

AKSHAY RAMESH

Tampa, Florida | akshay.ramesh.8693@gmail.com | 813-406-3905 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

UNIVERSITY OF SOUTH FLORIDA | M.S. Business Analytics and Information Systems

AUG 2022 - MAY 2024

TECHNICAL SKILLS

- BIG DATA:** Hadoop, Spark, Hive, Kafka, Flume, Sqoop.
- CLOUD:** Azure Databricks, Azure Synapse Analytics, Azure Data Factory, Azure ML, Azure DevOps, CI/CD.
- PROGRAMMING:** SQL, Python, Spark SQL, PySpark.
- MACHINE LEARNING:** Supervised and Unsupervised machine learning techniques.
- DATA ANALYTICS:** Statistics, Hypothesis Testing, ANOVA, Time series analysis, OLS Regression.
- DATA VISUALIZATION:** Power BI, Tableau, and OBIEE 11g.
- DATABASES:** SQL Server, MySQL DB, PostgreSQL DB, Pinecone vector database, MongoDB, Cassandra.

WORK EXPERIENCE

STUDENT ASSISTANT | USF IT, TAMPA, USA

AUG 2022 – PRESENT

- Introduced **incremental data loading** using **Azure Synapse Analytics**, reducing data load time by 87.5% with an engineered ELT data pipeline.
- Orchestrated the creation of compelling **Power BI** visualizations on university student academic data utilizing **Azure Synapse Analytics**; provided comprehensive insights into student performance and diversity, empowering informed action within the academic institution.

SENIOR SOFTWARE ENGINEER | ACCENTURE SOLUTIONS PRIVATE LTD | BENGALURU, INDIA

OCT 2015 - MAY 2020

- Spearheaded the utilization of **Agile principles** to convert business requirements into scalable programming logic, ensuring that structured data seamlessly loaded into **Impala target tables**, reducing data processing time by 40%.
- Leveraged **Azure Databricks** and **Azure Data Factory** to implement diverse data loading schedules, including hourly, monthly, and quarterly intervals, enhancing operational flexibility and efficiency.
- Designed and implemented data pipelines for **real-time streaming data ingestion** and processing using **Apache Kafka** and **Apache Spark**, enabling timely insights and decision-making.
- Directed the effective utilization of **machine learning** strategies on **geospatial data**, resulting in a \$5 million profit margin enhancement and optimized resource allocation, **reducing operational costs by 15%** while maintaining service quality standards.
- Developed interactive **Power BI** visualizations on telecom datasets using, empowering clients with insights into telecom signal strength for informed, data-driven decision-making.
- Performed a comprehensive data analysis and predictive modeling to forecast customer **churn rates**, resulting in a 15% reduction in customer attrition and increased revenue retention.
- Collaborated cross-functionally with stakeholders to identify **key performance indicators (KPIs)** and develop dashboards in **Power BI** for real-time monitoring, optimizing decision-making processes and operational efficiency.

PROJECTS

ON PREMISE STRUCTURED DATA MIGRATION TO AZURE CLOUD PLATFORM: ([GitHub](#))

- Created a comprehensive data migration solution leveraging **Azure Data Factory** to migrate on-premise structured data into **Azure cloud**, followed by data transformations using **Azure Databricks** and loading into the gold layer of **Azure Storage Gen 2**.
- Engineered custom views and reports in **Power BI** by leveraging **Azure Synapse Analytics**; stored access credentials securely in **Azure Key Vault** and incorporated a robust security protocol through **Azure Active Directory**.

AZURE DATABRICKS CI/CD PIPELINE IMPLEMENTATION AND ENVIRONMENT SETUP: ([GitHub](#))

- Achieved seamless **DEV-to-PROD** deployment for **Azure Databricks** notebooks via end-to-end **CI/CD pipelines**, integrating **Azure DevOps** and utilizing **YAML** for pipeline creation, reducing deployment time by 30% and ensuring 100% successful deployments.
- Demonstrated environment setup, branch protection, and permissions configuration for seamless dev-to-prod deployment.

TWITTER SENTIMENT ANALYSIS ON THE RUSSIA-UKRAINE CRISIS: ([GitHub](#))

- Optimized processing time by 40% with **Azure Databricks**, handling 3 million tweets; **logistic sentiment model** achieved **0.82 accuracy**, ensuring precise classification of positive and negative sentiments.

AZURE CLOUD BASED MOVIE RECOMMENDATION SYSTEM: ([GitHub](#))

- Orchestrated a comprehensive data pipeline in **Azure Data Factory (ADF)** to handle movie data, incorporating tasks such as metadata extraction, schema validation, file copying, and notebook execution.
- Developed a scalable **Azure Databricks notebook** utilizing **collaborative filtering** with the **ALS (Alternating Least Squares) algorithm**, enhancing movie recommendation accuracy by 25%.
- Integrated **Azure Data Factory (ADF)** with **Azure Logic Apps**, automating email notifications for pipeline events, ensuring timely responses to successful recommendations or errors, resulting in a 50% reduction in response time to issues and improved operational efficiency.

CERTIFICATIONS

- Exam DP-203: Data Engineering on Microsoft Azure **May 2024 (EXPECTED)**
- Databricks Certified Associate Developer for Apache Spark **May 2024 (EXPECTED)**