

RONIT RAI

LinkedIn: [linkedin.com/in/ronitrai](https://www.linkedin.com/in/ronitrai)
GitHub: github.com/ronitrai27

Email: ronitrai1237@gmail.com
Mobile: +91 6280620947

SKILLS SUMMARY

Languages: Python, TypeScript, C++, Java, JavaScript, SQL
Frameworks & Libraries: Next.js, React, TanStack, Node.js, Express, FastAPI
Databases: PostgreSQL, MongoDB, Convex, Redis, Neo4j, Pinecone, ChromaDB
AI / Agentic Systems: LangChain, LangGraph, Google ADK, MCP, LlamaCloud, Langfuse, Context Engineering
ML / DL: PyTorch, scikit-learn, Hugging Face Transformers, Unsloth, Vllm, MLOps
Cloud & DevOps: AWS, GCP, Docker, CI/CD, Jenkins, Apache Airflow, Temporal.io, RabbitMQ, Grafana, K6
Soft Skills: Leadership, Problem-Solving, Client Communication, Cross-functional Collaboration, Team Mentorship

WORK EXPERIENCE

Founder — VRSA Analytics Pvt Ltd Mar 2025 – Aug 2026

- **Founded** and led a technical agency, owning end-to-end system design, product engineering, client delivery, and technical execution across e-commerce, B2B, and B2C applications, **generating ₹3.5L+** revenue across 1+ years.
- Owned **full-stack architecture** and requirement gathering for multiple e-commerce/B2B platforms (incl. Upharvilla.in/ Inventory/Chatbots) — from client discovery through production deployment and inventory management systems.

Freelance Software Engineer — Think Vers Labs Feb 2026 – May 2026

- Led end-to-end system design and deployment for Bounty Monster, mapping full scenario coverage, engineering for failure modes, and delivering production-ready solutions that integrated with legacy customer environments.
- Engineered **multi-tenant SaaS** architecture with RBAC access control; owned **CI/CD pipelines** (GitHub Actions) for automated build/test/deploy; optimized LLM API usage, cutting inference cost and latency while shipping features end-to-end.

PROJECTS

Enterprise-Grade Bilingual Sales Agent | [GitHub](#) Jun 2026 – Aug 2026

- **Orchestrated** a stateful **Python/FastAPI** multi-agent runtime using LangGraph's Supervisor Pattern, orchestrating parallel RAG, Web Search, Critic, and Deep Memory sub-agents with Redis checkpointing for **HITL** approval gates, and Temporal.io durable background workflows for long-running, stateful execution.
- Instituted a sub-300ms 2-tier **Guardrail** pipeline integrating instant regex **PII redaction** (Indian PAN, credit cards, SSN) with a model-based semantic judge (openai/gpt-oss-safeguard-20b), enforcing prompt injection and persona-hijacking defenses and blocking adversarial jailbreaks before reaching downstream agents.
- Fortified 0-downtime fault tolerance & observability, featuring a 3-state circuit breaker (CLOSED/OPEN/HALF-OPEN) with automated OpenAI to Groq failover, full-stack Logfire & Langfuse distributed tracing, and containerized **Jenkins CI/CD** with DeepEval quality gates and **k6 load testing**.

Tech Stack: FastAPI, LangGraph, Redis, Temporal.io, Langfuse, CI/CD Jenkins, Docker, AWS, DeepEval, k6, Next.js

Wekraft | wekraft.xyz | [GitHub](#) Nov 2025 – Jul 2026

- Pioneered and **launched** an AI-first project execution platform for dev & PM teams — ranked **Top 8 Product of the Day** on Product Hunt, reaching 800+ sign-ups and 100+ active users in the first month, with validation from engineers at Oracle & Microsoft.
- Shipped Kaya, a multi-agent PM system integrating **20+ tools via MCP** (Linear, Slack, Sentry, Datadog, etc) for autonomous task orchestration; engineered **dynamic LLM routing** by query complexity and a **semantic caching layer**, cutting frontier-model inference costs by **60–70%**.
- Architected Harry, an autonomous senior-dev agent reviewing every PR/commit via GitHub webhooks, running on durable Inngest background jobs with automatic retry — auto-committing fixes and filing issues, eliminating third-party code review and scoring **60–70% F1 on CodeReviewBench** (precision/recall vs. human-verified golden comments).

Tech Stack: LangGraph, Python, FastAPI, Temporal.io, GitHub webhooks, MCP, Ably, Inngest, TanStack, Google Cloud Run

Clarío | [Clarío.Live](https://clarío.live) | [GitHub](#) Sep 2025 – Dec 2025

- Recognized by **Google**, receiving **\$5K in cloud credits**; onboarded 10+ mentors across industries and shipped a secure, real-time video-calling module for transparent mentor-mentee doubt resolution
- Conceptualized and deployed a hyper-personalized career recommendation engine by **fine-tuning Qwen2.5-3B** and deploying it with Vllm, leveraging continuous batching and optimized KV-cache management to achieve **3.2× higher throughput**.
- Engineered a LangGraph-powered career agent with persistent graph memory (**Neo4j**) that performs adaptive skill-gap diagnosis and generates dynamic, personalized career roadmaps.

Tech Stack: Qwen2.5-3B, Vllm, LangGraph, Neo4j, React, ReactFlow, Unsloth, Vllm, Supabase, Google Cloud

TRAINING

Data Structures & Algorithms Training | [Lovely Professional University](#) Jun 2026 - Jul 2026

- Implemented core data structures — arrays, linked lists, stacks, hash maps, trees, and graphs — with focus on time/space complexity and Big-O analysis.
- Practiced algorithmic problem-solving across recursion, dynamic programming, and sorting/searching techniques, with focus on optimizing time and space complexity for scalable solutions.

- Applied graph traversal algorithms (BFS/DFS, shortest-path) to model network structures like social-graph connections (e.g., LinkedIn-style relationship graphs) and used tree structures (BST, heaps) for efficient search and sort.
- Data Analytics Industry Training** | [Aadvi Business Pvt Ltd](#) May 2024 - Jul 2024
- Built end-to-end data analysis pipelines using Pandas, NumPy, and Matplotlib for data cleaning, EDA, and visualization; applied Scikit-learn for classical ML
 - Developed time-series forecasting projects — trend/seasonality decomposition and predictive modelling on real-world datasets.
 - Applied feature engineering and model evaluation (train/test split, cross-validation) to improve predictive accuracy, and communicated findings through data visualizations and reports.

CERTIFICATIONS / CERTIFICATES

- Basics of Data Structures & Algorithms — [Lovely Professional University](#) Jun 2026
- Hands-on Bootcamp on Artificial Intelligence — [BlockseBlock & OpenxAI](#) Aug 2025
- Introduction to Python — [Coding Ninjas](#) (*Top Performer*) Apr 2023
- The Complete JavaScript Course: From Zero to Expert - [Udemy](#) Jan 2023

ACHIEVEMENTS

- Launched Wekraft to **Top 8 Product of the Day** on Product Hunt and **#1 on forg.to**, reaching **100+ active users** in month one. Jun 2026
- Won the National Virtual Hackathon organized by **Hack India**, securing **2nd Prize** and demonstrating strong technical execution. Jan 2026
- Led a team of 3 to Runner-up at the National **Google GenAI Hackathon**, Bengaluru, competing among 2.75L+ participants. Dec 2025
- Garnered Social Impact & Most Innovative Project Awards at **LPU InnoTek 2025**, competing among 10K+ participants Dec 2025
- Secured **Top 10** and the **Social Impact Award** at CT University's National Hackathon, competing among **2K+ participants**. Sep 2025
- Captured 1st Prize at Build with Gemini, **Delhi University**, a Google-sponsored National Hackathon with 10K+ participants. Nov 2025
- Clinched 1st Prize at Zinnovatia 3.0, **Chandigarh**, a National-level Hackathon with 3.5K+ participants nationwide. Sep 2025
- Emerged as Finalist at Code for Bharat 2.0, **Microsoft** Pune, selected among 12K+ participants. Aug 2025
- Spearheaded teams to 8+ State/Inter-College Hackathon wins and 6 consecutive National-level hackathon victories.

EXTRACURRICULAR ACTIVITIES

Hackathon Mentor/Judge — Mentored/judged teams at Hackathons (Bug bounty by wekraft/ College hackathons), advising on technical architecture, feasibility, and pitch framing.

Speaker — Delivered a talk on "How to Use AI to Win Hackathons," reaching 100+ attendees virtually.

EDUCATION

Lovely Professional University	Punjab, India
B.Tech, Computer Science & Engineering - CGPA: 6.7	Mar 2025 – Mar 2028
Lovely Professional University	Punjab, India
Diploma, Computer Science & Engineering - CGPA: 7.1	Jun 2022 – Mar 2025
MGM Public School	Punjab, India
Class 10th – Grade – 8.4	Mar 2022