Mahaboob Sheik

Data Engineer

+91- 9182912647 | Mahaboobsheik26@gmail.com| linkedin.com/in/mahaboob-sheik |Hyderabad,Telangana

SUMMARY

Experienced Data Engineer adept at transforming complex data into valuable insights. Proficient in utilizing Azure cloud services such as **Databricks, Data Factory, and SQL** to collect, process, and analyze data. Skilled in **PySpark**, **SQL**, **Apache Spark** and **Python**. Demonstrated success in spearheading data engineering initiatives, leading cross-functional teams, and delivering impactful solutions. Seeking new challenges in a dynamic, growth-oriented environment.

WORK EXPERIENCE

Data Engineer
Modak Analytics

April 2021 - Present
Hyderabad, Telangana

- Managed and directed a team of 5 individuals to optimize product delivery timelines, resulting in a **20%** increase in on-time project completions.
- Designed Scalable big data processing pipelines using **Hadoop, Spark, Scala** and other distributed computing technologies to process **50+** Terabytes Monthly.
- Improved query performance by **20%** through analysis and optimization, enhancing data processing efficiency.
- Developed GLAPI certified **CI/CD** pipelines for automated deployment using SonarQube, SEC API on **Azure Devops**, leading to **30%** faster deployment and enhanced data processing capabilities.
- Orchestrated seamless data ingestion and storage across multiple sources into **Azure Data Lake Storage (ADLS Gen 2)** amplifying data accessibility by **30%.**
- Implemented efficient data pipelines using **Scala and SPARK**, reducing data processing time by **20%** .
- Conducted root cause analysis on **ETL** process failures and implemented 10+ changes that reduced recurrences by **38%.**
- Implemented CI/CD pipeline to automate exporting **StreamSets** pipelines/jobs from Dev to QA environment .

SKILLS

- **Programming Languages**: PySpark, Python, SQL, Scala, Data Structures and Algorithm.

- **Cloud Services and Technologies**: Azure Databricks, Azure Data Factory, Azure Synapse, ADLS, Azure DevOps and Delta Lake.

- **Data Processing**: Hadoop, Hive, Apache Spark, Kafka, StreamSets.

EDUCATION

Bachelor of Technology in Electronics and Communication Engineering
SRKR Engineering college

2017 – 2021

- Cumulative GPA : **8.7/10.0**

CERTIFICATIONS

- Microsoft Certified Azure Fundamentals (AZ-900)
- StreamSets WhiteBelt Certification