

MANVITH BELAME YOGESH

Boston, MA | +1 617-318-8506 | belameyogesh.m@northeastern.edu | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

PROFESSIONAL EXPERIENCE

Data Engineer Co-op, Plymouth Rock Assurance, Boston, MA

Aug 2024 – Present

- Streamlined migration of legacy **SAS** code to **Python**, cutting processing time by **90%** and enhancing scalability in **AWS**.
- Leveraged **Python's** concurrency capabilities along with **Polars** library to process **terabytes** of data efficiently, significantly boosting system performance.
- Optimized vendor data processing by transitioning from **SageMaker** to **AWS Glue**, achieving 80% cost reduction and cutting processing time from **300 hours to 100 hours**.
- Implemented **ETL** pipelines using **ECS Fargate** and **Step Functions**, automating data ingestion and enhancing operational efficiency.

Senior Data Analyst, Accenture Solutions, Bangalore, India

Jun 2022 - Jan 2023

- Developed web services for analysis and reporting using **AWS CloudWatch** and used **SQL** triggers for automation of business reporting, reducing manual work by **60%**.
- Crafted and visualized **25** action-driven **dashboards** on **Splunk** to visualize booking & application health metrics.
- Programmed **automation** of validating bookings using **Python scripts**, reducing manual effort by **80%**.
- Conducted **ad-hoc** analyses using **SQL** queries and reported various trends within 6 different customer-facing apps by aggregating datasets containing over **5 million records** per day.
- Integrated **KPI metrics dashboard** on **Splunk**, integrating data from multiple sources to monitor key performance indicators, resulting in a **40% reduction** in data processing time.

Data Analyst, Accenture Solutions, Bangalore, India

Nov 2020 - May 2022

- Crafted a **real-time dashboard** for **6** different apps to monitor peak load performance metrics during the initial 15 mins of ticket booking window, ensuring **zero downtime**.
- Led a team of 5, ensuring smooth launch of booking engine by **coordinating across cross-functional teams**.
- Analysed data from nearly **half a million** booking records using **SQL** querying to triage **high priority** requests.
- Deployed change requests on **AWS**, ensuring updates to critical systems within 30 mins of off-park window.

Data Engineering Intern, Nuotech Steels, Bangalore, India

Jan 2020 - Jun 2020

- Designed **ETL** processes to efficiently extract and transform sales data from multiple sources, leading to a **25% reduction** in data handling time and improved **data accuracy**.
- Demonstrated key trends to stakeholders with **PowerBI** by extracting sales data, showing **20% spike** in sales.

PROJECTS

World Economic Growth Forecast - MLOps, [Airflow, MLFlow, Docker, Flask, Streamlit]

Apr 2024 - Jun 2024

- Extracted data from IMF's API to build predictive models for GDP growth across 196 countries, utilizing **Python**, **SQL**, and **TensorFlow** to analyze macroeconomic indicators.
- Developed **MLOps** workflows using **Docker** and **Airflow** to automate data ingestion and model training.
- Created interactive dashboards and web application using **Streamlit** and **Looker**, presenting insights on GDP trends.

Realtime Data Streaming Using Airflow & Spark, [Airflow, Kafka, Spark]

Sep 2023 - Dec 2023

- Implemented real-time data streaming from **PostgreSQL** to **Kafka**, reducing latency by 25%.
- Orchestrated an end-to-end pipeline with **Airflow** for scheduling, monitoring, and execution of workflows.
- Executed **Spark** jobs to analyze and transform raw data, achieving a 30% reduction in processing time.

TECHNICAL SKILLS

Languages: Python, Java, HTML, CSS, JavaScript, Angular, R, C, C++

Developer Tools: Tableau, PowerBI, Google Analytics, MATLAB, Jira, Git, MS Office Suite, AirFlow, Splunk

Cloud Computing: AWS (S3, Glue, Athena, Sagemaker, RDS, Redshift, EMR, DynamoDB), Microsoft Azure, Snowflake

Database: MySQL, MS SQL, MongoDB (NoSQL), PostgreSQL, Google BigQuery, Neo4j

Libraries: Pandas, NumPy, TensorFlow, Matplotlib, Seaborn, Folium, ipython-sql, Scikit-learn, SciPy, PySpark

EDUCATION

Northeastern University, Boston, MA

Expected May 2025

Master of Science in Data Analytics Engineering,

3.9 / 4.0 GPA

Courses: Database Management, Data Visualization, Data Mining, Machine Learning, Statistical Methods, MLOps.

JSS Science and Technology University, Mysore, India

Jun 2020

Bachelor of Engineering