Jiaqi Duan

https://victoria-duan.vercel.app/

jd.victoria.work@gmail.com
408-438-7247

Education

University of California, Santa Cruz

Computer Science and Engineering, Master of Science (M.S.) — *Expected Dec 2025*

Computer Science and Engineering, Bachelor of Science (B.S.); Psychology, Bachelor of Art (B.A.) — *Dec 2022*

Publications

- S.R.Castro, R.Campbell, N.Lau, O.Villalobos, **J.Duan**, A.A.Cardenas. *"Large Language Models are Autonomous Cyber Defenders."* Presented at IEEE CAI Workshop on Adaptive Cyber Defense, 2025. Proceedings to appear. arXiv:2505.04843

Projects

Large Language Models (LLMs) are Autonomous Cyber Defenders — Python

Explainable AI (XAI) in the context of Cyber Security | LLMs + RL Agents as defenders for CAGE 4 Challenge

- Designed and implemented an adversarial simulation agent to stress-test the robustness of **LLMs** and **Reinforcement Learning (RL) based** autonomous cyber defense systems against real-time service disruptions
- Extracted and embedded action-reason statements using **OpenAI's Embeddings API**, converting LLM-generated rationales into high-dimensional vectors for downstream clustering
- Applied **unsupervised machine learning (K-Means, DBSCAN, PCA)** with feature standardization and dimensionality reduction to uncover interpretable behavioral clusters in agent decision-making
- Built a reasoning summarize driven by **OpenAI GPT-4o** that converts clustered behavior into human-readable defense strategies via **advanced prompting strategies**, advancing explainability and transparency in LLM-driven autonomous systems

Travel Agent — React Native (TypeScript), NativeWind, Redux Toolkit, FastAPI, Supabase, PostgreSQL, Python

Mobile Full-Stack Multi-Agent Travel Planner with Tool Use & Multi-Turn LLM Reasoning

- Architected a modular **multi-agent LLM pipeline** using **AutoGen** to generate travel itineraries based on content that is accurate, up-to-date, semantically relevant, and aligned with user preferences and constraints
- Built a robust **agentic web scraping** module by integrating **Perplexica** for search-based discovery, **Playwright** for dynamic content rendering, and **Trafilatura** for clean content extraction
- Combined **LLM-as-Judge** with **natural language processing (NLP), machine learning (ML)**, and **rule-based filtering** to evaluate content quality and ensure alignment with user travel needs
- Curated and annotated a custom **Constraint-Fulfillment and User-Preference Travel Planning Dataset** of real travel queries and content chunks to supervise and evaluate judgment layers across multiple quality dimensions
- Fine-tuned the Itinerary Generation Agent using **LoRA** to condition LLM outputs on user travel needs, reducing hallucinations and improving personalization across both structured JSON schemas and rich-text itinerary narratives

Experience

Full Stack Engineer — *June 2023 – Jan 2025*

Tech4Good Lab

- Led a cross-functional team of 10 engineers and designers to develop Pathways, an AI self-directed learning platform; increased weekly active users by **25%** in two months via dynamic UI workflows and adaptive content personalization
- Conducted **iterative user research** and **prompt optimization** grounded in self-directed learning literature, improving the quality of LLM-generated content and increasing recommendation relevance by **15%** based on user feedback
- Built a **retrieval-augmented generation (RAG) recommendation engine** using **vector embeddings** and **Pinecone**, increasing learning efficiency by **30%**, as measured by quiz accuracy, completion, and retention rates

Backend Developer Intern — *July 2021 - August 2021*

WayOps

- Collaborated with engineers and stakeholders to align backend deliverables with business goals, ensuring infrastructure changes supported product timelines and operational priorities
- Improved query performance by **40%** in a production relational database via **indexing, query restructuring**, and **strategic caching**, reducing latency and enabling high-throughput analytical workloads at scale
- Refactored database schema using advanced normalization techniques to reduce redundancy by **30%** and support long-term scalability, increasing write throughput under growing data volume and concurrent access.

Skills

- **Programming Languages:** Python, TypeScript, JavaScript, Java, C, C++
- **Web Development:** React Native (Expo), React, Next.js, SolidJS, Node.js, FastAPI, Django, Django REST Framework, Express.js, Tailwind CSS, NativeWind, Vite, HTML, CSS
- **AI/ML/NLP:** Hugging Face Transformers, TensorFlow, PyTorch, Keras, Scikit-learn, SentenceTransformers, OpenAI APIs, Gemini API, Ollama, WandB, LlamaIndex, Pinecone, Perplexica, AutoGen, AutoGen-MagneticOne
- **Data & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Chart.js
- **Databases & Infrastructure:** PostgreSQL, Supabase, Firebase, MongoDB