Determining the Optimal Number of GAT and GCN Layers for Node Classification in Graph Neural Networks

Humaira Noor[1], Niful Islam[1], Md. Saddam Hossain Mukta[1], Nur Shazwani Binti Kamarudin[2]
Mohaimenul Azam Khan Raiaan[1], Sami Azam[3]
[1]Department of CSE, United International University (UIU), Bangladesh
[2]Faculty of Computing, University Malaysia Pahang (UMP), Malaysia
[3]Faculty of Science and Technology, Charles Darwin University (CDU), Australia
hnoor222007@mscse.uiu.ac.bd, nislam201057@bscse.uiu.ac.bd, saddam@cse.uiu.ac.bd, nshazwani@ump.edu.my
mraiaan191228@bscse.uiu.ac.bd, sami.azam@cdu.edu.au

Abstract—Node classification in complex networks plays an important role including social network analysis and recommendation systems. Some graph neural networks such as Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT) have emerged as effective approaches for achieving high-performance classification in such tasks. However, constructing a graph neural network architecture is challenging particularly due to the complex task of determining the optimal number of layers. This study presents a mathematical formula for determining the optimal number of GCN and GAT hidden layers. The experiment was conducted on ten benchmark datasets, evaluating performance metrices such as accuracy, precision, recall, F1-score, and MCC for identifying the best estimation of number of hidden layers. According to the experimental findings, the number of GAT and GCN layers selected has a substantial impact on classification accuracy. Studies show that adding extra layers after the optimum number of layers has a negative or no impact on the classification performance. Our proposed approximation technique may provide valuable insights for enhancing efficiency and accuracy of the Graph Neural Network algorithms.

Index Terms—Node Classification, Graph, GAT, GCN, Neural Network

I. INTRODUCTION

Graph is one of the most powerful data structures that captures relationship and interaction between entities [1]. This structure is commonly seen in social media, biological networks, transportation networks, recommendation system and others [2], [3]. In graph, each entity is represented as a node and edges represent the interactions between the nodes. Some tasks require labeling the nodes based on their attributes and connectivity pattern which is known as node classification [4]. In recent times, Graph Neural Networks (GNN) have gained a significant attraction as a powerful tool for this complex task. Graph Convolutional Network (GCN) and Graph Attention Network (GAT) are two prominent GNN architectures due to their ability to effectively capture the structural information of complex graphs [5].

GCN learns expressive node representations by aggregating information from adjacent nodes and integrating it with attributes from the core node. In order to effectively extract and classify features from the graph, we use graph convolutional operations to capture local and global dependencies [6]. Due to its capacity to efficiently capture relational information of the topology of graph, GCN has shown great potential in a variety of applications [7] [8]. GAT, on the other hand, adds attention mechanisms to GNNs, allowing nodes to selectively attend to informative neighbors throughout the aggregation process [9]. However, for constructing a high performing GCN or GAT architectures, determining the optimal number of layers play a crucial role in the classification process. An excessive number of layers can cause overfitting, where the model becomes too narrowly focused on the training data and struggles to generalize well to new examples [10]. On the other hand, using too few layers may lead to underfitting, which can reduce expressive power and the ability to recognize complex patterns in the data.

In this paper, we propose a mathematical formula for determining the optimal number of layers in GCN and GAT architectures. We introduce the formula by observing the number of nodes and edges in a specific graph network. For conducting the experiment, we conduct experiment over ten benchmark datasets. By determining the optimal number of layers, we achieve a trade-off between model complexity and generalization. Our approximation ensures that the network can learn informative node representations and captures the underlying structure of the graph while avoiding over parameterization. However, in summary, the contributions of this paper are as follows:

- We propose a mathematical formula for determining the optimal number of GAT and GCN hidden layers.
- We compare the performance of classification problem between GCN and GAT architectures.

The rest of the paper is organized as follows. Section II presents studies related to node classifications and finding optimal number of layers for neural networks. Section III describes our approach followed by our findings in Section IV. The study concludes in Section V.