

Durgamalleswarao Ponnuru

Data Engineer

+1 (313) 478-2685 | dponnuru6@gmail.com | [LinkedIn](#)

Professional Summary:

Accomplished Data Engineer with over 5 years of diverse experience in designing, developing, and optimizing data-driven solutions across multiple industries, including technology, finance, and healthcare. Adept at building robust and scalable data pipelines, integrating complex data sources, and implementing advanced analytics to support business intelligence and data science initiatives. Proficient in leveraging cloud platforms like AWS and Azure, big data technologies such as Apache Spark, and modern ETL tools to streamline data workflows and enhance decision-making processes. Demonstrated expertise in data modeling, real-time data processing, and data warehousing, ensuring data accuracy, quality, and security across the enterprise.

Strong background in programming with Python, R, and SQL, complemented by hands-on experience with data visualization tools like Tableau and Power BI to deliver actionable insights. Skilled in working with relational and NoSQL databases, including MySQL, PostgreSQL, MongoDB, and Snowflake. Committed to continuous learning and staying updated with emerging technologies, contributing to innovative solutions that drive business growth.

Technical Skills:

- **Programming Languages:** Python, R, SQL, Bash, Shell Scripting
- **Big Data Technologies:** Apache Spark, Hadoop, Kafka, Flink, HDFS, HBase
- **Databases:** SQL Server, MySQL, PostgreSQL, MongoDB, Redshift, BigQuery
- **Data Visualization Tools:** Tableau, Power BI
- **Data Processing & Storage:** Azure Data Factory, Azure Synapse Analytics, Databricks, Snowflake, AWS Glue
- **Cloud Platforms:** AWS (S3, Lambda, VPC, Redshift), Azure (Data Lake, Data Factory, Synapse, DevOps)
- **ETL Tools:** SSIS, Informatica, Talend, Apache Airflow
- **Orchestration & Automation:** Airflow, Jenkins, Terraform, Ansible
- **Version Control:** GitHub, GitLab, Azure DevOps
- **Monitoring Tools:** Prometheus, Grafana, ELK Stack (Elasticsearch, Logstash, Kibana)
- **Project Management:** Agile, Scrum, Jira

Education:

Master of Science in Data Science

University of Michigan, Dearborn

Aug 2022 – Apr 2024

Bachelor's degree in Computer Science

Velagapudi Ramakrishna Siddhartha Engineering College, India

Aug 2016 - May 2020

Work Experience:

Data Engineer

April 2024 – Present

Molina Healthcare, Long Beach, CA

Roles & Responsibilities:

- Designing and implemented end-to-end data pipelines using AWS Glue and Apache Airflow, ensuring efficient data ingestion and transformation.
- Processing large-scale data using Apache Spark and Databricks, reducing data processing time by 50%.
- Developing data models in Snowflake to support complex analytics and reporting requirements.
- Automating ETL workflows using Python, improving data reliability and reducing manual efforts.
- Integrating data from various sources including APIs, cloud storage (AWS S3), and on-premise databases.
- Building interactive dashboards using Tableau and Power BI to deliver actionable business insights.
- Optimizing SQL queries for faster data retrieval and improved performance.
- Ensuring data quality through robust validation frameworks and error-handling mechanisms.
- Collaborating with data scientists to deploy machine learning models in production environments.
- Implementing security protocols to safeguard sensitive data in compliance with industry standards.
- Utilizing GitHub and Jenkins for version control and CI/CD pipeline management.
- Conducting data profiling, cleansing, and transformation to support analytics initiatives.
- Mentoring junior data engineers and provided technical guidance on best practices.
- Documented data architecture, workflows, and processes for knowledge sharing.
- Participated in Agile ceremonies, contributing to sprint planning and retrospectives.
- Utilized Apache Airflow to schedule and manage complex workflows effectively.
- Supported ad-hoc data requests from stakeholders, providing timely and accurate insights.

Data Engineer

June 2020 – August 2022

Tata Consultancy Services (TCS), India

Roles & Responsibilities:

- Designed and managed scalable ETL processes using SSIS and Informatica for enterprise-wide data integration.
- Utilized Azure Data Factory and Synapse Analytics for large-scale data processing and storage.
- Developed and optimized data pipelines in Databricks, improving processing efficiency.
- Created complex SQL queries for data extraction, transformation, and analysis.
- Automated data workflows with Python, reducing manual tasks and improving accuracy.
- Integrated structured and unstructured data from diverse sources to support business analytics.
- Built dashboards in Power BI to visualize key performance indicators and business metrics.
- Ensured data quality through validation checks and error-handling procedures.
- Worked closely with business stakeholders to understand data requirements and deliver solutions.
- Implemented data governance practices to maintain data integrity and compliance.
- Optimized performance of Spark jobs and SQL queries to handle large datasets efficiently.
- Managed version control using Git and facilitated code reviews for quality assurance.
- Developed REST APIs for data access and integration with third-party applications.
- Conducted data migration projects from on-premise to cloud platforms.
- Participated in Agile development processes, contributing to sprint planning and reviews.
- Participated in team meetings to discuss project progress and deliverables.

Data Engineer Intern

May 2018 – May 2020

Indian Servers Pvt Ltd, India

Project: Data Warehouse Development and Optimization

- Assisted in the development of ETL pipelines using Talend and Apache Nifi for data integration.
 - Wrote SQL queries to extract, transform, and load data into relational databases.
 - Supported data analysis activities using Python (Pandas, NumPy) to generate insights.
 - Created data visualizations with Tableau to support business decision-making.
 - Performed data cleaning and validation to ensure data accuracy and consistency.
 - Assisted in automating data workflows, reducing manual intervention.
 - Participated in database design and optimization to improve query performance.
 - Worked with APIs to ingest external data into internal systems.
 - Contributed to the development of data models for business intelligence applications.
 - Documented data processes, workflows, and technical specifications.
 - Supported ad-hoc data analysis requests from business stakeholders.
 - Learned and applied Agile project management practices using Jira.
 - Utilized Git for version control and code collaboration within the team.
 - Assisted in maintaining data security and compliance with organizational standards.
 - Collaborated with cross-functional teams to support data-driven projects.
-

Projects & Achievements

- **Cloud Data Migration:** Successfully migrated 10 TB of data from on-premise to AWS, improving data accessibility and performance.
- **Real-Time Data Processing:** Implemented a real-time data streaming solution using Apache Kafka and Spark, reducing data latency from hours to seconds.
- **Machine Learning Integration:** Collaborated with data science teams to deploy predictive models into production, enhancing business decision-making capabilities.