Divyaansh Seth

(413) 275-7143 | divyaanshset@umass.edu | LinkedIn | GitHub

Education

University of Massachusetts Amherst Sep 2023 – May 2027
B.S. Computer Science and Mathematics; Minor: Business *GPA: 3.6/4.0*

Skills & Coursework

Languages: Python, Java, JavaScript, TypeScript, R, SQL
Frameworks/Tools: MERN Stack, TensorFlow, PyTorch, Scikit-learn, Flask, Docker, GitHub Actions, REST APIs, Linux, Redis
AI/ML: LangChain, FAISS, Milvus, Qwen, LLMs, Statistical Modeling, Monte Carlo Simulation
Visualization/BI: PowerBI, Matplotlib, Seaborn, Streamlit
Core Coursework: Data Structures, Algorithms, OOP, Computer Systems, Probability & Statistics, Discrete Math, Linear Algebra, Multivariate Calculus, Differential Equations

Experience

Deloitte India Jul 2025 – Aug 2025
Summer Intern *Gurugram, India*

- Engineered a RAG pipeline to assign Gardner scale ratings to 2,000+ embryo images using DINOv2 and FAISS (HNSW), achieving 84% accuracy.
- Integrated the RAG output into an ensemble forecasting system (XGBoost + Random Forest + Linear Regression) for IVF success prediction, using 30+ clinical features.
- Used Boruta to rank key biomarkers; identified top 5 predictors influencing embryo viability.
- Collaborated with clinicians and Deloitte's healthcare clients to validate model fairness and interpretability across diverse demographics.
- Reduced inference latency to <2s/query and improved calibration across age groups by 18%.

National Informatics Centre (NIC) Jun 2025 – Jul 2025
AI Intern *New Delhi, India*

- Migrated NIC's chatbot VANI from DialogFlow to a 2-stage RAG stack with FAISS + NVIDIA NeMo for multilingual response generation.
- Implemented Redis-backed memory pipeline for ingestion, context storage, and modular retrieval logic; supported memory across multilingual prompts.
- Deployed on DGX A100 using TensorRT + Triton Inference Server, achieving sub-100ms latency.
- Improved user experience by reducing irrelevant responses by 30% and expanding language coverage.

Grant Thornton Bharat LLP Jun 2024 – Aug 2024
Data Analytics Intern *New Delhi, India*

- Forecasted India's Multidimensional Poverty Index (MPI) using time-series + ARIMA modeling; built simulation models for policy forecasting.
- Designed Scheme Performance Index (SPI) to benchmark scheme efficiency; proposed funding reallocations to maximize MPI reduction.
- Visualized outcomes and budget strategies using PowerBI dashboards; contributed to national policy insights aligned with Viksit Bharat 2047.

Projects

AI-Powered Financial Doc Search Apr 2025 – Jun 2025

- Built local RAG pipeline to semantically search/summarize >100 documents using LangChain + Qwen; vectorized via Milvus Lite.
- Integrated Llama 3.2 Instruct with grounding; reduced hallucinations by 60% and enabled <1s semantic search latency.

BUILD UMass Sep 2023 – May 2024

- Led 8-member team in Scrum sprints to build education foundation website with tiered workflows using MERN stack; helped raise $25K+ for local schools.
- Designed PCI-compliant workflows and responsive components via Figma + Tailwind CSS; conducted demos with Amherst Foundation.