

HIMANSHU NIMONKAR

530-204-8976 | hnimonkar@gmail.com | [linkedin.com/in/himanshunimonkar](https://www.linkedin.com/in/himanshunimonkar) | github.com/himanshu-nimonkar | himanshu-nimonkar.github.io

EDUCATION

University of California Davis <i>Master of Science in Computer Science, GPA: 3.7/4.0, TA (ECS 130) & 3 Research Fellowships (66,000\$)</i>	Davis, CA <i>Expected: December 2027</i>
K. J. Somaiya College of Engineering <i>Bachelor of Technology in Information Technology, Honors in AI, GPA: 3.91/4.00, 1,500\$ Scholarship, Top 5%</i>	Mumbai, MH <i>Aug 2020 – May 2024</i>
Courses: Distributed Systems, Operating Systems, Algorithms, Data Structures, Software Engineering, Database Systems, DevOps	

TECHNICAL SKILLS

Languages & Databases: JavaScript/TypeScript, Python, Go, SQL (PostgreSQL, MySQL), NoSQL (MongoDB), Supabase, HTML/CSS
Cloud & Infrastructure: AWS, GCP, Cloudflare (Workers, R2, D1, KV, AI, Tunnels, Wrangler, Durable Objects, Pages), Firebase
DevOps & CI/CD: Docker, Kubernetes, CI/CD (GitHub Actions, Gitlab, Jenkins), Terraform, Ansible
Data Engineering: Kafka, Apache Spark, Airflow, Databricks, Snowflake, Hadoop
Observability & Analytics: Redis, ClickHouse, Elasticsearch, Kibana, Logstash, Grafana, Prometheus, Sentry, Power BI, Datadog
Frameworks & Libraries: React, Next.js, Node.js, Express.js, Flask, Tailwind CSS, Redux, Vite, Nginx, Django
AI/ML, Testing & Tools: PyTorch, Hugging Face, LangChain, MCP, RAG, REST/WebSockets, Jest, Vitest, Playwright, axe-core, Figma

EXPERIENCE

Software Engineering Intern <i>Cloudflare</i>	May 2026 – September 2026 <i>Austin, TX</i>
- Worked on Frontend Platform team building Medic, real-time diagnostic tool scoring Web Vitals, accessibility, React runtime health, build data, deployed via Wrangler on Cloudflare's distributed edge network with R2 backed scheduled CI/CD snapshot pipeline, reaching 97.2% coverage on 2,030+ Cloudflare Dashboard routes, cutting code review time by 19%. - Resolved three critical production defects, a Rules promise-rejection, Rspack and Vite mismatches and Speed i18n failure, on 50 call sites of 9+ dashboard surfaces, protecting 85,000 users and clearing 228,000 Sentry events, driving 93.6% drop in related support tickets using Sentry, Grafana, ClickHouse. - Refactored React Router routing and Redux state management with object-oriented design, migrating 3 major dashboard pages and feature flag SDK to Kumo design, cutting 31.4% legacy code and 18.5% render performance boost. - Migrated static Dev registry to Vite plugin host, cutting scanner warm-rebuild time from 1.1 to 0.2 seconds (81.8% improvement), then shipped i18n/l10n, all, event-log dev toolkit improving defect resolution by 58% and payload efficiency by 13.4%. - Built AI-agent skills for feature-flag auditing with (AST-based route mapping), and for replacing manual accessibility QA with 3-layer agentic, MCP-based DOM walk (axe-core) and Chrome browser automation, cutting QA cycle time by 87%.	
Software Engineering Intern <i>Talent Questor</i>	June 2025 – August 2025 <i>Eugene, OR (Remote)</i>
- Delivered fault-tolerant Stripe based payment system raising payment success to 85% and reducing response time by 30%. - Revitalized user account management with secure Node/Next.js microservices, centralizing password recovery and email preferences (AWS SES), cutting API latency by 35%, and boosting recovery completion by 80%. - Streamlined cloud deployment and quality assurance with Dockerized CI/CD pipelines, enabling reproducibility and scalability planning by optimizing real-time throughput and observability monitoring systems via Elasticsearch health dashboards.	
Graduate AI Engineer Intern <i>VIA Lab</i>	January 2025 – June 2025 <i>Davis, CA</i>
- Engineered CATF, Python-based ML framework for context-aware multivariate time-series forecasting, with a manager-worker architecture, data structures and end-to-end data pipelines for feature engineering seven attention and optimal-transport models. - Implemented MLOps including unified ETL pipelines, automated model validation, hyperparameter tuning workflows, multi-dataset CLI support, CI/CD pipelines to enable reproducible experimentation and scalable model development. - Achieved SOTA result on 5/7 dataset, improved 26/35 experiment, reduced MSE on TimesNet, iTransformer by 49.3% and 32.8% using optimal-transport algorithms (Manuscript in ACM SIGKDD Data Mining Research Conference).	
Full Stack Software Engineer Intern <i>DeepCytes Cyber Labs (UK)</i>	June 2022 – June 2024 <i>Mumbai, MH</i>
- Built CVE/CWE tracker with email alerts for network/encryption vulnerabilities, making threat monitoring 19% faster. - Constructed CLI first red-teaming automation and OSINT toolkit on Linux to stress test apps and model threats, which detected 37 new attack vectors and boosted malware defense by 29%. - Engineered Android security testing pwa to exploit permission bugs in location, camera, mic, SMS, media system functions. - Developed LangChain-based legal assistant leveraging Hugging Face transformers, which increased engagement by 25%. - Architected AI threat intelligence platform with DarkBERT to analyze dark web content for proactive security insights.	

PROJECTS

llmao.AI (UC Berkeley AI Hackathon) <i>LLMs, Agentic AI, Letta Cloud agents, Flask, Cloudflare Workers AI, RAG</i>	
- Constructed stateful agentic platform that analyzes codebases for generating API/SDK documentation using GitHub MCP Server and repository-aware embeddings. Blog: devpost.com/software/llmao-tx5j8p - Deployed natural language querying agent for Cloudflare Agents SDK, leveraging serverless edge compute, RAG pipelines, Durable Objects, Workers AI to support low-latency, secure inference at the edge.	
Charitap (Published OSS Chrome Extension) <i>React, Node.js/Express, MongoDB, RESTful APIs, Tailwind CSS</i>	
- Engineered micro-donation platform with Google OAuth, Stripe Connect, cron automations, JWT authentication that rounds up users' everyday purchases and auto donates spare change to charities. - Integrated real-time dashboards (Charts.js) and blockchain-backed ledger for transparent, secure crypto exchanges.	
AgriBot (Won SacHacks VII) <i>AI Agent, AI Gateway, Supabase, Vapi, Deepgram, Vite, Morph, ChromaDB, GEE</i>	
- Launched offline first voice AI copilot over phone calls for agricultural guidance with Vapi/Deepgram STT, FastAPI gateway, Supabase storage, Nginx handling intents, sessions and WebSockets over Cloudflare Tunnels, multithreaded Python asyncio Tool Orchestrator with Cloudflare Workers AI citation-backed reasoning engine, Morph LLM for reranking and smart routing. - Prototyped telemetry HUD PWA for live field conditions, Google Earth Engine satellite imagery, pulling weather forecasts from OpenMeteo and research embeddings RAG pipelines backed by a ChromaDB vector store and containerized on Docker.	