Graph Minimally-supervised Learning

Kaize Ding  
Arizona State University  
kaise.ding@asu.edu

Nitesh Chawla  
University of Notre Dame  
nchawla@nd.edu

Jundong Li  
University of Virginia  
jundong@virginia.edu

Huan Liu  
Arizona State University  
huan.liu@asu.edu

ABSTRACT
To model graph-structured data, graph learning, in particular deep graph learning with graph neural networks, has drawn much attention in both academic and industrial communities lately. Prevaling graph learning methods usually rely on learning from "big" data, requiring a large amount of labeled data for model training. However, it is common that graphs are associated with "small" labeled data as data annotation and labeling on graphs is always time and resource-consuming. Therefore, it is imperative to investigate graph learning with minimal human supervision for the low-resource settings where limited or even no labeled data is available. In this tutorial, we will focus on the state-of-the-art techniques of Graph Minimally-supervised Learning, in particular a series of weakly-supervised learning, few-shot learning, and self-supervised learning methods on graph-structured data as well as their real-world applications. The objectives of this tutorial are to: (1) formally categorize the problems in graph minimally-supervised learning and discuss the challenges under different learning scenarios; (2) comprehensively review the existing and recent advances of graph minimally-supervised learning; and (3) elucidate open questions and future research directions. This tutorial introduces major topics within minimally-supervised learning and offers a guide to a new frontier of graph learning. We believe this tutorial is beneficial to researchers and practitioners, allowing them to collaborate on graph learning.

CCS CONCEPTS
• Information systems → Data mining; • Computing methodologies → Machine learning; Neural networks.

KEYWORDS
Graph Neural Networks; Weakly-supervised Learning; Few-shot Learning; Self-supervised Learning

1 INTENDED AUDIENCE
This tutorial mainly focuses on the problems and techniques of minimally-supervised learning (i.e., weakly-supervised learning, few-shot learning, and self-supervised learning) for graph-structured data. We expect the audience to have the general background knowledge of machine learning, graph mining, and graph neural networks. The intended audiences are students, researchers, and practitioners who are new to this topic or are interested in this topic. The tutorial will be presented as graduate-level lecture. Basic knowledge on deep learning, graph mining, and machine learning is preferred but not required. We will broadcast our tutorial information through website and social media.

2 PRESENTER BIOGRAPHY
The presenters and contributors include Kaize Ding (main contact), Jundong Li, Nitesh Chawla and Huan Liu. Their short biographies are as follows.

Kaize Ding  
is currently a Ph.D. student at Arizona State University. He received his master’s and bachelor’s degrees in Computer Science at Beijing University of Posts and Telecommunications in 2017 and 2014. His research interests are broadly in data mining and machine learning, with a particular focus on graph neural networks, few-shot learning, and self-supervised learning. He has published over 20 papers on top conferences and journals such as WWW, WSDM, EMNLP, AAAI, and TNNLS. More details can be found at http://www.public.asu.edu/~kding9/.

Jundong Li  
is an Assistant Professor in the Department of Electrical and Computer Engineering at University of Virginia. His research interests are in data mining and machine learning, with a particular focus on graph mining and causality learning. His work on feature selection and graph representation learning are among the most cited articles in ACM CSUR, WSDM, SDM, and CIKM within the past five years according to Google Scholar Metrics. He was selected for the AAAI 2021 New Faculty Highlights program. More details can be found at: http://www.ece.virginia.edu/~jl6qk/.

Nitesh V. Chawla  
is the Frank M. Freimann Professor of Computer Science and Engineering at the University of Notre Dame. His research is making fundamental advances in artificial intelligence, data science, and network science, and is motivated by the question of how technology can advance the common good through interdisciplinary research. He is the recipient of National Academy of Engineers New Faculty Fellowship. He also is the recipient of the...
In this tutorial, we aim to provide a comprehensive review on this emerging and important research topic: Graph Minimally-supervised Learning. We will first introduce the theoretical foundations on graph learning with a special focus on graph neural networks. We then discuss the three fundamental problems as well as the key literature of graph minimally-supervised learning in recent years. Aligned with the major graph mining tasks (e.g., node classification, link prediction, and graph classification) in different granularity levels, we introduce the techniques in each category for node-level, edge-level and graph-level tasks. Finally, we will introduce the applications of graph minimally-supervised learning in different domains, and also discuss the future directions in this research filed. In the graph learning community, we believe that graph minimally-supervised learning is an edge-cutting research topic with important social impacts, which will attract both researchers and practitioners from academia and industry.

4 TUTORIAL OUTLINE

The outline and schedule of this half-day tutorial (3 hours, 30 minutes, plus 30-minute break) are illustrated in Table 1. More detailed outline is as follows:

- **Introduction and Overview.** At first, we will briefly introduce deep graph learning and its challenges under the low-resource setting, then we will cover some basic knowledge about the related topics such as graph neural networks [30], meta-learning [13], and contrastive learning [2], and introduce the overview of graph minimally-supervised learning.

- **Graph Weakly-supervised Learning.** We will present methodologies and applications of graph learning with weak supervision. We will mainly focus on three types of weak supervisions,
i.e., incomplete supervision, indirect supervision, and inaccurate supervision. A series of learning techniques such as graph self-training [8, 18], graph active learning [5, 6, 14], and graph transfer learning [7, 11, 31] will be reviewed in this part.

- **Graph Few-shot Learning.** We will present methodologies and applications of graph few-shot learning. Specifically, we will cover two categories of approaches: meta gradient based methods [22, 28, 35, 36] and metric learning-based [10–12, 16, 31, 34] to show how to handle never-before-seen nodes, edges, and graphs. In addition, we will also discuss graph zero-shot learning [19, 20, 24].

- **Graph Self-supervised Learning.** We will present methodologies and applications of graph self-supervised learning. Specifically, we will cover three main paradigms, including graph generative modeling [3, 4, 17, 23], graph property prediction [15, 26] and graph contrastive learning [1, 25, 32, 33].

**Conclusion and Discussion.** We will summarize the covered topics and discuss their connections. Also we will discuss the future research directions of this topic.

## 5 ADDITIONAL INFORMATION

**Relevance to the Community.** The research community of web search and data mining cares about how to use machine learning techniques in the real-world problems. In the real world, however, we often face small-data challenges, especially on graph-structured data. The community seeks how to innovate the techniques so that the methods are more applicable in the real world. Many related approaches have been approached while there is still a great room for improvement. Giving this tutorial helps our community understand the state of the art and can better innovate the techniques in this topic as well.

**Related Tutorials.** Several similar tutorials have been delivered at recent conferences. For example, (i) Xiao Huang, Jundong Li, et al., Learning From Networks: Algorithms, Theory, and Applications, at KDD 2019; (ii) Huaxiu Yao et al., Learning with Small Data, at KDD 2020; (iii) Chuxu Zhang, Jundong Li, and Meng Jiang, Data Efficient Learning on Graphs, at KDD 2021. Those related tutorials either focus on the i.i.d. data or the general graph mining, while our tutorial covers the most recent techniques in graph minimally-supervised learning.

**Impacts.** This tutorial will attract a good number of audiences as the topic is hot and fast-developing. Since the tutorial will be held virtually this year, we will record tutorial videos with high quality, with intuitive examples. Besides, in order to optimize the participation experience of the audience and the interactivity of the tutorial, we will encourage the audience to ask questions during and after the tutorial. This tutorial will provide a good resource and inspirations for researchers and practitioners.

**ACKNOWLEDGEMENTS**

This work is partially supported by Office of Naval Research (ONR) N00014-21-1-4002 and Army Research Office (ARO) W911NF2110030.

**REFERENCES**


