ANIRUDHRAJ VEERABOINA

(573) 271-1873 | anirudhrajv27@gmail.com | St louis, Missouri 63146

PROFESSIONAL SUMMARY

As a Data Engineer, I worked with large-scale data sets and cloud technologies to deliver data solutions for various clients. I have successfully designed, developed, and deployed data pipelines, ETL processes, and data warehouses using tools and frameworks such as Azure DevOps Server, Snowflake, SQL Server, and Python. I have also contributed to data quality, data governance, and data security initiatives, ensuring compliance and accuracy across data sources and platforms.

EXPERIENCE

Ernst & Young(EY)   October 2024 – Present
Senior Data Analyst,
St Louis, MO

- Leading end-to-end data migration from PeopleSoft to Oracle HCM, ensuring data integrity and compliance with business rules.
- Conducting data cleansing, transformation, and validation to ensure accuracy and consistency before migration.
- Coordinating with data stewards, client leadership, and cross-functional teams to align on data requirements and resolve discrepancies.
- Developing and maintaining ETL workflows to handle large datasets using Snowflake, Oracle, and Collibra.
- Generating data quality reports and dashboards in Excel to track migration progress and highlight anomalies.
- Delivering weekly status updates and data migration insights to client leadership and stakeholders.
- Ensuring compliance with data governance and security standards throughout the migration process.
- Executed data integration and cleansing processes in Databricks, ensuring the consistency, quality, and availability of data across various stages of the data pipeline.
- Proficient in SQL optimization for complex queries, enabling faster data retrieval, improved ETL processing speeds, and efficient data warehousing.
- Advanced scripting skills in Python and SQL, applied extensively in data transformation, pipeline automation, and performance optimization.

Biztegy   September 2023 – September 2024
Data Engineer,
Arlington Heights, IL

- Developed and implemented data pipelines to ingest, store and process large datasets using Apache Spark and Hive.
- Created ETL jobs to migrate data from various sources such as Oracle, MySQL, MS SQL Server into HDFS.
- Utilized Python scripts to automate the loading of data from multiple sources into a centralized system.
- Streamlined batch processing times by 40% by implementing parallel data processing in Azure Data Factory.
- Reduced manual data inputs by utilizing Azure cognitive services for form scanning and data input, achieving 30% labor cost efficiency.

- Designed and maintained ETL/ELT pipelines using Azure Data Factory, implementing data transformation and integration solutions that optimize data flow across multiple systems.
- Utilized Delta Lake and Databricks for high-efficiency, large-scale data processing, leveraging Delta Lake's ACID compliance and data versioning to support reliable data lakes.
- Refactored SQL scripts for ETL processes, achieving 20% performance improvement.

Iflex   September 2022 - December 2022

Data Engineer Intern,

Fairlawn, OH

- Researched emerging technologies related to big data processing, analytics and reporting tools.
- Deployed Apache Spark clusters on AWS EC2 instances for distributed computing tasks.
- Developed machine learning algorithms using Python libraries such as Scikit-learn, Numpy and Pandas for predictive analytics purposes.
- Established efficient ETL jobs scheduling mechanisms via crontrol-m jobs.
- Developed backend service components, resulting in a 25% decrease in latency for user requests.
- Built effective code deployment pipelines, reducing deployment errors by 35%.

Atos Syntel   December 2019 - December 2021

Data Engineer,

Pune, India

- Developed custom applications using Java, Python, Scala programming language for data transformation, cleansing and validation.
- Created ETL jobs to migrate data from various sources such as Oracle, MySQL, MS SQL Server into HDFS.
- Utilized Python scripts to automate the loading of data from multiple sources into a centralized system.
- Deployed machine learning algorithms to identify patterns in large scale datasets for predictive analytics purposes.
- Developed data models for applications, metadata tables, views, and related database structures.

EDUCATION

Master of Science (M.S.) in Applied Computer Science

Southeast Missouri State University, Cape Girardeau, MO, US, May 2023

GPA: 3.6

BTech in Computer Science Engineering

CMR Technical Campus, Hyderabad, India, May 2019

GPA: 7.8

CERTIFICATIONS

- Microsoft Azure Fundamentals(AZ-900)
- Microsoft Azure Data Engineer Associate(DP-203)
- SnowPro

SKILLS

- **Cloud Platforms:** Azure Data Factory, Azure Databricks, Azure DevOps, Azure SQL Database, Azure Blob Storage, Azure Data Lake
- **Programming & Scripting:** Python (Pandas, PySpark), Scala, SQL, PowerShell, Bash, C#, Snowflake, Oracle
- **Data Processing & ETL:** Apache Spark, Hadoop, Informatica, SSIS, Talend, DataStage
- **DevOps & CI/CD:** Docker, Kubernetes, Jenkins, Git, Maven, Gradle