Ameerraja

Data Engineer | Data Platform Engineer

+91-8667692789 | ameerrajaeee@gmail.com | linkedin.com/in/ameerraja

EXPERIENCE

Software Engineer - Data
Jun 2023 – Present

Guidewire Software Solutions — *Bangalore, India*

- Built a comprehensive data analytics platform for the insurance domain, reducing time-to-insight by 65% through implementation of a medallion architecture and efficient data pipelines
- Engineered robust ETL pipeline frameworks with Spark and Apache Iceberg, creating a seamless data lake management system that substantially improved processing efficiency
- Established comprehensive data governance protocols and quality frameworks that dramatically reduced data incidents while ensuring strict regulatory compliance
- Modernized ETL workflows through Apache Airflow and dbt implementation, enhancing both processing speed and data quality validation capabilities
- Optimized PostgreSQL database performance through strategic query refactoring, advanced indexing methodologies, and SCD-2 techniques, reducing query execution time by 55%
- Developed cost-effective data pipelines in AWS, reducing cloud infrastructure expenses by 30% while maintaining performance benchmarks
- Integrated Kafka and Spark Structured Streaming to create real-time analytics capabilities, enabling business stakeholders to access insights 8x faster than previous batch processes
- Streamlined infrastructure management by implementing Terraform and CI/CD pipelines, transforming environment setup from a multi-day process to minutes while enhancing release stability
- Incorporated LLM technologies to enhance data classification and anomaly detection, improving accuracy of risk assessment models by 42%
  Leveraged Technologies - Python, SQL, Spark, Databricks, Airflow, Iceberg, dbt, PostgresQL, ETL, Data Warehousing, Kafka, AWS (Glue, Athena, Redshift, EMR), Terraform, Docker, Kubernetes, CI/CD, Data Governance, Agile Methodology

Data Engineer
Jan. 2021 – Jun 2023

Tata Consultancy Services — *Chennai, India*

- Architected and implemented ETL pipelines using Azure Data Factory for a leading hotel chain client, resulting in 70% reduction in manual data handling and streamlined integration across diverse systems
- Built Python-based ETL workflows that automated batch processing for over 200 applications, enabling real-time analytics via Power BI and significantly decreasing SLA violations
- Engineered data transformation solutions using Apache Spark in Azure Databricks, delivering consistent, high-quality datasets that empowered machine learning teams to develop more accurate predictive models

TECHNICAL SKILLS

Languages: Python, SQL, GoLang, Shell Scripting
Big Data Technologies: Apache Spark & Spark Structured Streaming, Iceberg, Kafka, Airflow, Flink, dbt, FastAPI
Cloud Platforms: AWS (Glue, Athena, Redshift, S3, Lambda, EMR), Azure (Data Factory, Databricks)
Databases: PostgreSQL, DynamoDB, Redshift, BigQuery, MongoDB, Glue Data Catalog, DuckDB
Data Modeling: Data warehouse, Dimensional modeling, Slowly Changing Dimension, Medallion Architecture
CI/CD: Git, Docker, Kubernetes, Terraform, Github Actions, TeamCity
Additional Expertise: REST APIs, LLMs, AI Agents, Prompt Engg, DataOps, SDLC, Power BI
Tools: Claude Code, Github Copilot

CERTIFICATION

- Microsoft - Azure Data Engineer (DP-203), Power BI Analyst (PL-300), Azure fundamentals (AZ-900)
- AWS Certified Cloud Practitioner
- Astronomer Certified - Apache Airflow Fundamentals, DAG Authoring for Apache Airflow

EDUCATION

Kamaraj College of Engineering and Technology
Madurai, India

Bachelors in Engineering, Major in Electrical — *Jun. 2016 – Mar 2020*