Chapter 7
Graph Convolutional Neural Networks for Link Prediction in Social Networks

Nimish Kumar
B.K. Birla Institute of Engineering and Technology, India

Himanshu Verma
B.K. Birla Institute of Engineering and Technology, India

Yogesh Kumar Sharma
Koneru Lakshmaiah Education Foundation, India

ABSTRACT

Social networks are complex systems that require specialized techniques to analyze and understand their structure and dynamics. One important task in social network analysis is link prediction, which involves predicting the likelihood of a new link forming between two nodes in the network. Graph convolutional neural networks (GCNNs) have recently emerged as a powerful approach for link prediction, leveraging the graph structure and node features to learn effective representations and predict links between nodes. This chapter provides an overview of recent advances in GCNNs for link prediction in social networks, including various GCNN architectures, feature engineering techniques, and evaluation metrics. It discusses the challenges and opportunities in applying GCNNs to social network analysis, such as dealing with sparsity and heterogeneity in the data and leveraging multi-modal and temporal information. Moreover, this also provides reviews of several applications of GCNNs for link prediction in social networks.

DOI: 10.4018/978-1-6684-6903-3.ch007

INTRODUCTION

Graph Convolutional Neural Networks (GCNs) are a specialized type of neural network designed for processing graph-structured data. They have become increasingly popular in recent years due to their ability to learn and represent nodes and edges within graphs. This technology has been applied to numerous graph-related tasks, such as link prediction in social networks.

Social networks are a vital part of modern society, serving as a medium for individuals and organizations to connect and interact with each other. Predicting links between nodes in a social network is a complex problem that involves determining the likelihood of connections between nodes. Accurately predicting these links is critical for various applications, such as recommender systems, social network analysis, and community detection. In social networks, nodes are representations of individuals or entities, and edges depict connections or interactions between them. Graphs typically represent social networks, where nodes represent individuals or entities, and edges represent connections or interactions between them. By utilizing GCNs, one can learn representations of nodes and edges within these graphs, thus being able to predict the likelihood of connections between them. Studies conducted by Kipf and Welling (2016) and Schlichtkrull et al. (2018) have demonstrated the effectiveness of GCNs in link prediction tasks within social networks. The former study used GCNs for link prediction in the Cora citation network, a citation network of scientific papers, and showed that GCNs outperformed traditional methods such as logistic regression and neural networks. Similarly, in the latter study, GCNs were used for link prediction in large-scale knowledge graphs, achieving state-of-the-art performance on the task. One of the significant benefits of GCNs is their ability to learn node and edge representations within graphs by applying convolutional operations to the graph structure. This enables the network to learn both local and global patterns within the graph. Traditional convolutional neural networks (CNNs) use convolutional operations on regular grid-structured image data. In contrast, graph data is irregular and features nodes with varying numbers of neighbors. Hence, GCNs use graph convolutions to operate on the adjacency matrix of the graph, allowing them to learn node representations.

Graph Convolutional Networks (GCNs) use a convolutional operation on the graph structure to propagate information from neighboring nodes to a central node. In the first layer of a GCN, a linear transformation is used to transform the input features of each node into a low-dimensional representation. Next, a graph convolution operation is performed to update the representation of each node by combining the representations of its neighbors. The aggregation is done using a weighted sum of the neighbor representations, and the weights are learned during training. A non-linear activation function, such as ReLU or sigmoid, is applied to the output of the convolutional operation to introduce non-linearity. GCNs can have multiple layers, allowing them to learn increasingly complex representations of the graph structure. Regularization techniques such as dropout and L2 regularization can be applied to prevent overfitting.

Various types of GCNs have been proposed in the literature. One such variation is the Graph Attention Network (GAT) which uses attention mechanisms to assign different weights to the neighbors of a node based on their relevance. Another variation is the GraphSAGE, which employs a sampling-based approach to aggregate node representations, enabling it to handle larger graphs. GCNs have shown promise in link prediction tasks for social networks, as they can learn node and edge representations, making it easier to predict the existence of links between two nodes in a network. Multiple layers can be used to capture increasingly complex features of the graph, while regularization techniques can help prevent overfitting.