ROHIT AKOLE

Hartford, CT | rohitakole5@gmail.com | rohitakole50.github.io | (959) 223 0185 | LinkedIn: 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
- Awarded merit-based scholarship twice (Spring 2024 and Spring 2025) for outstanding academic performance.
- *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 (E-Commerce), 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

- Automated ESG data integration for 6.7K+ hotels using APIs & Python, extracting data from 33 certifications.
- Validated 1.2M+ records and applied supervised ML to resolve inconsistencies, improved data accuracy by 97%.
- Reduced client onboarding time by 83% (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 81%.
- Collaborated with founders to translate business insights into data-driven strategies, improved product decision-making and increased 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, improving issue resolution processes.
- Developed an interactive Tableau dashboard to present historical trends, forecast outputs and text mining driven visualizations (e.g., word clouds) for cross-functional use.
- Achieved 1st place among five teams in data science competition by applying advanced data science techniques and solutions.

UNIVERSITY OF CONNECTICUT, Hartford, CT   January 2025 – May 2025
Graduate Teaching Assistant – 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 – May 2023
Business Operations Lead

- Restructured business strategy during COVID-19, stabilizing operations and adapting to shifting demand.
- Optimized transportation routes, resource allocation, 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.

KS WEBAPPS, Pune, India   May 2014 – November 2016
Software Developer (July 2015 – November 2016) & ***Software Developer Intern*** (May 2014 – June 2015)

- Improved application performance by 25% by optimizing Python workflows and backend logic in Agile Framework.

- Built and tested scalable application components in an Agile, Test-Driven Development (TDD) and CI/CD environment using Django REST Framework, ensuring high reliability and performance.
- Enabled seamless deployment across 10+ systems by tuning DML/DDL queries via MySQL and Python connector.
- Accelerated ERP integration for 5+ clients by designing RDBMS schemas and delivering 15+ cross-functional presentations.

PROJECTS

Travelers Modeling Competition (XGBoost, Lasso Regression, Gradient Boosting)   September 2024 – December 2024
Ranked Top 3 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.

S&P500 Stock Analysis: Evaluating Risk & Return (Modern Portfolio Theory, Integer Programming, Monte Carlo Simulation, YFinance)
- Built a stock portfolio optimization framework using Modern Portfolio Theory (MPT) across Utilities, Healthcare, and Real Estate sectors of the S&P 500, leveraging expected return and covariance matrices for asset selection.
- Simulated over 1,000+ portfolios using Monte Carlo techniques to evaluate Sharpe ratios and frontiers under different rebalancing strategies; applied integer programming for budget-constrained portfolio selection.
- Pulled historical pricing using Yahoo Finance API and visualized risk-return tradeoffs with heatmaps and frontier plots.

ADDITIONAL INFORMATION

- *Hackathons:* 1st & 2nd Place - Won three hackathons and placed 1st in two and 2nd 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*:
  - Archives of Current Research International (Certificate No: PRACRI121283ROH), August 2024.
  - Asian Journal of Economics, Business and Accounting (Certificate No: PRAJEBA113398ROH), March 2024.
- *IBM Professional Certifications:* IBM Data Analyst (Coursera, 2022), IBM Data Science (Coursera, 2021)
- *Languages*: Fluent in Hindi and Marathi
- *Skills*:
  - **Programming & Scripting:** Python, SQL, PySpark, CUDA, HTML5, CSS3, JavaScript, R
  - **Libraries & AI Tools:** Pandas, NumPy, Scikit-learn, PyTorch, Pyomo, Mistral, LLaMa3, PyMuPDF, Gemini AI, Prompt Engineering
  - **APIs and Integration:** OAuth 2.0, Postman, Rest API
  - **Data Science & Machine Learning:** Machine Learning, Deep Learning, Tensorflow, SHAP, Predictive Modeling, Time Series Forecasting (SARIMAX, Prophet), Natural Language Processing (NLP), Generative AI, Semantic Search, Sentiment Analysis, Feature Engineering, Statistical Modeling, Model Evaluation, Exploratory Data Analysis (EDA), A/B Testing, t-test, ANOVA
  - **Data Engineering & Automation:** Extract Transform Load (ETL), Data Pipelines, Automation, Data Integration, Linear Optimizations, Non-Linear Optimizations, GPU Programming, MLOps, Docker
  - **Business Intelligence & Reporting:** Tableau, Power BI (DAX), Excel, Access, SAS Studio, SAS Enterprise Miner, Lucid Chart
  - **Cloud & Database Warehousing:** Snowflake, AWS, SQL Server, Relational Databases, Big Data, Oracle, Data Warehousing
  - **Project Management & Collaboration:** Agile, Scrum, CI/CD, Kanban, JIRA, GitHub, KPI Reporting, MS Project, Sharepoint