지표에서는 모든 임베딩에서 0.02를 상회하는 성능을 확인하였으며, 특히 이미지와 의상 속성을 조합하였을 때 가장 높은 성능을 보였다(NDCG@10: 0.024). 다음으로는 원핫벡터(NDCG@10: 0.023), 의상 속성 텍스트(NDCG@10: 0.022), 이미지(NDCG@10: 0.020)가 순서대로 높은 성능을 보였다.

의상 속성 텍스트에 기반한 사용자 임베딩은 F1 score에서 가장 높은 성능을 보이며 의상의 형태, 텍스처, 패턴 등의 높은 수준의 의미 정보가 결과에 대한 설명력을 제공할 수 있음과 동시에 사용자 선호도 파악 성능에도 중추적인 역할을 수행할 수 있다는 가능성을 제시한다. NDCG 지표에서 의상 속성 텍스트와 포스팅 이미지가 결합 되었을 때 사용자의 선호를 가장 잘 나타낸 결과는 의상 속성에 대한 의미 정보에 더해진 의상을 착장한 배경과 그 시각적 분위기를 담은 전체적인 착장 이미지가 사용자의 선호를 파악하는 데에 부가적인 정보가 될 수 있음을 보여준다. 본 연구 결과는 소셜 미디어 포스팅 이미지와 이로부터 탐지된 의상 속성으로 사용자의 취향을 파악하여 개인화 된 아이템 추천 시스템으로 확장할 수 있는 가능성을 시사한다.

4.3 사용자 임베딩 시각화

t-SNE를 활용한 사용자 임베딩 차원 축소를 통해 사용자의 분포를 시각화하여 확인하였다(그림 3)[8]. 3차원

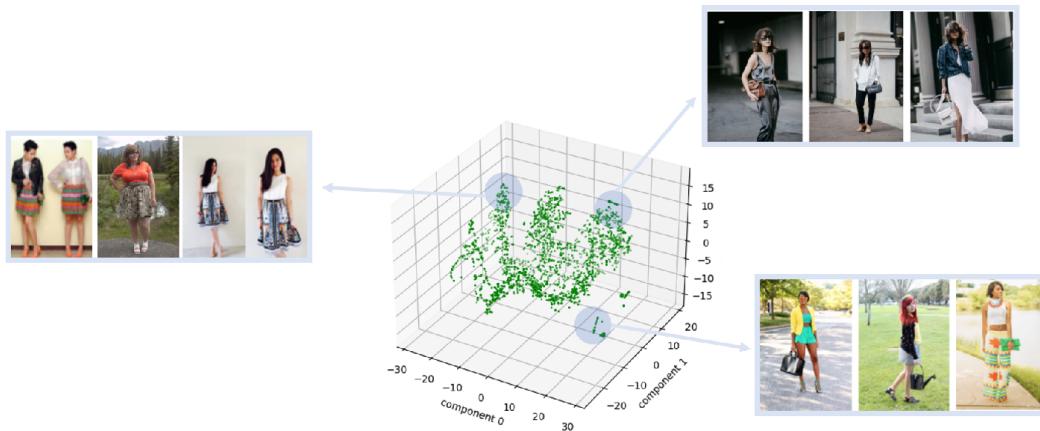


그림 3 패션 개인 선호 판별을 위한 메타패스 기반 이종 그래프 모델링 개요

Fig. 3 Overview of meta-path based heterogeneous graph modeling for fashion preference learning

으로 축소된 사용자 임베딩에서 우측 상단에 위치한 사용자는 주로 도시적이고 여성스러운 차장을 추구하며, 우측 하단의 사용자는 밝은 색감의 차장을 선호함을 확인할 수 있다. 좌측 상단의 사용자들은 주로 화려한 플레어 스커트(flare skirt)를 주로 착용하는 모습을 보였다. 본 연구에서 제안하는 모델은 의미론적인 패션 속성을 기반으로 사용자 임베딩 정보를 제공함으로써 추천 시스템에 설명력을 제공할 수 있다는 강점을 지닌다.

5. 결론 및 향후 연구

본 연구 결과를 통해 추천 시스템 관련 연구들에 사용자의 선호도 분석을 위한 통찰을 제안할 수 있을 것으로 기대한다. 제안된 멀티 모달 사용자 임베딩은 패션을 비롯한 다른 영역의 개인화 추천 시스템으로까지 확장하여 활용될 수 있을 가능성을 제시한다.

향후 연구에서는 선호 판별 성능을 고도화할 수 있는 메타패스 조합과 임베딩 방식에 대한 추가적인 실험을 진행할 예정이다.

References

- [1] Zheng, H., et al., "Personalized fashion recommendation from personal social media data: An item-to-set metric learning approach." *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021.
- [2] Li, X., et al. "Hierarchical fashion graph network for personalized outfit recommendation." *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2020.
- [3] Guan, W., et al., "Personalized Fashion Compatibility Modeling via Metapath-guided Heterogeneous Graph Learning." *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2022.
- [4] Jia, M., et al. "Fashionpedia: Ontology, segmentation, and an attribute localization dataset." *European Conference on Computer Vision*. Springer, Cham, 2020.
- [5] Fu, X., et al., "Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding." *Proceedings of The Web Conference 2020*. 2020.
- [6] Alexey D., et al., "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929*, 2020.
- [7] Devlin, J., et al., "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805*, 2018.
- [8] Van der Maaten, L., et al., "Visualizing data using t-SNE." *Journal of Machine Learning Research*, vol. 9, no. 11, 2008.
- [9] Khosla, Prannay, et al., "Supervised contrastive learning." *Advances in Neural Information Processing Systems*, vol. 33, 18661–18673, 2020.



김 은 지

2024년 한양대학교 일반대학원 인공지능
학과 졸업(석사). 관심분야는 인간-컴퓨터
상호작용, 그래프신경망, 데이터과학



여 해 인

2022년 동덕여자대학교 컴퓨터학과 졸업
(학사). 2022년~현재 한양대학교 인공지
능학과 석박통합과정. 관심분야는 인간-
컴퓨터상호작용, 자연어처리, 소셜컴퓨팅



한 경 식

2015년 펜실베니아 주립대학교 졸업(박사)
2015년~2017년 미국 퍼시픽 노스웨스트
연구소 선임연구원. 2017년~2021년 아주
대학교 컴퓨터공학과·인공지능학과 조교
수. 2021년~현재 한양대학교 데이터사이
언스학부 및 인공지능학과 부교수. 관심
분야는 인간-컴퓨터상호작용, 인공지능, 소셜컴퓨팅