ZHIYU NI

Email: zhiyuni@berkeley.edu Tel: (+1)323-633-1447

EDUCATION

University of California, Berkeley (UC Berkeley) *2025.1 - Present*
PhD student, *Computer Science*, Advisor: Pierluigi Nuzzo

University of Southern California (USC) *Los Angeles, CA*
PhD student, *Computer Engineering* *2022.8 - 2024.12*

University of Science and Technology of China (USTC) *Hefei, China*
B.S. *Physics*, Outstanding Graduates *2018.8 - 2022.6*

PUBLICATIONS

Analyzing Adversarial Vulnerabilities of Graph Lottery Tickets (ICASSP 2024 **Oral**)
Zhiyu Ni*, Subhajit Dutta Chowdhury*, Qingyuan Peng, Souvik Kundu, Pierluigi Nuzzo

Finding Adversarially Robust Graph Lottery Tickets (TMLR)
Zhiyu Ni*, Subhajit Dutta Chowdhury*, Qingyuan Peng, Souvik Kundu, Pierluigi Nuzzo

RESEARCH EXPERIENCE

Adversarially Robust and Light-weighted GNNs (github)

- Analyzed the robustness of pruned graph neural networks (GNNs) against adversarial attacks.
- Developed self-training techniques and a loss function that improved sparse models' robustness, achieving state-of-the-art (SOTA) robust GNNs with a 90% reduction in computational cost.

LLMs for Anomaly Detection (github)

- Investigated capabilities of LLMs (e.g., ChatGPT, Llama) in anomaly detection and designed in-context learning flows.
- Achieved SOTA precision on GPT-3.5-Turbo compared with GNN-based methods.

Defending against Prompt Injection on LLMs

- Examined prompt injection attack mechanics which mislead LLMs to generate unwanted answers.
- Generated positive and negative CoT-based prompt pairs to align LLMs to defend against the attack.

WORK EXPERIENCE

NLP Intern (Iflytek *AI Research Institute*)

- Independently carried out machine translation tasks from English to German and Italian.
- Enhanced translation models to achieve a 15% higher BLEU score compared to Google Translate.

SKILLS

Python, C, PyTorch, Linux