Ganyu Wang

London, ON, Canada | wangganyu0@gmail.com | (+1) 2265045633 | https://ganyuwang.github.io/

Summary

Highly motivated and results-oriented machine learning scientist and cloud-based full-stack developer with over three years of industry experience in distributed ML, federated learning, and cloud-based AI. Proficient in Python, PyTorch, TensorFlow, AWS, Kubernetes, React, and Node.js. Experienced in academic research with publications in **top-tier ML conferences (NeurIPS, ICLR, KDD)** and reviewing for leading AI/ML venues. Skilled in research leadership, mentoring, and cross-functional collaboration. **PhD** in Computer Science with a focus on Distributed ML Systems. Looking for a full-time ML scientist, software engineer, or data engineering position.

Professional Experiences

Machine Learning Researcher & Developer **Sept. 2021 - Present**

Western University

- Designed and implemented **scalable distributed machine learning system**, especially in the application of **LLM**, used black-box prompt tuning techniques for cloud-based LLMs, such as ChatGPT, optimizing system cost, improving adaptability and performance.
- Published peer-reviewed papers in **top-tier ML conferences (NeurIPS[1], ICLR[2], MLJ[3], KDD[4])** as **first author and project leader**, made cutting-edge contributions to the research of distributed ML system.
- Developed, deployed, and monitored ML models using **AWS, Kubernetes**, and cloud-based APIs (AWS, Azure), integrating SotA ML algorithms from academic papers, including Online Learning, Zeroth-Order Optimization, and Differential Privacy with PyTorch, TensorFlow, Hugging Face, and OpenAI API.

Full-Stack and Cloud Solutions Developer **Dec. 2023 - Present**

Asgard Alliance Inc.

- Designed and developed a **full-stack** application integrating RFID IoT devices for smart storage solutions, enabling real-time inventory tracking, automated management, and seamless user interactions. Built with **Vite, React, and Node.js**.
- Implemented **secure authentication** and **scalable cloud-based data processing** using **AWS services**, including Cognito for user authentication, **Lambda** for serverless processing, and **DynamoDB** for efficient data storage.
- Adapted quickly to new technologies and cloud architectures, optimizing performance and scalability while ensuring robust security.

Serves as Reviewer for Top-tier AI & ML Conferences **Oct. 2023 – Present**

AISTATS-2024, ICML-2024, KDD-2024, AAAI-2025, ICLR-2025, ICML-2025

- Provided comprehensive, in-depth reviews to advance the quality of ML research publications.
- Quickly **adapted to new research trends** and evolving methodologies in ML

Lecturer – Data Mining **Jan. 2022 – May 2022**

Wilfrid Laurier University

- Designed and taught a comprehensive **course on Data Mining**, covering theory and real-world applications.

Projects

Optimization Efficiency and Privacy in Vertical Federated Learning **Apr. 2022 - Jan. 2024**

- Published as the **first author** in the **top-tier conference (NeruIPS-2023)**[1] and **journal (MLJ)** [3].
- Developed a *novel VFL framework*, a large-scale distributed ML system, by pioneering a hybrid optimization approach that significantly improves efficiency while preserving privacy, addressing critical challenges in distributed ML system.
- Introduced *theoretical advancements* with novel analyses of optimization techniques and innovative implicit differential privacy guarantees, establishing new benchmarks in the field.
- Practically achieved a *substantial reduction in communication costs* through strategic algorithmic optimizations, paving the way for scalable AI solutions in resource-constrained large-scale distributed ML environments.

Federated Black-box Discrete Prompt Tuning (BDPL) for Cloud-based LLM **Dec. 2023 - Present**

- Proposed a novel federated framework, designed to optimize query efficiency for Federated BDPL with cloud-based Large Language Models (LLMs). Submitted to ICML-2025
- **Led research team**: planning and execution, overseeing architecture and experiment design, milestone tracking, and progress monitoring. **Managed version and branching strategies** on GitHub to ensure efficient collaboration.
- Conducted the *first theoretical analysis* of query efficiency in Federated BDPL, identifying the relationship between client activation strategies and cloud-based LLM service query costs.
- Demonstrated significant improvement of query-efficiency of our framework through experiments on both a benchmark model (RoBERTa) and a real-world scenario of cloud-based LLM (GPT-3.5 Turbo).

Event-Driven Online Vertical Federated Learning **Jan. 2023 - Oct. 2024**

- Published as the **first author** in **top-tier conference, ICLR-2025** (top **6%** review score) [2].
- Proposed a novel event-driven framework for online learning in VFL.
- Addressed the real-world challenges including asynchronous data reception and non-stationary environments.
- Established the framework as a *scalable and efficient solution* for VFL in practical applications, paving the way for real-time collaboration in VFL.

Technical Skills Summary

ML Tools: PyTorch, TensorFlow, Scikit-Learn, JAX, HuggingFace, OpenAI API.
ML expertise: Deep learning, Distributed system application, Federated learning, Parallel computation, Optimization, Differential privacy, Large Language Model (LLM).
Programming Languages: Python, C/C++/C#, R, Java, JavaScript, SQL, HTML, CSS, VHDL.
Cloud Development: AWS, Azure, Kubernetes, Docker, DynamoDB, MongoDB, Sealos Cloud, Git, Bitbucket.
Full-Stack: React, Vue, Vite, Amplify, Node.js

Education

Ph.D. in Computer Science **Sept. 2021 - May 2025**
Western University

M.Sc in Computer Science (Thesis-based) **Sept. 2019 - July. 2021**
Ontario Tech University

B.Sc in Computer Science (with Honor Bachelor's Degree) **Sept. 2015 - Jul. 2019**
University of Electronic Science and Technology of China
*Yingcai Honors College (Top **5%** of undergraduates)* *Overall GPA: **3.84/4.00** (87.02/100)*

Publication

[1] **Wang, Ganyu**, Bin Gu, Qingsong Zhang, Xiang Li, Boyu Wang, and Charles X Ling. A unified solution for privacy and communication efficiency in vertical federated learning. *Advances in Neural Information Processing Systems*, 36, 2024.

[2] **Wang, Ganyu**, Boyu Wang, Bin Gu, and Charles X. Ling. Event-driven online vertical federated learning. In *International Conference on Learning Representations (ICLR)*, 2025.

[3] **Wang, Ganyu**, Qingsong Zhang, Xiang Li, Boyu Wang, Bin Gu, and Charles X Ling. Secure and fast asynchronous vertical federated learning via cascaded hybrid optimization. *Machine Learning*, 113(9):6413–6451, 2024.

[4] Ke Zhang, **Wang, Ganyu**, Han Li, Yulong Wang, Hong Chen, and Bin Gu. Asynchronous vertical federated learning for kernelized auc maximization. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 4244–4255, 2024.

[5] **Wang, Ganyu** and Miguel Vargas Martin. Segmentperturb: Effective black-box hidden voice attack on commercial asr systems via selective deletion. In *2021 18th International Conference on Privacy, Security and Trust (PST)*, pages 1–12. IEEE, 2021.

[6] **Wang, Ganyu**, Miguel Martin, Patrick Hung, and Shane MacDonald. Towards classifying motor imagery using a consumer-grade brain-computer interface. In *2019 IEEE International Conference on Cognitive Computing (ICCC)*, pages 67–69. IEEE, 2019.