

Dikshant Rajput

Senior Full-Stack & AI Engineer | Tech Lead

Engineer with 4+ years building production systems, currently leading a team of four on AI commerce infrastructure across a 1.5M+ product catalog. Depth in search and agentic systems, PostgreSQL architecture, and cloud cost optimization.

dikshanraj2001@gmail.com | +91 7011869058 | Gurgaon, India

linkedin.com/in/dikshant-rajput-072228193 | github.com/dikshantraiput | dikshanraj2001.medium.com

WORK EXPERIENCE

Tech Lead (Contract) | [Jingo Technologies](#) | 05/2025 - Present | San Francisco (Remote)

AI-native personalized e-commerce platform

- Lead a team of four and own end-to-end technical decisions across the full stack, from AI product quality and evaluation strategy through infrastructure and vendor relationships with Pinecone and Google Cloud.
- Rebuilt product search around dataset attribution and deterministic filters instead of LLM-driven retrieval, resolving long-standing relevance failures and cutting query latency from over 5s to under 2s.
- Architected the AI-native shopping experience across voice, chat, and semantic search with multi-agent orchestration, conversational memory, and personalization, improving conversion and reducing search-to-purchase drop-off.
- Built backend services and APIs with NestJS, FastAPI, and PostgreSQL behind search, product, and order pipelines, plus the React Native mobile app, web frontend, and internal tooling for order management and analytics.
- Reduced cloud infrastructure costs from \$30K to \$8K per month, a 73% saving, by re-architecting services and optimizing resource utilization on GCP.

Full Stack Developer | [ONBO](#) | 08/2022 - 04/2025

- Architected full-stack SaaS applications and Python microservices across React, SvelteKit, NestJS, and Django, with Keycloak authentication and PostgREST over PostgreSQL.
- Led development of an employee onboarding platform, increasing new-hire productivity by 35%.
- Built an internal framework for managing PostgreSQL Row-Level Security, packaged as a database extension and reused across services for tenant and role-based data isolation.
- Orchestrated containerized microservice deployments with Docker and shipped AI-driven learning paths into the product.

Full Stack Developer + Intern (3M) | [Decentral System Labs](#) | 01/2022 - 08/2022

- Designed the high-level architecture of a Web3 analytics tool, ensuring scalability, performance, and secure data processing.
- Developed the backend architecture and MVP of a Web3 fantasy sports web app, integrating NestJS with smart contracts for ERC20 and ERC721.
- Built a rule engine microservice using Python for game logic processing.

PROJECTS

Jingo | **Commerce Search & Recommendations** | jingo.app

NestJS, FastAPI, PostgreSQL, Pinecone, GCP, React Native

- Built semantic search across a 1.5M+ product catalog using embeddings and Pinecone vector search, with recommendations personalized by conversational context and collaborative filtering.

- Developed production voice and chat shopping workflows with structured tool calls for search, filters, cart, and checkout, plus conversational memory for follow-up refinements.
- Built the catalog ingestion and enrichment pipeline, and owned the evaluation harness and prompt optimization with fallbacks for tool failures and weak result sets.

LocalClicky | On-Device Voice Agent | [GitHub](#) · [Product Hunt](#)

Python, Whisper.cpp, Ollama, openWakeWord, ffmpeg

- Built a Python macOS voice agent that runs offline by default, pairing Whisper.cpp speech recognition with local reasoning and vision models served through Ollama.
- Implemented on-device wake-word detection with openWakeWord and voice activity detection, plus multi-round tool calling for desktop automation, shell execution, and reminders.
- Added local video understanding and voice-controlled ffmpeg editing; launched on Product Hunt, reaching #9 Product of the Day.

Supabase Multi-Tenancy | Open-Source SaaS Reference | [GitHub](#) · [Write-up](#)

PostgreSQL, Supabase, SQL, Row-Level Security

- Built an open-source reference implementation for multi-tenant SaaS on Supabase, covering workspace isolation, Row-Level Security policies, Auth hooks, and per-tenant resource tracking.
- Published an accompanying write-up on multi-tenant database architecture and data isolation; the repository has 88 GitHub stars and 9 forks.

NeighborHoodShare | Peer-to-Peer File Sharing | [Live app](#) · [GitHub](#) · [Show HN](#)

SvelteKit, WebRTC, PeerJS

- Built a SvelteKit app for browser-to-browser file and message sharing over WebRTC and PeerJS, without centralized file storage.
- Implemented peer signalling, connection handling, and end-to-end encrypted transfers between browsers.
- Released it publicly and open-sourced it after community feedback; 132 upvotes and 105 comments on Show HN.

RECOGNITION

Toptal Verified Expert in Engineering | Top 3% talent network

#9 Product of the Day, Product Hunt | LocalClicky

250+ GitHub stars | across open-source projects

EDUCATION

Bachelor of Technology (B.Tech) | Maharshi Dayanand University, Rohtak | 08/2018 - 06/2022 | SGPA: 8.5

SKILLS

AI & Agentic: Agentic Harness, Agents Orchestration, RAG, Pinecone, Prompt Optimization, Evals

Frontend: React.js, Next.js, SvelteKit, React Native, JavaScript, PWA

Backend: Node.js, Express.js, NestJS, Python, FastAPI

Databases: PostgreSQL, MongoDB, Supabase

DevOps & Cloud: Docker, GitHub, Google Cloud Platform (GCP)

Languages: JavaScript, TypeScript, Python, C++

Interests: Chess, Cricket, Building products