Hui Wei

github.com/wll199566 davidhuiwei.github.io @HuiWei15 huiwei2@ucmerced.edu

EDUCATION

Electrical Engineering and Computer Science, University of California, Merced 2025 - Present
Ph.D. student in Computer Science *GPA: 4.0/4.0*

Courant Institute of Mathematical Sciences, New York University 2017 - 2019
M.S. in Computer Science *GPA: 3.8/4.0*

Beijing University of Posts and Telecommunications 2013 - 2017
B.Eng. in Telecommunication Engineering *GPA: 3.8/4.0*

RESEARCH INTERESTS

Machine Learning, Large Language Models, Time Series, AI for Healthcare, AIoT.

RESEARCH EXPERIENCE

Research Scientist Intern 05/2024 – 08/2024
PAII.Inc (Ping An Technology North America Research Institute, Sillicon Valley) *Palo Alto, CA*

- Worked on systematic evaluation on **LLM-as-a-Judge** methodology with summarization and conversation tasks.
- Improved **proximal policy optimization** algorithms for better LLM alignment with human preferences (**RLHF**).
- One paper [6] was accepted by ICLR workshop on Building Trust on LLMs and LLM applications 2025 and one paper [5] was submitted to EMNLP 2025.

Research Assistant 09/2022 – 01/2024
UMass Amherst *Amherst, MA*

- Worked on a **temporally sparse self-attention model** to address the **missing data problem** in large-scale longitudinal physical activity data from the *All of Us* research program.
- Worked on a **variational auto-encoder model** for irregularly sampled ICU data with various output distributions.
- Worked on a reconstruction-based **contrastive learning** approach to improve the quality of time series encodings.
- One paper [7] was accepted by CHIL 2024 and ICLR workshop on Learning from Time Series for Health 2024.
- One paper [8] was accepted by ICLR 2024.

Research Assistant 02/2019 – 06/2020
NYU Grossman School of Medicine *New York City, NY*

- Worked on evaluation of diagnostic accuracy and fairness for **Alzheimer's disease** and **Lewy body disease**.
- Worked on **CNN, LSTM and Transformer-based models** for **disease onset prediction** using NYU EHR data.
- One paper [9] was accepted by the journal *Frontiers in Aging Neuroscience* 2023.

PUBLICATIONS

* indicates equal contribution

1. **Hui Wei**, Dong Yoon Lee*, Shubham Rohal*, Zhizhang Hu, Shiwei Fang, Shijia Pan. "A Survey of Foundation Models for IoT: Taxonomy and Criteria-Based Analysis." *arXiv preprint arXiv:2506.02553* (2025).
2. Shenghua He*, Tian Xia*, Xuan Zhou*, **Hui Wei**. "Response-Level Rewards Are All You Need for Online Reinforcement Learning in LLMs: A Mathematical Perspective." *arXiv preprint arXiv:2506.02553* (2025).
3. Shangjie Du*, **Hui Wei***, Dong Yoon Lee, Zhizhang Hu, Shijia Pan. "Graph-Based Physics-Guided Urban PM2.5 Air Quality Imputation with Constrained Monitoring Data." *ACM Transactions on Sensor Networks* (**TOSN**), 2025.
4. **Hui Wei**, Zihao Zhang, Shenghua He, Tian Xia, Shijia Pan, and Fei Liu. "PlanGenLLMs: A Modern Survey of LLM Planning Capabilities." *The 63rd Annual Meeting of the Association for Computational Linguistics* (**ACL**), 2025. (***Oral Presentation***)
5. Jingyang Lin, Andy Wong, Tian Xia, Shenghua He, **Hui Wei**, Mei Han, and Jiebo Luo. "Facilitating Long Context Understanding via Supervised Chain-of-Thought Reasoning." *arXiv preprint arXiv:2502.13127* (2025).

6. **Hui Wei***, Shenghua He*, Tian Xia, Andy Wong, Jingyang Lin, Mei Han. "Systematic Evaluation of LLM-as-a-Judge in LLM Alignment Tasks: Explainable Metrics and Diverse Prompt Templates." *Building Trust on LLMs and LLM applications Workshop*, **ICLR**, 2025.
7. **Hui Wei**, Maxwell A. Xu, Colin Samplawski, James M. Rehg, Santosh Kumar, Benjamin M. Marlin. "Temporally Multi-Scale Sparse Self-Attention for Physical Activity Data Imputation." *Conference on Health, Inference, and Learning* (**CHIL**), PMLR, 2024.
8. Maxwell A. Xu, Alexander Moreno, **Hui Wei**, Benjamin M. Marlin, James M. Rehg. "Retrieval-Based Reconstruction for Time-series Contrastive Learning." *The 12th International Conference on Learning Representations* (**ICLR**), 2024.
9. **Hui Wei**, Arjun V. Masurkar, and Narges Razavian. "On Gaps of Clinical Diagnosis of Dementia Subtypes: A Study of Alzheimer's Disease and Lewy Body Disease." *Frontiers in Aging Neuroscience* 15 (2023): 1149036.

TEACHING EXPERIENCE

Head Teaching Assistant, UMass Amherst 09/2024 - 12/2024
COMPSCI 485: Introduction to Natural Language Processing *Amherst, MA*

- Advised student teams on final project ideas and provided extra weekly office hours for additional support.
- Managed grading for homework, exams, and in-class exercises, and handled online student questions.

Teaching Assistant, UMass Amherst 01/2024 - 05/2024
COMPSCI 250: Introduction to Computation *Amherst, MA*

- Led weekly discussion and lab sessions, held regular office hours to support students, and graded assignments.

AWARDS AND HONORS

Student Spotlight, Neuroscience at UMass Amherst, 2023.
Merit Student Scholarship, Beijing University of Posts and Telecommunications, 2013-2017.

ACADEMIC SERVICES

Reviewer, *workshop on Building Trust in LLMs and LLM Applications*, ICLR 2025.
Reviewer, *workshop on Learning from Time Series for Health*, ICLR 2024 and NeurIPS 2022.
Reviewer, *IEEE Journal of Biomedical and Health Informatics* (JBHI) 2024.
Evaluation committee, *workshop on DATA*, Sensys & Buildsys 2024, 2023, 2022.

TECHNICAL SKILLS

Languages: Python, C/C++, Bash, LaTeX, Markdown, SQL, HTML/CSS
Libraries: PyTorch, NumPy, Matplotlib, Pandas, Scikit-Learn, Seaborn, SciPy
Tools: Linux, Git/GitHub