

## [Review]

본 논문은 Pinterest에서 사용되는 추천 알고리즘인 PinSage에 대하여 소개된 논문이다. PinSage는 Pinterest와 GraphSage가 합쳐진 이름이다. GraphSage는 Graph Convolutional Network의 발전된 형태로 서로 연결된 노드, 혹은 정점들은 비슷한 정보를 가질 것이라는 가정하에 어떤 정점의 feature를 만들기 위해서 주변의 정보를 활용하는 형태의 알고리즘이다. 즉, 적절한 수의 주변의 노드를 샘플링하여 학습시킴으로써 GCN의 연산 속도를 개선한 알고리즘이라고 할 수 있다.

PinSage모델은 기존의 GCN의 연산 속도를 개선시키기 위해서 target Node를 선정 한 후 이웃을 샘플링하는 방식으로 학습이 진행되고, 그래프 데이터 생성 및 샘플링에는 CPU를 활용하고 모델 학습에는 GPU가 활용된다. 또한 MapReduce와 같은 파이프라인을 설계하여 반복적인 연산을 최소화하여 빠르게 임베딩 벡터를 생성한다. 그리고 효율적인 연산을 위해서 그래프 데이터를 미리 생성하며 이웃 생성시에는 Short Random Walk 방식을 활용한다. Random Walk를 통해서 노드들의 유사성도 capture할 수 있다. 또 학습 진행될 때 마다 조금 더 어려운 예시를 보여줌으로써 모델의 성능을 향상시켰다.

PinSage는 유저의 주 관심사인 pin과 pin의 collection인 board를 이분그래프 형식으로 설정해 데이터를 구성한다. 모델의 목적은 pin을 잘 임베딩 하는 것이다. 학습 시에 타겟 노드에게 상대적으로 많은 영향을 주는 노드들을 선별하여 aggregate/pooling을 진행한다. 이러한 방식으로 학습이 진행되면 중요한 이웃들만 고려하여 노드를 업데이트 시킬 수 있으며, 메모리를 줄일 수 있다는 장점을 가진다. Convolution이 종료된 후 생성된 target 노드의 Representation을 선형 레이어를 한번 더 거쳐서 최종 Embedding을 만든다. 만들어진 Embedding은 k-NN 알고리즘을 활용하여 가장 유사한 pin을 추천해주는 식으로 활용된다. 결론적으로 본 논문에서 제시된 알고리즘은 A/B Test 결과 좋은 성능을 보여줬고 유사한 pin을 추천해준다고 한다.

## [본문 URL]

<https://velog.io/@tobigs-recsys/Graph-Convolutional-Neural-Networks-for-Web-Scale-Recommender-System>