AIST4010 Foundation of Applied Deep Learning Kaggle Assignment 3 - Report

Name: Wong Shing Lok
SID: 1155156680

1 Introduction

The data preprocessing and model architecture are mainly relied on Pytorch Geometric[1] for the ease of data handling and model building.

2 Data Preprocessing

To handle the imbalance of the training, validation and test sets, I split 80% of the validation data to the training set for model training. All the edges, features and nodes are processed and mapped into the *Data* object of Pytorch Geometric[1], which a library built on top of Pytorch for better graph network training and graph data handling.

3 Model Architecture

The model is based on GCN architecture[2], which consists of 2 GCN convolutional layers for message passing, followed by ReLU and dropout after each GCN layer and a linear output layer. The architecture is similar to traditional CNN architecture that we have learned.

4 Hyperparameters

- Optimizer: Adam
- Loss Function: Cross-Entropy Loss
- Learning Rate: 5e-3
- Weight Decay: 5e-4
- Epochs: 5000
- Dropout Rate: 0.5
- Hidden Channels for GCN: 48

5 Training and Testing

The model was trained for 5000 epochs using GPU P100. The training involved forward passing the training set, loss calculation between true labels and predictions, backward propagation and updating the model weights. It would also save the model with the best validation accuracy to be the model for test prediction later on. The best validation accuracy in this case is 85%.

As for the testing, the best testing accuracy is about 82%.

References

[1] Matthias Fey and Jan E. Lenssen. "Fast Graph Representation Learning with PyTorch Geometric". In: *ICLR Workshop on Representation Learning on Graphs and Manifolds*. 2019.

[2] Thomas N. Kipf and Max Welling. *Semi-Supervised Classification with Graph Convolutional Networks*. 2017. arXiv: 1609.02907 [cs.LG].