NIHAL SHAH

+1 (862)-579-1974 | nihalshah4@gmail.com | linkedin.com/in/nihalshah4

Data Engineer | Data Scientist | AI / ML & Analytics | Transaction & Operational Data

PROFESSIONAL SUMMARY

Data Engineer with 6+ years of experience delivering enterprise-grade analytics, ETL pipelines, ML-driven workflows, and automated reporting systems across insurance (AIG), higher education (Drew University), and consulting environments. Strong background in SQL, Python, AWS, Snowflake, Tableau, NLP/LLMs, and financial/claims data management. Known for building scalable cloud-ready pipelines, designing KPI frameworks, and leading cross-functional initiatives that improve reliability, reduce downtime, and enable high-impact decision-making. Recognized as an SME in Data Engineering, AI-assisted automation, insurance systems, and analytics transformation.

TECHNICAL SKILLS

Languages: Python, R, SQL, Java
Databases: Oracle 19c, MySQL, Snowflake, Redshift, DB2
Cloud: AWS (EC2, S3, Lambda, Glue), FastAPI, REST APIs
Data Engineering: ETL/ELT pipelines, data modeling, feature engineering, scheduled ingestion, data validation
Machine Learning: Classification models, rule-based systems, forecasting, NLP, semantic search, LLM-assisted workflows
NLP & Research Systems: Text preprocessing, document ingestion, relevance scoring, summarization pipelines
Analytics & BI: Tableau, Power BI, KPI dashboards, operational and financial analytics
Frontend & Apps: Next.js, basic React UIs, Streamlit (execution & visualization)
Tools: Splunk, ServiceNow, Git, JIRA, Confluence, SPSS
Domains: Insurance (claims, payments, transactions), Higher Education analytics, operational and risk data

PROFESSIONAL EXPERIENCE

AA2IT – Data Engineer Remote | July 2025 – Present

- Built cloud-ready data pipelines in Python + AWS Lambda + SQL to unify CRM, sales, finance, and marketing datasets, improving reporting reliability and reducing manual effort by 40%.
- Designed executive-level cashflow, margin, and revenue dashboards using Tableau and Power BI, integrating financial KPIs such as MoM growth, forecast accuracy, churn, CAC, retention, and segment profitability.
- Developed forecasting and what-if models for revenue planning, enabling leadership to simulate pricing, volume, and margin impacts before rollout.
- Automated BI workflows using API integrations, AWS services, and scheduled ETL scripts, reducing cycle time for weekly and monthly reporting.
- Built cohort, funnel, and retention analytics similar to SaaS product performance frameworks, giving clients visibility into customer behavior and pipeline quality.
- Supported Salesforce Revenue Cloud, OmniStudio, and CPQ implementations by structuring revenue data models and enabling downstream analytics.
- Partnered with business development leadership to translate client requirements into scalable data products, creating value propositions that strengthened client trust and repeat business.

CampusWorks — AI Integration Engineer (Client: Drew University) Madison, NJ | Jan 2025 – May 2025

- Built an AI-powered Natural Language to SQL system using Python, Oracle 19c, and OpenAI GPT-4o, enabling non-technical university staff to run on-demand analytics through a conversational interface.
- Architected a microservice-style backend with input parsing, SQL generation via LLMs, validation layers, and secure execution, integrated into Oracle APEX through REST APIs.
- Standardized 3,400+ institutional reports, building metadata structures, data lineage, ownership models, and a governance framework to improve reporting accuracy and reduce duplication.
- Created dashboards for report usage, quality, and lifecycle metrics, giving IT leadership visibility into system-wide performance.
- Worked with institutional research, product teams, and IT engineers to streamline documentation, automate routine analytics, and improve operational decision-making across multiple departments.

Cognizant — Senior Data Analyst (Client: AIG) Chennai, India | Mar 2022 – Dec 2023

- Led a team of 8 analysts to design and implement SOPs for analytics workflows covering claims processing and financial transaction systems, improving operational efficiency by 20%.
- Coordinated root cause analysis (RCA) across engineering, QA, development, and product teams, reducing system downtime by 30% through proactive monitoring and structured triage.
- Developed automated alerting workflows in Splunk, SQL, and ETL tools, accelerating issue detection and minimizing manual intervention during high-severity incidents.
- Built Power BI dashboards tracking system health, SLA performance, claims transaction errors, and user impact, used by senior engineering and product leadership at AIG.
- Created centralized Confluence knowledge bases for processes, release validation, defect triaging, and SLA monitoring, reducing onboarding time and improving cross-team consistency.
- Worked closely with architects on release planning and pre-deployment data validation to maintain 99.9% system uptime for critical insurance applications.

Cognizant — Data Analyst (Client: AIG) Chennai, India | Sept 2018 – Mar 2022

- Supported claims and transaction platforms through analytics, incident triage, and defect management.
- Delivered SQL-based operational and ad-hoc reporting for claims and finance stakeholders.
- Supported release validation and regression testing for high-availability systems.
- Built repeatable data validation and diagnostics processes across policy, payments, and claims modules.

PROJECTS

LexiSQL — Natural Language to SQL AI Platform

- Built an end-to-end conversational analytics system converting natural language into validated SQL queries.
- Implemented full pipeline: prompt ingestion → LLM-generated SQL → validation → execution → formatted results.
- Integrated with Oracle APEX using REST APIs, enabling CSV and PDF report downloads.

AI News Research Recommender

- Built a full-stack AI research assistant using Python, FastAPI, and Next.js to ingest, analyze, and recommend relevant news and research articles.
- Designed backend pipelines for article ingestion, preprocessing, semantic relevance ranking, and summarization.
- Implemented NLP-based search and recommendation logic without OpenAI embeddings.
- Delivered a simple frontend UI enabling topic-based queries and summarized research outputs.

Bank Fraud Rule Explorer

- Built an analytics tool to explore and evaluate fraud detection rules on realistic transaction datasets.
- Implemented multiple rule types including velocity, threshold-based, behavioral, and geo-risk rules.
- Designed metrics to evaluate rule performance such as hit rate, false positives, and coverage.
- Executed and visualized analysis using Python and Streamlit to support explainable, analyst-friendly insights.

Travel Optimizer — Route & Mode Planning System

- Built a FastAPI backend and Next.js frontend to optimize multi-country travel routes.
- Implemented logic to auto-detect travel mode based on distance and geography.
- Visualized routes on a world map with connected paths and location pins.

EDUCATION

Drew University, Madison, NJ
Master of Science in Data Science — May 2025
Dr. A.P.J. Abdul Kalam Technical University, India
Bachelor of Technology in Computer Science — Jul 2018

CERTIFICATIONS

- Salesforce Revenue Cloud Consultant