Abhinav Allam

5427 Matthew Terrace, Fremont, CA (US CITIZEN)
510-890-8272 abhinavallo123@gmail.com Abhinav Allam github.com/kineticdirt

Skills

Programming & Development:Java, Kotlin, Python, JavaScript, C/C++, Scala, TypeScript, Android Development, iOS Development, Firebase, Node.js, Flask, Spring Boot, RESTful APIs, Microservices
Cloud, DevOps, & Tools: Google Cloud Platform (GCP), AWS EC2, Docker, Kubernetes, Git, Postman, Gradle, Maven, IDE's, Agile, Kanban, Scrum
Databases & Analytics: SQL, SQLite3, Postgres, Elasticsearch, Amazon Athena, Pandas, NumPy, RVC Models, Stable Diffusion Models, Neural Networks, PyTorch

Resume Summary

Dynamic, detail-oriented professional with proven ability to excel under pressure through strong problem-solving, troubleshooting, and communication. Passionate about continuous learning, improving myself, and proactively contributing in collaborative environments.

Experience

Celona, Inc. **March 2025 – Current**
Mobile Development Engineer *Cupertino, CA*

- Utilized **Kotlin (Android) and Swift (iOS)** with **geofencing APIs** to develop advanced mobile applications, thereby **enabling precise location-based data services and enhanced user experience** for enterprise clients.
- Leveraged **Celona's internal APIs** and explored **SIM and 5G-related technologies** to implement robust connectivity solutions, resulting in **optimized network performance** and streamlined device management within private mobile networks.
- Engineered and deployed custom **RESTful APIs** using **Firebase** to enable seamless data synchronization and real-time communication between mobile application and backend services.

Healthy Vibes **November 2024 – February 2025**
Member of Technical Staff *Remote*

- Developed a **headless Android app** to manage devices and achieve robust **Linux sync**, via integration with **Zoll devices**.
- Automated **ASP.NET API** calls and built Python scripts for real-time integration testing, via custom scripting on **Google Cloud Platform**.
- Created an **Agentic model prototype** to auto-generate and test code, via integration with the **Claude 3.5 Haiku** API.
- Developed a high-performance **HL7 scraper** in **Python** to parse and transform legacy clinical data into a structured **JSON** format, enabling its use in modern **AI** and analytics pipelines.
- Engineered a robust **AIJsonToPDF** utility to automatically generate human-readable clinical reports from model outputs, reducing manual report creation time from **15 minutes to under 2 seconds**.

Support Vectors **November 2024 – January 2025**
AI Research Intern *Fremont, CA*

- Utilized **Python, Langchain, and LlamaIndex** to construct and refine multiple **Retrieval-Augmented Generation (RAG) models**, thereby **enhancing information retrieval accuracy and the depth of contextual understanding** for complex query processing.
- Employed **Python and the Anthropic API** to advance the development of **Staff Agentic Reasoning models**, leading to improved autonomous decision-making capabilities and more efficient automation of complex analytical tasks.

Cequence Security, Inc. **June 2023 – October 2023**
Data Science Intern *Sunnyvale, CA*

- Developed and deployed a **Time-Delta model** to detect subtle bot patterns, via **PyTorch** with **CUDA**.
- Engineered preprocessing and feature extraction to lift accuracy from **18%** to **32%**, via **Pandas**, **NumPy**, cross-validation, and hyperparameter tuning.
- Automated daily reporting by compiling **XML** test results and reducing manual effort by **50%**, via streamlined scripting and **Google Cloud** integration.
- Developed a visual testing framework in **Python**, leveraging **Appium** for automated UI navigation and the **Pillow/pixelmatch** libraries for pixel-level screenshot comparison to safeguard mobile UI integrity.
- Engineered an automated quality gate that guaranteed **design fidelity** by catching visual regressions before release, reducing manual design review time by **100%** for all covered screens.

Celona, Inc. **June 2022 – September 2022**
App Development Intern *Cupertino, CA*

- Developed an **Android app** to streamline **Celona 5G Access Point installs** and cut setup time by **24%**, via **Kotlin** and **Android Studio**.
- Conducted requirements analysis, drafted **UML diagrams**, and integrated access point settings with secure **OAuth2 authentication**, via **REST API** integration.
- Added a **QR/barcode scanning** feature to automate data capture and boost productivity by **23%**, via the **ZXing** library.

Stealth Security, Inc. **June 2021 – September 2021**
Software Intern *Sunnyvale, CA*

- Developed a **JavaScript app** to auto-extract data from **Swagger specs** and support **Bot Defense**, via custom parsing logic.
- Automated **JSON pointer generation** to achieve a **12%** faster setup time for new clients, via precise algorithmic extraction.
- Built a **React UI** and adopted agile practices to streamline configuration, via **NodeJS REST APIs**, **Git**, and **JIRA**.
- Architected a microservices-based application by developing a **Python** worker service to handle compute-intensive image processing, offloading tasks from the primary **Node.js / Express** API.
- Improved main API responsiveness by **40%** by isolating long-running tasks, ensuring the **Node.js** event loop remained non-blocking and available to serve user requests.

Projects and Development

Rosehack @ UCR 2024 | ***Google BERT, Python, Machine Learning*** **April 2024**

- Developed **EQUALITY EYE** to detect hate speech and secure **first place** (Social Impact), via AI-driven analysis.
- Managed data collection, cleaning, feature extraction, and training to achieve high accuracy, via curated **Reddit datasets** and **Google's BERT**.
- Deployed a **browser extension** for real-time Twitter analysis and refined the system to **94%** accuracy, via user feedback and A/B testing.

Cequence Hackathon 2023 | ***Java, C/C++, LLM*** **August 2023**

- Developed an **AI chatbot** to simplify user interactions and secure **second place**, via integration with Cequence's **UAP Platform**.
- Researched and implemented **NLP techniques** to build a conversational interface, via leveraging **LLM backends**.
- Prototyped in **Java** and **C/C++** and refined the solution to streamline navigation and boost efficiency, via iterative testing and targeted improvements.

Education

University of California, Riverside **Sept 2020 – June 2024**
B.Sc. in Computer Science *Riverside, CA*

Continuing Education: Support Vectors **August 2024 – January 2025**
AI Research Intern *Fremont, CA*

- Completed advanced coursework to master **neural networks**, **LLMs**, and **generative AI**, via rigorous academic training.
- Built expertise in **CNNs**, **RNNs**, **Transformers**, **GANs**, and **Diffusion Models** via hands-on projects and labs.
- Optimized models to boost training efficiency and accuracy, via **back-propagation** and **regularization** techniques.