

# Prateek Gaur

📍 Noida, Uttar Pradesh, India 📩 gaur.prateek.1609@gmail.com

LinkedIn in/gaur-prateek Website pratekgaur.neocities.org/

## EDUCATION

### Artificial Intelligence and Data Science

Maharaja Agrasen Institute of Technology Delhi, India • Delhi, India • 2025 • 8.8

### Senior Secondary

Saini A.V.M. Eng. Med. School, College Bandikui, India • Bandikui, Rajasthan • 2020 • 7.9

## SKILLS

**Languages & Tools :** Python, C++, JavaScript, SQL, Docker, Git

**Frameworks :** PyTorch, TensorFlow, LangChain, FastAPI, Scikit-learn, Hugging Face, LlamaIndex

**AI & ML :** Generative AI, Agents, Deep Learning, Machine Learning, LLMs, RAG

**Core Strengths :** End-to-End AI Systems, API Integration, Model Deployment

## EXPERIENCE

### AI Engineer

DataAlchemy.AI

July 2025 - Present, Noida, India

- Developed an automated **GitHub code documentation generator** using Python and Fast API, converting repository structures into structured technical documentation with 95% accuracy and reducing manual documentation time by over 70%.
- Built a full-stack **AI-driven mock interview platform** enabling users to practice technical interviews with real-time feedback, integrated question generation, and performance analytics to improve user readiness by an estimated 60%.
- Architected modular microservices for event-driven automation and API integrations, ensuring scalable communication between GitHub webhook services and documentation pipelines while maintaining sub-second response latency.

### Generative AI Intern

Otomashen

February 2025 - June 2025, Noida, India

- Collaborated with 4+ cross-functional teams to design and deploy AI-driven applications, contributing to 6+ projects and gaining hands-on experience in big data, backend (Flask, SQLite), and frontend integration workflows.
- Engineered a Medical Chatbot handling 1000+ simulated patient interactions, leveraging an SQLite backend for appointment scheduling, Cancellation, Rescheduling and EHR (Electronic Health Record) management.
- Designed a Virtual Avatar system integrating Wav2Lip and SadTalker, achieving 90%+ lip-sync accuracy and realistic head-motion responses through LLM-based contextual conversation generation.

## PROJECT

### GitHub Code Documentation Generator | FastAPI · Kafka · GitHub API · LangChain · Docker

DataAlchemy.AI • July 2025 - Present

- Pioneered an AI-powered documentation generation system that automated extraction and summarization of GitHub repository codebases, and updates the documentation Repo — improving documentation efficiency by 65%.
- Programmed an event-driven pipeline leveraging Git APIs and NLP-based summarization models to automatically generate structured MARKDOWN files and module-level explanations for over 500+ code commits.
- Integrated Webhooks, Polling, and Apache Kafka as a distributed event-driven queue, enabling real-time communication across repository workflows and backend microservices — scaling to handle 10,000+ concurrent pull requests with zero message loss.

### Virtual Avatar Generation | Wav2Lip · SadTalker · GANs · LLMs · Python

E-Solutions • May 2025 - July 2025

- Implemented a virtual avatar generation pipeline leveraging text-to-image diffusion models and facial landmark detection, enabling lifelike 3D avatar synthesis with >92% identity accuracy.
- Created a FastAPI-based microservice for avatar generation, handling asynchronous requests and reducing API response time by 35% through optimized Input/Output and batch processing.
- Optimized inference and rendering workflow, achieving a 40% reduction in generation latency through model fine-tuning, FP16 and 8-bit quantization, caching mechanisms, and GPU resource optimization.

## CERTIFICATIONS

### Python for Data Science and Machine Learning Bootcamp (Udemy)

Udemy • 2024

### Google Analytics for Beginners (Google)

Google • 2022