(714) 516-0001
Irvine, CA
masamadi@uci.edu

Mohammad Amin Samadi

PhD Candidate

Website: aminsamadi.org
GitHub: aminsmd
LinkedIn: aminsamadi

Summary

Ph.D. candidate specializing in Large Language Models (LLMs) and Human–Computer Interaction (HCI). Experienced in applying AI, NLP, and Generative AI to study human–AI collaboration and learning. Skilled in developing and evaluating LLM-based systems, analyzing human interaction data, and applying deep learning methods across educational and behavioral domains.

Work Experience

Software Engineering Intern — **Summer 2026 (expected)**
Uber – Marketplace Signals Team — *Seattle, WA*

- Will collaborate with engineers, data scientists, and product managers to design and prototype ML solutions with real-world impact on Uber's supply-demand infrastructure.

AI Scientist Intern — **Winter 2026**
ROI-AI — *San Francisco, CA*

- Built end-to-end chargeback dispute automation system using Claude API with programmatic tool calling, achieving 95%+ extraction accuracy across different formats.
- Developed an agentic evidence gathering pipeline that autonomously searches ERP systems and email to locate documents, reducing manual document retrieval efforts.
- Implemented OCR service using AWS Textract for scanned PDF processing, with automatic fallback detection and content filtering for targeted document extraction.
- Designed REST API serving 6 pipeline stages (extraction, evidence gathering, verification, dispute generation) with LLM cost tracking, Docker deployment, and Terraform-managed AWS infrastructure.

Research Intern — **Summer 2024**
Lettercast.ai — *San Francisco, CA*

- Developed an agentic workflow for podcast transcript generation: Designed and implemented an efficient, automated pipeline for generating podcast transcripts using the CrewAI framework, improving modularity and scalability.
- Integrated Multi-modal LLMs for Content Generation: Designed and implemented a pipeline that generates podcasts from arXiv research papers, utilizing multi-modal large language models to transform complex research into engaging and accessible audio narratives.
- Built a Claim Verification System for AI Hallucination Detection: Designed a claim verification mechanism to identify and mitigate hallucinations in AI-generated data, ensuring accuracy and reliability in podcast content production.

Research Experience

TRAIL (Team Research and AI Integration Lab) — **May 2023 – Present**
School of Education, UCI — *Irvine, CA*

- https://the-language-and-learning-analytics-lab.github.io/TRAIL-landing-page/
- **TRAIL is a research platform that facilitates human–AI collaboration research through experimental design and data-driven evaluation.**
  - Designed and launched **TRAIL**, a research platform for controlled studies of human–AI collaboration, embedding configurable LLM teammates into real-time, multi-party chatrooms with research-grade logging.
  - Architected a **multi-agent workflow** for the AI teammate, integrating hierarchical memory (working → short-term → reflective) and a persona compiler (Big-Five based) to systematically vary behavior across experimental conditions.
  - Implemented a robust experimental framework enabling **A/B testing**, randomized team assignment, and automated data capture for large-scale behavioral analysis.
  - Developed analytics pipelines combining **chat/conversation logs**, survey responses, and performance outcomes, enabling mixed-effects modeling, ANOVA, and cross-condition comparisons.
  - Built pipelines for **statistical analysis** of collaboration outcomes, including conversational metric extraction (e.g., responsivity, newness) and comparative modeling across treatment arms.
  - Deployed end-to-end system on **AWS**, supporting experimental room isolation, reliable streaming, and over 100 concurrent participants using Flask + Socket.IO + React.
  - Directed empirical research: conducted a full pilot RCT comparing human-only vs. human–AI teams, uncovering measurable differences in inclusivity, peer learning, consensus formation, and trust patterns.
  - Managed a cross-disciplinary team (interns + MSWE Capstone developers), leading structured recruitment, agile workflows, and alignment of platform milestones with research studies.
  - **Tech Stack:** Python (Flask), React, Socket.IO, SQL/REST, Slack API, OpenAI/LangChain API, AWS

AI-Based Video Understanding Pipeline for Teacher Training **Jan 2025 – Dec 2025**
School of Education, University of California, Irvine *Irvine, CA*

- **Leveraging AI/LLM to evaluate teacher-student interactions to cultivate belonging in classrooms through video recordings**
- https://github.com/aminsmd/multimodal-transcription
  - Developed a step-by-step process for robust qualitative coding of classroom interactions ensuring high inter-rater reliability (>0.7 Cohen's Kappa).
  - Built a full video processing and multimodal transcription pipeline (audio transcription + visual cues) using Gemini API and to capture verbal and non-verbal dynamics in a classroom recording video.
  - Designed validation and evaluation processes, ensuring consistency and accuracy throughout the generated transcripts.
  - Implemented a semantic segmentation framework using temporal semantic similarity to divide videos into meaningful conversational segments.
  - GitHub actions setup that deploys Docker to AWS and runs the transcription task on AWS ECS.
  - **Tech Stack:** Google cloud, gemini-2.5-pro API, FFmpeg, AWS, Github actions

Skills

Tools and Languages	Python, R, Hugging-Face Transformers, Pytorch, Git, SQL, LangChain, crewAI, LaTeX
Technical Skills	Machine learning, Natural Language Processing, Prompt engineering, multi-agent design
Project & Research Skills	Project management, Team leadership, Experiment design, Technical mentorship, Collaboration in inter-disciplinary teams

Selected Projects

- **Real-Time Conversational AI Platform** (link)
  Built a low-latency voice-to-voice chat application using the OpenAI Realtime API, enabling continuous bi-directional speech streaming with sub-second latency. Integrated Whisper for automatic speech recognition and text-to-speech synthesis with dynamic buffering and queue control for uninterrupted dialogue. Optimized event handling via WebRTC and Socket.IO to support simultaneous sessions and resilient network recovery.
- **Crude Oil Price Prediction** (link): Constructed a predictive modeling pipeline using Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) architectures to forecast daily Brent crude oil prices. Processed and normalized multi-year EIA data, implemented sliding-window sequences (20-day inputs → 21st-day output), and visualized model performance using RMSE and MAPE metrics. Demonstrated the capacity of deep learning models to capture temporal volatility patterns in commodity markets.
- **Zoom Transcripts Name Detection and Anonymization** (link): Using pre-trained **huggingface transformers** for named entity recognition to remove identifiable names in audio transcripts.

Education

PHD, Education, University of California Irvine (UCI) **Winter 2021 - Jan 2026 (expected)**
PhD Advisor: Nia Nixon

B.Sc. in Computer Engineering, Isfahan University of Technology (IUT) **Fall 2015 - Fall 2019**
Advisor: Mohammad Hossein Manshaei

Selected Publications

- Mohammad Amin Samadi, Spencer Jaquay, Yiwen Lin, Elham Tajik, Seehee Park, Nia Nixon. "Minds and Machines Unite: Deciphering Team Social and Cognitive Dynamics with AI.", In Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24). Kyoto, Japan, (code)
- Mohammad Amin Samadi, Spencer JaQuay, Jing Gu, and Nia Nixon. "The AI Collaborator: Bridging Human-AI Interaction in Educational and Professional Settings.", Preprint - arxiv link: arxiv.org/abs/2405.10460.
- Harshita Chopra, Yiwen Lin, Mohammad Amin Samadi, Jacqueline Guadalupe Cavazos, Renzhe Yu, Spencer Jaquay, Nia Dowell. "Semantic Topic Chains for Modeling Temporality of Themes in Online Student Discussion Forums.", International Conference on Educational Data Mining, July 2023, Indian Institute of Science Campus, Bengaluru, India, (code)
- Mohammad Amin Samadi, Danielle Shariff, Raymond Ibarra, Nia Nixon. "Balancing the Scales: Using GPT-4 for Robust Data Augmentation.", Companion Proceedings of the 15th International Conference on Learning Analytics Knowledge (LAK25), March 2025, Trinity College Dublin, Dublin, Ireland, (code)
- Kia Karbasi, Kevin Hong, Mohammad Amin Samadi, and Gregory Pottie. "Multi-Agent Collaborative Framework For Math Problem Generation.", International Conference on Educational Data Mining, July 2025, Palermo, Italy, (code)
- Seehee Park, Danielle Shariff, Mohammad Amin Samadi, Nia Nixon, Sidney D Mello. "From Discourse to Dynamics: Understanding Team Interactions Through Temporally Sensitive NLP.", International Conference on Educational Data Mining, July 2025, Palermo, Italy **(Best Short Paper Award)**