

# Rohit Akole

[rohitakole5@gmail.com](mailto:rohitakole5@gmail.com) | [rohitakole50.github.io](https://github.com/rohitakole50) | (959) 223 0185 | [LinkedIn: rohitakole](https://www.linkedin.com/in/rohitakole) | Hartford, CT

## DATA SCIENTIST

Skilled in machine learning, deep learning, statistical modeling, and predictive analytics. Proficient in Python, R, SQL, & cloud platforms like AWS & Snowflake, with expertise in data analysis, visualization, & scalable solutions. Placed in the top 10% at Humana-Mays competition, demonstrating leadership in solving complex challenges.

## EDUCATION

**University of Connecticut**, Hartford, CT August 2023 – May 2025  
*Master of Science (M.S.), Business Analytics and Project Management (Data Science)* GPA: 4.00/4.00

- Coursework: Statistics in Business Analytics, Predictive Modeling, Data Mining & Business Intelligence, Data Science using Python, Agile Project Management & Methodologies, Data Management & Business Process Modeling, Business Decision Modeling, Deep Learning

**University of Bridgeport**, Bridgeport, CT January 2017 – May 2019  
*Master of Science (M.S.), Technology Management (IT and Big Data)* GPA: 3.89/4.00

- Coursework: Finance & Accounting, Strategic Sourcing & Vendor Management, Intro to Big Data & Data Science

**North Maharashtra University**, Jalgaon, India June 2012 – May 2015  
*Bachelor of Business Management*

- **Coursework:** Financial & Cost Accounting, Corporate Accounting, Managerial Economics, MS Access, Business Mathematics & Statistics, System Analysis & Designing, Business Management, MIS & ERP, Oracle

## EXPERIENCE

**Alo Index**, Rochester, NY January 2025 – April 2025

### Data Science Associate Intern

- Automated ESG data integration for 6,000+ hotels using APIs & Python, extracting data from 33 certifications.
- Validated 1.2M records and applied supervised ML to resolve inconsistencies, improving data accuracy by 97%.
- Pre-filled 300+ ESG onboarding questions using NLP, reducing client onboarding time by 80% (5+ hours per client).
- Created an LLM-powered assessment engine with Ollama, Mistral, & semantic search. It mapped 300+ ESG questions to certification text, automating evidence extraction from PDFs & cutting manual review time by 80%.
- Collaborated with founders to translate business insights into data-driven strategies, improving product decision-making and increasing operational efficiency by 25%; gained leadership exposure in fast-paced startup environment.

**Stanley Black & Decker**, Hartford, CT January 2025 – May 2025

### Data Scientist Consultant (Capstone)

- Forecasting 12-month warranty claim volumes & costs using time series modeling on 5 years of historical data (~375K claims, 1.6M+ rows) to support strategic planning.
- Leveraging LLM-based (Llama3 using Docker) text mining on dealer communications (diagnosis & repair fields) to identify recurring issues & extract nine actionable features.
- Developed an interactive Tableau dashboard to present historical trends, forecast outputs, & text mining driven visualizations (e.g., word clouds) for cross-functional use.

**University of Connecticut**, Hartford, CT January 2025 – May 2025

### Graduate Teaching Assistant – Introduction to Deep Learning

- Evaluated graduate-level assignments with a focus on code quality, model performance, & documentation.
- Delivered individualized, constructive feedback to support student learning, model implementation best practices.
- Collaborated with the teaching team to maintain academic standards & ensure timely grading cycles.

**Modlee AI/ML Student Club**, Hartford, CT September 2024 – May 2025

### President and Founder

- Founded & led the Modlee AI/ML Student Club (MAIC) at University of Connecticut, creating a collaborative platform for students of all backgrounds to explore AI & machine learning.
- Organized hackathons, competitive events & projects, providing members with firsthand experience & industry insights through events featuring AI/ML professionals.
- Featured twice in Student Highlights by the University for leadership & dedication to fostering a supportive, innovative learning environment in AI/ML.

## PROJECTS

**Travelers Modeling Competition** (XGBoost, Lasso Regression, Gradient Boosting) September 2024 – December 2024

- Developed and implemented ML models including XGBoost, Lasso Regression, and Gradient Boosting to forecast claim volumes for call center operations, enhancing resource planning.
- Performed data cleaning, exploratory data analysis, and data visualization using Python.

**Humana-Mays Healthcare Analytics Case Competition 2024** (XGBoost, LightGBM, CUDA) August 2024 – October 2024

- Developed CUDA optimized XGBoost model for a dataset of 300+ columns & 1.6 million rows, achieving 77% accuracy (AUC: 0.7686), identified KPI such as claims history & chronic conditions addressing preventive healthcare visit gaps.
- Implemented strategies such as mobile clinics, telehealth services, and outreach programs, which improved healthcare access and patient engagement, resulting in increased preventive care visits.

**Predictive Modeling of Adolescent Digital Overuse** (TSFresh, SMOTE, FNN, SHAP, Parquet Processing)

- Processed 986 time-series parquet files to engineer 300+ features from physical activity, heart rate, & sleep data, enabling predictive modeling using deep learning techniques of adolescent internet overuse.
- Achieved  $R^2$  of 0.72 using Feed-Forward Neural Network, with SHAP analysis revealing sedentary time & disrupted sleep as top predictors; applied SMOTE to improve model generalization & fairness.
- Delivered key correlations through multivariate visuals to support early interventions in pediatric behavioral health.

**Insurance Fraud Detection Using Machine Learning** (Decision Tree, Random Forest, Logistic Regression)

- Engineered machine learning models to detect insurance fraud, achieving a precision of 92% & a recall of 88%.
- Preprocessed & analyzed 50,000+ insurance claims data, performing feature selection & extraction which improved model accuracy by 15%.
- Implemented the final Random Forest model with hyperparameter tuning, reducing false positives by 20% & overall fraud detection rate by 25%, leading to projected annual savings of \$500,000.

**AI-Powered Gmail Assistant** (Gmail API, Gemini AI, NLP, TextBlob, Pandas, PyMuPDF, LLM, Agentic AI)

- Developed an AI-driven agent that reads, summarizes, & prioritizes unread Gmail messages, parsing both email content & attachments using NLP & LLM based summarization.
- Automated email replies with Gemini AI by generating sentiment-aware summaries & three personalized response options with dynamic placeholders & in-thread Gmail API integration.
- Demonstrated the end-to-end intelligent workflow & internal logic behind real-time, human-like email handling.

## ACHIEVEMENTS

- 1st Place, Data Science Capstone – Ranked 1st among 5 teams at Stanley Black & Decker.
- 1st & 2nd Place, Hackathons – Won 3 hackathons and placed 1<sup>st</sup> in 2 and 2<sup>nd</sup> in 1 hackathon organized by [Modlee.ai](#).
- Top 3, Travelers Modeling Competition – Ranked 3rd out of 22 teams for call center operations forecasting.
- Top 10%, Humana-Mays Case Competition – Ranked 26th out of 290+ teams in national healthcare analytics challenge.
- Invited Grader, Intro to Deep Learning – Personally selected by professor based on top performance.
- Featured twice in Student Highlights, University Weekly Scoop – Recognized for contributions to AI/ML community.
- AI Agent Demo Speaker, Hartford AI Day 2025 – Invited to deliver a live LLM-based agent demo at an event hosted by Launc[H], attended by 50+ industry professionals (technical and non-technical audience).
- Received strong endorsement from CEO & CTO of Alo Index for performance, speed, and dedication.

## SKILLS

**Programming:** Python, SQL, PySpark, Scala, CUDA, Pandas, NumPy, Scikit-learn, HTML5, CSS3, JavaScript

**Data Science & Analytics:** Machine Learning, Deep Learning, Predictive Modeling, A/B Testing, Time Series Forecasting, Natural Language Processing (NLP), Generative AI, Sentiment Analysis, Feature Engineering, Statistical Modeling, Model Evaluation, Exploratory Data Analysis (EDA), PyTorch, Snowflake, AWS

**Tools & Libraries:** GitHub, Tableau, Power BI, Django, Excel, Access, SAS Studio, JIRA, SAS Enterprise Miner, R

**Data Engineering & Infrastructure:** Extract Transform Load (ETL), Data Pipelines, Data Warehousing, APIs, REST APIs, SQL Server, Relational Databases, Big Data, Docker, GPU Programming, MLOps, Prompt Engineering

**Agentic AI & Large Language Models:** Semantic Search, Mistral, LLaMa3, PyMuPDF, Gemini, Prompt Engineering

**Business & Methodologies:** Business Intelligence, Data Analysis, Data Mining, Data Visualization, Dashboards, Agile, Scrum, Strategic Recommendations, Cross-Functional Collaboration, SaaS

**Certifications:** [IBM Data Science, professional certification by IBM on Coursera, 2021](#); [IBM Data Analyst, professional certification by IBM on Coursera, 2022](#)