Chengjie Ma

May 1999 Email: cjma@ramo.yonsei.ac.kr Phone: +86-15730038973
Research Interests: Efficient Machine Learning, LLM Fine-tuning, Federated Learning, AI-RAN

Education

Yonsei University, South Korea, QS:43 Sep 2024 – Present
Ph.D. Student in Electrical and Electronic Engineering GPA: 97.3/100
Supervisor: Prof. Seong-Lyun Kim Research: Efficient Machine Learning, AI-RAN

Beijing University of Posts and Telecommunications, China Sep 2021 – Jun 2024
M.S. in Computer Science and Technology GPA: 92.69/100
Supervisor: Prof. Junping Du (IEEE Fellow)
Thesis: Communication Optimization in Federated Learning Based on Asynchrony and Pruning

Chongqing University of Posts and Telecommunications, China Sep 2017 – Jun 2021
B.S. in Computer Science and Technology GPA: 87.11/100, Top 2%
Awarded Honours Bachelor's Degree Minor: Business Administration

Research Experience

AI-RAN: CPU–GPU Acceleration and PHY Scheduling Jul 2026 – Present *Yonsei University*

- Build a CPU–GPU heterogeneous computing platform using NVIDIA DGX Spark (GB10) and OCUDU 5G gNB, with PDSCH/PUSCH PHY processing and performance evaluation across CPU/GPU execution configurations.
- Analyze latency, throughput capacity, and resource overhead across CPU/GPU worker configurations, LDPC code block sizes, and PHY workloads; identify GPU bottlenecks and investigate Unified Memory data sharing, Static Routing, and Hybrid CPU–GPU PHY Scheduling.
- Plan technical exchanges during AI-RAN Summit 2026 activities (Nov 9–10, Seoul) and an initial demo at the AI-RAN Alliance All Member Meeting (Nov 11–12, Yonsei University), presenting the heterogeneous acceleration architecture and performance evaluation results.

TriBLoRA: Parameter-Efficient Fine-Tuning (Submitted) Sep 2025 – Present *Yonsei University*

- Propose TriBLoRA, a tri-factor blockwise low-rank adaptation method that separates public and private information through shared outer factors and interaction-specific middle factors, enhancing LoRA's expressive capacity.
- Achieve intra-layer rank diversity through non-uniform parameter allocation within layers under a fixed parameter budget, supported by theoretical analysis and experimental evaluation.

AI-RAN Sep 2025 – Jul 2026 *Yonsei University*

- Explore GPU-accelerated base station architectures in AI-RAN, where parts of the gNB computation can be offloaded from CPUs to GPUs.
- Investigated and implemented replacement of the gNB physical layer (L1) on the OCUDU platform using NVIDIA Aerial.

Federated Learning with Heterogeneous Clients Sep 2024 – Sep 2025 *Yonsei University*

- Studied federated learning in heterogeneous environments with diverse client devices, asynchronous training, and varying model capacities.
- Analyzed how system and model heterogeneity can introduce skewed updates and degrade training performance.
- Found that adaptive model capacity can improve the effectiveness of contributions from weak clients.

Communication Optimization in Federated Learning Sep 2022 – Jun 2024 *BUPT*

- Worked on a sub-project supported by the National Natural Science Foundation of China (62192784, U22B2038, 62172056).
- Studied communication-efficient federated learning with a focus on asynchronous training and model pruning.
- Explored federated topic modeling and its practical applications in decentralized learning scenarios.

Shared Mobility Evaluation and Sentiment Analysis Mar 2022 – Sep 2022 *BUPT–Didi Global*

- Conducted topic analysis on customer reviews for shared mobility platforms.
- Performed data analysis on taxi transfer feedback and user sentiment trends.

Publications and Patent

- Ma, Chengjie, et al. **Breaking the Capacity Bottleneck in Model-Heterogeneous Federated Learning via Gradual Model Restoration**. Proceedings of the 43rd International Conference on Machine Learning (ICML), vol. 306, pp. 84807–84845. PMLR, 2026.
- Ma, Chengjie, et al. **Efficient Federated Learning Based on Parameter Group Training**. Proceedings of the Korean Institute of Communications and Information Sciences (KICS) Conference, pp. 227–228, 2025.
- Ma, Chengjie, et al. **Federated Topic Model and Model Pruning Based on Variational Autoencoder**. Chinese Intelligent Automation Conference (CIAC), 2023.
- Ma, Chengjie, et al. **Topic Model Based on Co-occurrence Word Networks for Unbalanced Short Text Datasets**. IEEE International Conference on Data-driven Optimization of Complex Systems (DOCS), 2023.
- **Patent**: Training method, topic prediction method and device for federated variational autoencoder topic model, 202310826329, published on Jan 31, 2024. Student first author.

Skills

Programming	PyTorch, federated learning systems, model pruning, LLM fine-tuning, Docker
Research Areas	Efficient Machine Learning, Federated Learning, Parameter-Efficient Fine-tuning, AI-RAN
Languages	CET-4: 454, CET-6: 434, IELTS: 6.0

Honors and Scholarships

Outstanding Graduate of Beijing	Provincial Level	2024
First Class Graduate Scholarship (3 times)	School Level	2021–2023
Outstanding Graduate Cadre (2 times)	School Level	2022–2023
National Inspirational Scholarship (2 times)	Provincial Level	2018–2019
Guo Changbo Scholarship	School Level	2020
CQUPT Honours Bachelor's Degree	School Level	2021
Chongqing Advanced Individual in Enhancing Innovation Capability	Provincial Level	2021
1st Prize, May Day Mathematical Modeling Competition	Provincial Level	2020
2nd Prize, MathorCup Mathematical Modeling Challenge	Provincial Level	2020

Leadership and Activities

Class Monitor	Sep 2021 – Jun 2024
Volunteer, China Open Tennis Championship	Sep 27, 2023 – Oct 8, 2023