

# SOHAM MUKHERJEE

Kolkata, India | +91 8240771772 | [soumyapu31@gmail.com](mailto:soumyapu31@gmail.com) | [LinkedIn](#) | [Website](#) | [GitHub](#)

## SUMMARY

Aspiring Machine Learning & AI engineer with expertise in Deep Learning and Generative AI. Strong foundation in Neural Networks, Computer Vision, and NLP, with a focus on building intelligent and scalable AI solutions. Passionate about pushing the boundaries of AI-driven innovation. Also, a quick learner and a problem solver.

## SKILLS

- |                    |                    |                    |               |
|--------------------|--------------------|--------------------|---------------|
| • Python           | • Computer Vision  | • Transformers     | • RAG         |
| • MySQL            | • Object Detection | • LLMs             | • AI Agents   |
| • Machine Learning | • NLP              | • Fine-tuning      | • LangChain   |
| • Neural Networks  | • Embeddings       | • Vector Databases | • HuggingFace |

## PROJECTS

### Real-time Google Search AI Agent with Memory | [GitHub Repository](#)

- Developed a real-time AI-powered Google search agent using LangChain and Groq Inferencing Engine (Llama-3.1-8B model) to retrieve and summarize web data dynamically.
- Implemented conversational memory using LangChain's ConversationBufferMemory to retain context for follow-up queries, enhancing response relevance.
- Integrated Google Programmable Search API for structured web search results, formatted into markdown for improved readability in a Streamlit-based UI.

### PDF Question Answering RAG | [GitHub Repository](#)

- Built a PDF-based Q&A system using Google's Gemini API to extract and answer questions from documents.
- Integrated LangChain for efficient document processing and retrieval-augmented generation (RAG).
- Developed a user-friendly Streamlit UI for seamless interaction and question-answering.

### Facial Recognition Attendance System | [GitHub Repository](#)

- Used OpenCV's Haar cascades to detect faces and face\_recognition library to generate 128-dimensional embeddings for each face.
- Compared embeddings with a stored database using Euclidean distance to accurately identify students.
- Stores student data in Firebase Realtime Database and images in Firebase Storage for real-time attendance tracking.

### Multi-purpose AI Assistant | [GitHub Repository](#)

- Multimodal AI Q&A System using Google's Gemini API to handle text, image, and invoice-based queries.
- Streamlit Web Interface with dedicated pages for chatbot, text Q&A, vision-based queries, and invoice extraction..
- Gemini-Powered Features including a chatbot, invoice data extraction, vision-based answering, and text analysis in Markdown..

### Automatic License Plate Detector | [GitHub Repository](#)

- License plate detection after fine-tuning a pre-trained YOLOv8 model on our custom dataset.
- Text extraction using Tesseract OCR to recognize the license plate numbers.
- Visualizing detected plates and extracted text in the web interface built using Streamlit.

### Food Detector | [GitHub Repository](#)

- Initial classification with a CNN baseline model.
- Transfer learning with EfficientNet and ResNet, and enhanced via data augmentation.
- Fine-tuned pre-trained models to get better results.

## CERTIFICATIONS

- Python for Data Science & Machine Learning [*Udemy* ([URL](#))]
- TensorFlow Developer Certification [*Zero to Mastery* ([URL](#))]

## EDUCATION

**Master of Science (M.Sc.) in Data Science & Analytics**

*NSHM College of Management & Technology*

**(2023-present)**

**Bachelor of Science (B.Sc.) in Data Science**

*iLEAD*

**(2020-2023)**