Chinmay Arora

Data Scientist | Data Analyst | Machine Learning Engineer | Gen AI Engineer

Minneapolis, USA   737-340-7827   chinmayarora2001@gmail.com

Education

Master of Science (M.S.)   University of Minnesota, Twin Cities, MN, USA
Data Science, 3.59 GPA   09/2023-12/2025

Relevant Coursework: Data Mining, Machine Learning Fundamentals, Categorical Data Analysis, Multivariate Analysis, Spatial Data Science, Visualization with AI, Principles and Architecture of Database Systems, AI using Bayesian Optimization and RL for Decision Making

Graduate Research Volunteer: Data Analysis and Management Research Group, Dr. Jaideep Srivastava – work on Gen AI, RAG agentic systems and SLMs

Bachelor of Technology (B.Tech.)   SRM Institute of Science and Technology, Kattankulathur, India
Computer Science and Engineering, 8.98/10 GPA   06/2019-05/2023

Activities: Project Head – Enactus SRM (led student social innovation projects); Product Manager – Enactus Marketplace (campus startup initiative)

Relevant Work Experience

Graduate Research Assistant, CURA, University of Minnesota   **Minneapolis, MN, USA; 05/2025-12/2025**

- **Developed ML-driven resident segmentation and AI integration frameworks** with Hennepin County's Digital Experience and Innovation & Data Analytics teams, applying **demographic clustering, behavioral modeling, and automated ETL pipelines** to enhance digital accessibility and Power BI reporting.
- **Analyzed AI bias and** fairness in recommendation algorithms on public data, presenting findings that guide the integration of **responsible AI** checks (bias mitigation, transparency features) in the county's decision-making tools.

Data Science/ ML Intern, MarketMakerCRE (Real Estate Analytics)   **Remote/ Fl, USA; 07/2024-09/2024**

- Engineered a machine learning pipeline for real estate analytics, employing Azure cloud services to automate data ingestion and model deployment. Ensured model transparency by using SHAP and LIME for explainability in property price predictions, aligning the AI solution with stakeholder trust requirements.
- Collaborated with a multidisciplinary team to incorporate predictive AI insights into the product, balancing cutting-edge model performance with interpretability and user trust.

IT & AI Solutions Intern, Daikin India (R&D Division)   **Gurugram, India; 06/2022-07/2022**

- Developed an NLP-based chatbot for internal knowledge base queries, utilizing embedding techniques (Pinecone vector store) to enable semantic search and natural language responses. This early AI solution improved information accessibility for employees and demonstrated the potential of conversational AI in the company.
- Worked on integrating explainable AI techniques so that the chatbot could provide confidence levels or source context for its answers, promoting user confidence in the system's recommendations.

Projects/ Publications

Financial Balance Sheet RAG System – Engineered a **hybrid agentic RAG pipeline** for automated financial Q&A over balance sheets and filings using **LangChain, HuggingFace, ChromaDB, and Supabase**, deployed on a **private server** with containerized **FastAPI microservices**. Integrated **Text-to-SQL** for tabular reasoning and **multi-hop retrieval**. Implemented **MLOps workflows** with **Docker**, CI/CD automation, **MLflow**, and real-time evaluation via **RAGAS** and **Prometheus**, ensuring scalable, secure, and explainable retrieval performance.

Legal Document RAG System: Experimented with LLama, DeepSeek, and OpenAI APIs for retrieval-augmented generation. Integrated NVidia CUDA and OLLAMA to optimize performance using Crawl4AI, FastAPI, Flask, and Azure DB linkage for dynamic data extraction. Exploring vector databasing, embedding models, FAISS similarity and Langchain.MLOPs deployment using Docker, Hostinger, n8n and GCP for OCR

Generative AI Assistant Chatbot – Created a web-based AI assistant using multi-step reasoning and chain-of-thought prompting. Leveraged **LangChain** with HuggingFace transformers and a Supabase vector store to fetch relevant context for GPT-4 style responses. Implemented guardrails and user prompt filtering to ensure safe and reliable interactions, aligning with responsible AI best practices.

Interactive Pac-Man with XAI – (Academic Showcase) Developed a variant of Pac-Man game with an explainable AI agent. Integrated interpretable decision visualizations to demonstrate how the AI agent's neural network made decisions in real time, as an educational tool for AI transparency.

Relevant Skills

Generative AI: Large Language Models (GPT-3.5/Turbo, GPT-4, LLaMA family), Prompt Engineering, Few-shot prompting, Chain-of-Thought reasoning, LangChain framework, Hugging Face Transformers, RAG pipeline design.

AI/ML Expertise: Deep Learning (CNNs, RNNs, Transformers), NLP (spaCy, NLTK, seq2seq models), Computer Vision basics (OpenCV), Reinforcement Learning (fundamentals), Explainable AI (SHAP, LIME, interpretability techniques).

Development & Deployment: Python, FastAPI/Flask (API development), Docker & Kubernetes (containerization and orchestration), RESTful API integration, CI/CD for model updates.

Data & Cloud: SQL & NoSQL databases; Vector databases (Pinecone, FAISS); Azure and GCP cloud services for AI (Azure ML, GPU VM instances); experience with distributed computing for AI workloads.

Responsible AI & Research: Familiar with AI ethics principles, bias mitigation strategies, and human-centered design for AI. Experienced in writing research documentation and presenting technical findings to both technical and non-technical audiences.

Collaboration: Strong communication and teamwork skills, agile research & development approach, and a passion for mentoring and learning in the fast-evolving AI domain.