Nayan Deep

Cloud Data Engineer | Big Data Developer

(+91) 9123918500 | nayandeep3580@gmail.com| Delhi, IN | linkedin.com/in/deepnayan/

EXPERIENCE

Skilled Big Data Engineer with 3 years of excellent track record of designing and implementing efficient ETL processes. Proficient in Apache Spark, Python, SQL, ADF and Big Data Technology.

Tata Consultancy Services | *System Engineer*   *Gurugram 12.2021 – Present*

- Engineered an ETL pipeline following Medallion Architecture for a critical retail project, processing 1TB of transactional data weekly, which improved reporting efficiency by 20%.
- Implemented PySpark through Azure Databricks solutions in a large-scale retail project, reducing batch processing time by 30% and enhancing real-time data insights.
- Conducted in-depth analysis on MySQL databases for an e-commerce project, reducing query run time by 20% and speeding up customer insights delivery
- Collaborated with cross-functional teams in an Agile environment, contributing to a 95% on-time delivery rate.
- **Validated** data at a row level, successfully filtering out **10%** of records due to duplicates, null values, and datatype mismatches, with failed rows managed in an exception layer.

PROJECTS

Project – 2 (Retail Company)   ***Tech stacks: ADLS, Azure, Databricks*** *07.2023-Present*

- **Engineered** the entire ETL pipeline for country UK, which is one of their largest Revenue contributing market.
- **Performed** business transformations on silver layer data, preparing **50+ business-critical** datasets for downstream teams, enabling accurate dashboard creation and visualization.
- **Executed** reconciliation processes, logging and monitoring data movement across layers, ensuring **98%** data accuracy and integrity.
- **Automated** pipeline scheduling through Databricks workflow, reducing manual intervention by **60%**, and set up alerts via SMTP, ensuring prompt notification of any issues.
- **Integrated** data from SAP BW and AS400 into ADLS, handling the ingestion of **1TB** of data Weekly into the data lake.
- **Developed** a data pipeline using Data Frame Reader API, efficiently moving **50M+ Records** weekly from the landing zone to the gold layer via bronze and silver layer, ensuring Delta file format storage.

Project – 1 (Pharmaceutical Company)   ***Tech stacks: Hadoop, Hive, Sqoop*** *06.2022-05.2023*

- **Imported and exported** over 1TB of data into HDFS and Hive using Sqoop, efficiently managing and transforming complex JSON data into flat structures, resulting in a 30% reduction in data retrieval time.
- **Delivered** processed data to Hive by creating 20+ Hive tables, loading 1 TB of data, writing optimized Hive queries, and designing 10+ NiFi flows to ensure seamless data delivery to downstream systems, improving data accessibility by 25%
- **Optimized** Spark SQL queries, reducing query execution time by up to 40% through the implementation of advanced Spark optimization techniques, enhancing overall system performance.
- **Improved** Hive table performance by implementing Partitioning and Bucketing techniques, which resulted in a 35% reduction in query processing time for large datasets.
- **Conducted** over 15 code and peer reviews, performed unit testing for 50+ modules, and resolved critical bugs, leading to a 20% decrease in production issues and enhanced code quality.

EDUCATION & CERTIFICATION

Heritage Institute of Technology | *B. Tech (Mechanical Engineering)*   *Kolkata 07.2017 – 06.2021*

- Secured 8.24 CGPA overall in course work, demonstrating a strong understanding stats and Mathematics.

Guru Gobind Singh Public School | *Intermediate 12th (PCM)*   *Bokaro 06.2015 – 04.2017*

- Attained an **84%** aggregate in physics, chemistry and Mathematics, showcasing proficiency and excellence in core subjects.

TECHNICAL SKILLS

Big Data Technologies	Apache Sqoop, PySpark, Hadoop, Hive, Impala
Cloud Technology	ADLS, Azure Data Factory, Azure Databricks, Azure SQL Database, Azure Synapse
Data Engineering & Tools	ETL Data Pipeline, Data Analysis, Data quality Testing, DevOps
Database Technology	MySQL, Postgre SQL, Oracle DB
Soft Skills	Agile Development, Teamwork, Collaboration

ACHIEVEMENT

- **Earned the TCS On-the-Spot Award** from management for exceptional work performance, directly contributing to a 30% increase in team productivity within a single month.
- **Received the TCS Best Team Player Award** for outstanding collaboration and teamwork, recognized by management for significantly enhancing team efficiency and morale.
- Managed annual college fest - Heritage Utsav Apr'19, overseeing a budget of Rs.7,00,000 and a team of 25 volunteers.