Chi Xing

Dedicated and Results-Oriented AI Research Engineer

+44 07536371064 | chi.xing@liverpool.ac.uk | https://github.com/MartinRepo
Aug 2002 | https://martinspace.top & https://openchi.life | Edinburgh, United Kingdom

EDUCATION BACKGROUND

University of Liverpool & Xi'an Jiaotong-Liverpool University — **Liverpool, UK & Suzhou, China**
BSc Computer Science (First Class with Honors) — *Sep 2020 – May 2024*

- **Major Courses:** C/C++ Programming (90%), Java Programming (91%), Adv Artificial Intelligence (90%), Complexity of Algorithm (84%), Efficient Sequential Algorithms (81%), Optimization (83%), Intro to Theory of Computation (81%), Algorithmic Foundations and Problem Solving (81%), Thesis in First Class (80%)

University of Edinburgh — **Edinburgh, UK**
MSc Artificial Intelligence — *Sep 2024 – Sep 2025 (expected)*

- **Major Courses:** Accelerated Natural Language Processing, Machine Learning and Pattern Recognition, Machine Learning Practical, Natural Language Understanding, Generation and Machine Translation, Computer Vision

INTERNSHIP & OPEN-SOURCE EXPERIENCES

N8 CIR Summer Research Internship — **Liverpool, UK**
LLM Researcher, funded by EPSRC, Supervisor: Antony McCabe & Jianping Meng — *Jun 2024 – Sep 2024*

- Benchmark various LLMs for reading bio-medical literatures
- Use Llama.cpp to quantize Llama3.1-70B, Llama3.1-405B, Mixtral-8x22B etc. open-source models
- Design a summarization method to reduce the input size to fit models with smaller context window
- Develop an objective scoring system (using python and torch) for benchmarking various LLMs
- Compare models performance on different devices (GH200 vs A100, CPU reference vs GPU reference)
- Deploy LLMs on HPC architectures

Casibase (2700+ Stars on Github) — **Casbin Open-source Community, US**
Core Contributor, Mentor of 2024 OSPP (Open-Source Promotion Plan) — *Jan 2024 – Sep 2024*
Github Repository: https://github.com/casibase/casibase

- Expanded Casibase's capabilities by integrating support for a variety of LLMs (including open-source models and commercial models), including chat and embedding models.
- Enhanced Casibase with multimodal support, optimized output format, and addressed related bugs.
- Optimize the logic of the text splitter to facilitate better vectorized embedding of the AI knowledge base.
- Utilized Go for backend development and React.js for frontend development. Gained proficiency in applying the MVC design pattern in practical scenarios, deepening understanding of frontend-backend interactions. Recognized the importance of maintaining loose coupling to facilitate future code maintenance.
- Developing Instant Message for multi-agent function.
- Addressed more than 50 issues, details can be found at https://martinspace.top/en/about/

IFLYTEK — **Suzhou, China**
R&D Department, NLP Software Engineer — *Jun 2022 – Sep 2022*

- Assisted the department, worked mainly on natural language processing (NLP) in annotating and checking the address data to ensure the quality of the data.

"IFLYTEK Foresight" Police Super Brain System

- Checked and re-labelled indistinguishable address labels (Point of Interest), which were labelled by computer mechanically with an unexpected accuracy and a low efficiency.
- Tried to find regularity in the data structure and learned to write Python scripts to process the remaining 6,000 data, which greatly improved the labelling efficiency, thus receiving praise from my leader.
- Understood how to handle JSON data and improve my ability to write code in practice, unlike the paper-based approach taken in previous studies.

A Quiz Algorithm Competition in the Field of Heritage Culture

- Worked with the operations department to prepare the data for the algorithm competition.
- Learned about entity relationship extraction in natural language processing, such as SPO (Subject, Predicate, Object) extraction.
- Understood the methods of building knowledge graphs such as building knowledge graphs from structured data.

COMPETITION EXPERIENCE

2023 BMW Hackathon — **Shenyang, China**
2nd place in Channel: Energy-Saving Method in Production with HVB Reuse — *Aug 2023*
Design and implement the power scheduling algorithm.

- Designed a special algorithm to try to find the optimal solution for power scheduling with the idea of dynamic programming; found patterns in the peak and valley price of electricity, the amount of electricity generated during the period of the PV, and the intensity of solar radiation; listed the dynamic transfer equation and discovered the optimal solution finally.

Design and implement the battery scheduling algorithm.

- Energy storage cabinets form an integral component of electric power dispatch systems, incorporating a multitude of retired automotive batteries. The judicious scheduling of batteries within these cabinets can mitigate the rate of battery lifespan degradation. Centered around the greedy algorithm, we have devised a battery dispatch strategy. Furthermore, we have validated the correctness of this greedy algorithm, demonstrating that it satisfies the properties of greedy choice and recursion.

Docker engineering packaging

- Used Docker to engineering package and submitted the final scheduling solution, get familiar with Docker.

ACADEMIC & SCHOOL EXPERIENCES

Scheduling for Smart Grid **Liverpool, UK**

Researcher, Supervisor: Prudence Wong *Oct 2023 - May 2024*

Do research on some scheduling problems arising in demand response management, including FPT, Greedy, Genetic and Feasible Graph. Design a website to generate data reasonably and visualize performance indicator of these scheduling solutions. https://github.com/MartinRepo/SmartGrid

- Studied principals of 4 algorithms, including dynamic programming, graph theory, etc.
- Implemented 4 algorithms using C++, use CPPREST to interact with front-end and use CMake to build the project.
- Developed visualization website using React.js and use AntDesign/g2 to visualize performance metric data.
- Deploy the system on the cloud server (ubuntu os), which enhanced my cloud development skill.

Laplace Solver **Liverpool, UK**

Developer *Mar 2024*

Modelling heat transfer in a room using Laplace's equation (a second order partial differential equation)

- Implemented a multi-threaded Laplace Solver using OpenMP
- Distributes the radiator temperatures over multiple MPI processes
- Running data at scale on HPC architecture, testing the program running on (1, 2, 4, 8, 16, 32) threads. Observe its speedup efficiency
- Test the fastest running instance over 1, 2, 4, 8, 16 and 32 ranks as well, observe its strong scaling ability

AI Adversarial Training **Liverpool, UK**

Developer, Supervisor: Xiaowei Huang *Nov 2022 – Dec 2022*

This is an AI adversarial training based on FashionMinst dataset. The development uses the Pytorch framework and uses a few algorithms for training.

- During the training process, studied many algorithms (e.g. FGSM, PGD and C&W, etc.) to generate adversarial samples; delved into algorithm principles and did a bit of optimization specific to the task.
- After comparing the results of many training results, the group finally chose two algorithms (FGSM and PGD) for integration and optimization to achieve higher attack ability.
- Familiarize with the Pytorch framework and gain some understanding of many adversarial algorithms. In addition, realized that it is especially important to check the robustness of deep learning models after learning a variety of attack algorithms.

Real-time and Distributed Temperature Monitoring Network Based on AIoT **Suzhou, China**

Research Assistant, Funded by University's Summer Undergraduate Research Fellowship *Dec 2021 – Aug 2022*

This is a large-scale distributed temperature monitoring system based on the ZigBee communication protocol and AIoT (the combination of Artificial intelligence technologies with the Internet of things (IoT) infrastructure).

- Participated in the back-office management system development, including receiving temperature data in real time and providing alarms for abnormal temperatures.
- Assisted in the back-end development, mainly writing API to grab the temperature data sent by the coordinator, and other data used on the APP side and encapsulating them into APIs for the front-end to call.

OTHERS

Language: Chinese (Native), English (Proficient)

Certificate:

- The 2023 UK & Ireland Programming Contest Honorable Mention (2nd place in University of Liverpool)
- Asian-Pacific Mathematic Modelling Competition Third Prize (Top 25%)

Skills:

- Proficient: C/C++, Python, Go, JavaScript, Java, BeeGo, React.js, Git, Linux
- Semi-skilled: Pytorch, Tensorflow, MPI, C#, R, SQL, MATLAB, Next.js, Node.js