

Ali Fayed

San Jose, CA | 209-637-3465 | alifayed.h@gmail.com | linkedin.com/in/alifayed |

Education

San Jose State University

Bachelor of Engineering, Software Engineering

Relevant Coursework: Artificial Intelligence, Machine Learning w/ Big Data, Data Science, Operating Systems, Data Structures & Algorithms, Information Security

May 2025

San Jose, CA

Technical Skills

Languages: Python, Typescript, Javascript, C/C++, Java, SQL, C#

Libraries: Tensorflow, Scikit Learn, OpenAI, Gemini, MongoDB, PostgreSQL, Firebase, React

Tools: Git, AWS, Docker, Kubernetes, GCP, Linux

Experience

NetCoro

May 2024 – Sep 2024

Software Engineer Intern

Santa Clara, CA

- Automated configuration of **150+ network devices** by scripting Cisco DevNet API calls in Python, trimming manual setup time \approx **80% and cutting errors to <1%**
- Fixed a production logout bug in **under two hours** using Python's pdb, restoring stable auth flow for all active users
- Wrote a concise 10-page run-book & API guide that shrank new-contributor ramp-up from one week to two days

Projects

Resumate | *Node.js, MySQL, OpenAI*

Apr 2024 - Present

- Built LLM-driven agent that parses job posts and auto-crafts targeted resumes, delivering an **18% uplift** in interview call-backs for early adopters
- Designed a custom serialization layer (SQLAlchemy \rightarrow S3) that streams resume versions for 100+ concurrent users/request, achieving **90% storage-allocation efficiency** and zero data-loss incidents
- Implemented an NLP pipeline (spaCy + custom word-cloud generator) with **98% keyword-extraction accuracy**, surfacing skills gaps and generating data-driven feedback in **<3s per resume**

2m | *Node.js, OpenAI, GoLang, HTML/CSS/JS*

Jan 2024 - Present

- Engineered a high-throughput, multithreaded Python/Go service that aggregates real-time CS2 skin pricing from 6 public APIs, sustaining **50K req/s** with **<150 ms latency** during peak traffic
- Architected and deployed a web marketplace on AWS (PostgreSQL + Redis) that serves **200+ active traders** and processes **20K+ daily transactions** across a multi-billion-row dataset
- Integrated an LLM-powered support chatbot (OpenAI / LangChain) with persistent SQL chat-history, cutting average customer-issue resolution time by **42%**

Ecobite | *Python, Node.js, Tensorflow, Gemini*

Feb 2025 - Feb 2025

- Fine-tuned a transfer-learning CNN in TensorFlow on a 5GB image dataset, achieving **83% weight-estimation accuracy** through k-fold cross-validation and stratified testing.
- Augmented the vision model with Google Gemini to infer food density, outperforming a purpose-built ML baseline and enabling seamless horizontal scaling for real-time API calls.
- Accelerated training throughput **240x** by orchestrating distributed GPU jobs on brev.dev, slashing wall-clock training time from hours to minutes while trimming cloud-compute spend

Mai | *Python, Firebase, OpenAI, Typescript, React*

Dec 2024 - Feb 2025

- Fine-tuned a 7 B-parameter GPT-4-tier model with LoRA adapters on a curated mentorship corpus, increasing answer precision from **78% \rightarrow 94%** while trimming per-query inference cost **1.5x**
- Architected a sharded MongoDB + Redis stack that persists 12K+ dialog turns for 100+ monthly users, sustaining **<80 ms** median read latency and **99.995% uptime** during peak load
- Shipped a mobile-first React/TypeScript UI with live WebSocket hints, hitting **96 Lighthouse** and **40% week-1 retention**