

ARPAN MISHRA

Applied AI Engineer

New Delhi • +91-9958325792 • mishraarpan6@gmail.com • [Portfolio](#)

SUMMARY

Applied AI Engineer with 4.5 years of experience building production-grade LLM systems in regulated, high-stakes environments. I specialize in translating complex business problems into scalable AI architectures—spanning multi-agent pipelines, RAG-based retrieval, prompt engineering, and evaluation frameworks. At ZS Associates, I delivered clinical AI tools across multiple enterprise engagements, working closely with domain experts, medical writers, and cross-functional stakeholders to ensure AI outputs met real-world standards. I'm equally comfortable in deep technical implementation and in communicating architectural decisions to non-technical audiences.

SKILLS

Languages & Tools: Python, R, SQL, JavaScript, Docker, Streamlit, LangGraph, LangChain

LLM & Agent Systems: Multi-Agent Architectures, RAG, Prompt Engineering (Chain-of-Thought, Few-Shot, Structured), Agent Orchestration, Context Engineering, LLM Evaluation & Validation Frameworks, MCP Servers

Cloud & Infrastructure: AWS (SageMaker, Agent Runtime, Bedrock, EKS, DocumentDB), Azure Document Intelligence, Google Vertex AI

Domains: Clinical Trials, Healthcare AI, Regulated AI Deployment, Document Intelligence, Safety-Critical Systems

WORK EXPERIENCE

Data Science Consultant | ZS Associates

Dec 2025 – Present

Clinical Trial Document Authoring Platform

- Built and delivered a \$5M enterprise-grade **multi-agent document authoring system** supporting 12 clinical trial document types for a Fortune 500 client—a production-ready platform that reduced manual drafting time and earned trust from cross-functional clinical and regulatory stakeholders.
- Designed the core **agent orchestration layer** using a Planner-based framework with **RAG-powered context grounding**, enabling complex document workflows to execute with improved factual accuracy and compliance with regulatory standards.
- Built scalable **digitization and indexing pipelines** across AWS and Azure, enabling structured ingestion of diverse clinical document formats—creating a reliable foundation for downstream AI reasoning.
- Developed a centralized knowledge layer as a service (**MCP server on AWS Agent Core**) and established an **LLM evaluation framework**, ensuring high-quality, auditable AI outputs in a regulated clinical environment.

Clinical Query Analysis System

- Designed and deployed a **GenAI-powered clinical query intelligence system** that reduced redundant queries by **20%**, directly improving operational efficiency across clinical workflows.
- Architected a **multi-agent LLM framework** comprising an Executor Agent, Redundancy Classification Agent, and Query Category Detection Agent—each with well-defined responsibilities for modular iteration and evaluation.
- Processed and benchmarked insights from **~1.3M historical clinical queries**, establishing a performance baseline and improving classification accuracy through targeted prompt optimization and model tuning.
- Integrated the system into the live clinical query portal, enabling **real-time redundancy detection** and query-type tagging at the point of user input.

Data Science Associate Consultant | ZS Associates

Dec 2023 – Dec 2025

Safety Narrative Document Authoring

- Built an end-to-end AI pipeline that ingested **SDTM and ADaM clinical trial datasets** and automatically generated chronological **safety narratives** for patients—cutting a multi-hour manual process down to minutes.
- Engineered a full **ETL pipeline** to transform structured clinical datasets into **LLM-compatible formats**, enabling seamless downstream reasoning and summarization.
- Applied **structured prompting techniques**—including **chain-of-thought, few shot and temporal reasoning**—to produce coherent, medically aligned summaries that met clinical writer standards.
- Scaled the pipeline to generate safety narratives for **600+ patients** across 3 clinical trials; built automated Gantt chart visualizations from trial data to help medical writers validate event timelines.

Auto Document Redactor

- Designed and deployed an automated **PII redaction pipeline** for confidential clinical documents, achieving **98% recall and 90% precision**—meeting strict compliance requirements for regulated clinical environments.
- Implemented a hybrid architecture combining rule-based heuristics and an **ML-driven spaCy NER pipeline** with **chain-of-thought prompting**, accurately identifying nuanced and context-sensitive PII entities.
- Built **evaluation benchmarks and validation workflows** to ensure ongoing reliability and compliance, establishing a repeatable testing framework for the team.

Data Science Associate | ZS Associates

Nov 2021 – Dec 2023

ISR Entity Detection Pipeline

- Built an **entity and relationship extraction pipeline** from clinical protocol documents to support Industry Sponsored Research (ISR) decision-making—turning unstructured text into structured intelligence.
- Extracted key entities (drug, dosage, cycle, endpoints, inclusion/exclusion criteria) using **spaCy NER, custom fine-tuned entity recognition models**, and XGBoost-based classifiers, enabling structured protocol intelligence at scale.
- Enabled structured protocol intelligence to help stakeholders make **data-driven research sponsorship decisions**.

Patient Discontinuation Prediction

- Developed an **XGBoost-based predictive model** to identify patients at risk of therapy discontinuation; engineered features from claims data, patient demographics, and census datasets using **RFE and forward/backward elimination**.
- Packaged the solution as a **reusable internal asset**, enabling client-specific model training and feature generation across engagements.

Research Intern | Inria, University of Lille

Jun 2021 – Sep 2021

Recidivism Prediction — Mental Health & Suicide Risk

- Worked with medical data for mental health patients with **history of suicide attempts**.
- Modeled the recurrence of suicide attempts using both **parametric and non-parametric statistical methods**, incorporating medical survey data from VigilianS and identifying factors affecting re-attempt probability.
- Conducted **spatial analysis** of patient data and incorporated **geostatistical spatial autocorrelation** into the model to account for regional patterns.

Machine Learning Engineer (Part Time) | Omdena

Aug 2020 – Feb 2021

Asset Wealth Index Prediction with Satellite Imagery

- Part of a global team of 50, built an AI solution to model economic well-being in India using publicly available **satellite imagery and ground truth census data**.
- Used **Landsat 7 & 8 Satellite Images** and census data to fine tune a CNN based model predicting district-level census variables using a **multi-model, multi-task learning approach**.
- Used DHS data and sentinel images to classify the **Asset Wealth Index** of clusters across India.
- Project hosted by **World Resources Institute (WRI)** under UN Sustainable Development Goal 8 (Decent Work & Economic Growth).

EDUCATION

BSc (Hons) Statistics — Kirori Mal College, University of Delhi *Graduated: July 2021 | CGPA: 8.6*

PERSONAL PROJECTS

CalorieBot — Nutrition Tracking Agent (2025) • [GitHub](#)

- Built a **LangGraph-powered 6-phase nutrition agent** deployed as a **WhatsApp bot**—users describe meals in natural text or voice (“had 2 eggs and a bowl of oats”) and the agent runs item parsing → parallel food search → best-match selection → macro scaling → diary logging → daily summary, end-to-end in a single message.
- Beta tested with 5 users to collect real-world feedback on agent accuracy and **food parsing edge cases** (colloquial food names, multi-item messages, voice ambiguity)—iterating on **prompts and search scoring** based on their input.

Sentiment-Extraction-using-Bert (06/2020 - 06/2020) • [GitHub](#)

- Used BERT to detect sentiment in text and extract sentiment-bearing words; built a Flask API endpoint hosted locally using Pytorch and Hugging Face.

Anime-Generation-using-Deep-Learning (07/2020 - 07/2020) • [GitHub](#)

- Fine-tuned GPT-2 and trained LSTMs for anime-style text generation using Pytorch and Hugging Face.