

Nikhil Doye

Boston, MA | +1 (857) 214-9074 | doye.n@northeastern.edu | [GitHub](#) | [LinkedIn](#)

SUMMARY

Data Engineer with **2+ years of work experience** designing scalable **ETL pipelines**, optimizing **cloud data warehouses**, and automating **business intelligence reporting**. Skilled in **SQL, Python, and AWS (Glue, Redshift, S3)**, with expertise in workflow orchestration (**Airflow**), big data processing (**Spark, Hadoop**), and financial data analytics. Ranked as the **World No.1 SQL Developer** (Practice) on HackerRank. Achieved a **Silver Medal** in NeuroHack by developing a BERT-based model for IT service ticket classification, later integrating HDBSCAN for automated clustering. Earned a **Bronze Medal in a Kaggle Competition** for predicting student dropout rates using random forests, optimizing model performance through hyperparameter tuning.

WORK EXPERIENCE

Fix-It 24/7 Group LLC

Remote

Data Engineer Intern

July 2024 – December 2024

- Designed and optimized SQL-based ETL pipelines to consolidate customer and service request data, reducing query execution time by 40%.
- Developed Power BI dashboards for KPI tracking, replacing manual reporting and saving 30+ hours per week for business teams.
- Built an automated scheduling optimization model in Python, improving technician utilization efficiency by 35% and reducing customer wait times.
- Implemented data quality validation scripts, reducing data discrepancies by 25% and improving reporting accuracy.

Northeastern University

Boston, MA

Graduate Teaching Assistant (MKTG3501 – Marketing Analytics)

September 2024 – December 2024

- Led workshops on SQL-based data transformation and BI reporting, demonstrating real-world marketing data analysis in Power BI and Python.
- Developed and presented an Airflow-based automated ETL pipeline project, integrating real-time campaign performance data into a data warehouse.
- Provided technical mentorship to 40 students, enhancing their ability to work with SQL queries, Python data pipelines, and visualization tools.

LTI-Mindtree

India

Senior Data Engineer

December 2021 – June 2023

- Designed and deployed ETL pipelines using Airflow and Redshift, improving system efficiency by 50% for processing ITSM incident data.
- Developed an automated problem classification pipeline in Python, reducing manual ticket triage time by 40% and saving \$200K annually in support costs.
- Optimized MongoDB queries and migrated workloads to Snowflake, reducing query execution time by 25% and lowering infrastructure costs.
- Built an NLP-based incident clustering solution (BERT + HDBSCAN), improving IT problem identification accuracy by 30%.

Software Engineer

July 2021 – December 2021

- Developed APIs for an in-house data orchestration tool, improving data pipeline integration for automated problem management workflows.
- Implemented ETL automation scripts in Python and SQL, reducing data processing errors by 20% and improving response times.
- Created technical documentation and best practices for Airflow DAGs, enhancing maintainability of automated workflows.

Tarah Technologies

India

Machine Learning Intern

May 2020 – July 2020

- Created predictive models for image classification using CNN with an F1-score of 0.96, enabling real-time decision-making.
- Built a segmentation model for biomedical imaging using advanced techniques, achieving an F1-score of 0.89.

EDUCATION

Northeastern University

Boston, MA

Master of Science in Artificial Intelligence (GPA: 3.9/4.0)

September 2023 – March 2025

Indian Institute of Technology Tirupati

India

Bachelor of Technology in Electrical Engineering

August 2017-June 2021

TECHNICAL SKILLS

- Data Engineering & Workflow Automation:** ETL Pipelines, Data Modeling, Airflow, Batch & Real-Time Processing
- Big Data & Cloud Technologies:** AWS Glue, Redshift, Snowflake, Databricks, Apache Spark, Hadoop, Kafka
- Programming & Query Optimization:** Python, SQL, Scala, PL/SQL, HiveQL, SparkSQL
- Database & Data Distribution:** PostgreSQL, NoSQL (MongoDB, DynamoDB), Real-Time Data Processing
- Business Intelligence & Reporting:** Power BI, Tableau, Automated Dashboards, Financial Data Analytics
- Certifications:** AWS Machine Learning Engineer - Associate(Ongoing)