

AMEY MEHER

+1 (919) 897-0346 | avmeher@ncsu.edu | [LinkedIn](#) | [GitHub](#)

EDUCATION

North Carolina State University

Master of Science, Computer Science

August 2022 - May 2024

GPA: 4.0/4.0

University of Mumbai

Bachelor of Technology, Computer Engineering

August 2015 - May 2019

GPA: 9.1 / 10.0

PROFESSIONAL EXPERIENCE

IBM

Software Engineer Intern

Yorktown Heights, NY, US

May 2023 - August 2023

- Microservice API development:** Tech stack (Python, Flask, MongoDB, InfluxDB, Kafka, Docker, Kubernetes). Automated synchronization of metadata **reducing manual efforts of 100+ hours per week.**
- Time-Series Data Analysis:** Developed optimized InfluxDB queries on **3 years** of minute-level data for **efficient job scheduling**
- Machine learning:** Developed **clustering model** to analyze **daily 50+** incidents and **automate resolution steps**

Barclays

Data Engineer

Pune, MH, India

July 2019 - July 2022

- Data pipeline development:** Led **4+ ETL projects** for the complete software development lifecycle. Tech stack (Spark, Hadoop, AWS, Hive, Python, Ab Initio, Jenkins, Airflow, Oracle, Teradata, Tableau, PowerBI)
- Optimization:** Optimized pipelines by **99%**, reducing execution time from **8+ hours to 4-5 minutes** on **19M+ records.**
- Data governance:** **GDPR** complied data pipelines; Implemented **audit controls** like completeness, consistency, and accuracy.
- Data modeling:** Designed efficient data models, integrated data from **50+ tables** for creating JSON of over **12000 nested data**
- DevOps:** Developed automation Python framework to test **500+** jobs concurrently, aiding in code migration DevOps pipeline.
- Mentorship:** Mentored new joiners with scripting and **coding standards** for efficient data pipeline creations and modeling
- Certifications:** AWS Cloud Practitioner

SKILLS

Programming Languages: Python, Java, SQL, Scala, Rust, C++, JavaScript, C#

Data Engineering: Spark, Hudi, Hadoop, Hive, Kafka, Airflow, ETL/ELT, Oracle, Power BI, MongoDB, Superset, Postgres

Web Services: Microservice Architecture, GraphQL, REST APIs, Flask, Spring, React, Angular, Docker, Node, Kubernetes, Git

AWS: S3, EC2, Lambda, Glue, DynamoDB, Athena, Redshift, API Gateway, IAM

Azure: Data Factory, Databricks, Datalake storage Gen 2, Synapse Analytics, HDInsight

Courses: DBMS, Automated Learning & Data Analysis, Machine learning on graphs, Data Structures & Algorithms, OOP Methodology

PROJECTS

• Apache Hudi Lakehouse ETL

Created ETL processes for source agnostic data ingestion, CDC workload, and async clustering job for removing small file issue

• Stock Price Prediction using Graph Neural Networks

Developed a stock price prediction system using GCN and LSTM models. Implemented using PyTorch and TensorFlow frameworks

• Android application based on Proximity reminder: Oncue

Awarded first place for amassing 500+ downloads on Playstore for a Hackathon in collaboration with Facebook developers' circle

PUBLICATIONS

- "Error mitigation of Hamiltonian Simulation from an Analog based compiler (SimuQ)" paper accepted at the 2024 QCE IEEE conference
- "Application for Counterfeit Detection in Supply Chain using Blockchain Technology" paper published at the 2021 ICAC3 conference
- "Canteen Automation System" paper published in 2021 IJRASET journal