

Ziyang Yu

✉ ericyu8857@gmail.com

☎ [+1 548-577-0264](tel:+15485770264)

🌐 github.com/Ziyang-Yu

Education Background

University of Waterloo

M.Eng, Electrical and Computer Engineering

GPA: 89.86/100.00 | Coursework: Deep Learning, Computer Network

Ontario, Canada

Sep. 2023-Present

Southern University of Science and Technology (SUSTech)

B.Sc. (Hons), Mathematics and Applied Mathematics

GPA: 3.75/4.00 | Coursework: Convex Optimization, Linear Algebra

Guangdong, China

Aug. 2019-Jun. 2023

Research Experience

Distributed Co-training of LLM and GNN

Research Assistant | Advisor: [Liang Zhao](#)

Emory University

Apr. 2024- Present

- **Algorithm Implementation:** Used Huggingface and Pytorch to implement synchronous training of BERT/LLaMA-7b and GraphSAGE, significantly improving the accuracy of paper classification tasks.
- **Model Optimization:** Applied DeepSpeed technology for pipeline parallelism, accelerated model training and inference, and used LoRA and Offloading techniques to reduce training parameters and increase batch size.
- **Experimental Design:** Conducted experiments using NVIDIA A100 80GB, achieving 74.06% accuracy on the OGB-Arxiv dataset and 78.19% on the Cora dataset.

Systematic Survey of Resource-Efficient Large Language Models

Research Assistant | Advisor: [Liang Zhao](#)

Emory University

Sept. 2023-Oct. 2023

- **Paper Research:** Investigated low-rank decomposition techniques in large language models, conducted mathematical theoretical analysis, and wrote research papers.

Staleness-Alleviated Distributed Graph Neural Network Training

Research Assistant | Advisor: [Liang Zhao](#)

Emory University

Jul. 2022-Jan. 2023

- **Algorithm Implementation:** Used PyTorch to implement distributed training for Graph Neural Networks, significantly accelerating training on large-scale graph data, and utilized Plasma technology to speed up reading and updating.
- **Algorithm Optimization:** Proposed using an LSTM-GCN model to capture the temporal evolution of node embeddings, reducing data staleness issues, and employing pre-training techniques to improve model transferability across different datasets, significantly reducing training time.
- **Experiment Verification:** The proposed framework achieved an F1 score of 97.02% on the Reddit dataset and 80.21% on OGB-Products, significantly enhancing the performance and convergence speed of distributed GNN training.

Publications

Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models

Guangji Bai, Zheng Chai, Chen Ling, Shiyu Wang, Jiaying Lu, Nan Zhang, Tingwei Shi, **Ziyang Yu**, Mengdan Zhu, Yifei Zhang, Carl Yang, Yue Cheng, Liang Zhao.

Staleness-Alleviated Distributed Graph Neural Network Training via Online Dynamic-Embedding Prediction

Guangji Bai*, **Ziyang Yu***, Zheng Chai, Yue Cheng, Liang Zhao.

Efficient Imputation for Patch-based Tabular Missing Data via Cluster-regularized Optimal Transport

Yuyu Liu, Jiannan Yang, **Ziyang Yu**, Weishen Pan, Fei Wang, Tengfei Ma

Professional Experience

Machine Learning Engineer Intern

Beijing, Onsite

Baidu Inc.

Jan. 2024-Mar. 2024

- **Data Collection:** Collected and annotated ten days of lane change data on Beijing's Sixth Ring Road, creating a high-quality dataset.
- **Algorithm Implementation:** Conducted in-depth research and analysis of BDS (Beidou Navigation Satellite System) and GPS positioning services, using RandomForest and XGBoost to extract residual information from time series data, significantly improving the accuracy of lane change recognition.
- **Experimental Validation:** Integrated state-of-the-art lane line deep learning algorithms to simplify real-time lane change detection and accelerate data annotation processes, achieving a precision rate of 80% for Baidu's lane recognition.

Skills

- **Computer Languages:** C/C++, Python, Assembly Language, Java, JavaScript, Bash, Matlab.
- **Machine Learning Libraries:** NumPy, Sci-kit Learn, HanLP, Pandas, Matplotlib, Pytorch, Pytorch Geometric.
- **Mathematics:** Algebra, Analysis, Geometry, Topology.