

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

## EXPERIENCE

Applied AI & Data Developer / onAir Tech Networks	July 2025 - Present
- Engineered scalable <b>Python data pipelines</b> integrating <b>Hugging Face LLMs</b> (Pegasus, BART-MNLI) to ingest and transform <b>RSS feed</b> content in <b>MongoDB</b> .	
- Automated <b>WordPress</b> content distribution through pipeline integration, reducing manual effort by <b>80%</b> .	
- Developed <b>RESTful APIs</b> to deliver summarization and classification outputs to live WordPress hubs, supporting production workloads with minimal downtime.	
- Provisioned and optimized <b>AWS infrastructure</b> (EC2, S3, IAM, CloudWatch) to scale AI data pipelines, improving deployment reliability and efficiency by <b>30%</b> .	
- Exploring <b>Airflow-based</b> <b>agentic AI</b> workflows to modularize data ingestion and transformation for scalable orchestration.	
Graduate Teaching Assistant – Database Systems   George Mason University	Jan 2025 - May 2025
- 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.	
Data Engineer: ETL Developer / NTT DATA Services	Jan 2021 - Dec 2023
Client: Tokio Marine Kiln	
- Led migration of <b>60+ TB</b> of insurance/finance data from on-prem <b>MS SQL Server</b> to <b>Snowflake</b> and shifted ETL workloads from <b>SAP BODS</b> to <b>Talend</b> , improving scalability and reducing costs by <b>40%</b> .	
- Designed and optimized <b>ETL/ELT</b> pipelines processing <b>1M+</b> daily records from AWS S3 (Parquet, CSV, JSON) and relational databases for curated downstream analytics.	
- Built <b>event-driven</b> pipelines with <b>MSK (Kafka)</b> and <b>Spark Streaming</b> to process <b>15K+</b> transactions/min for near real-time analytics use cases.	
- Re-engineered workflows to reduce batch runtime by <b>50% (12 hrs to 7 hrs)</b> , enabling faster reporting cycles.	
- Monitored <b>1,000+</b> Snowflake and Talend workflows daily, resolving failures and bottlenecks within SLA timelines.	
- Developed <b>DBT models</b> on Snowflake to standardize transformations and automated quality checks.	
- Assisted in integrating <b>Talend</b> with <b>AWS IAM, S3, and Snowflake</b> for secure and auditable data flows.	
- Collaborated with stakeholders and analysts in an <b>Agile</b> environment to deliver actionable data solutions.	
- Implemented <b>data governance</b> and <b>quality validation checks</b> , ensuring reliable and compliant reporting datasets.	
- Created <b>Power BI dashboards</b> to visualize data insights, enabling senior management decisions.	
- <b>Mentored</b> a team of 2 junior engineers by guiding best practices, conducting code reviews and troubleshooting issues.	

## 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.