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**
- As part of the *Hennepin-University Partnership (HUP),* collaborated with Hennepin County's Digital Experience and Innovation & Data Analytics teams to create **data-driven resident personas** to enhance digital service accessibility by identifying local community segments. Applied demographic clustering, behavioral analysis, and segmentation modeling on census and marketing data to inform inclusive, user-centered design improvements for county-wide digital platforms.
- Developing an **AI integration and evaluation framework** for responsible use of AI in county operations to define **impact and risk metrics**, assessing fairness, and exploring **human-in-loop and bias mitigation strategies**.

Data Science Intern, MarketMakerCRE (Real Estate Analytics)   **Remote/ Fl, USA; 07/2024-09/2024**
- Developed and deployed **scalable data pipelines** using Python (BeautifulSoup, Selenium, aiohttp) and Azure (Data Factory, Functions) to collect, transform, and store real estate data in MongoDB and Azure SQL, improving acquisition speed by 4x.
- Applied machine learning models (XGBoost, Random Forest, KMeans) and statistical techniques (PCA, SHAP, LIME) for property valuation and segmentation, enhancing predictive insights and model transparency.

Data Analyst - Marketing Intern, Confetti AI (EdTech/Careers Startup)   **Remote/ Fl, USA; 05/2024-07/2024**
- Performed data-driven marketing analytics using **A/B testing**, engagement analysis, and statistical modeling for optimizing targeted campaigns and improving conversion rates by 25% while enhancing customer segmentation accuracy.
- **Developed interactive analytics dashboards** to track key KPIs and collaborated with DevOps to refine AI-driven algorithms for boosting user engagement with career tools by 20% and accelerating feature deployment.

IT/ Data Intern, Daikin India (Manufacturing Division)   **Gurugram, India; 06/2022-07/2022**
- Designed and deployed QlikSense Sales Analytics Dashboards and Python-based NLP chatbots (Flask, Pinecone, spaCy, NLTK), aiming to boost sales visibility by 35%, improve decision speed by 25%, and reducing customer response time by 40%.
- Automated employee workflow and reporting systems using ASP.NET (C#), enhancing operational efficiency by 50% and enabling data-driven insights through real-time monitoring.

Projects

Business Process Optimization Using Social Media Data: Leveraged **NLP, sentiment analysis, Reinforcement Learning** and predictive modeling to extract insights from social media data, enabling virality detection for advertising optimization, **time series forecasting** for trend analysis, and data-driven decision-making to enhance marketing ROI and business strategy using Influence Maximization Optimization

Designed and implemented ML and data analysis projects using **Python, R, SQL, and NLP**, including **Stock Market Forecasting**, **Sports Performance Prediction**, and **Manufacturing Cost Optimization**, improving model accuracy and business insight generation.

Developed **OCR-based document parsing** and **cross-platform ETL middleware**, automating data extraction and synchronization across systems, reducing manual data entry and enhancing real-time analytics capability.

Relevant Skills

Programming & Data: Python (Pandas, NumPy, Scikit-learn), R, SQL, C++; Data Cleaning & EDA, Web Scraping (BeautifulSoup, Selenium); **Database:** SQL Server, MongoDB.

Machine Learning: Supervised & Unsupervised Learning; Predictive Modeling; **ML Frameworks:** scikit-learn, TensorFlow, PyTorch, XGBoost; **Deep Learning:** CNNs, basic RNNs; **NLP:** spaCy, HuggingFace Transformers.

Data Visualization & BI: Tableau, Power BI, QlikSense; matplotlib, Seaborn; Dashboard Design & KPI Reporting.

Cloud & Big Data: Azure (Data Factory, ML Studio), Google Cloud (BigQuery); Data Warehousing (Snowflake, Redshift); Big Data Tools (Spark, Hadoop basics).

Other Tools: Git/GitHub, Docker; FastAPI & Flask (API development); **MLOps:** Model Deployment and lifecycle management; **Explainable AI:** SHAP, LIME; **Soft Skills:** Collaboration, presentation, problem-solving, agile methodology familiarity.