Zhaoxiang (Aaron) Feng

San Diego, CA | zhf004@ucsd.edu | (336) 337-7139 | aaronzhfeng | aaronzhfeng.github.io

Education

University of California San Diego (UCSD) *Sep. 2022 – Apr. 2026*
B.S. in Data Science & B.S. in Probability & Statistics *(expected)*

- GPA: 3.83 (Major GPAs: 3.95 in Data Science, 3.93 in Statistics)
- Honors: Provost Honors (multiple quarters)
- Graduate-level Coursework: Advanced Time Series Analysis, Machine Learning, Applied Statistics, Privacy-sensitive Data Systems

Research Interests

I am interested in self-evolving AI agents capable of test-time continual learning: systems that accumulate concepts, verify what they have learned, and improve after deployment. My focus areas include concept-level memory, calibrated uncertainty, and efficient inference.

Publications

[1] Matthew Ho, Chen Si, **Zhaoxiang Feng**, Fangxu Yu, Yichi Yang, Zhijian Liu, Zhiting Hu, Lianhui Qin. "ArcMemo: Abstract Reasoning Composition with Lifelong LLM Memory."
***Runner Up**, ARC Prize 2025 Paper Awards — Top 8 of 90 paper submissions.* *[paper] [code]*

[2] **Zhaoxiang Feng**, David Scott Lewis, Enrique Zueco. "LabMemo: Concept-Level Memory for Autonomous Scientific Discovery."
Accepted at IEEE IROS 2025 Workshop on Embodied AI and Robotics for Future Scientific Discovery (AIR4S) (non-archival), 2025. *[paper] [code]*

[3] **Zhaoxiang Feng**, David Scott Lewis. "SOKRATES: Distilling Symbolic Knowledge into Option-Level Reasoning via Solver-Guided Preference Optimization."
Accepted at AAAI 2026 Bridge Program on Logic & AI: Logical and Symbolic Reasoning in Language Models (LMReasoning), 2026. *[paper] [code]*

---

[4] Jinzhou Tang, Yufan Zhou, Zixuan Wang, **Zhaoxiang Feng**, Xinle Yu, Steven Ngo, Zhengding Hu, Luoshang Pan, Lianhui Qin, Yufei Ding, Tianmin Shu, Jingbo Shang, Zhiting Hu, Zhen Wang. "S^3-Sim: Simulating Humans for Personalized Language Modeling."
Under review at International Conference on Machine Learning (ICML), 2026.

[5] **Zhaoxiang Feng**, Zhongyan Luo, Mingyang Yao, Charlie Sun. "Explanation-Consistency Graphs: Neighborhood Surprise in Explanation Space for Training Data Debugging."
Under review at ACL Rolling Review (ARR), 2026.

[6] **Zhaoxiang Feng**, Zhongyan Luo, Mingyang Yao, Charlie Sun, Yuwen Zhang, David Scott Lewis. "EP-Prior: Interpretable ECG Representations via Electrophysiology Constraints."
Under review at IJCAI-ECAI 2026 AI and Health Track, 2026.

[7] **Zhaoxiang Feng**, Mingyang Yao, Shuheng Cao, Young Su Ko, Wei Wang. "TrustPPI: Deformation Stability as a Model-Agnostic Trust Signal for Protein-Protein Interaction Prediction."
Under review at KDD 2026 AI4Sciences Track, 2026. *[code]*

Research Experience

Research Assistant | Q-Lab, UC San Diego *Jan. 2025 – Present*
Advisors: Prof. Lianhui Qin and Matthew Ho

- Contributed to ArcMemo by leading the design of a program-synthesis-style memory ontology, developing many-to-one puzzle-to-feature mappings and manually curating concept parameterizations to enable concept-level reasoning.

- Engineered a reasoning-based retrieval mechanism (System-2 exploration) to resolve embedding failures, achieving a 7.5% relative gain on ARC-AGI-1 (59.33% official score).
- Built a complete concept dataset generation pipeline transforming hand-written concepts into validated helper puzzles through multi-stage LLM-based generation, code synthesis, and automated testing.
- Extended the framework to AIME math problems by designing a metacognitive self-assessment pipeline, improving accuracy by 9.3% via self-reflective memory usage.

Research Assistant | Wang Lab, UC San Diego *Jun. 2025 – Present*
Advisors: Prof. Wei Wang and Young Su Ko

- Developing TrustPPI (Ongoing): Domain-specific trust signals for protein-protein interaction prediction; showed deformation stability and interface plausibility achieve 0.70–0.80 AUROC vs ∼0.50 for generic confidence.
- Architected a heterogeneous Mixture-of-Experts system for chemical reaction prediction on USPTO datasets (1M+ reactions), integrating four specialized expert models with learned routing mechanisms.
- Implemented graph neural network encoders using directed message-passing architectures with shortest-path positional encodings, enabling permutation-invariant molecular representations for stereochemistry-sensitive reaction modeling.
- Engineered training pipelines with teacher-forcing, load-balancing losses, and router warmup; designed evaluation frameworks for top-k accuracy and per-expert ablation.

Student Researcher | Halıcıoğlu Data Science Institute, UC San Diego *Sep. 2025 – Present*
Advisors: Prof. Hao Zhang and Prof. Yian Ma (Senior Capstone)

- Built an end-to-end distributed training stack for a 1.8B-parameter language model on 8 NVIDIA B200 GPUs, implementing pipeline parallelism and scaling analysis.
- Developing TPU-optimized speculative decoding to reduce latency for test-time reasoning systems.
- Developing AR-Bench, an interactive reasoning benchmark where agents decide when to ask clarifying questions versus commit to an answer.
- Designing mutual-information-based uncertainty signals to detect epistemic uncertainty for risk-controlled stopping decisions.

Research Assistant | U.S. Immigration Policy Center, UC San Diego *Sep. 2023 – Jun. 2024*
Advisors: Prof. Tom K. Wong and Dr. Gabriel De Roche

- Built forecasting models (ARIMA, exponential smoothing) on 20+ years of naturalization data to project immigration trends for 2024 election policy briefs.
- Developed automated data pipelines and interactive Tableau dashboards to communicate findings to non-technical stakeholders.

Teaching & Service

Tutor — DSC 40A: Theoretical Foundations of Data Science I *Summer 2025*
University of California San Diego

- Conducted tutoring sessions covering set theory, probability, and algorithmic thinking; led review sessions and supervised oral exams.

Grader — Mathematics Department *Fall 2024 – Fall 2025*
University of California San Diego

- Graded for MATH 180A/C (Probability & Stochastic Processes), MATH 181A (Mathematical Statistics), and MATH 185 (Computational Statistics) a total of 7 times.

Technical Skills

Languages: Python, R, SQL, Java, JavaScript, HTML

ML & LLM: PyTorch, PyTorch Geometric, Transformers, scikit-learn, LLM fine-tuning and post-training, RAG systems, memory-augmented LLMs, reasoning-based retrieval

Data & Infrastructure: Pandas, Spark, AWS, Docker, distributed training, vector databases

Scientific Computing: Physics-based ML (neural operators, inverse problems), continual learning