

# SOMDEV DAS

+91 6295874389 ◇ Kolkata, WB, India

[somdevdas@protonmail.com](mailto:somdevdas@protonmail.com) ◇ [Portfolio](#) ◇ [GitHub](#) ◇ [LinkedIn](#) ◇ [LeetCode](#)

## EXPERIENCE

### Junior Software Engineer

SentientGeeks

Feb 2024 – Present

Kolkata, WB, India

- Built a Generative AI RAG Chatbot using **LangChain** and **OpenAI API**. Implemented document chunking, vector embeddings, and context-aware retrieval to deliver high-accuracy responses for enterprise use cases.
- Built and trained **YOLO** models from scratch for MRZ detection across multiple document types (passports, Emirates IDs). Handled the full pipeline from data annotation through deployment, integrating with **Azure Document Intelligence** for text extraction.
- Developed a Face Recognition API for Visitor Management using **Computer Vision** models. Designed the backend with clean architecture principles in Python, currently serving production traffic.
- Optimized a Music Recognition platform using **Flask**, **NumPy**, and **Pandas**. Refactored audio signal-processing algorithms to achieve a **10x performance improvement**.
- Built scalable REST APIs using **Python**, **FastAPI**, and **SQL**. Implemented data validation workflows and maintained SDLC best practices across multiple backend services.

### Software Engineering Intern

SentientGeeks

Jul 2023 – Feb 2024

Kolkata, WB, India

- Developed Passport MRZ Detection service using Python for image preprocessing and rule-based validation. Contributed to initial model training pipeline and data extraction logic.
- Built REST APIs using **Django REST Framework** and contributed to frontend development with **Next.js** and **TypeScript**. Focused on API design patterns, database optimization, and bug fixes across the stack.

## PROJECTS

### Study Buddy (GenAI Assistant) — [Live Demo](#) — [GitHub](#)

A production-deployed Generative AI platform built with **FastAPI**, **Next.js**, and **LangChain**. Auto-generates quizzes and summaries from uploaded documents using LLMs, vector search, and semantic embeddings. Features **Docker** containerization, **CI/CD** pipeline, **Cloudflare R2** storage, **Nginx** reverse proxy with **Certbot SSL**, and custom domain configuration.

### Lung Cancer Detection (Healthcare Analytics) — [GitHub Link](#)

Developed a Python-based ML prediction service for early-stage detection. Designed the data ingestion pipeline, API endpoints, and model inference logic for real-time classification.

## SKILLS

Generative AI	LangChain, OpenAI API, RAG Pipelines, Vector Databases
Computer Vision	YOLO, OpenCV, Azure Document Intelligence
Languages & Data	Python, SQL, JavaScript, TypeScript, NumPy, Pandas
Backend & Cloud	FastAPI, Flask, Django REST, PostgreSQL, AWS (EC2, S3), Docker
Development Tools	Git, CI/CD, Postman, VS Code

## EDUCATION

Master of Computer Science, Midnapore College, Midnapore

2021 – 2023

CGPA: 9.2

Bachelor of Computer Science, Prabhat Kumar College, Contai

2018 – 2021

CGPA: 7.48