Yug Jindal

New Delhi, India
yugjindal1234@gmail.com
Github LinkedIn Portfolio



Education

Jaypee Institute of Information Technology, Noida
Bachelor of Technology in Computer Science (B.Tech CSE) *Jul 2023 - Jul 2027*
CGPA: 8.24/10
Key Coursework: Data Structures & Algorithms, Object-Oriented Design, DBMS, Operating Systems, Computer Networks

Work Experience

Forward Deployed Engineer, Hiregram (AI Mock Interview Platform) *Jul 2026 - Present*

- Shipped **43 merged PRs** in three weeks as users grew **192%** to **1.26K**, owning problems from user feedback to deploy.
- Sustained **100% success across 1,549 jobs** with **zero dead-letters** at **57% higher volume** by hardening retries and failure handling in the **PostgreSQL job queue** behind every AI interview report.
- Built the startup's first **automated testing platform** (350+ tests) and **GitHub Actions CI gates**, 5 days in.

Product & Software Engineer, ShiftsDeal (B2B Venue-Booking Marketplace) *Oct 2025 - Feb 2026*

- Owned the **entire stack as sole engineer** (**Next.js, TypeScript, Supabase/PostgreSQL**), building listings, search, and a **4-step booking wizard** (event details, KYC, **AI-generated contract** with e-signature, confirmation).
- Prevented **double bookings under concurrent traffic** with 15-minute expiring slot holds plus a **UNIQUE constraint** on (venue, date, slot), resolving simultaneous requests at the **PostgreSQL** level.
- Delivered **role-based dashboards** (renter/owner/admin) covering bookings, payment tracking, and venue approvals, supporting **20+ venues** on the platform.

Software Engineering Intern, DM Integrators *May 2025 - Jul 2025*

- Cut client requirement-analysis time from ∼**1 week to a few hours** for **20 engineers** by building a **Multimodal RAG** platform (text + vision) over hundreds of industrial catalog PDFs.
- Designed a **4-stage AI pipeline** (retrieval → reasoning → pricing → synthesis) with grounded recommendations and exact PDF references.
- Deployed a containerized **FastAPI** backend with **11 REST APIs**, health checks, and graceful **MongoDB** degradation.

Projects

Aura: Voice & Video Avatar AI Agent for SaaS Automation *May 2026 - Jun 2026*
Tech Stack: FastAPI, LangGraph, Gemini, Next.js, Supabase, OAuth integrations (Slack, GitHub, Notion, etc.), Playwright

- Built a **bilingual** real-time agent executing **54 LLM tools** across **7 OAuth-connected SaaS apps** via an **OpenAI-compatible SSE streaming API**, holding first-spoken-token latency to **187ms p50** over 340 instrumented turns.
- Reduced multi-tool voice turns from **21.6s to 4.17s** with an **embedding-based tool router** (binds top-6 of 54 tools per utterance) and **three-layer caching**; shipped **116 unit, integration & E2E tests** that run fully offline.

Multimodal Satellite Image Analysis (ISRO Hackathon - Team Lead, Placed Top 5) *Nov 2025 - Dec 2025*
Tech Stack: GPT-OSS-20B, Earth Observation, Image Captioning, VQA, LoRA, PyTorch, NVIDIA A100 Cloud GPU

- Fine-tuned a self-hosted **vision-language pipeline** (InternViT-300M, MLP projector, **GPT-OSS-20B**) with **LoRA** on **VRSBench**, reaching **71.26% VQA accuracy** and **BLEU-1 0.415** on one **NVIDIA A100**.
- Engineered a **Sentinel-2 raster pipeline** (Flask, Rasterio, PyProj) that **downsamples 10,980px scenes 4x** and windows reads to the visible map region, replacing desktop **GIS** workflows with a browser upload.

Achievements

Smart India Hackathon 2024 Winner *Dec 2024*

- Won **UIDAI**'s problem statement among **300,000+ participants** (INR 1,00,000 prize) with a secure **Aadhaar authentication platform** on **AWS**.

Google CCD Project Exhibition *Sep 2024*

- Selected from **100+ applicants** to present an **AI Large Action Model** at **Google Cloud Community Day** to **1,000+ attendees**.

Technical Skills

Languages: Python (Proficient), C++, TypeScript/JavaScript, SQL
Frameworks, Cloud & AI: Next.js, FastAPI, Supabase/PostgreSQL, MongoDB, REST APIs, WebSockets, GitHub Actions, CI/CD, Docker, Playwright, Vitest, LangChain, LangGraph, Natural Language Processing (NLP), Pinecone, AWS
Systems & Development: Object-Oriented Design, API Development & Integration, Test Automation & Observability, Scalable Architecture, Unix/Linux Environments