

Sean(Haojin), WANG

Email: seanwang027@gmail.com | Tel: +86-15828330768
Address: 4800 Cao'an Highway, Shanghai 201804, P.R. China

EDUCATION

-
- | | |
|--|-------------------|
| Tongji University, China | 09/2021 - 06/2025 |
| ● Bachelor of Engineering in Computer Science and Technology (Expected) CGPA: 4.90/5.0 | |
| -Core modules: C/C++ Programming, Advanced Mathematics (Calculus), High-Level Language Programming, Linear Algebra, Discrete Mathematics, Introduction to Computer Science, Advanced Language Programming Practice, Data Structures, Probability and Mathematical Statistics, Web Technology, Formal Languages and Automata, Assembly Language Programming, Artificial Intelligence Principles and Technologies, Computer Architecture, Computer Organization, Algorithm Analysis and Design, Cocoa based Apple Software Development, Machine Learning, Brain and Cognitive Computing, Principles of Database Systems, Operating Systems, etc. | |
| National University of Singapore, Singapore | 01/2024 - 05/2024 |
| ● Exchange Student | |
| -Core modules: Introduction to Media Computing, Information Retrieval, Data Engineering | |

RESEARCH INTERESTS

-
- Natural Language Processing; Information Theory; Generative Models; Language Modeling; Representation vectors; Computational linguistics; AI Safety.

RESEARCH EXPERIENCE

-
- | | |
|--|-------------------|
| ● Research on Next Token Distribution Analysis | 05/2024 - Present |
| <i>Advisor: Prof. Freda Shi, David R. Cheriton School of Computer Science, University of Waterloo</i> | |
| - The project focuses on analyzing the behavior of large language models in producing target probability distributions with varying levels of information entropy through the use of soft prompts and textual prompts. Deepend understanding of transformers, especially understanding of word embeddings and representations. | |
| -Currently working as a Visiting Student at the University of Waterloo, conducting research on language learning, prefix tuning and gradient-based adversarial attacks in NLP. | |
| -Currently planning to publish a research paper based on this work as the first author and will be submitted to ACL'2025. | |
| ● Research on Estimation of Uncertainty on NLG tasks | 07/2024 - 12/2024 |
| <i>Advisor: Prof. Hanjie Chen, Department of Computer Science, Rice University</i> | |
| -As the primary researcher responsible for this project, conducted research on uncertainty quantification via conformal prediction, leading to the development of several well-formed ideas; | |
| -Gained foundational knowledge in uncertainty estimation, with a particular focus on adapting conformal prediction methods from classification tasks to natural language generation tasks; | |
| -Finished a research paper and will be submitted to ACL'2025. | |
| ● Knowledge Graph Construction for ChangYuan, ADMIS Lab, Tongji University | 07/2024 - 09/2024 |
| <i>Advisor: Professor Wengen Li, College of Electronics and Information Engineering, Tongji University</i> | |
| -In charge of LLM based work, utilizing LLM(Qwen model) to extract entity sets from unstructured data for the construction of a knowledge graph; | |
| -Refined the methodology to enhance accuracy based on the company's requirements. | |
| ● Research in Explainable and Controllable AI Lab, Stevens Institute of Technology | 04/2024 - 07/2024 |

Advisor: Professor Zining Zhu, Stevens Institute of Technology

-Acquired foundational knowledge in the concepts of Machine Unlearning;

-Developed an idea to apply machine unlearning to datasets in pretraining corpora to investigate the effects on model performance; however, a similar concept was published before our work could be completed.

- **Research in StatNLP Lab, Singapore University of Technology and Design** 01/2024 - 05/2024

Advisor: Dr. Lu Wei, Singapore University of Technology and Design

-Learned basic knowledge in NLP;

-Attended weekly meetings and discussed the fundamentals of LLM alignment, with a particular focus on RLHF area.

- **Research on Duplicate Text Detection, ADMIS Lab, Tongji University** 04/2023 - 09/2023

Advisor: Professor Wengen Li, College of Electronics and Information Engineering, Tongji University

-Gained basic knowledge in prompt engineering, taking a leading role in a project focused on detecting duplicate text in data associated with private company information;

-Constructed a Knowledge Graph for the company's information, and developed a search engine based on that knowledge graph.

EXTRACURRICULAR ACTIVITIES

- **Main Code Writer** in Project License Plate Detection and Recognition at Large Angles 01/2023 - 12/2023 (National Training Program of Innovation and Entrepreneurship for Undergraduates)

- **Project Leader** in Sarcasm Detection in Headlines based on BERT (HuaWei Mindspore 08/2023 - 09/2023 Training Program)

-Led the team to migrate the training process from pytorch to mindspore, gained great knowledge in architecture of Transformer and fundamentals in LLMs;

-Obtained the second prize in Innovation.

- **Team Leader** in China Collegiate Computing Contest: 2023 Mobile Application Innovation 03/2023 - 06/2023 Contest

-Served as the HCI designer and Idea Provider for the team work;

-Obtained the second prize in East China.

AWARDS & HONORS

- **Excellent Student Scholarship Second Prize** in December 2023.
- **Second Prize** in 2023 9th China Undergraduate Statistical Modeling Competition in August 2023.
- **Second Prize** in China Collegiate Computing Contest: 2023 Mobile Application Innovation Contest in July 2023.
- **National Scholarship** (as the top 5 student in the department) in July 2022.

SKILLS

- **Languages:** Mandarin (native), English(Proficient): TOEFL 109; GRE 328+3.5
- **Programming Languages:**
 - Proficient: C, C++, Python(2/3), MATLAB
 - Capable: JavaScript, HTML/CSS, Bash, SQL
- **Hobbies:** Fitness, Music, Movie