

ANIRUDH RAGHAVENDRA MAKULURI

(602)880-5334 | anirudh.makuluri@gmail.com | www.anirudh-makuluri.xyz | linkedin.com/in/anirudh-makuluri/ | Tempe, AZ, USA

Work Authorization: Eligible for full-time employment in the US for up to 36 months under OPT/STEM OPT beginning May 2026.

SUMMARY

Full-stack software engineer with 1.5+ years of experience building production AI features and scalable distributed systems. I help teams ship reliable features quickly by owning them end to end—from React UI to Node.js APIs and LLM integrations. Focused on reducing release friction, improving performance at scale, and delivering stable production software. Passionate about AI-driven solutions and agentic systems. Currently pursuing M.S. Computer Science at ASU (3.96 GPA).

EDUCATION

Arizona State University	Tempe, AZ
<i>Master of Science in Computer Science (GPA: 3.96/4.00)</i>	<i>August 2024 – Present</i>
Birla Institute of Technology and Science, Pilani	Hyderabad, India
<i>Bachelor of Engineering in Electronics and Instrumentation</i>	<i>August 2018 – May 2022</i>

TECHNICAL SKILLS & CERTIFICATIONS

Programming Languages: JavaScript, TypeScript, Java, C#, Python, C++, Kotlin, SQL

Frontend: React, React Native, Next.js, Redux, HTML, CSS, Tailwind

Backend & Databases: Node.js, Express.js, REST APIs, PostgreSQL, NoSQL (Firebase Firestore, MongoDB), MSSQL

Cloud & DevOps: Google Cloud Platform (GCP), Amazon Web Services (AWS), Docker, Kubernetes, GitHub Actions

Machine Learning & AI: TensorFlow, PyTorch, Scikit-Learn, LangChain, Prompt Engineering, RAG, LLMs (Gemini, Bedrock, Mistral), NLP

Tools & Others: Git, GitHub, Jupyter Notebooks, Agile/Scrum, OAuth, Linux/Unix, Data Structures

Certifications: Oracle Certified Generative AI Professional (August 2025), Machine Learning Specialization (May 2024)

PROFESSIONAL EXPERIENCE

Shardings

June 2022 – August 2023

Software Engineer

- Architected and operated full-stack platform (React, React Native, Node.js, Firebase Firestore) serving 12,000+ concurrent users; achieved 99.5% uptime through blue-green deployments enabling zero-downtime releases, real-time monitoring dashboards tracking active rooms and user sessions, and on-call incident response reducing MTTR to under 15 minutes.
- Optimized image loading performance by implementing Google Cloud CDN with cache invalidation policies; reduced image load latency from 3 seconds to 800 milliseconds, improving perceived performance across 12,000+ users.
- Integrated ChatGPT with Google Cloud Speech-to-Text to automatically transcribe and summarize live standup meetings, reducing post-meeting overhead and saving teams approximately 5 hours per week.
- Established CI/CD pipelines using GitHub Actions with automated testing and deployment processes; implemented automated rollback on error threshold, ensuring 99%+ deployment success rate and reducing release friction.
- Mentored 5+ junior engineers through code reviews, pair programming, and technical design discussions for real-time features and deployment strategy.

Shardings

July 2021 – December 2021

Software Engineer Intern

- Implemented full payment gateway with Razorpay and Stripe integrations supporting 3 pricing tiers (Basic, Standard, Pro); designed secure webhook handling, idempotency, and error recovery for reliable transactions.
- Developed Google Workspace Add-on using JavaScript and Google Apps Script APIs to integrate with external calendar services; built real-time room/resource scheduling interface with seamless data synchronization.

PROJECTS

SmartDeploy — Agentic DevOps Platform

Next.js, Firebase Firestore, Docker, Google Cloud Run, Amazon Bedrock (Mistral), AWS EC2

- Architected full-stack agentic deployment platform (Next.js, Node.js WebSocket server, Amazon Bedrock) deployed on AWS EC2 t3.micro using Docker Compose; enabled one-click deployment of 20+ GitHub repos to Google Cloud Run, reducing developer setup friction by 80%.
- Engineered AI-powered code analysis pipeline using Amazon Bedrock (Mistral) with GPU-optimized inference; achieved 90% accuracy in build configuration generation through systematic prompt engineering and LLM fine-tuning.
- Designed Docker multi-stage builds for Next.js frontend and WebSocket server with optimized image sizes; implemented Docker Compose orchestration enabling local development environment parity with production deployment architecture.

Chatify — Enterprise-Grade AI Chat Platform

React, React Native, Node.js, Gemini API, Zep, Firebase

- Integrated Gemini API (Gemini 2.5 Flash) to power AI chatbot with personalized responses; designed knowledge graph architecture using Zep to maintain user context and conversation history for coherent RAG-based interactions across sessions.
- Built real-time cross-platform chat application with Socket.IO messaging supporting one-on-one and group chats, media sharing, scheduled message delivery with timezone support, and optimized infinite scroll pagination; achieved sub-second message delivery and seamless user experience across React and React Native clients.