JAYANTH SUDDAMALLA
Data Engineer

📞 +1 (417) 633-2302

✉️ jayanthsuddamalla48@gmail.com

📍 Dallas, TX

in LinkedIn

Skills

- **Programming Languages:** Python, SQL, PySpark, Scala
- **Big Data Technologies:** Apache Spark, Hadoop, AWS EMR, AWS Kinesis, AWS EC2, AWS S3, Google Cloud Platform
- **BI Tools:** Tableau, Power BI
- **Data Warehousing:** Amazon Redshift, Google Big Query, Azure Data Factory, Azure Databricks, Azure Synapse, Snowflake
- **Database Management:** MySQL, PostgreSQL, HBase, Cosmos DB, SQL Server, MongoDB, Cassandra
- **Data Skills:** Data Visualization, Data Modeling, Data Normalization, Data Warehousing, Data Mining, Data Analysis, Data Quality, Data Integration, Data Transformation, Data Cleansing
- **Soft Skills:** Critical Thinking, Written & Communication Skills, Presentation Skills, Problem-Solving, Team Workflows, Leadership Skills, Project Management, Fast-Paced Work Skills, Integration Skills, Risk Management, Marketing Skills
- **ETL Tools:** Apache Airflow, Talend
- **Tools and Software:** GitHub, MS Excel, MS PowerPoint, MS Word, SharePoint, JIRA, Confluence
- **Cloud Platforms:** AWS, AWS GLUE, AWS Redshift, AWS Lambda, Azure, Google Cloud
- **Streaming Analytics:** Apache Kafka, Apache Flink
- **Containerization and Orchestration:** Docker, Kubernetes, Terraform
- **Version Control Systems:** Git, SVN
- **Data Governance and Compliance:** GDPR, HIPAA, CCPA
- **Agile Methodologies:** Scrum, Kanban
- **Testing Frameworks:** Unit Testing, Integration Testing

Education

Master Of Science in Computer Science and Data Science
Missouri State University
Aug 2021 – Dec 2022

Bachelor Of Technology in Computer Science and Engineering
Lovely Professional University, India
Aug 2017 – May 2021

Summary

Experienced data engineer with 4 years of expertise in data warehousing, ETL processes, and cloud platforms. Skilled in Snowflake schema optimization, Azure Data Lakes, and Databricks. Proficient in PySpark for complex data manipulation and SQL window functions for pipeline efficiency. Hands-on experience with AWS Kinesis and EMR for real-time streaming and large-scale processing. Strong in managing Redshift clusters and AWS Lambda integration. Dedicated to Agile and SDLC methodologies, adept at creating real-time Tableau dashboards. Collaborative team player, proactive in staying updated with industry trends.

Experience

Data Engineer, Doublene, USA · Jan 2023 – Present

- Implemented real-time data streaming solutions with AWS Kinesis, facilitating instant data insights.
- Utilized Amazon EMR for large-scale data processing, implementing Spark and Hive jobs for data transformation.
- Integrated AWS Lambda with other AWS services to trigger data processing tasks in response to real-time events.
- Managed and administered Amazon Redshift clusters, fine-tuning performance and ensuring data integrity.
- Utilized Agile and SDLC methodology to deliver end to end continuous integration/ continuous delivery (CI/CD) pipeline by integration of tools like Jenkins and AWS for VM provisioning
- Created real-time dashboards in Tableau to visualize and monitor key metrics and A/B test processing using both external and internal data.
- Developed mappings using Source Qualifier, Expression, Filter, Lookup, Update Strategy, Sorter, Joiner, Normalizer, and Router transformations.
- Optimized and maintained databases (mention specific databases, e.g., PostgreSQL, MySQL) to enhance overall data quality, performance, and reliability.
- Conducted data quality checks to ensure accuracy, consistency, and completeness.
- Stayed updated with industry trends, new tools, and techniques to enhance data analysis skills and stay at the forefront of the field.
- Provided strategic guidance and recommendations based on data analysis to drive business growth, improve operational efficiency, and optimize decision-making.
- Developed and maintained documentation for data models, processes, and workflows, facilitating knowledge transfer within the team.

Data Engineer, Tec Artists, India · Jan 2019 - Jul 2021

- Designed and implemented data warehousing solutions on Snowflake, optimizing schema design for performance and scalability.
- Worked on domain such as Procure to Pay, Account Payable used Tcodes Such BSAK, LFA1, EKKO& EKPO.
- Managed Azure Data Lakes (ADLS) and Data Lake Analytics to integrate with other Azure Services.
- Utilized Azure Databricks for complex manipulation and transformation of data in PySpark.
- Used complex SQL window functions queries to analyze and improve data accuracy of existing data pipelines by 20%
- Optimized SQL queries and data models to improve database performance and reduce query execution time, enhancing overall system efficiency, responsible for 20% of job responsibilities.
- Created data dictionary, data mapping for ETL and application support, DFD, ERD, mapping documents, metadata, DDL, and DML as required.
- Created clear and visually appealing data visualizations, such as charts and graphs, to effectively communicate findings to stakeholders.
- Collaborated with teams to collect and clean data from various sources, ensuring data accuracy and integrity.
- Collaborated with data scientists to implement machine learning models into production, enabling data-driven decision-making across the organization.
- Used the best algorithm and parameters to implement, adjust, and test the model on AWS EC2.jan.