

Anubhav Santra

+91 8617407813 | iamanubhav2001@gmail.com | [in LinkedIn](#) | [Github](#)

WORK EXPERIENCE

AI Engineer — Thomson Reuters, Bangalore

June 2026 - Present

- Built an **autonomous interviewing agent** that parses corporate meeting transcripts, conversationally identifies and collects missing information, auto-fills confirmed data, and generates end-to-end process reports with **Lucid swimlane diagrams** via MCP, reducing manual effort by **90%** and earning the **Q2 Hall of Fame Award for Innovation Trailblazer & Certificate**
- Architected and deployed** production-grade **agentic AI harnesses** using **Claude Code**, **Model Context Protocol (MCP)**, **LangGraph**, and **subagent orchestration** — engineering self-improving, self-probing agents via custom skill files and hooks
- Automated enterprise knowledge workflows**, eliminating **50,000+ hours** of manual effort annually via **Agentic RAG** pipelines and agent orchestration contributing to an estimated **45% increase** in targeted business-role revenue

AI Engineer — ZS Associates, Pune

June 2024 - June 2026

Marketing Asset Generation Assistant for US based pharma Company

- Built a **multi-agent** AI system with agent handoffs, tool orchestration, and **RAG** over OpenSearch to auto-generate pharma marketing assets (emails, SMS, etc) reducing manual effort by **66%** and boosted revenue of **1M\$** - **🏆 Won 2025 ZS Impact award**
- Spearheaded batch ingestion pipelines (**Apache Airflow** DAGs) to extract unstructured data from **S3** using **Amazon Textract**
- Engineered entity extraction using **few-shot**, **ReAct**, and **Chain-of-Thought (CoT) prompting** across multiple entity types, enabling **hybrid retrieval** on **OpenSearch** with **context precision/recall above 80%**.
- Led architecture design discussions on **scalable AI systems** using **LangGraph**, **Langchain** with **agent evaluation** and feedback loops to monitor response quality with solutions to **control hallucinations** in production resulting upto **50% error reduction**

Databricks Brand Assistant - for US based pharma Company

- Built a **multi-agent routing system** using **LangGraph (ReAct)** to orchestrate query routing across custom **Databricks AI/BI Genie** spaces (market share, customer engagement), **automating business analysis by 50%**
- Configured Databricks Genie room API** with custom business-logic configurations and text-to-SQL-to-visualization generation, **deployed as a model-serving endpoint** on Databricks with unity-catalog controlled model security and governance
- Engineered CI/CD pipelines** (LLMOps) via **GitHub Actions** and **Databricks Asset Bundles (DABs)**, with model versioning across **1M+ records** spanning across multiple **Gold** tables using **Apache Spark**
- Built **automated LLM evaluation frameworks** tracking **10+ metrics** — faithfulness, relevance, correctness, precision, recall — driving a **38% improvement** in response quality

DocGen - Knowledge Layer

- Developed the knowledge layer** of a larger **agentic AI system**, exposing modular **document-**, **section-retriever** tools via MCP using **Azure Document Intelligence** and custom **XML parsing** to process **1000+ page** multi-modal scientific Docx files, with **Langfuse** enabling end-to-end agent trace observability
- Engineered parsing pipelines** for complex multi-modal content — tables, nested tables, formulas, charts, diagrams, and images — achieving high-accuracy section retrieval within 5 minutes per document
- Deployed and scaled** retrieval infrastructure on **Elastic Kubernetes Service (EKS)**, using **Celery workers** and message-queue architecture to support **10,000+ concurrent users**

INTERNSHIPS/PROJECTS

AI/ML Engineer - Intern — ZS Associates, Pune

Dec 2023 - Feb 2024

- Engineered **Graph Neural Networks** for linking text statements and endpoints for **70% faster suggestions**
- Optimized computation time involving **link-prediction** and **anomaly-detection** of graphs for clinical data by 50%
- Scaled **NLP/ML** pipelines from research to production and boosted *GNN models* performance up to 98% accuracy.

Machine Learning Summer Intern at SURGE, IIT Kanpur

May 2023 - July 2023

- Accelerated **Phase Field Simulations** via advanced **Machine Learning (ML)** systems with **10k+ images**
- Developed a fusion based **forecasting ML model** using an "encoder-decoder" network along with **ConvLSTM** that accelerates phase-field simulation by **50%** [Github](#) | **🏆 Certificate**
- Relevant Skills: **Python**, **Computer Vision**, **Forecasting**, **ConvLSTM**, **Deep Learning**, **Classification**

Undergraduate ML Intern — CMATER Labs, Jadavpur University

Nov 2022 - Dec 2023

- Optimized **CNN models** for skin-cancer classification HAM10000, ISIC datasets reducing computation time by **20%**
- Implemented Python based ML pipelines, **Feature Engineering** using Model ensembling, **Attention Mechanism** to achieve **10% higher** accuracy than **SOTA**

ACADEMIC BACKGROUND

Jadavpur University

Kolkata, West Bengal, India

Bachelor of Engineering (Metallurgical Engineering), CGPA: 8.32

2020-2024

ACHIEVEMENTS

- Intel OneAPI Hackathon — "Predicting News Sentiment Analysis."** Developed an ML Model that predicts sentiment based on news information. Ranked **7th** among **5000** participants | **🏆 Certificate**
- Amazon ML Challenge 2023** — Built a regression-based **ML Model** with **NLP** that predicts product length harnessing product data information. **Finished among top 200 candidates.**
- 🏆 Google Hashcode 2021: AIR-924, World Rank-3380**

TECHNICAL SKILLS

Areas of Expertise: GenAI, Agentic AI, Model Context Protocol (MCP), Machine Learning, LoRA/PEFT fine-tuning, Microservices

Languages, Databases and Version Cntrol: Python, C/C++, PyTorch, MySQL, Aurora, Postgres, Git, Azure Devops

Cloud Services: Databricks, A.WS, Azure, Docker, Kubernetes, EKS, Lambda, S3, Jira

Libraries: Langchain, LangGraph, MCP (Model Context Protocol), FastAPI, VectorDB, MLFlow, Tensorflow, JSON