ANDY (ZEHAO) ZHAO

609-937-3646 | zhaozeha@usc.edu | linkedin.com/in/zehao-zhao | zehao.org | **No Sponsorship Needed**

EDUCATION

UNIVERSITY OF SOUTHERN CALIFORNIA   Los Angeles, CA
PhD Candidate in Electrical Engineering & Computer Science, Center of Autonomy & Artificial Intelligence, GPA:4/4   2024 –
- Awards: Ming Hsieh Fellowship, Viterbi Graduate Fellowship, Annenberg School Award, Early Career Innovator Award

MASSACHUSETTS INSTITUTE OF TECHNOLOGY   Cambridge, MA
Master of Science in Business Analytics & Operations Research, GPA:5.0/5.0   2022 –2024
- Awards: 2nd place in MIT Analytics Lab Contest out of 40 teams, Top 10 in MIT Hackathon out of 1000+ attendees

UNIVERSITY OF CALIFORNIA, BERKELEY *(4 Majors with High Honors, First in UC Berkeley 150 years)* Berkeley, CA
Bachelor of Science in Computer Science, Statistics, Math, and Data Science, GPA:3.9/4.0, Dean's List 8 Semesters 2018 – 2022
- Awards: 1st place in California actuarial league competition; Finalist for Mathematics Contest in Modeling
- Honors Thesis: Data Science Honors: Time series modelling on Finance; EECS Honors: PCA & Facial Recognition

TECHNICAL SKILLS

Python (TensorFlo, Keras, Pytorch, pandas), AWS (Sagemaker, Bedrock, Kendra), LangChain/LangGraph, Prompt Engineering, RAG, MATLAB, R (ggplot, dplyr), Julia (gurobi), C, C++, Java, Rust, LaTeX, RISC-V, SQL, CUDA, Git, ArcGIS, AutoCAD

EXPERIENCE

SAMSUNG RESEARCH AMERICA   Mountain View
Generative AI Research Intern, Advanced ML Lab   May 2025 –
- Built and optimized LLM-based internal search engine; improved top-k document relevance by 35% and latency by 40%
- Designed and deployed LLM-powered recommender system for semantic ad matching, boosting CTR by 20%

UNIVERSITY OF SOUTHERN CALIFORNIA   Los Angeles
Research Assistant to Professor Rahul Jain and Professor Bhaskar Krishnamachari   Aug 2024 –
- Developed Multi-agent General-Sum and Zero-Sum game algorithms that can be used in Large Scale Autonomous System
- Built reasoning circuits thru autoencoder to interpret LLM reasoning to complex math problems and decision-making tasks

LLM STEALTH STARTUP (Accepted by MIT Momentum Accelerator)   Boston & San Francisco
Founding Engineer (AI Branch, LangChain, LangGraph)   Aug 2023 – Aug 2024
- Fine-tuning of large language models with 100+ users for a social platform to foster friendships and dating experiences
- Developed personalized AI Agent for tailored event recommendation using LangGraph and increased user retention by 500%

AMERIPRISE FINANCIAL   Boston, MA
Data Scientist (Python, SKlearn, Deep Learning, Reinforcement Learning)   Feb 2023 – Aug 2023
- Built predictive models to forecast geographical location growth (500m grid level) within Europe for real estate investment
- Employed DNNs to increase the model performance by 50%, and employed Boosting Algorithms for interpretability
- Developed data-specific training metrics to enhance model performance

MIT OPERATIONS RESEARCH CENTER   Boston, MA
Research Assistant to Professor Dimitris Bertsimas (Python, Julia: Gurobi, IAI, Pytorch)   Sep 2022 – Feb 2023
- Proposed Mixed integer programming algorithms that performs better than state-of-the-art solver for some problems
- Developed ML scheme to maximize cancer patient survival time for various treatments and improved AUC to 80%

ALPHA SQUARE GROUP   New York, NY
Quantitative Research Intern (Python, Java, TensorFlow, Pytorch)   Jun 2022 – Aug 2022
- Created NLP model to summarize alternative data including customer review and quantify keywords
- Constructed smart scoring system for primary market investment based on 500+ historical investment data

PROJECTS AND PUBLICATIONS

- AI Powered Automatic Coastal Erosion Detection   (Published in Marine Geology)
- Computer Vision based Image Contamination Detection   (Published in Journal of Physics, First-Author)
- Optimized Computation for Quantum Random Number Generator   (Publish in arXiv/ResearchGate for review, First-Author)
- Deep-learning models for refugee relocations in Columbia   (MIT Operation Research Center & United Nations)
- Optimal allocation of Shared Bikes at Boston   (MIT Operation Research Center)
- On-Device Image Training and Inference on Microcontroller   (MIT EECS Department & Han's Lab)
- Actor Critic Reinforcement Learning Recommendation System   (MIT EECS Department)
- Uber/Lyft Demand Forecast in Different Weather Conditions   (MIT Sloan School of Management)