

JIAXING WANG

AI Research Lead | Data-Centric LLM Training

+1 (332) 250 5872 · WeChat: jiaxingw94 · wjiaxing94@gmail.com · Google Scholar

PROFILE

I study **when and how data curation can genuinely improve LLM capabilities**. My recent work develops a research line from domain-level mixture optimization, to token-level adaptive selection, to diagnosing whether curation provides true advancement or only faster convergence. I also deploy these ideas in industrial-scale LLM data flywheels at JD.com and led a 7B/1.5T-token pretraining project.

SELECTED RESEARCH: DATA-CENTRIC AI

Research agenda: When and how should training data be reallocated to improve an LLM's attainable capability, rather than merely accelerate learning?

TANDEM: Domain-Level Data Mixture Optimization

NeurIPS 2025, First Author

- Formulated LLM data mixture optimization as a bi-level problem and introduced synchronized twin networks to learn domain weights without Hessian computation, with first-order convergence guarantees.
- Going beyond the algorithmic design, identified data-restricted pretraining and multi-task SFT as regimes where reweighting matters most.

BLADE: Token-Level Adaptive Data Selection

NeurIPS 2026, First Author

- Extended the bi-level framework from domains to individual tokens through a Hessian-free online algorithm with a dynamic validation-informed reference model and adaptive learning trajectory.
- Using only 60% of candidate tokens, improved average math reasoning accuracy from 41.1 to 43.8 over RHO-1 on 7B models, and from 52.5 to 54.4 on 13B models.

Advancement or Acceleration? How Data Curation Improves LLM Training

Manuscript under review. Draft available upon request

- Distinguished genuine capability advancement from faster convergence in data curation. Showed that persistent gains require a correctable train-target gap and introduced a lightweight diagnostic to identify when curation is likely to improve attainable performance.

PROFESSIONAL EXPERIENCE

JD.com Tech Genius Team (TGT) Core Tech Owner

Aug 2021-Present

Self-Evolving Data Flywheel for Product-Understanding LLMs

Jan 2025-Present

- Built a closed-loop data-model system across hundreds of millions of product records, integrating data selection, mixture optimization, synthesis, and expert feedback for continuous defect correction.
- Achieved **94.2% precision**, **80.4% search-query coverage**, and a **55% improvement** in new-category construction efficiency on JD.com's e-commerce platform.

Large-Scale LLM Pretraining

Jan 2024-Jan 2025

- Led pretraining of a **7B general-purpose LLM on 1.5T tokens** using hundreds of GPUs; outperformed Qwen1.5-7B by 8.52 average points across MMLU, C-Eval, CMMLU, BBH, and GSM8K.

Product Matching and Model Compression

Mar 2022-Jan 2024

- Developed an LLM-based identical-product recognition system and distilled the model from **7B to 1.2B**, achieving a **4× inference speedup** with minimal accuracy loss.

Tech Genius Team Rotations

Aug 2021-Mar 2022

- Selected for JD.com's elite leadership track and completed rotations across e-commerce, smart warehousing, and automated logistics.

SELECTED AI INFRA COLLABORATIONS

1. **Rollout-Training Co-Design for Efficient LLM-Based Multi-Agent Reinforcement Learning.** ATC 2026.
2. **FlowTrain: Flow-Based Decoupled Training for Industrial-Grade Vision-Language Models.** EMNLP 2026 Industry Track.
3. **MixEmb: Importance-Aware Quantization and Caching Co-Design for Industrial Decentralized Embedding Training.** EMNLP 2026 Industry Track.

ADDITIONAL SELECTED PUBLICATIONS

1. **Beyond Logits: Aligning Feature Dynamics for Effective Knowledge Distillation.** ACL 2025. *Equal first author.*
2. **DynaMS: Dynamic Margin Selection for Efficient Deep Learning.** ICLR 2023. *First author.*
3. **Revisiting Parameter Sharing for Automatic Neural Channel Number Search.** NeurIPS 2020. *First author.*
4. **M-NAS: Meta Neural Architecture Search.** AAAI 2020. *First author.*
5. **Bayesian Automatic Model Compression.** IEEE Journal of Selected Topics in Signal Processing, 2020. *First author.*
6. **RaFM: Rank-Aware Factorization Machines.** ICML 2019.

EARLIER RESEARCH EXPERIENCE

Tencent AI Lab

AI Research Intern

Mar 2017-Apr 2020

- Researched neural architecture search and automated model compression using industrial-scale compute and datasets, resulting in publications at NeurIPS, ICML, AAAI, and IEEE JSTSP.
- Selected for the Tencent Rhino-Bird Elite Talent Program (acceptance rate below 10%).

Ant Financial

Research Intern

May 2020-Aug 2020

- Developed a meta-reinforcement-learning pricing model for cold-start and rapid strategy adaptation in short-cycle promotional campaigns.

EDUCATION

Institute of Automation, Chinese Academy of Sciences

2015-2021

Ph.D. in Pattern Recognition and Intelligent Systems. Advisor: Jian Cheng

University of Michigan

2014-2015

Visiting Student

North China Electric Power University

2011-2015

B.S. in Automation

SERVICE

- Reviewer: NeurIPS (2021-2023, 2025-2026), ICML (2023-2026), ICLR (2022-2024, 2026, 2027), ACL (2026), and AAAI (2023-2024).
- NeurIPS 2026 Top Reviewer.
- More than 10 national invention patents; 4 granted, including 3 as first inventor.