

Abhiram Kakinada

abhiramsss33@gmail.com | +91 7382391510 | linkedin.com/in/abhiramkakinada | github.com/Abhiram1510

Summary

Data Engineer with around **3 years of experience** in building scalable data architectures and integrating AI into workflows. Skilled in Azure technologies (**Databricks, ADF, ADLS Gen2**), **ETL pipelines**, and performance optimization. Experienced in fine-tuning large language models (**LLaMA, BERT, GPT**) and using **Python, PySpark**, and **SQL** to improve data processing and analysis.

Technical Skills

Programming Languages: Python, SQL, PySpark, Cypher Query Language (CQL), Dart

Data Engineering Tools: Databricks, Apache Spark, Streamlit, VS Code, Power BI

Cloud Technologies: Microsoft Azure (Azure Data Factory, Azure Data Lake Storage Gen2, Azure Synapse Analytics, Azure Log Analytics)

Databases: MySQL, Microsoft SQL Server, Neo4j (Graph Database)

AI and Machine Learning: LLaMA, BERT (Fine-tuning and Optimization), Generative AI (GPT), Prompt Engineering, Large Language Models (LLM), Retrieval-augmented generation (RAG)

Development Frameworks: Flask, Streamlit, Modular Programming, Agile Development

Work Experience

Tata Consultancy Services | Data Engineer | July 2022 – Present

- **Azure Flow OptiConnect (Financial Sector):**
 - Designed a **medallion architecture**-based data processing framework using **Databricks** and **ADF**, managing raw, processed, and curated data layers.
 - Migrated SFTP-based ingestion pipelines from NiFi to Azure, enabling high-throughput data processing with optimized performance.
 - Automated data synchronization with Hive tables via **HBase**, improving accessibility by 40%.
 - Implemented **Git** for version control to ensure traceability and collaboration.
- **Modernizing Data Workflows:**
 - Migrated legacy ETL pipelines to **Azure Databricks** and **ADLS Gen2**, streamlining integration processes across multiple environments.
 - Enhanced workflow reliability with **Azure Log Analytics**, reducing error resolution time by 30%.
 - Improved data processing speed by tuning **SparkSQL** queries, achieving a 20% reduction in runtime.
- **AI-Driven Data Mapping (Proof of Concept):**
 - Fine-tuned **LLaMA**, **BERT**, and **PHI3** models for extracting and structuring complex data relationships, increasing mapping efficiency by 35%.
 - Designed ETL pipelines with **PySpark** and **Pandas** to preprocess raw data for AI-driven workflows.
 - Developed a Streamlit-based interface for interactive querying, providing stakeholders with AI-enhanced outputs.

Certifications & Achievements

- **Databricks Certification:** Databricks Certified Data Engineer Associate
- **A-Grade Rating:** Consistently received an A-grade performance rating for two consecutive years in my professional career.
- **State Topper:** Achieved the highest rank in the International Mathematics Olympiad, showcasing exceptional problem-solving skills.
- **65th All India Rank :** Got 65th rank in GATE 2022
- **Mentorship and Leadership:** Served as a student leader for my department, fostering collaboration and organizing key academic and extracurricular initiatives.

Education

Graduation: Graduated in Naval Architecture and Marine Engineering from **Andhra University** - Visakhapatnam (2018 – 2022) - CGPA:8.02