



Nuwa Xi

Nationality: Chinese **Date of birth:** 22/05/1999

Phone number: (+86) 18581608925 **Email address:** nvxxx6021@gmail.com

Work: Harbin Institute of Technology, Xida Street, 150001 Nangang, Harbin, Heilongjiang (China)

EDUCATION AND TRAINING

Bachelor of Engineering in Bioinformatics

Harbin Institute of Technology [30/09/2017 – 20/06/2022]

City: Harbin

Country: China

Website: <http://en.hit.edu.cn/>

Master of Computer Science and Technology

Harbin Institute of Technology [01/09/2022 – Current]

City: Harbin

Country: China

Website: <http://en.hit.edu.cn/>

RESEARCH EXPERIENCE

Unified Cognitive Signal Reconstruction for Brain Decoding (Team Leader)

[01/09/2022 – Current]

- Introduced a novel task called fMRI2text, which is the first open-vocabulary task that decodes fMRI time series into human language in a naturalistic context.
- Proposed a unified framework UniCoRN (Unified Cognitive signal Reconstruction for brain decoding) to translate cognitive signals into human language, and validate its potential on both EEG and fMRI.
- Primary results published on ACL2023 and received **Honorable Mention**.

AS-ES Learning: Efficient CoT Learning in Small Models

[01/09/2023 – 01/02/2024]

- Introduced AS-ES learning, a novel data-efficient training paradigm that exploits the intrinsic value of existing CoT data.
- Showed that the limitations in CoT learning previously attributed to the inherent capabilities of small models can be substantially mitigated through an improved data utilization strategy without additional data.
- Provided a theoretical foundation for AS-ES learning, offering insights into the underlying dynamics of CoT.
- Research results currently under review.

Low-Resource Molecule Discovery with Pseudo Data from LLM (Team Leader)

[01/04/2023 – 01/09/2023]

- Proposed two methods to leverage pseudo data constructed from LLM for low-resource molecule discovery, data augmentation and domain adaptation
- Designed the experiments to study the effect of the amount of pseudo data, and analyze the experimental results
- Research results published on AAAI2024.

Tuning-Free Manifold Re-embedding for Enhanced Prompt-Based NLP Classification (Team Member)

[01/12/2022 – 01/09/2023]

- Collaborated on a research that proposes a tuning-free, manifold-based space re-embedding technique, Locally Linear Embedding with Intra-class Neighborhood Constraint (LLE-INC), to refine the representation of verbalizer embeddings.
- Contributed to applying the strategy to leveraging the output embeddings of PTMs with no parameter updating for tuning-free applications for LLMs.
- Co-authored a paper published on AAAI2024.

Paraphrase-Aided Fine-Tuning of Pre-trained Models (Team Member)

[01/02/2022 – 01/06/2022]

- Collaborated on a research that proposes a method to enhance the understanding of rare biomedical words in pre-trained models, addressing the challenge of limited data and rare biomedical words hindering performance
- Contributed to designing and conducting experiments to study the effect of the number of paraphrases of rare biomedical words involved in the samples
- Co-authored a paper published on COLING2022.

PROJECTS

Huatuo-Llama-Med-Chinese (Team Member)

[01/03/2023 – Current]

- Utilize GPT3.5 API to interface with the medical knowledge base using various prompt formats to fully leverage our resources
- Fine-tuned LLaMA&GLM to improve its performance in medical question-answering tasks
- Received **3.9k+ star** on github

Link: <https://github.com/SCIR-HI/Huatuo-Llama-Med-Chinese>

Let there be LIGHT: Independent Game (Team Leader)

[10/06/2022 – 20/08/2022]

- A single-player 2D roguelike game with up to eight main characters and a wide range of collection
- Designed and implemented the game architecture and database, created the game's graphics and assets
- Came in the third place in NUS summer workshop

PUBLICATIONS

[UniCoRN: Unified Cognitive Signal Reconstruction bridging cognitive signals and human languages](#)

Nuwa Xi, Sendong Zhao, Haochun Wang, Chi Liu, Bing Qin and Ting Liu

The 61st Annual Meeting of the Association for Computational Linguistics (ACL2023)

[From Artificially Real to Real: Leveraging Pseudo Data from Large Language Models for Low-Resource Molecule Discovery](#)

Yuhan Chen, **Nuwa Xi**, Yanrui Du, Haochun Wang, Chen Jianyu, Sendong Zhao, Bing Qin

The 38th Annual AAAI Conference on Artificial Intelligence (AAAI2024)

[Manifold-based Verbalizer Space Re-embedding for Tuning-free Prompt-based Classification](#)

Haochun Wang, Sendong Zhao, Chi Liu, **Nuwa Xi**, Muzhen Cai, Bing Qin, Ting Liu

The 38th Annual AAAI Conference on Artificial Intelligence (AAAI2024)

[Prompt Combines Paraphrase: Teaching Pre-trained Models to Understand Rare Biomedical Words](#)

Haochun Wang, Chi Liu, **Nuwa Xi**, Sendong Zhao, Meizhi Ju, Shiwei Zhang, Ziheng Zhang, Yefeng Zheng, Bing Qin, Ting Liu

The 29th International Conference on Computational Linguistics (COLING2022)

[Global Prompt Cell: A Portable Control Module for Effective Prompt](#)

Chi Liu, Haochun Wang, **Nuwa Xi**, Sendong Zhao, and Bing Qin

The 12th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC2023)

[HuaTuo: Tuning LLaMA Model with Chinese Medical Knowledge](#)

Haochun Wang, Chi Liu, **Nuwa Xi**, Zewen Qiang, Sendong Zhao, Bing Qin, Ting Liu

Technical Report for Huatuo-Llama-Med-Chinese

[AS-ES Learning: Towards Efficient CoT Learning in Small Models](#)

Nuwa Xi, Yuhan Chen, Sendong Zhao, Haochun Wang, Bing Qin, Ting Liu

Under Review

[Beyond Direct Diagnosis: LLM-based Multi-Specialist Agent Consultation for Automatic Diagnosis](#)

Haochun Wang, Sendong Zhao, Zewen Qiang, **Nuwa Xi**, Bing Qin, Ting Liu

Under Review

[Knowledge-tuning Large Language Models with Structured Medical Knowledge Bases for Trustworthy Response Generation in Chinese](#)

Haochun Wang, Sendong Zhao, Zewen Qiang, Zijian Li, **Nuwa Xi**, Yanrui Du, MuZhen Cai, Haoqiang Guo, Yuhan Chen, Haoming Xu, Bing Qin, Ting Liu

Under Review

LANGUAGE SKILLS

Mother tongue(s): **Chinese**

Other language(s):

English

LISTENING C2 READING C2 WRITING C2

SPOKEN PRODUCTION C2 SPOKEN INTERACTION C2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

SCHOLARSHIPS

Top-Level Scholarship of Graduate Student

[01/09/2023 – Current]

Samsung Scholarship

[01/12/2023 – Current]