

■ 연구과제 요약문

과제명(기간)	Deep metric learning and Deep Reinforcement learning (2017.09.01. ~ 2019.08.31.)
연구책임자	송 현 오 (hyunoh@snu.ac.kr)
개요	Representation learning은 데이터를 저차원의 공간으로 추상화 시켜 유사한 데이터 탐색, 데이터 생성 및 보간 등 의 다양한 작업에 활용되는 머신러닝의 근간이 되는 기술 중 하나이다. 그리고 이러한 representation이 실제 환경에서 사용되려면 학습된 representation이 효율적이어야 하고 설명가능 해야 한다. 이를 위해 효율적이고 설명가능한 최적의 representation을 찾을 수 있는 머신러닝 연구를 수행한다.
연구개발 결과	설명가능한 representation 연구는 최근 많이 진행되고 있는 연구중 하나 이지만 representation이 continuous 변수 뿐 아니라 discrete 변수가 있을 때 어떻게 학습할 지는 아직 연구된 바가 거의 없다. 이에 본 과제에서 최소비용흐름 그래프 문제와 total correlation의 lower bound를 통해 근사해를 구하는 최적화 알고리즘 [1] 을 고안하여 ICML2019에 게재되었다. 효율적인 representation 연구는 학습된 representation을 이진화된 hash code 로 최적화 하여 hashing을 통해 정확도 손실없이 수백배의 탐색 속도가 증가하는 조합최적화 알고리즘을 고안하였다. 이 연구성과는 우선 ICML 2018에 long talk (상위 8.6%) 으로 선정되어 게재되었고, 후속 연구로 계층 구조를 이용한 조합최적화 알고리즘을 통해 효율을 더욱높혀 그 성과가 CVPR 2019에 게재되었다. ----- [1] Learning Discrete and Continuous Factors of Data via Alternating Disentanglement, Yeonwoo Jeong, Hyun Oh Song, ICML 2019 [2] End-to-End Efficient Representation Learning via Cascading Combinatorial Optimization, Yeonwoo Jeong, Yoonsung Kim, Hyun Oh Song, CVPR 2019 [3] Efficient end-to-end learning for quantizable representations, Yeonwoo Jeong, Hyun Oh Song, ICML 2018 Selected long talk, top 8.6%
활용분야 및 기대효과	앞서 설명한 바와 같이 representation learning 기술은 데이터 유사도 탐색, 생성, 보간 및 다양한 응용 범위를 가지고 있고, 전 세계적으로 여러 학교와 산업연구소의 연구진들이 활발히 연구하고 있는 분야 이다. 구글과 페이스북 에서도 이미 representation learning 기술을 활용하여 이미지 및 음성 데이터 검색에 활용하고

있다. 본 과제로 진행된 연구로 앞으로 representation learning 기술이 더욱 진보하여 더 많은 산업 영역에 영향을 미칠것으로 예상된다.