Sami Halabieh

925-964-7401 | samisam@berkeley.edu | samisam.me

EDUCATION

University of California Berkeley — Berkeley, CA
Bachelor of Arts in Computer Science, Data Science, and Applied Mathematics — *Aug. 2023 – May 2027*
- GPA: 4.0/4.0

EXPERIENCE

CVS Health — May 2025 – August 2025
Incoming Software Engineering Intern — *Scottsdale, AZ*

Honda Research Institute — January 2025 – Present
Undergraduate Researcher — *Berkeley, CA*
- Achieved a top-100 ranking in the ARC AGI Challenge by designing a dual-headed CNN with symbolic logic verification in Python and PyTorch, reducing training time by 41% compared to peer models.
- Engineered a data pipeline to process raw JSON into tensors and automated training data augmentation using domain-specific language generator functions, tripling the training dataset size.

California Resource Board — January 2025 – Present
Software Engineering (Machine Learning) Intern — *Berkeley, CA*
- Building an interactive UI (React) with a Node.js/Express backend to handle routing, leveraging NLP and LangChain (LLMs) to interpret user inputs, generate search queries, and retrieve literature on chemical contaminants in wastewater.
- Integrating REST APIs (CrossRef, Springer) and scraping automation tools (Selenium) to retrieve and process academic articles, utilizing PDF extraction tools and storing key information in Firebase and vector databases (Pinecone) in the backend.
- Cross-comparing machine learning models to predict wastewater contaminant levels under extreme climate conditions, achieving 96 percent validation accuracy.

Topcon — February 2023 - August 2024
Data Science and Automation Intern — *Livermore, CA*
- Developed interactive dashboards displaying KPIs in PowerBI and Tableau powered by regression models deployed on AWS EC2, accurately predicting part failures. Reducing part testing, saving over $200,000 in one fiscal year.
- Automated an end-to-end ETL pipeline from Excel to PostgreSQL hosted on AWS RDS, streamlining data ingestion and improving data reliability for predictive modeling.
- Embraced agile methodologies to iteratively develop data models, collaborating with cross-functional teams to deliver actionable, data-driven insights.

PROJECTS & OUTSIDE EXPERIENCE

EigenFood | *Python, React Native, Expo Go, FastAPI, Uvicorn*
- Developed a fullstack iOS, Android, and web app allowing users to find healthy food using React Native with Expo Go.
- Built a FastAPI and Uvicorn backend, integrated with Google Places API and Gemini for real time data retrieval, geolocation, and personalization.
- Deployed the web frontend on Vercel, scaling to support over 1,000 concurrent users with zero downtime.

Big Gem | *Go, TypeScript, gRPC-Web, Node.js, OpenAI API*
- Developed an interactive browser game using React/Next.js with TypeScript, leveraging gRPC-Web for real-time communication and integrating a LLM (OpenAI API) to generate adaptive narratives and dynamic in-game dialogue.
- Created scalable microservices using Go/Node.js with gRPC, optimizing concurrency and throughput for game logic.

Guitar Hero | *Java, Maven, StdAudio, JUnit, Git*
- Developed a Java musical instrument simulator using the Karplus-Strong algorithm to mimic a plucked guitar string.
- Implemented a double-ended queue as a ring buffer that uses noise initialization, iterative averaging, and a decay factor to capture harmonic nuances.
- Integrated an interactive GUI with real-time audio playback (StdAudio).

TECHNICAL SKILLS

Languages: Java, Python, C++, SQL (Postgres), JavaScript, HTML/CSS, GoLang, LaTeX
Technologies: React, Node.js, Flask, JUnit, LangChain, Three.js, Rest API, CrossRef, Springer, Selenium, Firebase, Pinecone, Tensorflow, Pandas, Kubernetes, SQLite
Developer Tools: Git, Docker, Unix, Azure, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse
Courses: Deep Learning for Computer Vision, Optimization Models, Data Structures and Algorithms, Object-Oriented Programming, Probability Theory, Abstract Linear Algebra, Putnam