Jinglin (Ollie) Jian

(+1) 217-819-7884 jj50@illinois.edu jianjinglin.github.io https://www.linkedin.com/in/olliejian1999/

Education

University of Illinois Urbana-Champaign, School of Information Sciences Illinois, US
M.S. in Information Sciences/Management *Aug 2023 - May 2024*
Course: Large Language Models, Text Information System, Machine Learning, Data Mining

Peking University, National School of Development Beijing, China
B.Econ. - Double Major *Sep 2021 - Jul 2023*

Beijing Normal University, Faculty of Education Beijing, China
B.S. in Educational Technology, GPA: 3.76/4.0 *Sep 2017 - Jul 2021*
Course: Data Structure, Database, Software Engineering, WebDev, OOP, Intelligent System

Publications and Conference

[1] **Jian, J.**, Jiao, Z., Chen, J. & Yang, Z. (2024). **Parallel Computing-Tree Search for Optimized Drug Sequence Design**. In *2024 IEEE International Conference on Big Data* (under review)

[2] Liu, H., Li, Y., **Jian, J.**, Cheng, Y., Lu, J., Guo, S., ... & Wang, H. (2024). **Toward a Team of AI-made Scientists for Scientific Discovery from Gene Expression Data**. arXiv preprint arXiv:2402.12391 [Paper][Code]

[3] Xiao, Y. and **Jian, J**. (2024). **Which Animal Would You Like to See on Your Flashcards? Designing Visual Aids Together with Kids Using GIMs**. In *25th International Conference on Artificial Intelligence in Education* (Interactive Event) [Web]

Research Experience

Parallel Computing-Tree Search for Optimized Drug Sequence Design *Jan 2024 – Present*
Advisor: Prof. Yang Zhang (UIUC) and Dr. Jin Chen (Cleveland Clinic)

- Developed an enhanced **Monte Carlo Tree Search** framework, prioritizing simulation iteration over classifier accuracy, achieving **16-fold** computational efficiency and improved sequence quality through integrated simple but precise chemical simulations.
- Built a high-concurrency, fault-tolerant system on **AWS** for molecular docking simulation, employing RESTful API (**ApiGateway**), computation and load balancing (**EC2**) and parallel processing (**MPI**).

Team of AI-made Scientists (TAIS) [**Paper**][**Code**] *Aug 2023 – Feb 2024*
Advisor: Prof. Haohan Wang (UIUC)

- ML can discover disease-predictive genes from gene expression data. We introduced **TAIS**, a pioneering LLM-based framework for streamlining ML analysis, outperforming **GPT-4/MetaGPT/AutoGPT**.
- Fetched data from the **GEO/TAGC** database, stored meta-/raw-data for **efficient I/O**.
- Created the **GenoTEX**, an expert-made benchmark for evaluating the exploration of genomics data, with manually aligning gene symbols (mygene lib.), logging, statistical corrections, and ML.
- Created several **agents** for autonomous analytics, via creating codes (**template-based prompting**), execution (**subprocess**), outputs/errors capture (**logger**), and built communication within (**Data Engineer**, **Code Reviewer**, and **Domain Expert** agents).

Semi-automatic Knowledge Graph Construction [Web] *Sep 2020 - Jul 2021*
Advisor: Prof. Qinhua Zheng (Beijing Normal University)

- Developed a **semi-automatic** paradigm for **knowledge graph creation** for addressing time-consuming issues by combining **supervised ML** with **human-in-the-loop** incorporation.
- Created video-to-text transcription (NetEase API) and did text annotation (**BIO tagging**).
- Iterated a supervised **BiLSTM-CRF** model for **entity recognition** and dynamic term re-ranking (**mutual information** and **human feedback**), improving F1-score (0.54 → **0.76**).

Using Hypervideo to Facilitate Online Interactions [Paper] *Sep 2019 – Aug 2020*
Advisor: Prof. Jingjing Zhang (Beijing Normal University)

- Addressing **low peer interaction** and **enhancing knowledge acquisition** in online learning: Developed a **hyper-video** environment with on-screen commenting and tested **productive failure** methodology in a randomized controlled trial.
- Created an **MVP** for the online learning platform featuring video streaming and real-time commenting.
- Assisted in **pre-test** and **post-test** experiments, and contributed to data-driven behavior pattern analysis via (**Gephi**), resulting in publication.

Professional Experience

Software Developer Intern *May 2024 – Aug 2024*
Supervisor: David Bachtler and Ian Cowen, Redirect

- Engineered subscription functionality using Flutter framework and implemented unit testing.

Research Assistant *Aug 2023 – Aug 2024*
Supervisor: Prof. Mackenzie Alston, University of Illinois Urbana-Champaign

- Conducted literature review (randomized controlled trials) using Zotero and scraped 2000+ emails.

Head on Online Learning Department and Teacher Volunteer *May 2019 – Aug 2022*
China Starry Night (non-profit organization) [Web]

- Scaled our technical team from 5 to 30, achieving a 22k view increase on the Bilibili channel.
- Presented at the 5th China Education Innovation Expo (**national award - Top 1%**).

Selected Projects

- **Demo of ChemTutor: AI Q&A system with Chemistry Textbooks** [Code] *2024*
  #LLM #Q&A system #RAG (Retrieval-Augmented Generation) #LangChain
- **HMM-DRL Model for Data-driven Auto-Trading** [PDF] *2022*
  #Reinforcement Learning #Hidden Markov Model #Time Series Data #Financial Index
- **Evolution of Key Themes in Learning Sciences** [Web] *2020*
  #Topic Model #TF-IDF #Text Mining #Data Visualization

Honors and Awards

- **National Innovation and Entrepreneurship Training Award (1%)**, Ministry of Education *2021*
- **Jianghaiziqiang Scholarship (1%)**, Beijing Normal University *2020*
- **First-class Scholarship for Competition Excellence (1%)**, Beijing Normal University *2019*
- **Outstanding for Academic Excellence (10%)**, Beijing Normal University *2017 - 2021*

Skill Set

Machine Learning & NLP	TensorFlow, PyTorch, LangChain, SciPy, sklearn, gensim, NLTK
Programming Language	Python, Java, C, JavaScript, HTML, CSS, Matlab, Stata
Framework & Database	React, Node.js, RESTful API, MySQL, MongoDB, Neo4j
Cloud	AWS - EC2, API Gateway
Code Management & Others	Git/Github, Docker, Tableau, Unit Testing, LaTeX