

Yousuf Golding

San Francisco Bay Area

+16507137793 | yousuf@goldin.gs | <https://yoolee.xyz/> | [yoolee](#) | [yousuf-golding](#)

Personal Profile

I'm a recent graduate with a master's in Natural Language Processing from UC Santa Cruz, focusing on LLM evaluation methodology, applied multimodal systems, and AI in educational technology. I am searching for software engineering roles building AI applications, with an interest in evaluations of model behavior in deployment contexts.

Education

University of California, Santa Cruz

San Jose, California

MS in Natural Language Processing (NLP)

Sept 2024 - Dec 2025

- **Courses:** Deep Learning for NLP, Data Science and Machine Learning Fundamentals, Advanced Machine Learning for NLP, Conversational Agents, Projects in AI, Natural Language Processing (1, 2, 3)

University of California, Santa Cruz

Santa Cruz, California

BS in Applied Mathematics

Sept 2020 - June 2024

- **Courses:** Intro to Probability Theory, Intro to Proof and Problem Solving, Intro to Number Theory, Intro to Dynamical Systems, Mathematical Methods for Engineers, Foundations of Scientific Computing, Mathematical Modeling (1 & 2), Cryptography

BlueDot Impact

Remote

AGI Strategy (course certificate) 

June 2026

- 30-hour intensive program. Mapped technical AI trends to future capabilities, developed threat models through kill chain analysis, and applied defence-in-depth frameworks to design protective interventions.

Skills

Languages Python, SQL, TypeScript

ML/NLP PyTorch, Hugging Face Transformers, scikit-learn, NumPy, pandas

LLM development OpenAI and Anthropic APIs, LangChain, LangGraph

Backend and databases FastAPI, Flask, SQLite

Development tools Git, Docker, pytest

Projects

Orion – AI-Assisted Project Tracking and Reporting

June – August 2026

- Designed and developed a progress-reporting application with AI coding assistance, turning Git activity, task checklists, and coding-session summaries into updates for reviewers.
- Built a Python reporting pipeline with conditional LLM summarization, secret redaction, report previews, and SQLite state tracking. Integrated delivery through a hosted dashboard, Discord, and Slack.
- Developed and deployed a React/TypeScript dashboard with project-scoped access and two-way reviewer discussions. Added automated backend and frontend checks through GitHub Actions.
- Stack: Python, SQLite, Anthropic API, React, TypeScript, Docker, Fly.io.

Dual Use Navigator (Apart Research AlxBio Hackathon)

April 2026

- Built an open-source web application that ingests dual-use export control lists from three jurisdictions (US eCFR XML, EU Annex I and CELLAR XHTML, Australian DSG 2024 EPUB) into a unified Pydantic schema with side-by-side regulatory comparison, multilateral regime overlays (Australia Group, Wassenaar, NSG, MTCR), and EU General Export Authorisation applicability.
- Covered 3,384 regulatory entries with primary-source attribution; identified a cross-source EU regulatory inconsistency (1C352 EUGEA case) not visible in any single source.
- Stack: Python, Flask, Pydantic, BeautifulSoup, Leaflet. Documented in a 10-page technical report covering methods, limitations, dual-use risks, responsible disclosure, and ethical considerations.

QUD Empowered LLM Co-Writing Tool (Capstone)

Sep 2025 – Dec 2025

- Built, with a capstone team, a full-stack NLP system (FastAPI, React, PostgreSQL) that generates diverse LLM summaries, extracts implicit Questions Under Discussion (QUDs), and classifies student coverage using embedding similarity and LLM-as-a-judge verification.
- Designed the relevance-classification study and led a human evaluation with 100 participants across 10 articles, demonstrating statistically significant alignment between AI-ranked questions and human importance judgments ($r = 0.28$, $p = 0.002$).

- Co-developed multi-agent system using LangChain/LangGraph and SIM-RAG for appliance repair support.
- Integrated Llama-3 for orchestration, BLIP for vision analysis, and a Flan-T5 critic trained on iFixit synthetic data; evaluated on MyFixit dataset and synthetic multi-turn dialogues.

EduMUSE – Multimodal Educational Assistant

Apr 2025 – Jun 2025

- Built interactive study system using CrewAI, where users upload PDFs, highlight text, and trigger AI workflows (summarization, quiz generation, contextual QA, knowledge retrieval, podcast-style audio).

Comparative Analysis of LLMs for Mental Health Counseling

Jan 2025 – Mar 2025

- Evaluated proprietary and open-source LLMs (ChatGPT, Claude, LLaMA, Deepseek) on counseling capabilities using classification and generation tasks with comprehensive metrics.
- Developed annotation guidelines and analyzed inter-rater reliability to systematically assess therapeutic response quality across models.

Work Experience

HigherSummit

Oakland, California

Machine Learning Research Intern

Sept 2023 - 2024

- Researched educational frameworks and theories compatible with AI/NLP methods: <https://arxiv.org/abs/2402.01770>
- **Technical Skills:** Python, NumPy, Git, LaTeX, Prompt Engineering, GPT

Frontier AI

San Francisco, California

Machine Learning Engineering Intern

June 2023 - Sept 2023

- Built full-stack quiz software using LLMs as part of framework for enterprise micro-learning.
- **Technical Skills:** Python, PyTorch, NumPy, Git, LaTeX, Prompt Engineering, GPT, Open AI APIs

Chabot Space and Science Center / NASA Ames Visitor Center

San Francisco, California

Science Educator

2016 - 2020

- Galaxy Explorers program @ NASA Ames Visitor Center. Designed and built exhibits for science education and outreach (physics and chemistry). Worked as an Exhibit Presenter (676 hours), earned the President's Volunteer Service Award (gold and silver).

Publications

Extending Interactive Science Exhibits into the Classroom using Anthropomorphized Chatbots and Bloom's Taxonomy

Yousuf Golding

arXiv:2402.01770, 2024