

# Deep Learning on Graphs for Natural Language Processing

**Lingfei Wu**  
IBM Research AI  
lwu@email.wm.edu

**Yu Chen**  
Facebook  
hugochen@fb.com

**Heng Ji**  
University of Illinois Urbana-Champaign  
hengji@illinois.edu

**Yunyao Li**  
IBM Research AI  
yunyaoli@us.ibm.com

**Keywords:** Natural Language Processing, Deep Learning, Graph Learning, Graph Neural Networks

**Type of the tutorial:** cutting-edge

**Suggested duration:** 3.5 hours

**Estimated number of participants:** 300

**Estimate of what % of the tutorial covers work by the tutorial presenters vs. work by other researchers:** 40% vs. 60%.

**Open access:** We agree to allow the publication of our slides and video recording of our tutorial in the ACL Anthology. Other teaching material (i.e., data and code) will be openly available.

## 1 Description

Deep learning has become the dominant approach in Natural Language Processing (NLP) research today, especially when applied on large scale corpora. Conventionally, sentences are typically considered as a sequence of tokens in NLP tasks. Hence, popular deep learning techniques such as recurrent neural networks (RNN) and convolutional neural networks (CNN) have been widely applied for modeling text sequence.

However, there is a rich variety of NLP problems that can be best expressed with a graph structure. For instance, the structural and semantic information in sequence data (e.g., various syntactic parsing trees like dependency parsing trees and semantic parsing graphs like abstract meaning representation (AMR) graphs) can be exploited to augment original sequence data by incorporating the task-specific knowledge. As a result, these graph-structured data can encode complicated pairwise relationships between entity tokens for learning more informative representations. However, it is well-known that deep learning techniques that were disruptive for Euclidean data such as images or sequence data such as text are not immediately applicable to graph-structured data. Therefore, this gap

has driven a tide in research for deep learning on graphs, especially in development of graph neural networks (GNN).

This wave of research at the intersection of deep learning on graphs and NLP has influenced a variety of NLP tasks. There has seen a surge of interests in applying/developing various types of GNNs and achieved considerable success in many NLP tasks, ranging from classification tasks like sentence classification, semantic role labeling and relation extraction, to generation tasks like machine translation, question generation and summarization. Despite these successes, deep learning on graphs for NLP still face many challenges, namely,

- Automatically transforming original text sequence data into highly graph-structured data. Such challenges are profound in NLP since most of the NLP tasks use the text sequence as the original inputs. Automatic graph construction from the text sequence to take into account underlying structural information is a critical step in the use of graph neural network models for NLP problems.
- Effectively modeling complex data that involves mapping between graph-based inputs and other highly structured output data such as sequences, trees, and graph data with multi-types in both nodes and edges. Many generation tasks in NLP such as SQL-to-Text, Text-to-AMR, Text-to-KB are emblematic of such challenges.

This tutorial of **Deep Learning on Graphs for Natural Language Processing (DLG4NLP)** is timely for the computational linguistics community, and covers relevant and interesting topics, including automatic graph construction for NLP, graph representation learning for NLP, various advanced GNN based models (e.g., graph2seq, graph2tree,

and graph2graph) for NLP, and the applications of GNNs in various NLP tasks (e.g., machine translation, natural language generation, information extraction and semantic parsing). The intended audiences for this tutorial mainly include graduate students and researchers in the field of Natural Language Processing and industry professionals who want to know how the state-of-the-art deep learning on graphs techniques can help solve important yet challenging Natural Language Processing problems.

In addition, hands-on demonstration sessions will be included to help the audience gain practical experience on applying GNNs to solve challenging NLP problems using our recently developed open source library – Graph4NLP, the first library for researchers and practitioners for easy use of graph neural networks for various NLP tasks. After attending the tutorial, the audience are expected to 1) have a comprehensive understanding of basic concepts of deep learning on graphs for NLP; 2) learn major recent advances of research in the intersection of NLP and GNNs; and 3) explore novel research opportunities of GNNs for NLP, and learn how to use or even design novel algorithms with GNNs for effectively coping with various NLP tasks.

We will start with a broad overview of various NLP problems that deal with graph structured data, and highlight some challenges of modeling graph-structured data in the field of NLP with traditional graph-based algorithms (e.g., random walk methods, spectral graph clustering, graph kernels). We will then introduce the general idea as well as some commonly used models of GNNs, which have been an emerging popular tool to deal with graph structured data. After the introduction of NLP tasks on graph data and graph neural networks, we will describe some important yet challenging techniques for deep learning on graphs for NLP, including automatic graph construction for NLP, graph representation learning for NLP and various advanced GNN based models (e.g., graph2seq, graph2tree, and graph2graph) for NLP. Some representative NLP applications are introduced following the methods. When introducing the NLP applications, we also include hands-on demonstration sessions on how to quickly build GNN-based models for solving NLP tasks using our recently developed open source library Graph4NLP, which was designed for the easy use of GNNs for NLP. We will summarize the tuto-

rial and highlight some open directions in the end of this tutorial. The Introduction, Methodologies, Applications, and Conclusion and Open Directions form the four segments of this tutorial.

## 2 Prerequisites

The audience is expected to have some basic understanding of natural language processing and deep learning. However, the tutorial will be presented at college junior/senior level and should be comfortably followed by academic researchers and industrial practitioners.

## 3 Tutorial Outline

The intended duration of this tutorial is 3.5 hours, including a half an hour break.

### I. (30 minutes) Introduction

1. Natural Language Processing: A Graph Perspective
2. Graph Based Algorithms for Natural Language Processing
3. Deep Learning on Graphs: Graph Neural Networks
  - i. Foundations
  - ii. Methodologies
  - iii. Applications in Natural Language Processing: An Overview

### II. (70 minutes) Methodologies

1. Automatic Graph Construction
  - i. Static Graph Construction
  - ii. Dynamic Graph Construction
2. Graph Representation Learning
  - i. Graph Neural Networks for Improved Text Representation
  - ii. Graph Neural Networks for Joint Text & Knowledge Representation
  - iii. Graph Neural Networks for Various Graph Types
3. GNN Based Encoder-Decoder Models
  - i. Graph-to-Sequence Models
  - ii. Graph-to-Tree Models

### III. (70 minutes) Applications

1. Machine Translation
  - i. Methodologies
  - ii. Hands-on Demonstration
2. Natural Language Generation

- i. Methodologies
- ii. Hands-on Demonstration
- 3. Machine Reading Comprehension
  - i. Methodologies
  - ii. Hands-on Demonstration
- 4. Information Extraction
  - i. Methodologies
  - ii. Hands-on Demonstration
- 5. Semantic Parsing
  - i. Methodologies
  - ii. Hands-on Demonstration

IV. (10 minutes) Conclusion and Open Directions

## 4 Reading List

We aim to make the tutorial self-contained. For trainees interested in reading important studies before the tutorial, we recommend the following papers regarding GNNs (Kipf and Welling, 2016; Li et al., 2015; Hamilton et al., 2017), automatic graph construction for NLP (Bastings et al., 2017; Chen et al., 2020b,a), joint text and knowledge representation learning (Feng et al., 2020; Lin et al., 2020), modeling directed graphs (Xu et al., 2018; Chen et al., 2020b) and heterogeneous graphs (Bastings et al., 2017; Chen et al., 2020c), and GNN based encoder-decoder models (Xu et al., 2018; Chen et al., 2020b; Li et al., 2020).

## 5 Diversity

Beside the state-of-the-art deep learning on graphs techniques we are planning to cover, we will discuss how these graph-based Deep Learning techniques can be used in several NLP applications that exploit multilingual data, including but not limited to Machine Translation and Information Extraction. Our tutorial lectures are full of diversities from many perspectives. Our team have male and female researchers (two female tutors and two male tutors); We are from three different institutions including IBM Research, UIUC, and Facebook; We are senior, middle-career, and junior-career researchers. We are researchers and professors from academic institutions and industrial labs.

## 6 Presenters

**Lingfei Wu** earned his Ph.D. degree in computer science from the College of William and Mary in 2016. He is a research staff member at IBM Research and is leading a research team (10+ RSMs)

for developing novel Graph Neural Networks for various tasks, which leads to the #1 AI Challenge Project in IBM Research and multiple IBM Awards including Outstanding Technical Achievement Award. He has published more than 70 top-ranked conference and journal papers and is a co-inventor of more than 30 filed US patents. Lingfei Wu has given many tutorials/keynote presentations in deep learning on graphs for natural language processing in multiple workshops in KDD'20, CVPR'20, AAAI'20, and Machine Learning & Artificial Intelligence'20.

Email: [lwu@email.wm.edu](mailto:lwu@email.wm.edu)

Homepage: <https://sites.google.com/a/email.wm.edu/teddy-lfwu/>.

**Yu Chen** is a Research Scientist at Facebook. He got his PhD degree in Computer Science from Rensselaer Polytechnic Institute. His research interests lie at the intersection of Machine Learning (Deep Learning), and Natural Language Processing, with a particular emphasis on the fast-growing field of Graph Neural Networks and their applications in various domains. His work has been published in top-ranked conferences including but not limited to NeurIPS, ICML, ICLR, AAAI, IJCAI, NAACL, KDD, WSDM, ISWC, and AMIA. He was the recipient of the Best Student Paper Award of AAAI DLGMA'20. He has served as PC members in many conferences (e.g., ACL, EMNLP, NAACL, EACL, AAAI, IJCAI and KDD) and journals (e.g., TNNLS, IJIS, TKDE, TKDD and DAMI).

Email: [hugochen@fb.com](mailto:hugochen@fb.com)

Homepage: <http://academic.hugochan.net>

**Heng Ji** is a professor at the Computer Science Department of University of Illinois at Urbana-Champaign. She has received many awards including "AI's 10 to Watch" Award by IEEE Intelligent Systems in 2013, NSF CAREER award in 2009, PACLIC2012 Best paper runner-up, "Best of ICDM2013" paper award, and "Best of SDM2013" paper award. She has served as the Program Committee Co-Chair of NAACL-HLT2018, NLP-NABD2018, NLPCC2015, CSCKG2016 and CCL2019, and senior area chair for many conferences. She has given a large number of keynotes/tutorials presentations in Event Extraction, Natural Language Understanding, and Knowledge Base Construction in many

conferences including but not limited to ACL, EMNLP, NAACL, NeurIPS, AAAI and KDD.

Email: [hengji@illinois.edu](mailto:hengji@illinois.edu)

Homepage: <http://blender.cs.illinois.edu/hengji.html>

**Yunyaoli Li** is a Principal Research Staff Member and Senior Research Manager with IBM Research - Almaden, where she manages the Scalable Knowledge Intelligence Department. She is a Master Inventor and a member of the IBM Academy of Technology. Her expertise is in the interdisciplinary areas of natural language processing, databases, human-computer interaction, and information retrieval. She is interested in designing, developing, and analyzing large-scale systems that are usable by a wide spectrum of users. Her current research focuses on scalable natural language processing. She regularly gives talks and tutorials at conferences and universities across the globe, including a MOOC on information extraction. She received her Ph.D. from the University of Michigan, Ann Arbor.

Email: [yunyaoli@us.ibm.com](mailto:yunyaoli@us.ibm.com)

Homepage: <https://researcher.watson.ibm.com/researcher/view.php?person=us-yunyaoli>

## References

- Joost Bastings, Ivan Titov, Wilker Aziz, Diego Marcheggiani, and Khalil Sima'an. 2017. Graph convolutional encoders for syntax-aware neural machine translation. *arXiv preprint arXiv:1704.04675*.
- Yu Chen, Lingfei Wu, and Mohammed J Zaki. 2020a. Iterative deep graph learning for graph neural networks: Better and robust node embeddings. In *Proceedings of the 34th Conference on Neural Information Processing Systems*.
- Yu Chen, Lingfei Wu, and Mohammed J. Zaki. 2020b. Reinforcement learning based graph-to-sequence model for natural question generation. In *Proceedings of the 8th International Conference on Learning Representations*.
- Yu Chen, Lingfei Wu, and Mohammed J Zaki. 2020c. Toward subgraph guided knowledge graph question generation with graph neural networks. *arXiv preprint arXiv:2004.06015*.
- Yanlin Feng, Xinyue Chen, Bill Yuchen Lin, Peifeng Wang, Jun Yan, and Xiang Ren. 2020. Scalable multi-hop relational reasoning for knowledge-aware question answering. *arXiv preprint arXiv:2005.00646*.
- Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pages 1024–1034.
- Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
- Shucheng Li, Lingfei Wu, Shiwei Feng, Fangli Xu, Fengyuan Xu, and Sheng Zhong. 2020. Graph-to-tree neural networks for learning structured input-output translation with applications to semantic parsing and math word problem. *arXiv preprint arXiv:2004.13781*.
- Yujia Li, Daniel Tarlow, Marc Brockschmidt, and Richard Zemel. 2015. Gated graph sequence neural networks. *arXiv preprint arXiv:1511.05493*.
- Ying Lin, Heng Ji, Fei Huang, and Lingfei Wu. 2020. A joint neural model for information extraction with global features. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7999–8009.
- Kun Xu, Lingfei Wu, Zhiguo Wang, Yansong Feng, Michael Witbrock, and Vadim Sheinin. 2018. Graph2seq: Graph to sequence learning with attention-based neural networks. *arXiv preprint arXiv:1804.00823*.