

Namita Chougule

Ph no: +1-571-564-7751 | namitac9817@gmail.com | [LinkedIn](#) | [Github](#) | [Portfolio](#)

SUMMARY

AWS Certified Data Engineer with 3+ years of experience building, migrating, and operating large-scale ETL and streaming data pipelines in production environments. Proven expertise in Snowflake, Talend, Spark, Kafka, and AWS, with hands-on ownership of multi-terabyte data migrations, real-time processing systems, and cost-optimized cloud architectures supporting finance, insurance, and AI-driven applications. Strong background in SQL, Python, and data modeling, with a focus on reliability, performance, and data quality.

EDUCATION

Master of Science in Computer Science George Mason University GPA: 3.87/4.00	Jan 2024 - May 2025
Bachelor of Engineering in Computer Science & Engineering Visvesvaraya Technological University GPA: 8.18/10.00	Aug 2016 - Aug 2020

EXPERIENCE

Applied AI & Data Developer / onAir Tech Networks - Engineered scalable Python data pipelines integrating Hugging Face LLMs (Pegasus, BART-MNLI) to ingest and transform RSS feed content in MongoDB . - Automated WordPress content distribution through pipeline integration, reducing manual effort by 80% . - Developed RESTful APIs to deliver summarization and classification outputs to live WordPress hubs, supporting production workloads with minimal downtime. - Provisioned and optimized AWS infrastructure (EC2, S3, IAM, CloudWatch) to scale AI data pipelines, improving deployment reliability and efficiency by 30% . - Exploring Airflow-based agentic AI workflows to modularize data ingestion and transformation for scalable orchestration.	July 2025 - Present
---	---------------------

Graduate Teaching Assistant – Database Systems George Mason University - Assisted instruction and grading for a graduate-level course on Database Systems, covering relational schema design, dimensional modeling, normalization, indexing, DDL, DML, DCL, SQL, NoSQL, and transaction management.	Jan 2025 - May 2025
---	---------------------

Data Engineer: ETL Developer / NTT DATA Services Client: Tokio Marine Kiln - Led migration of 60+ TB of insurance/finance data from on-prem MS SQL Server to Snowflake and shifted ETL workloads from SAP BODS to Talend , improving scalability and reducing costs by 40% . - Designed and optimized ETL/ELT pipelines processing 1M+ daily records from AWS S3 (Parquet, CSV, JSON) and relational databases for curated downstream analytics. - Built event-driven pipelines with MSK (Kafka) and Spark Streaming to process 15K+ transactions/min for near real-time analytics use cases. - Re-engineered workflows to reduce batch runtime by 50% (12 hrs to 7 hrs) , enabling faster reporting cycles. - Monitored 1,000+ Snowflake and Talend workflows daily, resolving failures and bottlenecks within SLA timelines. - Developed DBT models on Snowflake to standardize transformations and automated quality checks. - Assisted in integrating Talend with AWS IAM, S3, and Snowflake for secure and auditable data flows. - Collaborated with stakeholders and analysts in an Agile environment to deliver actionable data solutions. - Implemented data governance and quality validation checks , ensuring reliable and compliant reporting datasets. - Created Power BI dashboards to visualize data insights, enabling senior management decisions. - Mentored a team of 2 junior engineers by guiding best practices, conducting code reviews and troubleshooting issues.	Jan 2021 - Dec 2023
---	---------------------

CERTIFICATIONS

AWS Certified Data Engineer – Associate | Agile & Scrum (NLCI) | Oracle Data Engineer (NLCI) | Data Warehousing (NLCI)

TECHNICAL SKILLS

Programming & Development: Python, SQL, Java, Javascript, Web Development, Spring Boot, React.
Data Engineering: Talend, ETL/ELT, SSIS, SSRS, DataBricks, SAP BODS, DBT, AWS Glue, Spark, Kafka, Athena, Airflow.
Databases & Warehousing: Snowflake, MSSQL, Oracle, Postgres, MongoDB, Redshift, DynamoDB, AWS RDS.
DevOps & CI/CD: Docker, Kubernetes, Rancher, Jenkins, GitHub, JIRA.
Cloud & Visualization: AWS, GCP, Azure, Power BI, Quicksight.

PROJECTS

Real-Time Data Pipeline for Smart City Data Kafka, Spark, AWS Redshift, Quicksight - Built an event-driven Kafka + Spark pipeline to process IoT sensor data, enabling real-time analytics in AWS Redshift and visualized insights via Quicksight dashboards.
AI-Powered Hybrid Chatbot Python, React, GPT, Pinecone, RAG - Developed an LLM-based RAG chatbot with semantic memory and web search fallback, enhancing contextual awareness for user interactions.
Customer Support Call Analytics & Insight Platform MySQL, Python, FastAPI, Supabase, SQL, Analytics, React - Built an end-to-end customer support analytics platform with a FastAPI backend aggregating KPI metrics in PostgreSQL (Supabase) and serving rule-based, AI-ready insights via REST APIs to an interactive dashboard.
Cloud-Native Survey Application with CI/CD Deployment Spring Boot, Docker, Kubernetes, Jenkins, AWS RDS - Deployed a Spring Boot-based survey application with a CI/CD pipeline on AWS using Jenkins, Docker, and Kubernetes integrated with an Oracle RDS backend.