

그래프 신경망 하이퍼 파라미터 연구

민연아*, 전진영^o

*한양사이버대학교 응용소프트웨어공학과,

^o한양사이버대학교 미래융합공학대학원 기계IT융합전공

e-mail: yah0612@hycu.ac.kr*, 2022201681@hycu.ac.kr^o

A Study on Hyper Parameters of Graph Neural Network

Youn-A Min*, Jin-Young Jun^o

*Dept. of Applied SW Engineering, Hanyang Cyber University,

^oGraduate School of Mechanical & IT Convergence Engineering, Hanyang Cyber University

● 요약 ●

본 논문에서는 인공지능 신경망의 하이퍼 파라미터들이 그래프 신경망 모델의 성능에 미치는 영향을 알아보기 위하여 대규모 그래프 데이터를 기반으로 이진 분류 문제를 예측하는 그래프 합성곱 신경망 모델(Graph Convolution Network Model)을 구현하고 모델의 다양한 하이퍼 파라미터 중 손실함수와 활성화 함수를 여러 가지 조합으로 적용하며 모델 학습과 예측 실험을 시행하였다. 실험 결과, 활성화 함수보다는 손실함수의 선택이 모델의 예측 성능에 좀 더 큰 영향을 미치는 것을 확인하였다.

키워드: 그래프 알고리즘, 그래프 신경망, 손실함수, 하이퍼 파라미터

I. Introduction

사회 관계망 데이터나 분자구조 데이터와 같이 샘플과 샘플 사이의 복잡한 관계도 객체로 나타내고 연산할 수 있는 그래프 알고리즘에 관한 관심이 높아짐에 따라, 그래프를 신경망에 연결하는 그래프 신경망에 관한 연구가 활발히 진행되고 있다[1][2]. 그러나, 그래프 데이터는 일반적인 데이터셋보다는 더 복잡한 구조를 가지므로, 그래프 신경망 모델이 일반적인 인공지능 신경망과 동일한 하이퍼 파라미터로 설계되는 것이 언제나 최선이라고 가정하기는 어렵다. 본 논문에서는 하이퍼 파라미터들이 그래프 신경망 모델의 성능에 미치는 영향을 알아보기 위하여 몇 가지 손실함수와 활성화 함수를 조합하여 모델 학습 실험을 하고 실험 결과를 분석하였다.

패턴을 발견하기 위해 새롭게 정의된 인공지능 신경망 기법들도 빠르게 개발되고 있다[1]. 그래프 신경망은 그래프에서 노드의 정보를 임베딩하는 방식에 따라 다양하게 연구되고 있는데, 대표적인 그래프 신경망 기법으로는 이웃 정보를 반복적으로 집계하기 위해 그래프 리플라시안의 1차 고유 분해를 근사하는 GCN(Graph Convolution Networks)와 새롭게 추가된 노드에 대해서도 학습할 수 있도록, 노드 임베딩이 아닌 aggregation function을 학습하는 GraphSAGE, 중요한 노드에 더 큰 가중치를 주는 GAT(Graph Attention Network), RNN에서 자주 사용되는 GRU를 적용한 GGNN(Gated Graph Neural Networks), 그리고 하이퍼 그래프 구조에서 고차 데이터 상관관계를 인코딩하는 HGNN(Hyper GNN) 등이 있다[2].

II. Preliminaries

1. Related works

1.1 그래프 신경망

서로 독립된 축으로 이루어진 유클리드 공간의 데이터는 기본적으로 샘플들이 서로 독립적이라고 가정하는 데 반해 노드와 엣지로 표현되는 그래프 데이터는 상호 작용이나 위치 관계 등으로 서로 관련된 복잡한 데이터이다. 이러한 복잡한 그래프 데이터에서 잠재

III. The Proposed Scheme

본 연구에서는 손실함수와 활성화 함수가 그래프 신경망 기반 이진 분류 모델의 성능에 미치는 영향을 알아보기 위해 세 가지의 손실함수와 세 가지의 활성화 함수를 조합하여 총 9개의 모델을 학습하고 데이터셋에 대한 분류를 예측하였다. 모델의 기본 구조는 합성곱 신경망 구조를 2단으로 적층하고, 그 중간에 활성화 함수와 한 개의 dropout layer를 추가하여 최종적으로 [Fig. 1]과 같이 설계하였다.



Fig. 1. 그래프 신경망 기반 이진 분류기 모델

최적화 알고리즘과 learning rate, hidden feature 크기, output feature 크기, 학습횟수는 각각 Adam, 0.1, 64, 2, 100으로, 그리고 dropout은 0.5로 고정하였다. 실험 대상 손실함수와 활성화 함수는 [Table 1]과 같다.

Table 1. 실험대상 손실함수와 활성화 함수

손실함수	활성화 함수
Cross Entropy(CE)	ReLU
Negative Log Likelihood(NLL)	Tanh
Multi-class Hinge(MM)	LeakyReLU

연구에서 사용한 모델은 output feature의 크기인 2개의 열벡터를 갖는 $n \times 2$ 크기의 행렬을 반환한다. n 은 노드의 크기이다. 정확도는, m 개의 값을 갖는 벡터에서 가장 큰 값의 인덱스를 반환하는 argmax 를 이용하여 노드별로 0 또는 1의 예측 결과를 반환하도록 하고, 반환된 1들의 총합을 노드의 총 개수로 나눈 값이다. 1회 학습이 완료될 때마다 정확도와 손실 값을 산출하고 100회의 학습이 완료된 후 각 데이터셋마다 총 학습횟수에 대한 정확도 평균을 구하여 [Table 2]에 정리하였다.

Table 2. 손실함수(Loss F.)와 활성화 함수(Act. F.)별 예측 정확도

Loss F.	Act. F.	총 학습횟수 평균		
		train_acc	val_acc	test_acc
CE	ReLU	0.8590	0.8492	0.8467
	Tanh	0.8522	0.8450	0.8427
	LeakyReLU	0.8690	0.8631	0.8601
NLL	ReLU	0.5798	0.5811	0.5794
	Tanh	0.7384	0.7436	0.7364
	LeakyReLU	0.7159	0.7198	0.7135
MM	ReLU	0.8551	0.8506	0.8456
	Tanh	0.8443	0.8371	0.8345
	LeakyReLU	0.8664	0.8578	0.8552

Table 3은 손실함수별 평균 예측 정확도를 나타낸 것이다.

Table 3. 손실함수별 평균 예측 정확도

Loss F.	train_acc_mean	val_acc_mean	test_acc_mean
CE	0.860067	0.852433	0.849833
NLL	0.855267	0.848500	0.845100
MM	0.678033	0.681500	0.676433

활성화 함수와 관계없이 손실함수별 각 데이터셋에 대한 예측 정확도 평균을 나타내었고 해당 내용에 대하여 Fig.2와 같이 시각화 하였다.

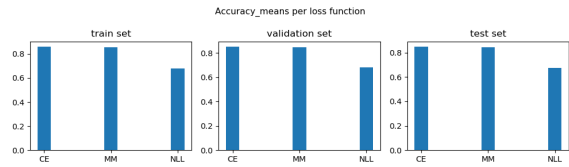


Fig. 2. 손실함수별 평균 예측 정확도

CE(Cross Entropy)와 MM(Multi-class Hinge) 손실함수를 적용한 경우, 소수집 들쳐 자리에서 반올림했을 때 train, validation, test set에 대하여 모두 활성화 함수와 관계없이 서로 같은 수준의 정확도를 보였으나, NLL(Negative Log Likelihood) 손실함수를 적용했을 때는 Cross Entropy나 Multi-class Hinge에 보다 0.18 ~ 0.17 정도의 낮은 정확도를 보였다.

IV. Conclusions

37,700개의 노드와 4,005개의 노드 특성, 그리고 578,006개의 엣지로 이루어진 대규모 그래프 데이터를 기반으로 이진 분류 문제를 예측하기 위한 그래프 합성곱 신경망 모델을 구현하고 세 가지 손실함수와 세 가지 활성화 함수를 각각 적용하며 모델 학습과 예측 실험을 시행하였다. 실험 결과, 활성화 함수보다는 손실함수의 선택이 모델의 예측 성능에 좀 더 큰 영향을 미치는 것을 확인할 수 있었다. 모델의 예측 성능을 개선하기 위해서는 효율적인 손실함수의 적용이 중요한 요소임을 알 수 있는 결과이다. 그러나, 본 실험은 매우 잘 알려진 몇 개의 손실함수와 활성화 함수만을 대상으로 진행되었으며, 모델의 구조나 그 밖의 다양한 하이퍼 파라미터에 대해서는 고려하지 못하였다. 특정 그래프 알고리즘에 효율적으로 적용 가능한 신경망 구조와 손실함수를 포함한 다양한 하이퍼 파라미터에 관한 추가 연구가 필요하다.

REFERENCES

- [1] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S., "A Comprehensive Survey on Graph Neural Networks," IEEE Transactions on Neural Networks and Learning Systems IEEE Trans. Neural Netw. Learning Syst. Neural Networks and Learning Systems, IEEE Transactions on. 32(1):4-24 Jan, 2021.
- [2] Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, Bin Cui, "Graph Neural Networks in Recommender Systems: A Survey," ACM Computing Surveys, Vol. 55, Iss. 5, pp. 1-37, 2023.
- [3] SNAP, "Github Social Network," <https://snap.stanford.edu/data/github-social.html>.
- [4] GNN4SocialNWTutorial, <https://github.com/Awadelrahman/GNN4SocialNWTutorial>