Sasidhar Guthi

6203 Love Dr, 3132, Irving, TX, 75039

📞 +1-571-899-1692 ✉ guthisasidhar@gmail.com 🔗 https://www.linkedin.com/in/sasidhar-guthi-777a25104/

summary

Data Engineer with extensive experience in optimizing data processing and analytics solutions. Notably increased data processing efficiency and reduced query times while developing secure, large-scale data pipelines at Vizio TV.
- Enhanced dashboards and automated workflows to support key business decisions with timely insights.
- Improved data accuracy and cut reporting time substantially through robust ETL and visualization platforms. Looks to leverage solid technical expertise and collaborative acumen to ensure secure, efficient data access and insightful reporting tailored to organizational needs.

TECHNICAL SKILLS

- **Data Engineering:** AWS, GCP, Pyspark, Big Data, SQL Server, MongoDB, Snowflake, PL/SQL, REST API, ETL Pipelines, Data Modeling, Data Warehousing, AWS Redshift, AWS S3, AWS Glue, EMR, Kinesis, FireHose, AWS Lambda, IAM Roles, Non-relational Databases
- **Programming Languages:** Python, SQL, Java, Shell, VB Scripting, Flask, Streamlit, NodeJS, React
- **Machine Learning:** Scikit-learn, TensorFlow, Pytorch, Mlib, NLTK, Spacy, Text Blob
- **Data Visualization:** Power BI, Tableau, Excel, Looker, Quicksight
- **Version Control & Collaboration:** GIT

WORK EXPERIENCE

Vizio TV — Apr 2024 - Present
Data Engineer — *Dallas, TX*

- Designed, developed, and maintained scalable data pipelines using AWS Glue and EMR clusters to support Vizio's data analytics and business intelligence operations, resulting in a 25% increase in data processing efficiency. Utilized Google BigQuery and AWS S3 to manage and analyze large datasets efficiently, reducing query execution time by 40% and optimizing cost-efficiency by 15%.
- Built automated ETL workflows in PySpark within AWS Glue and EMR, handling complex joins, aggregations, and filtering to prepare data for downstream analytics and reporting.
- Implemented robust ETL processes using AWS Lambda, S3, Glue, Snowflake, focusing on data integrity and validation checks. Developed continuous integration/continuous deployment (CI/CD) pipelines, improving data reliability and reducing errors by 25%, achieving reduced pipeline cost and data latency.
- Developed and maintained Tableau and Looker Dashboards to support key business decisions, achieving 80% automation in daily, weekly, and monthly reporting workflows. Integrated Looker queries into automated Slack reporting, significantly reducing manual labor.
- Collaborated with data scientists, analysts, and software engineers to understand and fulfill data requirements, leading to a 20% improvement in stakeholder satisfaction by providing reliable data solutions.
- Designed and developed an analytics event validation website using Python (Flask) for backend processing and React.js with HTML for a responsive interface, streamlining the validation process for developers and QA teams by enabling real-time identification of bottlenecks and data discrepancies, reducing manual intervention by 60%, and automating workflows to ensure data consistency across releases, improving time-to-market efficiency by 25%.

Intelligenie — Jun 2023 - Mar 2024
Data Engineer — *Charlotte, NC*

- Enhanced sales forecasting models in Python for advanced data wrangling and analysis, improving predictive accuracy by 18%. Leveraged statistical and machine learning techniques for deeper business insights.
- Optimized Tableau dashboard performance through data aggregation and extract optimization, increasing responsiveness and scalability with large datasets. Developed standardized Tableau templates, cutting report creation time by 30%.
- Reduced data processing time by 30% by designing automated ETL pipelines with AWS Glue, Step Functions, and Python. Focused on efficient data integration and loading, ensuring timely availability and boosting platform performance.
- Facilitated secure data exchange by integrating Python's Requests library with RESTful APIs, providing stakeholders with real-time access to critical datasets and enhancing system interoperability.

Infosys — Sep 2019 - Aug 2021
Data Engineer — *Hyderabad,India*

- Reduced resource consumption by 15% in SAP HANA environments through database performance tuning and optimization, demonstrating advanced database management prowess and cost efficiency.
- Elevated SAP system functionality by providing expert technical support and resolving performance bottlenecks, contributing to seamless SAP-related functions and enhancing user satisfaction.
- Improved database performance by 20% through regular SQL and Shell scripting, underscoring robust database administration and maintenance capabilities, directly impacting production support efficiency.

EDUCATION

University of Maryland, Baltimore County — **May 2023**

Master of Professional Studies, Data Science
  • **Achievements:** Major GPA: 3.8

ACADEMIC PROJECTS

University of Maryland, Baltimore County

Histopathology Cancer Detection
  • Transformed histopathologic cancer detection by leveraging CNNs and autoencoders, enhancing image clarity and, improving diagnostic accuracy.

University of Maryland, Baltimore County

Speech Emotion Recognition
  • Elevated Speech Emotion Recognition accuracy by applying machine learning algorithms (MLPC, KNN, Decision Tree, Logistic Regression, CNN), surpassing standard models and significantly enhancing emotion classification precision.