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*

Experience

Founding Engineer @ Ripplet | *June 2024 – Present*

- Partnered with therapists and domain experts to co-design **user-facing LLM** features and define product goals, translating clinical insights into actionable UI workflows and system boundaries
- Refactored backend schema and query logic for **40% faster retrieval latency**, and integrated responsive frontend updates to streamline therapist access to real-time client insights during live sessions
- Architected a **HIPAA-compliant, multi-agent** AI system with **multi-modal retrieval augmented generation (RAG)** pipelines to surface client narratives and **evidence-based** psychology
- Built **mobile- and web-accessible features** that deliver LLM responses in high-stakes, real-time therapist sessions

Full Stack Engineer @ Tech4Good Lab | *June 2023 – Jan 2025*

- Led a **cross-functional** team of 10 engineers and designers to develop Pathways, an AI self-directed learning platform
- Engineered a high-performance full-stack architecture (**Solid.js** + **Express.js** + **Firebase**) for real-time **LLM-driven** recommendations, reducing latency and improving frontend responsiveness for seamless user interaction
- Achieved **25% weekly active user growth** in two months through dynamic UI workflows and personalized content
- Designed and tested prompt variants through **iterative A/B testing and user research**, improving LLM-generated content quality and **increasing recommendation relevance by 15%** based on engagement metrics

Projects

Large Language Models (LLMs) are Autonomous Cyber Defenders (ACD) | Python

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

- Co-authored a peer-reviewed research paper presented at IEEE CAI 2025; published on **arXiv:2505.04843**
- 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 summarizer 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, NativeWind, Redux, FastAPI, PostgreSQL

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, travel domain specific, and aligned with user preferences and constraints in real-time
- 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 dataset of real travel queries and content chunks to supervise and evaluate judgment layers across multiple quality dimensions, including **constraint-fulfillment and user preference alignment**
- Fine-tuned the Critic 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

Skills

- **LLM Systems & AI Tooling:** AutoGen, OpenAI APIs, Gemini APIs, Google AI Studio, Hugging Face Transformers, Ollama, LlamaIndex, Pinecone, Chroma, Weights & Biases (W&B), LastMile AI, Scikit-learn, PyTorch, TensorFlow, Keras
- **Web & Mobile Development:** React Native (Expo), React, Next.js, SolidJS, Node.js, FastAPI, Express.js, Django, Django REST Framework, Vite, Tailwind CSS, NativeWind, HTML, CSS/Sass
- **Backend & Infrastructure:** PostgreSQL, Firebase (Firestore), Supabase, MongoDB, Docker, NGINX, Google Cloud Platform (GCP), AWS
- **Programming Languages:** Python, TypeScript, JavaScript, Java, C, C++
- **Data & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, Chart.js
- **DevOps & Tools:** Git, GitHub Actions, CI/CD, Postman, NPM, Jira
- **Experimentation & Evaluation:** A/B Testing, Prompt Evaluation, User Feedback Analysis