Adil Inthiyaz Shaik

+91 9390212052 | adil.inthiyaz@gmail.com | linkedin/adilinthiyaz | github.com/Inthiyaz184

EXPERIENCE

Data Engineer   Sep 2023 – Present
Cognizant (Client: eBay)   *Hyderabad, IND*

- Designed and maintained scalable **ETL** pipelines (**Spark/Airflow** on **Hadoop/Hive**) processing **50M+** records daily at **85**% uptime, automating compliance reports for legal and risk stakeholders.
- Rearchitected batch pipelines to real-time (**Kafka/PySpark**) systems, reducing data processing latency by **80**% and accelerating compliance enforcement.
- Optimized Spark jobs via partitioning, caching, and query tuning, slashing runtime by **35**% for faster risk insights.
- Implemented efficient Hive data warehouse models and bucketing, boosting query performance by **25**%.
- Collaborated with cross-functional teams to translate business needs into technical solutions, reducing data delivery timelines by **30**%.

Data Engineer Intern   Mar 2023 – Jul 2023
Cognizant   *Hyderabad, IND*

- Built ETL workflows using **ADF/Databricks**, processing **1M+** records/day, enabling faster business insights.
- Engaged in knowledge-sharing sessions and hackathons, contributing to team innovation and presenting solutions that improved data processing efficiency by **20**%.

TECHNICAL SKILLS

Languages: Python, SQL, Java, Shell Scripting
Big Data & Platforms: Spark (PySpark, Spark SQL), Hive, Hadoop, Kafka, PostgreSQL, MySQL, NoSQL, Linux
Orchestration & DevOps: Airflow, UC4, Azure Data Factory, CI/CD, Git, Docker
Cloud & Platforms: Azure (ADLS Gen2, Databricks), Snowflake, HDFS, Delta Lake
Data Engineering: Data Modeling, Data Validation, Lineage, Data Governance

PROJECTS

Netflix Data Engineering Pipeline | *Azure Cloud, Databricks, PySpark, Delta Lake*

- Developed scalable ETL pipeline using **ADF**, **ADLS**, and **Databricks** to process **100K+** records with **99.99**% accuracy via Medallion Architecture.
- Implemented Databricks **Autoloader** and **Delta Live Tables**, reducing processing time by **40**% and boosting data integrity to **99.9**%.
- Orchestrated Databricks workflows with metadata-driven logic, reducing pipeline failures by **50%**.

Realtime Data Streaming | *Airflow, Kafka, Spark, Zookeeper, Docker, Cassandra*

- Engineered real-time pipeline (**Airflow**) ingesting **1M+** daily API streams into **PostgreSQL**.
- Streamed **500K+** messages/day to **Kafka** via **Zookeeper** & processed **200K+** events/min with **Spark** into **Cassandra**.
- Containerized solution (**Docker**) cutting deployment time by **30**% while ensuring performance consistency.

ACHIEVEMENTS & CERTIFICATIONS

- **LeetCode** : Tackled **300+** DSA problems. **HackerRank** : Solved over **200** SQL problems.
- Achieved **Databricks Certified Data Engineer Associate** (Delta Lake optimizations applied to reduce Netflix pipeline runtime by 25%).
- Earned **Microsoft Certified - Azure Data Engineer Associate** (ADLS Gen2 & Data Factory solution ensured 99.99% accuracy across 100K+ records).
- **Snowflake SnowPro Core** Certified, covering 50+ topics including virtual warehouses, query optimization, data sharing, and access control.

EDUCATION

Gayatri Vidya Parishad College of Engineering   Sep 2019 – May 2023
Bachelor of Technology in Electronics & Communication Engineering - CGPA: 8.52   *Vizag, IND*