

# ROHIT AKOLE

Hartford, CT | [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)

## DATA SCIENTIST

Data Scientist with a passion for transforming data into actionable insights. Skilled in programming, data analysis and visualization, project management, machine learning, deep learning, statistical modeling, and predictive analytics.

## EDUCATION

**UNIVERSITY OF CONNECTICUT**, Hartford, CT

**Master of Science (M.S.), Business Analytics and Project Management (Data Science)**, May 2025

- GPA: 4.00/4.00 | Key Courses: 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
- **Leadership**: President and Founder of Modlee AI/ML Student Club. Organized hackathons, competitive events and projects, providing members with firsthand experience & industry insights through events featuring AI/ML professionals.
- **Speaking Engagement**: 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).
- **Recognition**: Featured twice in Student Highlights, University Weekly Scoop – recognized for contributions to AI/ML community.

**UNIVERSITY OF BRIDGEPORT**, Bridgeport, CT

**Master of Science (M.S.), Technology Management (IT and Big Data)**, May 2019

- GPA: 3.89/4.00 | Key Courses: Finance & Accounting, Strategic Sourcing & Vendor Management, Intro to Big Data & Data Science

**NORTH MAHARASHTRA UNIVERSITY**, Jalgaon, India

**Bachelor of Business Management**, May 2015

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

## EXPERIENCE

**ALO INDEX**, New York, NY

January 2025 – April 2025

**Data Science Associate Intern**

- Automated ESG data integration for 6K+ 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%.
- Reduced client onboarding time by 80% (5+ hours per client) by pre-filling 300+ ESG onboarding questions using NLP.
- Created an LLM-powered assessment engine with Ollama, Mistral, & semantic search. It mapped 300+ ESG questions to certification text, automating evidence extraction from PDFs and cut 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)**

- Forecasted 12-month warranty claim volumes & costs using time series modeling on five years of historical data (~375K claims, 1.6M+ rows) to support strategic planning.
- Leveraged 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 and text mining driven visualizations (e.g., word clouds) for cross-functional use.
- Earned 1st Place – ranked 1st among five teams.

**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 and documentation.
- Delivered individualized, constructive feedback to support student learning and model implementation best practices.

**GREETWORLD HOLIDAYS**, Pune, India

January 2020 – December 2022

**Business Operations Lead – Tourism Startup**

- Restructured business strategy during COVID-19, stabilizing operations and adapting to shifting demand.
- Optimized transportation routes and tour schedules, reducing logistics costs by streamlining operations.
- Led customer outreach initiatives to re-engage existing clients and attract new segments.
- Negotiated partnership with a booking platform to secure improved commission rates and increase booking visibility.
- Applied routing and cost minimization techniques to enhance tour profitability and resource allocation.

**KS WEBAPPS**, Pune, India

May 2014 – November 2016

**Software Developer** (July 2015 – November 2016) & **Software Developer Intern** (May 2014 – June 2015)

- Built and tested scalable application components (85–90% success rate) in Agile, Test-Driven Development (TDD) environment using Django REST Framework.

- Optimized Python-based queries and workflows, improving application performance by 25%.
- Executed and tuned DML/DDI queries using MySQL and Python-MySQL connector for backend integration.
- Designed RDBMS models for small business clients using MS Visio ERP tools.
- Delivered 15+ client presentations with a cross-functional team to support implementation and deployment.

## PROJECTS

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

*Ranked 3rd out of 22 teams for call center operations forecasting.*

- 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

*Ranked 26th out of 290+ teams, top 10% in national healthcare analytics challenge.*

- Developed CUDA optimized XGBoost model for a dataset of 300+ columns & 1.6M rows, achieving 77% accuracy (AUC: 0.7686), identified KPI such as claims history and 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 and 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 and 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% and a recall of 88%.
- Preprocessed/analyzed 50K+ insurance claims data, performing feature selection & extraction improving model accuracy by 15%.
- Implemented the final Random Forest model with hyperparameter tuning, reducing false positives by 20% and overall fraud detection rate by 25%, leading to projected annual savings of \$500K.

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

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

## ADDITIONAL INFORMATION

- **Hackathons:** 1st & 2nd Place - Won three hackathons and placed 1<sup>st</sup> in two and 2<sup>nd</sup> in one hackathon organized by [Modlee.ai](#).
- **Publications:**
  - [Rohit Vikas Akole \(2021\): Operation and Significance of Supply Chain Management \(SCM\) in Business; International Journal of Scientific and Research Publications \(IJSRP\) 11\(6\) \(ISSN: 2250-3153\)](#)
  - [Akole, R. V. \(2019\) Problems and Solutions for Project Management Information Systems. International Journal of Science and Engineering Investigations \(IJSEI\), 8\(87\), 80-85.](#)
- **Peer Reviews:** Reviewed manuscripts for *Asian Journal of Economics, Business and Accounting* (Certificate No: PRAJEB113398ROH), March 2024 and for *Archives of Current Research International* (Certificate No: PRACRI121283ROH), August 2024.
- **IBM Professional Certifications:** IBM Data Analyst (Coursera, 2022), IBM Data Science (Coursera, 2021)
- **Languages:** Fluent in Hindi and Marathi
- **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
  - **Methodologies:** Agile, Scrum, Kanban