

Rayid Ali

New York, NY | 864-765-4848 | rayidali3@gmail.com | rayidali.github.io | github.com/rayidali | linkedin.com/in/rayidali

Skills

Languages: Python, TypeScript, JavaScript, Swift, Go, SQL, C/C++

Frontend & Mobile: React, Next.js, TailwindCSS, native iOS (Swift, Capacitor), Angular, HTML/CSS, RxJS

Backend & APIs: Node.js, Express, FastAPI, Flask, REST, GraphQL, WebSockets

AI/ML: OpenAI API, LangChain, RAG, Prompt Engineering, PyTorch, TensorFlow, HuggingFace, scikit-learn

Data: PostgreSQL, MongoDB, BigQuery, Snowflake, Pinecone, pgvector

Cloud & DevOps: AWS, GCP, Vercel, Docker, Kubernetes, Terraform, GitHub Actions, GitLab CI, Airflow, MLflow

Experience

AI Engineer | Gensol Inc | New York, NY

Dec 2024 – Present

- Cut bid preparation from 2 hours to 5 minutes per drawing across 50+ contractor drawings by building a Claude Vision pipeline that parses DXF/DWG geometry and resolves detected device symbols into a structured takeoff
- Shipped the product end to end as sole engineer across 13 REST endpoints and 4 AI features by building a React and TypeScript frontend on a FastAPI service layer
- Handled legends that differ on every project by decomposing each sheet into regions and matching symbols against that drawing's own legend rather than a fixed library, routing low-confidence matches to human review
- Caught extraction regressions before every release by building a labeled ground-truth set from production drawings and replaying each prompt and model change against it to track accuracy

Research Assistant | Clemson University | Clemson, USA

Aug 2023 – May 2024

- Beat the baseline navigation algorithm by 15% on search-and-rescue routing in complex terrain by implementing a multi-agent reinforcement learning approach (Deep Q-Learning, graph neural networks) with MIT Lincoln Laboratory
- Validated flight-path planning across 10+ coordinated agents before any hardware trial by building OpenAI Gym simulation environments with GNN-based inter-agent communication

Software Engineer Intern | Atkins Global | Bengaluru, India

Feb 2022 – Jul 2022

- Replaced static PDF reporting for survey teams with an interactive 3D terrain interface by building a React and Flask application over ML-classified LiDAR point-cloud data in PostgreSQL
- Cut manual contract review 70% and saved 200+ analyst hours per quarter across 1,000+ contracts by deploying a BERT clause-extraction service that fed a reviewable queue

Software Engineer Intern | Gulf Marvel | Hyderabad, India

Jan 2021 – Jun 2021

- Shipped the company's public website end to end by building responsive pages in HTML, CSS and JavaScript with cookie consent, session handling, and analytics tracking wired in
- Consolidated data from multiple source systems into a single reporting warehouse by building ETL workflows in Informatica and SQL, replacing manual spreadsheet exports

Projects

Cinechrony | AI Film-Discovery App, iOS + Web

cinechrony.com

- Cut worst-case latency 88%, from 223s to 26s, by replacing serial model retries with parallel execution that takes the first completion, in a Gemini pipeline that names every film in a shared TikTok or Reel
- Cut database reads ~99% per repeat session, from ≈ 270 to near zero, by adding write-invalidated caches, write-time denormalization, and client-direct queries
- Shipped solo to the App Store across web and native iOS (Next.js 15, Capacitor, Swift share extension) with 858 automated tests gating every release

Pdf2Video | AI Research Paper to Video Converter

pdf2video.onrender.com

- Cut a 4+ hour manual editing workflow to under 10 minutes by architecting a 6-stage async pipeline with FFmpeg preprocessing, stage-level checkpointing, and fault-tolerant retries

Network-Based Recommendation System

Published | IEEE ICETCI 2022

- Improved prediction accuracy 18% and cut RMSE 12% against collaborative filtering baselines by designing a graph-based recommendation model over large Amazon datasets

Education

Clemson University, Clemson, SC – MS in Computer Science

May 2024

Research: Multi-agent reinforcement learning for autonomous systems | Collaborated with MIT Lincoln Laboratory

Jawaharlal Nehru Technological University, Hyderabad, India – B.Tech in Computer Science

May 2022