Graph Neural Networks: Models and Advances

Yiqi Wang, Wei Jin, Yao Ma, Jiliang Tang
Michigan State University

{wangy206,jinwei2, mayao4,tangjili}@msu.edu

Abstract

Graphs are essential representations for relational data where data points (nodes) are related to each other with links. To perform downstream machine learning tasks on graph structured data, it is important to design advanced algorithms to extract appropriate features from them. Because of the strong representation learning ability, deep neural networks have been shown to advance many research areas involving regular grid data such as image classification, speech recognition, and natural language processing. Adapting deep neural networks model for representation learning on graph structured data has been attracting increasing attention. In this tutorial, we provide a comprehensive overview of recent advances in graph neural networks including both models and advances. In particular, we introduce some basic concepts, review four main streams of graph neural networks, discuss the robustness of the graph neural networks, and illustrate a variety of recent advances. Finally, we summarize the tutorial with discussions on open issues and challenges about graph neural networks.

1 Tutorial Outline.

There are two major streams of Graph Neural Networks research based on their designs. One stream of the GNNs, which is built on the spatial domain, involves transforming, propagating and aggregating node features over the graph through links. The other stream of works is based on defining graph convolutional operations in the spectral domain through the graph Fourier analysis. On top of these GNNs designs, pooling schemes are necessary to be designed to learn graph level representations. One big challenge of training GNNs is the scalability issue on large graphs and thus various methods have been proposed to address this issue. GNNs have also encouraged new development of encoder-decoder models for taking graph-structured inputs and outputting sequence or structured objects. Meanwhile, various adversarial attack strategies have been designed to investigate the robustness of the GNNs. We discuss all the aforementioned contents in the model part of the tutorial. To smoothly introduce these models, we provide some basic concepts and foundations on graph theory and graph Fourier analysis prior to them. We also introduce some advanced topics on GNNs. The foundations, models and advanced topics together with introduction and summary form the 5 segments of this tutorial.

2 Target Audience and Prerequisites.

The intended audience for this tutorial mainly includes researchers, graduate students, and professionals who are new to this area or who already have some experience with data mining and machine learning. The audience is expected to have some basic understanding of data mining, machine learning, linear algebra, and optimization. However, the tutorial will be presented at college junior/senior level and should be comfortably followed by academic researchers and practitioners from the industry.

3 Contents.

The tutorial mainly consists of 5 segments: introduction, foundations, models, advanced topics and summary.

3.1 Introduction Graph structured data are ubiquitous in the real world. Here, we briefly illustrate graph structured data with some real-world examples. Mining graph structured data has attracted much attention from various domains including chemistry, natural language processing and social networks.We briefly review various tasks and. To handle the aforementioned tasks on graph structured data, it is essential to learn appropriate representations. GNNs have been an emerging popular tool to deal with graph structured data. In this part, we briefly describe the GNNs in a general sense and introduce some history and background of the development of the research on GNNs [1, 2, 3, 4, 5].

3.2 Foundations We first introduce basic concepts of graphs including definitions and notations of graph structured data. We will introduce the matrix representations of graphs, the attributed graph, and how to regard the attributes as functions or signals [6] defined on the graph. Also, we will present the concepts of graph Fourier analysis, which lays the foundation for the spectral based GNNs design. Then, we further introduce the convolution operation in the spectral domain, which is the fundamental building component of the spectral based GNNs.

3.3 Models In this part, we describe Graph Neural Network models, including the spectral based GNN layers, spatial based GNN layers, pooling schemes for graph level representation learning, the scalabilty of the GNNs, GNN based encoder-decoder, the robust analysis of GNNs and self-supervised learning for GNNs .

3.4 Advances In previous parts, we have discussed the most established methods of deep learning on graphs. In this part, we will introduce some limitations identified for existing GNNs including potential discrimination and expressive power. Also we will include some advanced strategies that have been adapted GNNs from traditional DNNs. We package these recent efforts into this part about advanced topics in GNNs with two goals. First, we aim to bring our audience near the frontier of current research on GNNs. Second, these topics can serve as promising future research directions.

3.5 Summary We summarize the tutorial with discussions on basic concepts, fundamental principles, open issues, and future directions about GNNs.

4 Tutor Information

Yiqi Wang is a Ph.D. student at Michigan State University. She is working on graph neural networks including fundamental algorithms and applications. She has published her research in top conference proceedings (e.g., WSDM and WWW). She is one of the key tutors for the tutorial at KDD'20.

Wei Jin is a CSE Ph.D. student at Michigan State University. He works on the area of graph neural networks including the theory foundations, model robustness and applications. He has published his research in top conference proceedings (e.g., KDD and WWW). He is one of the key tutors for the tutorials at AAAI'20 and KDD'20.

Yao Ma is a PhD student of the Department of Computer Science and Engineering at Michigan State University (MSU). He is the recipient of the Outstanding Graduate Student Award and FAST Fellowship at MSU. He is the first author of the book "Deep Learning on Graphs". He has published papers in top conferences such as WSDM, ICDM, SDM, WWW, IJCAI, SIGIR and KDD, which have been cited hundreds of times. He is the leading organizer and presenter of tutorials on GNNs at AAAI'20 and KDD'20, which received huge attention and wide acclaim. He has served as Program Committee Members/Reviewers in many well-known conferences and magazines such as AAAI, IJCAI, TWEB, TKDD and TPAMI.

Jiliang Tang is Assistant Professor in the Department of Computer Science and Engineering at Michigan State University. Previously, he was a research scientist in Yahoo Research. He received the 2020 SIGKDD Rising Star Award, 2020 Distinguished Withrow Research Award, 2019 NSF Career Award, the 2019 IJCAI Early Career Invited Talk and 7 best paper (runnerup) awards. He has organized top data science conferences including KDD, WSDM and SDM, and is associate editor of the TKDD journal. His research has been published in highly ranked journals and top conferences, and received more than 12,000 citations with h-index 55 and extensive media coverage.

References

[1] Scarselli, Franco and Gori, Marco and Tsoi, Ah Chung and Hagenbuchner, Markus and Monfardini, Gabriele. *The graph neural network model*,IEEE Transactions on Neural Networks, 2008.

[2] Bruna, Joan and Zaremba, Wojciech and Szlam, Arthur and LeCun, Yann. *Spectral networks and locally connected networks on graphs*, arXiv preprint arXiv:1312.6203.

[3] Li, Yujia and Tarlow, Daniel and Brockschmidt, Marc and Zemel, Richard. *Gated graph sequence neural networks*, arXiv preprint arXiv:1511.05493.

[4] Defferrard, Michaël and Bresson, Xavier and Vandergheynst, Pierre. *Convolutional neural networks on graphs with fast localized spectral filtering*, Advances in neural information processing systems, 2016.

[5] Gilmer, Justin and Schoenholz, Samuel S and Riley, Patrick F and Vinyals, Oriol and Dahl, George E. *Neural message passing for quantum chemistry*,Neural message passing for quantum chemistry, 2017.

[6] Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. and Vandergheynst, P. *The emerging field of signal pro-cessing on graphs: Extending high-dimensional data analy-sis to networks and other irregular domains*,arXiv preprintarXiv:1211.0053.