Kuber Jain

+91-8975033034 kuberjain144@gmail.com linkedin.com/kuberjain144 github.com/Kuber144

Education

Indian Institute of Information Technology Allahabad | 2021-2025 CGPA (8 Semesters): **8.38**

Bachelor of Technology in Information Technology

Experience

Edra Labs **Mumbai, India**

Software Engineer *July 2025 – Present*

- Built and shipped **Havi, an AI-powered presentation platform**, including its **ChatGPT App Plugin** end-to-end using the **OpenAI Apps SDK and MCP**, enabling users to generate and edit presentations across the web and directly within **ChatGPT**.
- Engineered Havi's **AI presentation-generation pipeline** spanning outline generation, structured slide generation, layout selection, image retrieval, slide continuity, themes, and document-to-presentation workflows.
- Built **agentic presentation editing** workflows using **LangGraph, LangChain, and LLMS**, enabling natural-language editing with contextual awareness, image actions, layout changes, element-level operations, and viewport context.
- Designed and iterated on **LLM evaluation systems** for AI-generated presentations, improving internal slide-evaluation scores from **48% to 75%** while reducing **cost per slide by 82%** through evaluation-driven prompt and workflow optimization.
- Benchmarked and optimized **AI generation models** through A/B experimentation, multi-file document extraction, and image reuse, achieving **6.5× lower image-generation cost** and **4 seconds faster latency** while maintaining generation quality.
- Built **automated product analysis loops** to identify user drop-off points and surface opportunities for improving Havi's core product and AI-generation workflows, enabling faster product iteration.
- Built and executed an **AI UGC growth strategy** for Havi, generating **100K+ organic Instagram views in under 3 weeks**, including a Reel with **52K+ views**, through iterative short-form product content experimentation.

Software Engineer Intern *Jan 2025 – Jun 2025*

- Built a granular **RBAC system from scratch**, designing role-resource-action permissions with hierarchical resources and a denormalized permission cache for constant-time authorization checks at login.
- Built **LLM-powered content generation and validation workflows** for **10,000+ production records**, automating description and FAQ generation, Markdown/URL validation, and large-scale data cleanup.
- Engineered a distributed **Python/Celery crawling and extraction pipeline** with Redis-based asynchronous fan-out, Playwright/Scrapy web automation, and **LLM-powered extraction and vector search** for large-scale web data processing.

Projects

Collaborative Code Editor-Project Demo | *ExpressJS, ReactJS, WebSockets, Google Generative AI* **March 2024**
Github Repository

- Developed a solution for developer collaboration by creating private rooms with shared coding environments, enabling real-time editing and code review.
- Implemented an online compiler that supports multiple programming languages, allowing users to compile and run code within the platform.
- integrated a chat feature and PDF sharing capabilities to facilitate seamless communication and resource sharing among users alongwith the ability to download and upload code.

CCTV Based Face Detection-Project Demo | *Python, RetinaFace, CodeFormer, Facenet* **April 2024**
Github Repository

- Developed a system for face detection, extraction, restoration, and facial recognition against a pre-existing database.
- Employed RetinaFace ResNet50 for face detection and the CodeFormer ML Model for enhancing low-quality faces.
- Used Facenet for feature extraction and MongoDB Atlas Vector Search to identify individuals.

Achievements

LeetCode Knight: Max Rating 1923, Top 3.9%
Codeforces: Max Rating 1416
Codeforces Round 928: Rank 583/26,000+
CodeChef April Long Two 2022: Rank 24/12,400+

Technical Skills

Languages: Python, TypeScript, JavaScript, C++, SQL
AI / ML: LLMs, Multimodal Models, OpenAI, Claude, LangChain, LangGraph, MCP, LLM Evaluation, RAG, Vector Search
Backend / Systems: Node.js, Express, FastAPI, Flask, Celery, Redis, PostgreSQL, Prisma, SQLAlchemy
Infrastructure / Tools: Docker, AWS, Playwright, Git