

Zhang Haichao

Phone : 18279159611 | Email : zhang_haichao@163.com | Address : Suzhou
Website : <https://zhanghaichao520.github.io>

EDUCATION

Ph.D. in Computer Science (AI & Large Language Models)	Jun 2024 - Present
Xi'an Jiaotong–Liverpool University	Suzhou & Liverpool
Research focus: Large Language Models, Foundation Models, AI Systems, Recommendation system.	
M.Res. in Computer Science	Sep 2022 - Mar 2024
Xi'an Jiaotong–Liverpool University	Suzhou
<i>Mres Computer Science</i> GPA 77/100(distinction) Ranking: Top 1%	
Major course: Data Mining and Big Data Analytics, Research Methods, Object Oriented Programming, Technological Project Management etc.	
B.Eng. in Computer Science and Technology	Sep 2014 - Jun 2018
East China Jiaotong University	Nanchang
<i>Computer Science and Technology</i> GPA 87/100 Ranking: Top 5%	
Major course: C Language Programming Foundation, Data Structure, Operating System, C++ Programming, JAVA Language Programming Technology, Linux Foundation, Computer Network, Computer Composition Principle, JAVA advanced Programming Technology, Software Engineering, Compilation Principle, etc.	

PROFESSIONAL WORKING EXPERIENCE

Beijing Alibaba Software Technology Co. Ltd.	Nov 2018 - Aug 2022
<i>Development Engineer</i>	
Project : Alibaba's recommendation platform for head merchants. DAMOPAN(DMP)-Help Sellers Establish Personalized User Segmentation and Precision Marketing Tools	
Dingfu Data Technology Beijing Co., Ltd.	Jul 2018 - Nov 2018
<i>Development Engineer</i>	
Project: CV algorithm. Image processing, the main project is to study ways to remove image watermarks	

ACADEMIC RESEARCH

Machine Vision-Enabled Octahedral Network Reconstruction and Structural Analysis of Perovskite Quantum Dots	Mar 2024 - Mar 2025
Co–first author	under review at ACS Nano (IF 15.6)
Developed and implemented a machine-vision–enabled framework for automated reconstruction and structural analysis of PbX ₆ octahedral networks in perovskite quantum dots, enabling self-supervised denoising, atomic species classification, and high-precision lattice parameter extraction from low-dose STEM images, establishing atomic-scale structure–property relationships. I mainly responsible for all AI-related technical design and implementation.	
Customized Retrieval-Augmented Generation with LLM for Debiasing Recommendation Unlearning	Jan 2025 - May 2025
First author	Accepted by ICDM 2025 (CCF-B)
Proposed CRAGRU, a retrieval-augmented generation (RAG) framework that enables fine-grained, user-level recommendation unlearning with LLMs, effectively mitigating unlearning bias while preserving recommendation performance without model retraining. Extensive experiments on multiple benchmarks demonstrate near-retraining accuracy with up to 7× unlearning speedup.	
Explain-then-Forget: Causal Explanation-based Unlearning for Efficient and Precise Recommendation	Sep 2025 - Nov 2025
First author	under review at DASFAA 2026 (CCF-B)
Proposed CEU (Explain-then-Forget), a causal explanation–driven recommendation unlearning framework that identifies minimal causal support sets and performs targeted post-hoc ranking adjustment to precisely remove the influence of forgotten interactions without retraining. Experiments show higher unlearning precision, significantly lower computational cost, and near-original recommendation performance compared with state-of-the-art baselines.	

Recommender systems struggle to delete specific user data (e.g., when a user asks to be forgotten) without retraining the entire model, which is slow and can harm other users’ recommendations.

This work makes recommender systems “forget on demand” by adding a small, learnable prompt that erases targeted interactions while keeping the original model and other users’ recommendations intact, enabling fast and privacy-compliant updates.

Modern recommender systems often push popular items simply because “everyone clicks them,” not because users truly like them, which hurts personalization and fairness.

This work separates users’ real interests from herd-following behavior, so the system recommends what each user actually prefers rather than what is just popular.

PUBLICATION

Accept by ICDM2025

Zhang, Haichao, et al. "Customized Retrieval-Augmented Generation with LLM for Debiasing Recommendation Unlearning." arXiv preprint arXiv:2511.05494 (2025).

Published in Knowledge-Based Systems

Liu, Y., Du, G., Yin, C., Zhang, H., & Wang, J. (2024). Clustering-based incremental learning for imbalanced data classification. Knowledge-Based Systems, 111612.

Accept by ICANN2024

Yin, C., Wang, J., Zhang, H., Feng, K., Shi, L., & Ma, Q. (2024, September). Counterfactual Contrastive Learning for Fine Grained Image Classification. In International Conference on Artificial Neural Networks (pp. 169-183). Cham: Springer Nature Switzerland.

Accept by CSCWD2024

Shi, L., Li, Y., Han, Y., Wang, J., Wu, F., Yin, C., & Zhang, H. (2024, May). Two-branch Network with Feature Fusion for Time Since Deposition Estimation of Bloodstains. In 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 2191-2196). IEEE.

Recommended for submission to KDD 2026 (Second Review Cycle)

Yin, C., Fan, L., Wang, J., Hu, D., Zhang, H., Zhang, C., & Xiang, Y. (2025). Uncertainty-Aware Semantic Decoding for LLM-Based Sequential Recommendation. arXiv preprint arXiv:2508.07210.

PROJECT EXPERIENCE

Project leader

Project Description: DMP computing engine control platform users face internal products, algorithms, operations, etc.

Main responsibilities: independently responsible for product design, interaction and front-end development of the engine control desk. Cooperate with the engine leaders of the engine to design the required interfaces and join the landing and iteration of the overall promotion project

Business result: the control platform has become an indispensable part of the DMP control platform engine, which visualizes the production process of the engine and improves the efficiency of problem troubleshooting. It covers all the operation and maintenance control operations of the engine.

Project leader

Project Description: Build a quality assurance system for one million people in DMP Ad Serving System. A system connected in series with all links is needed to inspect the crowd at high frequency and quickly locate and repair problem.

Main responsibilities:

1. Responsible for the overall design of the system and the development of the system framework . 2. Coordinate the responsible persons of all modules in the whole link of DMP to build an attribution system for abnormal people.

Business results:

1. After the system landed, it supported dozens of business lines of DMP, reduced the customer complaint rate by more than 90%, cleared the failure rate, and liberated the personnel on duty to investigate and solve the problem.2. Become Alibaba's internal technological innovation project.3. The process of applying for software copyright is in progress.4. The project is recognized as a high-tech enterprise.

HONORS & AWARDS

1. East China Jiaotong University 2015 ACM Programming Competition First Prize	2015.11
2.Jiangxi Province Computer Works Competition and "Internet" Innovation and Entrepreneurship Competition (Provincial) Second Prize	2016.06
3. New national line transportation college students excellent thesis (national) first prize	2016.09
4.The third prize of the "Internet Transportation" Innovation and Entrepreneurship Competition in the Creator Competition (National) of the School of Transportation of China University	2016.11

ADDITIONAL INFORMATION

- **Skills:** Programing language (Python, JAVA) , Deep learning (Pytorch) Big data computing (flink, mapreduce, hbase) application development(spring cloud ,mysql)
- **Certifications:** Intermediate Professional Title: Software Designer
- **Languages:** LanguageCert IESOL C1 Pass