Lei Wang

+1 (424) 402-6563 | leiwang1@ufl.edu | lei-wang-link.github.io | Gainesville, FL

Education

University of Florida — Gainesville, FL
Ph.D. Candidate in Electrical and Computer Engineering | GPA: 4.00/4.00 — *Aug. 2022 – May. 2027 (Expected)*

University of California, Los Angeles — Los Angeles, CA
M.S. in Electrical and Computer Engineering | GPA: 3.93/4.00 — *Sep. 2021 – Jun. 2022*

University of Electronic Science and Technology of China — Chengdu, China
B.E. in Electronic and Information Engineering | GPA: 3.97/4.00 — *Sep. 2016 – Jul. 2020*

Professional Experiences

PayPal — San Jose, CA
Machine Learning Scientist Intern — *May. 2025 – Aug. 2025*

- Identified critical limitations in existing ACH fraud detection systems where traditional single-transaction models failed to predict ACH return fraud in loyal customer segments, particularly for NSF (Non-Sufficient Funds) cases with sudden high-impact loss spikes despite strong historical performance.
- Developed novel contextual-temporal two-stage sequence modeling approach combining DistilBERT foundation model for semantic feature understanding with transformer-based temporal pattern recognition to address behavioral shift and fraud trend detection over past several transactions.
- Validated early fraud detection capability through individual driver analysis, demonstrating model correctly identified fraud pattern development before actual losses occurred while production ACH risk models showed erratic signals that failed to capture sequential fraud evolution in loyal customer segments.
- Achieved significant performance improvements on 1.52M+ production ACH transactions across 30K ACH accounts: 92% relative improvement in fraud catch rates 5% action rate, 31% reduction in false positive rates, and prevented $1.25M additional net losses with $3.12M gross loss reduction under $35M Total Payment Volume compared to production ACH risk models, demonstrating superior loss prevention across all operational thresholds.

Research Experiences

Adaptive LoRA Experts Allocation and Selection for Federated Fine-Tuning — *NeurIPS 2025*

- Proposed **FedLEASE**, a novel framework addressing two critical challenges in federated LoRA fine-tuning: determining optimal number and allocation of LoRA experts across heterogeneous clients, and enabling adaptive expert selection based on client-specific data characteristics.
- Implemented data-driven clustering approach using silhouette coefficient and cosine similarity of LoRA matrices to identify optimal expert allocation, combined with innovative adaptive top-M Mixture-of-Experts(MoE) mechanism for dynamic expert selection.
- Achieved significant performance improvements on GLUE and FLAN benchmarks with RoBERTa-Large and LLaMA-2-7B models, demonstrating 5.53% average improvement over strongest baselines while maintaining communication efficiency in heterogeneous federated settings.

Federated Fine-Tuning through Adaptive Local Training with Rest-of-World LoRA — *AAAI 2026*

- Presented **FedALT**, a personalized federated LoRA fine-tuning framework that addresses cross-client interference in heterogeneous settings by departing from FedAvg-based aggregation and enabling adaptive local training with frozen global components.
- Designed a decoupled dual-LoRA architecture with an Individual LoRA for client-specific adaptation, a frozen Rest-of-World (RoW) LoRA for shared knowledge transfer, and an adaptive Mixture-of-Experts mixer for balancing global and local contributions.
- Demonstrated up to 4.69% improvement in ROUGE-1 scores across eight heterogeneous NLP tasks on FLAN benchmarks using LLaMA2-7B and Bloom-560M models, validating superior personalization and stability under data heterogeneity.

Federated LoRA Fine-Tuning with Aggregation and Initialization Refinement — *ICCV 2025*

- Developed **LoRA-FAIR**, an innovative solution to address two key challenges in federated fine-tuning of foundation models with LoRA: server-side aggregation bias and client-side initialization drift.
- Designed a residual-based correction mechanism to enhance LoRA module aggregation on the server, ensuring global updates closely approximate ideal model parameters.
- Evaluated the approach on ViT and MLP-Mixer foundation models using diverse and challenging non-IID datasets (DomainNet and NICO++), achieving up to 5.03% accuracy improvement and enhanced training stability in both feature and label non-IID settings compared to baseline methods.

Taming Cross-Domain Representation Variance in Federated Prototype Learning — *NeurIPS 2024*

- Designed **FedPLVM**, a dual-level prototype clustering algorithm with a novel α-sparsity prototype loss to address performance gaps in federated prototype learning across clients with heterogeneous data domains.
- Implemented the algorithm in Python using PyTorch, employing FINCH and cosine similarity metrics to ensure efficient training and model deployment across multiple environments.
- Achieved up to 9.88% higher accuracy on domain-shift benchmarks such as Digit-5, Office-10, and DomainNet, outperforming SOTA methods while optimizing communication efficiency and enhancing data privacy.

Technical Skills

General Skills: Machine Learning, Federated Learning, Computer Vision, Data Structure, LLM Fine-tuning
Languages/Frameworks: Python, Java, MATLAB, SQL, PyTorch, Keras, TensorFlow, Linux