Harsh Gaur

Ghaziabad, Uttar Pradesh, India | Phone: +91 7017986885 | Email: harshgaur017@gmail.com
LinkedIn: linkedin.com/in/harsh-gaur-520a73258 | GitHub: github.com/HarshGaur017 | Portfolio: harshgaur017.github.io

Summary

Full-Stack AI Developer with 2+ years building production systems for B2B commerce. Designs distributed Node.js backends – event-driven queues, GraphQL APIs, PostgreSQL with pgvector and production LLM pipelines – and the React Native clients that consume them. Experienced in semantic search, retrieval-augmented generation, OCR pipelines and bidirectional ERP integration, with mobile applications live on Google Play and the App Store.

Technical Skills

Languages: JavaScript (ES6+), TypeScript, Node.js, Python, SQL, Java, Dart, HTML, CSS
Backend: Express.js, GraphQL, Apollo Server, REST APIs, FastAPI, Sequelize, Prisma, Mongoose, JWT, Casbin RBAC, Zod
Databases & Search: PostgreSQL, pgvector, Redis, MongoDB, MySQL, Meilisearch, Elasticsearch, Firebase, Query Optimization
AI & ML: OpenAI API, LLM Integration, Prompt Engineering, RAG, Vector Embeddings, Semantic Search, OCR, Recommendation Systems
Queues & Events: BullMQ, Agenda, Redis Pub/Sub, Socket.IO, Cron Workers, Event-Driven Architecture, Microservices
Mobile: React Native, Expo, EAS Build, Redux Toolkit, React Navigation, Reanimated, FlashList, MMKV
DevOps & Tools: Docker, Git, CI/CD, Jest, Supertest, Sentry, SonarQube, Mixpanel, Postman

Experience

Mulltiply Tech India Pvt. Ltd. **Feb 2025 – Present**
Software Engineer – Full-Stack & AI *Noida, Uttar Pradesh*

- Built an **LLM-powered multilingual search parser** that compiles Hindi, Hinglish and Devanagari product queries into structured Meilisearch filter expressions, with price-intent detection and constraint validation to prevent hallucinated filters.
- Owned the **ERP integration platform** connecting the product to **Marg, Tally, Busy, Zoho and Shopify**, including full OAuth 2.0 flows, bidirectional item/pricing/order/ledger/invoice sync, and batch-wise processing for oversized payloads.
- Designed a **hybrid recommendation engine** on **PostgreSQL and pgvector**, blending vector similarity with market-basket co-occurrence, category and attribute signals; built the similarity-computation worker and cron-driven embedding jobs.
- Built an **OCR and LLM document pipeline** (PaddleOCR, Tesseract and ONNX Runtime behind FastAPI, structured by OpenAI) converting photographed invoices into order lines, with a training-data capture loop.
- Shipped **voice ordering** – speech to transcript to LLM parsing with automatic language detection – removing the keyboard from the ordering path entirely.
- Replaced hardcoded notification triggers with a **configurable rules engine** (event registry, recipient-type resolution and validation), removing deploys from routine operations changes.
- Migrated marketing-template scheduling from Agenda to **BullMQ**, resolving a duplicate-send defect; diagnosed and fixed **event-loop blocking** caused by synchronous filesystem and logging calls in a high-traffic API.
- Created **16 backend modules from scratch** – recommendations, vouchers, notification workflow, marketing templates, ERP connectors, user preference scoring, OCR training data and audit logging.
- Integrated **Razorpay and MargPay** for payment links, webhooks, refunds, QR payments and ledger statement generation.

Mulltiply Tech India Pvt. Ltd. **Sept 2024 – Feb 2025**
Software Engineer Intern – React Native *Noida, Uttar Pradesh*

- Delivered **16 screens and 12 reusable components** in React Native and TypeScript – payments, wallet, credit applications, bulk quotes, seller comparison and shop onboarding.
- Implemented a **Server-Driven UI architecture** for the home experience, letting an admin panel control which modules render and in what order, so merchandising changes ship without an app release.
- Improved **app cold-start time by approximately 40%** through lazy-loaded home surfaces, shimmer placeholders and render profiling.
- Built the app's **analytics layer** (8 Mixpanel modules covering home, onboarding, wallet, payments, profile, deals and disputes) alongside Sentry and Crashlytics instrumentation.

Projects

MulltiplyR & MulltiplyD – B2B Commerce Apps | *React Native, Expo, TypeScript, Redux Toolkit* **2025 – 2026**

- Retailer marketplace app serving **10,000+ SKUs** with real-time stock, GST invoicing, multilingual UI, deep linking, RBAC and OTA updates; **second-highest all-time contributor of 14** with 1,023 commits.
- Live on Google Play and the App Store; distributor app also live on Google Play and the App Store.

AI Business Operations Assistant – Multi-Tenant RAG Platform | *TypeScript, Express, PostgreSQL, pgvector, OpenAI* **2026**

- Multi-tenant SaaS platform letting organizations upload internal documents and query them in natural language through a **retrieval-augmented generation pipeline**.
- Implemented PDF parsing, intelligent chunking and **1536-dimension embeddings in pgvector with an IVFFlat cosine index**; enforced tenant isolation at the schema level with organization-scoped foreign keys.
- Built document versioning, JWT authentication with role-based access, and a **job queue on PostgreSQL** driving the embedding worker without an external broker.

DevFlow – AI Engineering Workspace | *Next.js, Express, Prisma, PostgreSQL, Redis, Turborepo* **2026**

- Engineering dashboard that harvests GitHub history, including private organization repositories, and synthesizes it into a career and impact report.
- Walks the **GitHub GraphQL API** with author-filtered commit pagination, exponential backoff honoring rate-limit headers, and a fixed-size concurrency pool so multi-year scans survive throttling.

Education

Raj Kumar Goel Institute of Technology (AKTU) **2020 – 2024**
Bachelor of Technology in Computer Science and Engineering (IoT) *CGPA: 8.0*