

# Debabrata Das

☎ +91 9957085950 | ✉ debodas1333@gmail.com | 🐙 [Github](#) | 🔗 [LinkedIn](#) | 🌐 [Website](#)

## PROFESSIONAL SUMMARY

---

AI-focused data professional with experience building **LLM-powered applications, RAG pipelines, and AI-driven data systems**. Skilled in **prompt engineering, vector embeddings, and backend development**, with a proven ability to transform unstructured data into scalable solutions.

## TECHNICAL SKILLS

---

**Programming:** Python, SQL

**AI/ML:** LLMs (Claude, LLaMA, Mistral), Prompt Engineering, RAG, Vector Embeddings

**Data:** Data Pipelines, ETL, Data Cleaning & Preprocessing, EDA

**Cloud & Backend:** AWS Bedrock, S3, Vector Databases, FastAPI

**NLP & Extraction:** Information Extraction, Unstructured Data Processing, OCR + LLM

## PROFESSIONAL EXPERIENCE

---

**FACCTUM — Data Analyst (AI/ML Product Development)**

Feb 2024 – Present

*Bengaluru, India*

- Architected and deployed **15+ LLM-powered data pipelines** using **AWS Bedrock (Claude, Mistral)**, automating structured extraction from PDFs, web sources, and regulatory documents.
- Designed **multi-stage AI workflows** combining web scraping (**Playwright, Selenium**), OCR (**Mistral OCR, Tesseract**), and LLM-based structured extraction, reducing manual processing from **days to under 1 hour**.
- Engineered **prompt strategies** (few-shot, chain-of-thought, schema enforcement) to reduce hallucinations and improve extraction consistency across diverse document formats.
- Built **delta detection systems** comparing live data against **10M+ records**, triggering LLM processing only for new or modified entries, optimizing API costs.
- Developed a reusable **Python automation library** for PDF parsing, table extraction, web scraping, and AI-powered address parsing, adopted across team pipelines.

**ZIGRAM — Analyst**

Jun 2023 – Feb 2024

*Gurugram, India*

- Built data cleaning pipelines, achieving **99% data accuracy**.
- Improved data accessibility, increasing efficiency by **30%**.
- Streamlined workflows for faster reporting and analytics.

## PROJECTS

---

**AI Blog Generator – LLM + RAG System**

- Built an **LLM-powered blog generation platform** using AWS Bedrock, embeddings, and FastAPI.
- Developed **data ingestion and embedding pipelines** for context-aware, SEO-optimized content.
- Implemented **prompt templates and keyword scoring** for consistent, brand-aligned output.
- Built backend workflows for **real-time content generation and refinement**.
- **Impact:** Reduced blog creation time by **80%**.

**LocalMind – Local LLM Interface for Private AI Usage**

[GitHub](#)

- Built a **web-based interface for locally hosted LLMs**, enabling **privacy-preserving AI usage** without external APIs.
- Designed backend workflows in Python for **model interaction, prompt processing, and response generation**.
- Developed a **user-friendly interface** to make local LLMs accessible for a broader set of users.
- Built a **modular architecture** supporting multiple open-source models and inference backends.
- Continuously improving the system with features for better **usability, extensibility, and real-world adoption**.

## EDUCATION

---

**Bachelor of Science (Physics)**

Aug 2019 – Sept 2022

Assam University, Silchar, India