

Mutian (Austin) He

San Jose, CA | 408-548-1060 | mhe3@scu.edu | [LinkedIn](#) | [Github](#) | [Portfolio](#)

EDUCATION

Santa Clara University, Leavey School of Business

Jan 2025 - Jan 2027

Master of Science, Information Systems

Santa Clara, US

• **Coursework:** Cloud Computing, Database Systems, Machine Learning, Deep Learning, Reinforcement Learning, NLP

University of Glasgow, School of Computing Science

Sep 2021 - Jun 2024

Bachelor of Science, Computer Science

Glasgow, UK

• **Coursework:** Data Structures & Algorithms, Database Systems, Software Engineering, Human-Computer Interaction

EXPERIENCE

AI Engineer Intern | Amazon Web Services (AWS)

Jan 2026 – Jun 2026

Project Description: Led a team to build an AI-driven platform called RiskLens AI that transforms unstructured SEC 10-K filings into structured risk insights, enabling cross-year risk comparison and decision support. Processed 348 filings across 76 companies and 11 industries. **Explore Product here: risklensai.org**

- Designed and implemented an end-to-end AWS data pipeline to parse, clean, and structure SEC EDGAR filings using S3 and Bedrock for downstream LLM workflows.
- Built a multi-step LLM agent using AWS Bedrock Converse API with custom tools for risk retrieval, cross-company comparison, keyword search, and financial/news analysis.
- Developed a full-stack AI application (React + FastAPI) with hybrid LLM-based risk classification, three-dimensional risk scoring, and interactive dashboards for comparative risk analysis.

Research Assistant | Santa Clara University

May 2026 - Present

Project Description: Researching how large language models can function as reliable AI agents for complex reasoning and decision-support tasks using real-world EHR and structured datasets.

- Evaluated open-source LLMs (7B–70B) across reasoning, planning, and risk assessment tasks using EHR-based evaluation workflows and multi-step agent interactions.
- Designed prompt engineering and agent evaluation frameworks to benchmark reasoning quality, hallucination control, tool usage, and response reliability.
- Conducted comparative analysis across model architectures (Dense vs MoE), parameter scales, and prompting strategies to study agent performance and consistency.

PROJECTS

Goal-Conditioned Maze Navigation with Curriculum Learning | [Github](#)

Apr 2026 - May 2026

Project Description: Designed and implemented a goal-conditioned RL environment in MuJoCo PointMaze, training a PPO agent to navigate configurable mazes with curriculum learning and reward shaping.

- Implemented a PPO agent in MuJoCo PointMaze with goal-conditioned observations, improving maze navigation success rate from 6% to 38%.
- Designed a multi-stage curriculum learning framework with automatic progression across increasing maze complexity levels.
- Conducted comparative experiments on reward shaping and training strategies, including trajectory visualization and performance analysis.

Image Search System | [Live Demo](#) | [Github](#)

Mar 2026 - Apr 2026

Project Description: Built a content-based image retrieval system that enables visual similarity search over large-scale image datasets, using deep learning feature extraction and efficient vector indexing.

- Extracted high-dimensional image embeddings using a pre-trained ResNet-50 model and indexed them with FAISS for approximate nearest-neighbor search, achieving sub-second retrieval on 100K+ images.
- Designed a Flask-based web interface with drag-and-drop upload, real-time similarity ranking, and configurable distance metrics (cosine, L2).
- Implemented PCA-based dimensionality reduction to compress feature vectors from 2048-d to 256-d, reducing index size by 87% while maintaining top-10 retrieval accuracy above 95%.

SKILLS

- **Programming:** Python, SQL (MySQL, PostgreSQL), Java Script, HTML/CSS, MongoDB
- **Machine Learning & AI:** PyTorch, Reinforcement Learning (PPO, Policy Gradient fundamentals), Model Evaluation, Feedback-driven Optimization, Multimodal Learning, Representation Learning, LLM Agents, Prompt Engineering
- **Simulation & RL Tools:** Gymnasium, Stable-Baselines3 (SB3), MuJoCo / Isaac Gym, Reward Design, Training & Evaluation Loops
- **Data Engineering:** Apache Kafka, Apache Airflow, ETL/ELT Pipelines, Streaming Architecture, Data Modeling
- **Cloud & Platforms:** AWS (EC2, S3, Glue, Athena, Lambda), Docker, Railway, Cloudflare