

YIKUN WANG

☎ +86 19121764250 | @ g1547246193@gmail.com 🐙 GitHub | 📁 Portfolio

EDUCATION

Fudan University

Ph.D. Candidate in Artificial Intelligence;

Shanghai, China

Aug 2024 – Present

Fudan University

B.Sc. in Computer Science;

Reclared major from Electronic Science and Engineering (ESE) to Computer Science (CS);

Shanghai, China

Sept 2019 – Jun 2024

April 2021

- **Relevant coursework:** Program Design (A), Introduction of Computer System (I) (A) Machine Learning (A), Natural Language Processing (A), Computer Vision (A), Computer Network (A)
- Cumulative GPA 3.52
- **During 2021-2023 GPA 3.75/4.0**, ranked 8/102 out of the CS major.
- **2023 Xiaomi Scholarship** (Top ≈5%)

PUBLICATIONS & TECHNICAL REPORTS

RRescue: Ranking LLM Responses to Enhance Reasoning Over Context

[\[ACL 2024 SRW\]](#) Yikun Wang, Rui Zheng, Haoming Li, Qi Zhang, Tao Gui, Fei Liu

Uncertainty Aware Learning for Language Model Alignment

[\[ACL 2024 Main\]](#) Yikun Wang*, Rui Zheng*, Liang Ding, Qi Zhang, Dahua Lin, Dacheng Tao

LLM-DA: Data Augmentation via Large Language Models for Few-Shot Named Entity Recognition

[\[ArXiv preprint\]](#) Junjie Ye, Nuo Xu, Yikun Wang, Jie Zhou, Qi Zhang, Tao Gui, Bingning Wang, Xuanjing Huang

RESEARCH EXPERIENCE

Shanghai Collaborative Innovation Center on Intelligent Visual Computing

Undergraduate Research Assistant

Shanghai, China

Jun 2021 – Mar 2022

- Engaged in advanced research within the Trustworthy Artificial Intelligence group, working on novel transferable attacks against convolution-based neural networks under Dr. Wei Zhipeng and Assoc. Prof. Chen Jingjing's mentorship.
- Devised and proposed an innovative generative method targeting the perturbation of intermediate network layers; conducted a series of rigorous experiments to establish the method's efficacy in compromising neural network integrity.

Natural Language Processing Laboratory, Fudan University

Undergraduate Research Assistant

Shanghai, China

Sept 2022 – Jun 2023

- Conducted empirical studies into the phenomena of memorization and robust overfitting within adversarially trained neural networks, focusing particularly on the influence of dataset characteristics, under the supervision of Dr. Zheng Rui.
- Innovated a data augmentation methodology that leverages the textual rewriting capabilities of Large Language Models (LLMs), augmenting selected examples through strategic entity substitution and context rewriting tailored by prompt engineering.

Emory University

Remote Research Assistant

Atlanta, GA, USA

Jun 2023 – Oct 2023

- Collaborated on a cutting-edge initiative to refine Large Language Model (LLM) optimization via ranking metrics, focusing on contextually-grounded candidate responses and promoting a robust partial ordering approach under Dr. Fei Liu's guidance.
- Participated in the creation of a comprehensive multi-document question answering dataset, with the goal of testing and enhancing the proposed system's contextual comprehension; prepared research findings for submission to the prestigious ICLR 2024 conference.

Computational Evolutionary Intelligence Lab, Duke University

Summer Research Assistant

Durham, NC, USA

Jul 2023 – Sept 2023

- Explored the vulnerability of diffusion models, particularly focusing on the stability of latent signatures within these models, in collaboration with Jingyang Zhang, under the mentorship of Dr. Yiran Chen.
- Developed a robust watermarking system aimed at fortifying the protection of digital intellectual properties and enabling the reliable identification of AI-generated content sources within the burgeoning field of AI-generated content (AIGC).

WORK EXPERIENCE

Shanghai Artificial Intelligence Laboratory
LLM R&D Intern

Shanghai, China
Mar. 2024 – Oct. 2024

- Implemented an autonomous framework extracting informative content from HTML pages via text/multimodal LLMs and iterated several versions for improved performance.
- Explored and developed advertisement preference models to filter advertisements in pretraining corpus, and compared their performance with scorer models.
- Deployed reward model for large-scale data annotation and accelerated inference speed through hacking popular VLLM project. Constructed large scale data (especially long context documents) recall module for data synthesis.

AWARDS & ACHIEVEMENTS

2019-2020, 2021-2022 Fudan Outstanding Student Scholarship Second Prize ($\approx 25\%$)

2022-2023 Fudan Outstanding Student Scholarship First Prize (Xiaomi Scholarship) ($\approx 5\%$)

Wangdao Undergraduate Researcher Funding Candidate Awarded to undergraduate students who take part in research projects carried out by Fudan University Undergraduate Research Plan (FDUROP)

TOEFL 103, Reading(30/30), Listening(25/30), Speaking(22/30), Writing(26/30)

SKILLS

Programming: C, C++, Java, Python, MATLAB, R, MySQL, System Verilog

Technologies: Git, Xilinx ISE, Raspberry Pi

Languages: Chinese (Native), English (Professional), Korean (Amateur)