JUNJIE XU

+1(814) 996-8596 ◇ University Park, PA
◇ junjiexu@psu.edu ◇ Homepage ◇ Google Scholar ◇ LinkedIn

SUMMARY

Ph.D. candidate in Machine Learning with 3+ years of R&D experience in the field.

- **Research Interests**: Machine Learning, Graph Learning, Large Language Model, Geometric Deep Learning
- **6 peer-reviewed papers** published or under submission, with **300+ citations** as of 03/2024.
- Active **reviewers of top-tier conferences** and participants in **cutting-edge open-source projects**.

EDUCATION

The Pennsylvania State University, University Park, USA 08/2021 - Present
Ph.D. candidate in Informatics
Advisor: Dr. Suhang Wang & Dr. Xiang Zhang

Huazhong University of Science and Technology, Wuhan, China 09/2017 - 06/2021
B.E. in Software Engineering
GPA: 3.91/4.00

University of California, Berkeley, Berkeley, USA 01/2020 - 06/2020
Exchange student in Computer Science

PEER-REVIEWED PAPERS

Preprints

[1] **Shape-aware Graph Spectral Learning**
Junjie Xu, Enyan Dai, Dongsheng Luo, Xiang Zhang, Suhang Wang
Submitted to ICML 2024

[2] **A Comprehensive Survey on Trustworthy Graph Neural Networks: Privacy, Robustness, Fairness, and Explainability**
Enyan Dai, Tianxiang Zhao, Huaisheng Zhu, **Junjie Xu**, Zhimeng Guo, Hui Liu, Jiliang Tang, Suhang Wang
Submitted to TIST

[3] **Self-Explainable Graph Neural Networks for Link Prediction**
Huaisheng Zhu, Dongsheng Luo, Xianfeng Tang, **Junjie Xu**, Hui Liu, Suhang Wang
Submitted to ICDE 2024

Peer-Reviewed Papers

[4] **HP-GMN: Graph Memory Networks for Heterophilous Graphs**
Junjie Xu, Enyan Dai, Xiang Zhang, Suhang Wang
ICDM 2022, IEEE International Conference on Data Mining

[5] **Revisiting Time Series Outlier Detection: Definitions and Benchmarks**
Kwei-Herng Lai, Daochen Zha, **Junjie Xu**, Yue Zhao, Guanchu Wang, Xia Hu
NIPS 2021, Neural Information Processing Systems, Datasets and Benchmarks Track

[6] **TODS: An Automated Time Series Outlier Detection System**
Kwei-Herng Lai, Daochen Zha, Guanchu Wang, **Junjie Xu**, Yue Zhao, Devesh Kumar, Yile Chen, Purav Zumkhawaka, Mingyang Wan, Diego Martinez, Xia Hu
AAAI 2021, AAAI Conference on Artificial Intelligence, Demo track

EXPERIENCE

Research Assistant, Penn State University, USA 08/2021 - Present
Advisors: Suhang Wang & Xiang Zhang

Multi-modal Knowledge Distillation with LLM for Molecule Property Prediction

- Explored the impact of **multi-modal** inputs, including textual descriptions, diagrams, node features, and graph structures, on molecule property prediction, and assessed the LLM's capability to capture multi-modal information.
- **Fine-tuned** smaller language models to enhance the encoding of Large Language Model outputs, and **augmented the input space** for superior representation learning performance.
- Implemented **knowledge distillation** from Large Language Models by aligning their outputs with smaller models, like GNNs and MLPs, boosting **efficiency** and reducing **costs** during reasoning.

Heterophily-aware GNNs and Trustworthy GNNs

- Developed Graph Memory networks to tackle heterophily in graphs, enhancing prediction with local and global pattern learning, and introduced regularization for improved global information capture, achieving top performance.
- Investigated the impact of filter frequency in **Spectral GNNs**, using theoretical and empirical analyses.
- Introduced shape-aware regularization and a Newton Interpolation-based spectral filter in GNN, customizing filters to match homophily levels and enhancing performance on various datasets.
- Surveyed trustworthy graph neural networks to improve **privacy, robustness, fairness, and explainability**.

Research Assistant, Texas A&M University, USA 05/2020 - 06/2021
Advisor: Xia "Ben" Hu

Automated Time-series Outlier Detection System

- Developed **full stack and automated** machine learning system with preprocessing, feature extraction, detection algorithms, and human-in-the-loop interfaces.
- Integrated a wide range of algorithms including PyOD. Revisited the definition of the time-series anomalies and proposed a taxonomy for point-wise, piece-wise, and pattern-wise anomalies.
- Implemented Automated Machine Learning **(AutoML)** for knowledge-free pipeline construction and automatic optimization of module combinations. Developed graphical user interfaces **(GUI)** to improve usability.
- [Code] (Github with 1.2k+ stars, 100+ forks); [Website]; [Video].

Research Assistant, HUST, China 04/2019 - 10/2019
Subject Analysis of News Texts

SERVICE

Reviewer

The International Conference on Machine Learning (ICML)	2024
The International Conference on Learning Representations (ICLR)	2024
The Learning on Graph Conference (LoG)	2023
Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)	2023
The IEEE International Conference on Data Mining (ICDM)	2023
The ACM Conference on Knowledge Discovery and Data Mining (KDD)	2023
Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS)	2022
The IEEE International Conference on Data Mining (ICDM)	2022
The ACM Conference on Knowledge Discovery and Data Mining (KDD)	2022

The ACM Web Conference (WWW) 2022
The ACM Transactions on Knowledge Discovery from Data (TKDD) 2022

TEACHING EXPERIENCE

Teaching Assistant, PSU Spring 2024
IST 597: Machine Learning on Graphs

Teaching Assistant, PSU Spring 2024
HCDD 364W: Methods for Studying Users

Teaching Assistant, PSU Fall 2023
DS 305: Algorithmic Methods & Tools

HONORS & AWARDS

IST Travel Award IST, PSU, 2022
ICDM Student Travel Award ICDM, 2022
Graham Endowed Fellowship PSU, 2021
Mitacs Globalink Research Scholarship China Scholarship Council, 2020
Scholarship for Academic Excellence HUST, 2018-2019
Scholarship for Academic Excellence HUST, 2017-2018

SKILLS

Languages English (Fluent), Mandarin (Native)
Programming Python, Java, Matlab, C
Deep Learning Pytorch, Pytorch Geometric, DGL, Tensorflow, PyTorch Lightning