

Biruk Abebe

baap2023@mymail.pomona.edu | (951) 902-8085 | [linkedin.com/in/biruk-abebe-aa3b68212/](https://www.linkedin.com/in/biruk-abebe-aa3b68212/) | github.com/abebebiruk

EDUCATION

Pomona College, Claremont, CA

Class of 2027

B.A. Computer Science, Math (GPA: 3.7)

Relevant Coursework: Python Programming; Data Structures; Algorithms; Database Systems; Neural Networks; Advanced Algorithm; Computer Systems; Haskell; Theory of Computation; Linear Algebra; Statistical Linear Model; Combinatorial Mathematics

SKILLS

Languages: Python, Java, SQL, Ruby, Haskell, R, TypeScript, HTML, CSS

Frameworks & Tools: SpringBoot, Ruby on Rails, FastAPI, React, Next.js, REST APIs, LangChain4J, Server-Sent Events, GitHub

Databases & DevOps: AWS(Lambda, ECS, DynamoDB, CloudWatch, SQS), Terraform, Kubernetes, Docker, PostgreSQL, SQLite

AI/ML & Data: ReAct agent & LLM tool-calling, RAG, Vector DB (Pinecone), TensorFlow, scikit-learn, Random Forests, Neural Networks, Decision Trees, kNN, Matplotlib, NumPy, Prompt engineering

EXPERIENCES

Stealth Startup

Jan 2026 – Present

Backend Engineering Intern (AI Systems)

Remote

- Re-architected an LLM content-generation pipeline from a multi-step, slot-based workflow into a single batched call, cutting API calls by ~50% and lowering latency and cost.
- Designed a competing-model pipeline for content generation, using independent AI models for quality control, where one AI model produces content and a separate model independently reviews and scores it, filtering weak outputs to improve accuracy without added manual review.
- Building a RAG (retrieval-augmented generation) pipeline that grounds AI outputs in trusted external sources instead of the model's own memory, keeping generation factually accurate and far less prone to hallucination.

JPMorgan Chase

June 2026 – August 2026

Software Engineer Intern

Palo Alto, CA

- Built a distributed message-replay service that re-sends failed data updates to external partner systems, keeping internal and external records in sync by ID or date range across three downstream services.
- Engineered an AI chatbot as a ReAct agent (LangChain4J) that interprets user requests and calls the app's backend tools to act on them, running independent tool calls in parallel to cut response time and streaming answers in real time.
- Deployed services on AWS (Lambda, ECS, SQS) with 100% unit-test coverage and end-to-end tests, tracking execution through CloudWatch and Datadog to trace and debug issues across environments.

Heyy Property

July 2025 – August 2025

Software Engineer Intern

Washington, DC (Hybrid)

- Built a booking reschedule/cancellation service (Ruby on Rails, PostgreSQL, AWS) with real-time Google Calendar sync that automatically creates and updates events when a booking changes.
- Improved AI-generated property listings by grounding descriptions in real location data via the Google Places API, cutting inaccurate content ~70%.
- Developed an internal task-management service supporting task creation, assignment, and status tracking over REST APIs, streamlining the team's operational workflow.

PROJECTS & AWARDS

1st Place – JPMorgan Chase Code for Good Hackathon

October 2025

Digital community platform for Team IMPACT

Plano, TX

- Built a full-stack community platform (Flask + PostgreSQL backend, Next.js + TypeScript frontend, AI-generated newsletters via OpenAI API) enabling groups, updates, and threaded discussions.
- Won 1st of 11 teams in the Team IMPACT division (40+ teams event-wide) for innovation and community impact.

p-TeamSync – P-ai

September 2025

Scheduling platform for Pomona-Pitzer athletics

Claremont, CA

- Built the backend of a scheduling platform for Pomona-Pitzer student-athletes: RESTful APIs for filtering and scheduling with secure role-based access control, on a 5-member cross-functional team.