# Zhaoning Yu

202-807-7892 | zn.yu1029@gmail.com | linkedin.com/in/zhaoningyu | zhaoningyu1996.github.io

## Education

**Iowa State University** — Ames, IA
*Ph.D., Department of Computer Science* — *Jan. 2021 – May 2026 (Expect)*

**The George Washington University** — Washington, DC
*M.S., Department of Computer Science* — *Aug. 2018 – May 2020*

**Wuhan University** — Wuhan, Hubei
*B.E., School of Computer Science* — *Aug. 2014 – May 2018*

## Experience

**Ph.D. Software Engineer Intern**, **Google**, Sunnyvale, CA — *Aug. 2025 - Nov. 2025*

- Create a pipeline to evaluate slide animation **Agent** based on **Gemini**
- Create a high-quality animation sync dataset for **Automatic Prompt Optimization**

**Ph.D. Software Engineer Intern**, **Meta**, Bellevue, WA — *May 2025 - Aug. 2025*

- Evaluated **RL post-training** of **LLMs** (Qwen2.5, Qwen3, Llama3) using gold-labeled datasets
- Proposed a new **RL** algorithm that can make **LLMs** robustly self-train without explicit label
- Used **verl** and **Pytorch** to implement the algorithm, trained models on **AWS cluster** with **S3** storage
- Outperformed baselines on **Pass@1** accuracy cross nine benchmarks by average **8.8%**

**Ph.D. Machine Learning Intern**, **Genentech**, South San Francisco, CA — *May 2024 - Aug. 2024*

- Used **Pytorch Lightning** to build a multi-task **Graph Neural Network model** to learn ADME properties
- Studied how different graph pooling methods and loss functions affect the efficiency and effectiveness
- Compared extremely randomized trees with MTNN model to estimate model's uncertainty and prediction accuracy
- Enhanced prediction accuracy by **19%** and **tripled** the training and inference efficiency

## Selected Projects

**Large Language Model on Molecule Graph Alignment** — *Jan. 2024 – Current*

- Proposed a graph to tree text encoding method to help molecule generation using LLM
- Fine-tuned a Llama3.1-8B model using **TorchTune** and **LoRA** with encoded dataset to generate Molecules
- Achieved competitive performance with most SOTA molecule generative model

**Motif-based Graph Neural Networks Explainer** — *Oct. 2021 – Dec. 2023*

- Proposed a **model-level** and a **instance-level** explanation methods to explain the Graph Neural Networks
- Used **attention mechanism** to select motifs and generated explanations by **variational autoencoder**
- Achieved **100% validity** for molecule datasets and **5.1-19.0%** improvements on five real-world datasets

**Molecular Representation Learning** — *Mar. 2021 – Sept. 2023*

- Proposed and implemented a **motif**-based **heterogeneous** graph to help learn a molecule representation
- Outperformed other molecular representation learning baselines by up to **10.6%** on five TUDatasets
- Outperformed other motif extraction method by **0.44-4.10%** on six MoleculeNet datasets

## Selected Publications [Google Scholar]

**MAGE: Model-Level Graph Neural Networks Explanations via Motif-based Graph Generation**
**Zhaoning Yu**, Hongyang Gao – Published at International Conference on Learning Representations (**ICLR**) 2025

**Molecular Representation Learning via Heterogeneous Motif Graph Neural Networks**
**Zhaoning Yu**, Hongyang Gao – Published at International Conference on Machine Learning (**ICML**) 2022

**G2T-LLM: Graph-to-Tree Text Encoding for Molecule Generation with Fine-Tuned LLM**
**Zhaoning Yu**, Xiangyang Xu, Hongyang Gao – Preprint