

EDUCATION

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

Aug 2022 - May 2024

WORK EXPERIENCE

STUDENT ASSISTANT | UNIVERSITY OF SOUTH FLORIDA | TAMPA, USA

Aug 2023 – May 2024

- Designed a robust **data model in Snowflake**, tailored for handling **Type 2 slowly changing dimensions (SCD 2)**, resulting in a 20% improvement in historical data accuracy and analysis efficiency.
- Initiated and optimized **batch data loading** processes in **Snowflake** from diverse sources, cutting loading time by 30% and improving data availability and reliability for critical analytical insights.
- Integrated **Snowflake with Power BI** to create interactive reports, boosting user engagement by 25%. Established **row-level security in Power BI**, ensuring compliance with access control policies.

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

Oct 2015 - May 2020

- Spearheaded the utilization of **Agile Methodology** to convert business requirements into scalable programming logic, ensuring that semi-structured data seamlessly loaded into target tables, reducing data processing time by 40%.
- Leveraged **Amazon Redshift, AWS Glue, and AWS Lambda** to implement diverse data loading schedules, including hourly, monthly, and quarterly intervals, enhancing operational efficiency by 100%.
- Architected and managed **near real-time** data ingestion and processing pipelines using **AWS Kinesis and Databricks**, boosting data availability by 35% and empowering timely and informed decision-making processes.
- 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.
- Crafted detailed **Power BI** reports on telecom signal strength; empowered clients with critical insights, resulting in a 30% increase in data-driven decision-making.
- Performed a comprehensive data analysis and **predictive modeling** to forecast **customer Churn Rates** using **AWS SageMaker**, resulting in a 15% reduction in customer attrition and increased revenue retention.
- Pioneered the creation of real-time monitoring dashboards in **Power BI**, collaborating with stakeholders to identify critical **KPIs**, optimizing operational efficiency, and boosting decision-making speed by 20% within the first quarter.
- Leveraged **Apache Airflow** to fully **automate data workflows**, resulting in a 100% improvement in overall process efficiency.
- Architected an end-to-end data migration framework using **Azure Data Factory** and **Azure Databricks**, transitioning on-premise data to Azure cloud; enhanced data processing speed by 50%.
- Engineered custom views and reports in **Power BI** by leveraging **Azure Data Factory**, stored access credentials securely in **Azure Key Vault** and incorporated a robust security protocol through **Azure Active Directory**.
- 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.

PROJECTS

AMAZON SALES DATA ANALYSIS FOR MOBILE PHONES USING AWS AND SNOWFLAKE: ([GitHub](#))

- Engineered a data pipeline to process Amazon sales data for mobile phones over a year, ingesting **CSV, JSON, and Parquet formats via AWS Lambda into S3**.
- Incorporated **Snowflake** to transform and analyze the data and developed 3 visualizations using **Power BI**.
- Automated the entire workflow with **Apache Airflow** on an **EC2 instance**, ensuring efficient data processing with 100% accuracy.

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** utilizing **Collaborative Filtering with the ALS (Alternating Least Squares) algorithm**, enhancing movie recommendation accuracy by 25%.

TWITTER SENTIMENT ANALYSIS: ([GitHub](#))

- Performed comprehensive **exploratory data analysis** on a dataset consisting of 3 million tweets.
- Improved processing efficiency by 40% using **Azure Databricks**, enabling swift handling of large-scale data.
- Implemented a **logistic sentiment classifier model** achieving an accuracy of 0.82.

TECHNICAL SKILLS

- BIG DATA:** Hadoop, Spark, Data Warehouse, Avro, Unix, Data Lake, Delta Lake, Data Architecture.
- CLOUD:** Snowflake, Amazon S3, Visual Studio Code, Exploratory Data Analyses, Spark Structured Streaming, ADLS, ETL/ELT, DynamoDB, Docker, Athena, OLTP, OLAP, Azure Data Lake.
- PROGRAMMING:** Advanced SQL, Python, PySpark, T-SQL.
- MACHINE LEARNING:** Supervised, Unsupervised Machine Learning techniques, Statistics, REST API.
- DATA VISUALIZATION:** Power BI, Tableau, JIRA reporting.
- DATABASES:** MySQL DB, NoSQL.