DEBABRATA DAS

📞 +91-9957085950   ✉ debodas1333@gmail.com   🐙 Github   in LinkedIn   🌐 Website
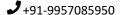
Data and AI professional with experience building AI-based data pipelines and LLM-powered applications. Strong understanding of machine learning concepts, embeddings, and backend API development, with proven ability to translate business requirements into AI-driven solutions.

Skills

- Programming & Analysis: Python, SQL
- Machine Learning & AI: Large Language Models (Anthropic Claude, LLaMA, Mistral), Vector Embeddings, Prompt Engineering, RAG
- Analytics: EDA, Data Cleaning & Preprocessing, Data Warehousing
- Cloud & Tools: AWS (Bedrock, S3, S3 Vector DB), FastAPI
- Business & Research: Company Profiling, Market Research, Data-Driven Decision Making, Cross-functional Collaboration

Experience

FACCTUM   **Feb 2024 – Present**

Data Analyst   Bengaluru, India

- Contributed to the development of a production AI-driven data intelligence platform that ingests data from public and official sources, processes unstructured content, and delivers structured outputs to multiple enterprise clients.
- Designed and implemented **LLM-based information extraction pipelines** using **AWS Bedrock**, converting unstructured and semi-structured text into **schema-compliant, machine-readable formats**.
- Engineered **source-specific extraction strategies** to handle heterogeneous publishing formats, document structures, and data layouts, significantly improving extraction robustness and consistency.
- Actively worked on **prompt engineering and output refinement**, improving field-level accuracy, reducing hallucinations, and ensuring deterministic, client-ready outputs.
- Built reusable **Python-based extraction tools and automation frameworks** that reduced manual processing effort and **significantly decreased end-to-end data processing time**.
- Collaborated with engineering and product teams to define **data schemas, validation rules, and quality checks**, ensuring consistent data flow across downstream systems.
- Utilised **AWS services (Bedrock, S3)** to support scalable ingestion, storage, and processing of AI-extracted data.
- Contributing to the development of an in-house OCR pipeline tailored to internal use cases, focusing on document preprocessing, text extraction, and integration with AI-based extraction workflows.

ZIGRAM   **Jun 2023 – Feb 2024**

Analyst   Gurugram, India

- Standardized data cleansing, validation, and integration processes, **achieving 99% data accuracy** and enhancing overall data quality
- Led a database overhaul, resulting in a **30% boost in data accessibility** and efficiency
- Established streamlined data management and optimization processes, strengthening project communication and analytics capabilities
- Established optimized data management workflows, enabling faster reporting and analytics across teams.

Projects

AI Blog Generator (Internal Hackathon Project):

- Designed and developed an AI-powered blog generation platform for the Sales & Marketing department using **AWS Bedrock (Anthropic Claude 3.7 Sonnet, LLaMA 70B Instruct), AWS S3 Vector DB**, **embedding models**, and **FastAPI**.
- Engineered data ingestion pipelines to process company-specific blogs and solution documents, converting them into vector embeddings for contextually accurate and SEO-optimised content generation.
- **Developed keyword scoring and matching algorithms**, **tone-specific templates**, and an **interactive refinement system** to ensure brand-consistent, high-quality blog output.
- Built API integrations for seamless backend–frontend communication and real-time content refinement.
- **Impact:** Reduced blog creation time by **80%**, enabling scalable, AI-driven content production.

Education

Assam University   **Aug 2019 – Sept 2022**

Bachelor of Science with Physics   Silchar, India