

ROHIT AKOLE

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

WORK EXPERIENCE

Technology Solutions Architect | CIS Labs, Inc.

January 2026 – Present

- Built and optimized cloud-native data platforms by designing Snowflake schemas, writing performant SQL transformations, and developing Airflow DAGs to reliably ingest, validate, and serve data on GCS and Snowflake for application and analytics use cases
- Maintained and enhanced CI/CD pipelines using Bitbucket and Jira to support automated builds, testing, and deployments, while refactoring services and pipeline code to improve maintainability, reduce defects, and accelerate release cycles
- Engineered AI-driven financial document processing system using Ollama to split, classify, and organize multi-document files, persisting outputs to structured directories in GCS to improve accessibility and reduce manual document handling for client data
- Developed backend services, integrating database-backed data access layers while supporting data migrations and implementing validation and reconciliation checks across GCS and Snowflake to ensure correctness, performance, and secure data exchange

Data Engineer | University of Connecticut | Hartford, CT

July 2025 – December 2025

- Built distributed PySpark ETL pipelines for climate and energy data, processing over 240K records adding hourly Cloud Run scraper that collects and stores raw XML files in Bronze layer of GCS data lake
- Designed robust data transformation flows: cleaned, validated, and organized Silver layer parquet data within the data lake before loading the curated outputs into Gold layer tables in BigQuery, using Spark to parse and flatten XML fields for downstream use
- Automated ingestion and transformation using scheduled queries and Cloud Run jobs, orchestrating movement of parquet files from data lake into BigQuery, with reproducible deployments handled through GitHub Actions for consistent pipeline updates
- Developed RNN/LSTM anomaly detection pipelines, integrating their outputs into Gold layer as additional columns so monitoring tasks and analytics workflows can access everything directly

Data Engineer Intern | Alo Index | New York, NY

January 2025 – April 2025

- Automated ESG data ingestion using NLP-driven autofill on Cloud Run, converting unstructured documents into structured datasets and loading them into a GCS bucket, processed downstream through BigQuery
- Built and maintained CI/CD pipelines (GitHub Actions, Docker) for BigQuery jobs and API services, and documented end-to-end architecture to ensure audit readiness and reproducible data operations
- Deployed an in-house extractive LLM ingestion pipeline (Mistral, Ollama, PyMuPDF) with RAG to automate document processing, reducing manual validation costs by \$5.9K per month and improving data quality
- Accelerated client onboarding by 40 percent through A/B-tested segmentation pipelines and created data quality monitors flag data drift and maintain reliable ingestion
- Boosted client acquisition by 22%, presented data-backed growth strategies to C-suite, shaped GTM initiatives that drove sales

Data Engineer | Stanley Black and Decker | Hartford, CT

December 2024 – May 2025

- Achieved 83% accuracy by productionizing a forecasting model using 5 years of housing and mortgage-related data; performed cash flow modeling and ROI analysis to support compliance and investment decisions
- Monitored model performance using precision, recall, and F1 metrics with automated alerts for data drift, enabling retraining trigger, ensuring continuous model accuracy; drafted pipeline documentation for architecture, CI/CD workflows for reproducibility
- Cut analysis time by 30% and eliminated manual errors by engineering a generative LLM pipeline with prompt engineering using Ollama, LLaMa3, Docker, Python, and APIs to extract 9 features from 3M+ unstructured dealer logs
- Shortened manual analysis time by 40% for monthly reviews by building a Python-based ETL pipeline that integrated forecasts and LLM-derived features into a centralized Tableau dashboard for executive decision-making
- Ranked 1st for delivering \$1.5M+ projected value with production-ready ML solutions for the capstone project; Led team of 4 using Agile (standups, sprints, RACI) with JIRA, Python, Tableau to shape roadmap and cross-functional strategy

Machine Learning Engineer | Greetworld Holidays | Pune, India

October 2019 – July 2023

- Decreased operational costs by 15% and amplified NPS by 12% by running A/B tests on 5+ service workflows, analyzing 50K+ customer feedback records with SQL & Python, identifying process bottlenecks and prioritizing high-impact service enhancements
- Maximized repeat bookings by 37% by training segmentation models (K-Means, decision trees) on demographic and behavioral data from 120K+ travelers using Python, Pandas, and Scikit-learn, enabling targeted outreach to high-value cohorts
- Built financial reporting dashboards in Tableau and Excel, reconciling transactional data across booking, payment, and accounting systems to support accurate month-end closing, budget tracking, forecast variance analysis, revenue recognition
- Increased revenue by 20% and elevated campaign ROI by 18% by partnering with marketing and operations leaders to translate analytics into personalized travel packages and dynamic pricing strategies

Data Analyst | Greetworld Holidays | Pune, India

May 2015 – November 2016

- Increased website traffic by 18% and revenue by 8% by analyzing booking data in SQL and Excel, later migrated to Redshift-based data warehouse for scalable reporting

- Improved onboarding efficiency by 30% and trimmed service downtime by 25% by building Python and SQL-based workflows (SSIS) to monitor real-time system metrics, automate ranking/filtering of 10K+ property listings, and resolve issues proactively
- Optimized client engagement and minimized churn by 10% by working with sales and support teams to segment customers based on travel history and purchase behavior, enabling targeted campaigns and personalized recommendations

EDUCATION

M.S. in Business Analytics & Project Management | *University of Connecticut* | *Hartford, CT* | GPA: 4.0/4.0

- **Graduate TA – Deep Learning:** Graded 30+ students, evaluated Deep Learning & Neural Network assignments

M.S. in Technology Management (IT and Big Data) | *University of Bridgeport* | *Bridgeport, CT* | GPA: 3.9/4.0

Bachelor of Business Management (E-Commerce) | *North Maharashtra University* | *Jalgaon, India* | GPA: 6.9/10.0

SKILLS

- **Programming and Libraries:** Python, SQL, PySpark, Pandas, NumPy, Scikit-learn, Seaborn, Matplotlib, SciPy
- **Frameworks & Cloud:** GCP, Airflow, Snowflake, PyTorch, AWS, TensorFlow, Docker, REST APIs
- **Machine Learning & AI:** Causal Inference, Natural Language Processing, Sentiment Analysis, Recommender Systems
- **Modeling:** Classification, Regression, Predictive Modeling, Time Series Forecasting, Dimensionality Reduction (PCA), SHAP, ETL, Text Mining, Feature Engineering, Sampling, Scaling, Model Evaluation, Exploratory Data Analysis (EDA), CUDA, ML DAGs
- **LLMs:** Prompt Engineering, Generative AI, RAG, LLMs (LLaMa3, Mistral, Gemini), Ollama, PyMuPDF, Semantic Search
- **Analytics:** A/B Testing, t-test, Statistical Modeling, Probability, Optimization (Linear, Non-Linear, Integer), Tableau, Power BI (DAX)
- **Version Control & Workflow:** GitHub, CI/CD, JIRA, Confluence, Jenkins, Lucidchart, KPI Reporting

TECHNICAL PROJECTS | [Full Project Portfolio](#)

Machine Learning Engineer | *Humana-Mays Healthcare Analytics Case Competition 2024* | *National Finalist (Top 10%)*

- Led team of 4 to achieve 0.7686 AUC on 1.6M+ claims by training CUDA-accelerated Machine Learning Models with hyperparameter tuning, choosing best performing XGBoost model with advanced feature engineering in Python
- Identified high-risk patients and designed targeted interventions projected to increase preventive visit rates by 37%

Fantasy NFL Lineup Optimization | [GitHub Link](#)

- Faced the challenge of constructing optimal fantasy football rosters where manual selection was inefficient and subject to bias
- Designed and implemented a Integer Linear Programming (MIP) model in Pyomo, applied salary cap and positional constraints, and solved using the CBC solver with no-good cuts to compute both best and second-best lineups
- Delivered a reproducible optimization pipeline that maximized projected fantasy points and demonstrated how mathematical programming can outperform manual lineup creation, providing a clear, data-driven roster selection strategy

S&P 500 Stock Analysis: Evaluating Risk & Return | [GitHub Link](#)

- Compared four portfolio rebalancing strategies in Python (NumPy, Pandas, Matplotlib) using Sharpe ratios and efficient frontiers, identifying the strategy with the highest risk-adjusted return
- Modeled risk-return profiles with Pyomo and Yahoo Finance API, simulating 1,000+ portfolios across three S&P sectors and optimizing allocations under budget constraints, improving projected portfolio return by 15% compared to baseline allocation

AWARDS & LEADERSHIP

- **Scholarships:** Received merit-based scholarships in Spring 2024 and Spring 2025 for top academic performance
- **Leadership:** President & Founder, [Modlee AI/ML Club](#). Hosted events and hands-on projects upskilled 250+ members
- **Keynote:** Delivered live [LLM email agent demo](#) at AI Day, engaged 80+ technical and non-technical professionals
- **Recognition:** Featured twice in University Scoop for contributions to AI/Machine Learning advancement at UConn
- **Hackathons:** Won three hackathons and placed 1st in two and 2nd in one hackathon

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 \(Certificate No: PRAJEB113398ROH\), March 2024](#)