SHAIK RAHAMA THULLA | Data Engineer
Saint Louis | +1 (929)493-9704 |shaik.rahamathulla@aogjob.com | LinkedIn

Professional Summary

- Dynamic and results-oriented Data Engineer with a 4+ years proven track record of driving innovation and efficiency in data processing and analytics. With extensive experience spanning both international and domestic markets, I have spearheaded transformative initiatives at PTC in the USA and Sigma Info Solution in India, leveraging cutting-edge technologies and agile methodologies to deliver impactful results.
- At PTC, I led the development and optimization of scalable ETL pipelines in a cloud environment, managing diverse data streams and achieving a remarkable 30% reduction in data processing time. My expertise in designing and optimizing relational database management systems resulted in a 25% improvement in query performance and database efficiency through the implementation of over 200 stored procedures, triggers, and indexes.
- Moreover, my strategic leadership in automating Continuous Integration/Continuous Deployment pipelines using Git for version control streamlined deployment processes by 50%, driving efficiency and reliability in data-driven applications. Additionally, I played a pivotal role in driving technical feasibility studies for innovative data solutions, contributing to a 40% increase in project approval rates based on proof-of-concept outcomes.
- During my tenure at Sigma Info Solution, I led initiatives to optimize SAP HANA Cloud instances, implementing sophisticated data modeling and query techniques to enhance query efficiency by 25%. I orchestrated comprehensive integrations of SAP HANA Cloud with various data sources, reducing data latency by 20% and enabling seamless data fluidity across the organization.
- Furthermore, my commitment to data security and compliance was demonstrated through the successful overhaul of security protocols for SAP HANA Cloud, achieving 100% compliance with GDPR and internal governance standards. By directing cross-departmental teams and aligning operations with corporate goals, I accelerated project delivery by 30% and fostered collaboration between IT, business analytics, and development teams.
- My proficiency in implementing comprehensive AWS solutions, including Glue, S3, Lambda, and Kinesis, resulted in a 60% increase in data throughput and supported predictive modeling initiatives that improved data-driven decision-making accuracy by 35%.

Education

Master of Science in Cybersecurity | southeast Missouri State University, Cape Girardeau Missouri, USA   **Dec 2023**
B-Tech in Computer Science |KLEF, University, India   **May 2020**

Skills

Languages: Python, R, SQL
Python Libraries/Packages: NumPy, Pandas. Matplotlib, SQLAlchemy, Beautiful Soup, Py Query
Tools: Tableau, Power BI, Advanced Excel (Pivot Tables,VLOOKUP)
Statistical Analysis Skills: A/B Testing, Time Series Analysis
Machine Learning and Analytical Tools: Supervised Learning (Linear Regression. Logistic Regression, Decision Tree. Random Forest. SVM, Classification. Unsupervised Learning (Clustering, KNN, Factor Analysis, PCA), Natural Language Processing, Tableau
Cloud Computing: AWS, Azure
AWS Services: Amazon EC2. Amazon S3. Amazon Simple DB. Amazon MQ. Amazon ECS, Amazon Lambdas. Amazon Sagemaker
Databases/Servers: MySQL, PostgreSQL, MongoDB, Teradata, Apache Web Server, Tomcat
ETL: Informatica, Data Stage. SSIS.
Web Services/Protocols: TCP/IP, UDP, FTP, HTTP/HTTPS, SOAP, Rest, Restful
Build and Cl tools: Docker, Kubernetes, Maven, Gradle, Jenkins
SDLC/Testing Methodologies: Agile/ Scrum, Waterfall, TDD
Certification: IOT - Internet of things

Work Experience

PTC, USA | Data Engineer   **May 2023 - Current**
- Spearheaded the development and maintenance of scalable ETL pipelines in a cloud environment, managing over 50 data streams and reducing data processing time by 30% through advanced techniques and agile methodologies.
- Designed and optimized relational database management systems, implementing over 200 stored procedures, triggers, and indexes, which enhanced query performance and database efficiency by 25%.
- Automated Continuous Integration/Continuous Deployment pipelines for data-driven applications, enhancing deployment efficiency by 50%. Utilized Git for version control across 100+ deployments in both testing and production environments.
- Led over 10 technical feasibility studies for innovative data solutions using AWS, REST APIs, and Python, significantly increasing project approval rates by 40% based on proof-of-concept outcomes.
- Drove the data engineering team's candidate evaluation and recruitment process, boosting team capabilities by 30% through the strategic hiring and training of professionals skilled in big data technologies like Spark and AWS.

- Implemented comprehensive AWS solutions involving Glue, S3, Lambda, and Kinesis to enhance data handling and analytics, boosting data throughput by 60% and supporting predictive modeling that improved data-driven decision-making accuracy by 35%.
- Developed and maintained an advanced monitoring system for ETL processes, employing custom Python scripts and AWS CloudWatch, which improved system reliability and performance visibility by over 40%. This initiative also reduced incident response times by 50%.
- Enhanced data security measures by implementing encryption, access controls, and audit trails using AWS IAM and KMS. These improvements have ensured compliance with industry standards and regulations, leading to a zero-breach record during the tenure.
- Collaborated with cross-functional teams to integrate machine learning models into the data pipeline, using Python and TensorFlow. This integration allowed for real-time data analytics and insights, increasing operational efficiencies and enabling more informed business decisions by up to 45%.

Sigma Info Solution, India | Data Engineer   **Jun 2019 - Jun 2022**
- Led the optimization of 15+ SAP HANA Cloud instances by implementing sophisticated data modeling and advanced query techniques, resulting in a 25% increase in query efficiency and enhancing the real-time analytics capabilities essential for critical business decisions.
- Orchestrated a comprehensive integration of SAP HANA Cloud with various data sources, including on-premise and cloud-based platforms, through strategic use of SAP Data Services. This initiative enhanced data fluidity across the organization and reduced data latency by 20%.
- Overhauled security protocols for SAP HANA Cloud, focusing on strengthening data encryption and access controls, aligning with GDPR and internal governance standards. Successfully achieved 100% compliance, significantly minimizing organizational risk.
- Directed cross-departmental teams to align SAP HANA Cloud operations with broader corporate goals, accelerating project delivery by 30% through enhanced collaboration between IT, business analytics, and development teams.
- Advanced the real-time analytics framework by improving the efficiency of complex data queries and analytics processes by 25%, thereby enabling more informed and timely business decision-making.
- ETL Process Improvement Using SAP Data Services: Utilized SAP Data Services to refine ETL processes, improving data synchronization and operational responsiveness by 20%, which supported enhanced data handling and analysis capabilities.
- Developed and implemented a data governance framework for SAP HANA Cloud that standardized data management practices across the organization. This framework ensured data quality and consistency, leading to a 15% improvement in data reliability and support for strategic business initiatives.
- Spearheaded the adoption of predictive analytics techniques within SAP HANA, integrating Python and R for data modeling. This initiative improved forecasting accuracy by 30%, enabling proactive business strategies and operational adjustments.
- Conducted comprehensive training sessions for over 100 team members on the use of SAP HANA Cloud and SAP Data Services. These educational efforts resulted in a 40% increase in team productivity and streamlined data operations, enhancing overall team performance and data management capabilities.