# MANIKANTA POCHAMALLA

manikantapochamalla85@gmail.com | +1 984 329 1618 | UNITES STATES

## PROFESSIONAL SUMMARY

Data Engineer with 4+ years of experience designing and implementing scalable data pipelines and ETL processes in cloud environments. Expertise in data modernization, optimizing processing workflows, and ensuring high data quality and reliability. Proficient in Python, SQL, and big data technologies including PySpark, Databricks, and Apache Spark, contributing to efficient data integration and analytics. Experienced with AWS services like S3, EMR, Glue, Lambda, and Step Functions for orchestration, alongside Terraform for infrastructure as code. Skilled in deploying containerized applications using Amazon ECS and Kubernetes EKS, and managing streaming data with Kafka.

## WORK EXPERIENCE

**Gatik**, Data Engineer **Feb 2023 - Present**

- Developed scalable data ingestion pipelines for vehicle sensor data and telematics using Python, AWS S3, and Kafka, supporting real-time analytics for autonomous operations.
- Implemented data models in Databricks and SQL for analyzing autonomous truck performance and logistics, enabling critical insights for operational improvements.
- Built robust ETL workflows using Airflow to process operational logs and vehicle data, ensuring consistent data delivery to downstream analytical and machine learning systems.
- Created custom Python scripts for data quality checks, monitoring, and alerting systems, maintaining high accuracy of operational data across the data platform.
- Designed and deployed big data processing jobs on AWS EMR with Spark for transforming vast volumes of raw sensor data, preparing it for machine learning training.
- Managed and optimized cloud-based data infrastructure on AWS using Terraform, ensuring secure, scalable, and reliable environments for data engineering operations.
- Developed data access solutions for autonomy and operations teams, facilitating faster retrieval of crucial vehicle data to support system development and debugging.
- Implemented data orchestration using AWS Step Functions and EventBridge for complex data pipelines, automating workflows for enhanced efficiency and reliability.
- Contributed to data modernization initiatives by optimizing data processing workflows in Databricks, improving efficiency for analytical and reporting needs.
- Maintained and improved data platform components, including containerized data processing applications on AWS ECS, ensuring stable operations for logistics data.

**Deliverr**, Data Engineer **Oct 2021 - Aug 2022**

- Designed and developed scalable data pipelines using Python and Databricks, integrating diverse e-commerce order and inventory data to power real-time fulfillment and logistics analytics.
- Implemented robust data quality checks within Python-driven data pipelines, ensuring the reliability and accuracy of key metrics for e-commerce logistics performance tracking and reporting.
- Optimized cloud-based data processing workflows utilizing AWS services, enhancing the scalability and resilience of the core fulfillment data platform for rapidly growing e-commerce operations.
- Created complex data transformation logic using advanced SQL, generating aggregated datasets for data scientists to improve predictive models for e-commerce route optimization.
- Developed and maintained containerized data processing applications on Amazon ECS, ensuring stable operations for high-volume e-commerce fulfillment data ingestion and processing.
- Automated complex data workflows using AWS Step Functions and EventBridge, ensuring timely data availability for real-time analysis of shipping events and critical inventory updates.
- Configured and managed cloud data infrastructure on AWS using Terraform, provisioning secure and scalable environments for e-commerce fulfillment data engineering operations.
- Integrated Kafka streams for real-time data ingestion and processing, enabling immediate analysis of shipping events and critical inventory updates across the platform.
- Built robust ETL workflows using Apache Airflow to process large-scale logistics data, ensuring timely delivery of aggregated insights for business intelligence dashboards.
- Processed large-scale e-commerce logistics data using Spark on Databricks, enabling complex analytical models for enhanced route optimization and predictive planning.

## TECHNICAL SKILLS

**Cloud Platforms:** AWS S3, AWS EMR, AWS Glue, AWS Lambda, AWS Step Functions, AWS EventBridge, AWS CloudWatch, AWS IAM, AWS VPC, AWS EC2, AWS SQS

**Big Data Processing:** Spark, Databricks

**Data Formats & APIs:** Delta Lake, Parquet, JSON, REST APIs, FastAPI

**Programming Languages:** Python, SQL, PySpark

**DevOps & Infrastructure:** Terraform, Git, CI/CD, Docker, Amazon ECS, Kubernetes EKS

**Databases & Data Warehousing:** AWS Aurora, AWS OpenSearch, PostgreSQL

**Data Orchestration & Streaming:** Apache Airflow, Kafka

## EDUCATION

**M.S., Computer Science**, Computer Science Aug 2024

Campbellsville University