

INDRAJEET SINGH LODHI

Contact: +918435301423

Email: indrajeetsinghthakur750@gmail.com

Location: Indore, MP, 452010

PROFESSIONAL SUMMARY

With 3.2 years of experience Java Developer at Rakuten Symphony India Pvt. Ltd., specializing in ETL tool development & Data Engineering, Apache Spark, Kafka, big data technologies, Streaming & microservices. Experienced in developing Spring Boot applications, Hibernate JPA, and used CI/CD pipelines with Jenkins for cloud-based environments. Familiar with Kubernetes for managing and troubleshooting services. Looking forward to advancing my career in a Data Engineering role.

TECHNICAL SKILLS

- **Programming Languages** - Java, Scala (specific to spark-shell) & Python (specific to airflow DAG)
- **Big Data Ecosystem** – Apache Spark, SQL, Kafka, NIFI, Airflow for ETL tool & data pipelines development.
- **Messaging System:** Kafka (used in Spark Streaming for real-time data pipelines) & Event handling
- **Microservice Frameworks** - Spring Boot, RESTful APIs & Hibernate JPA
- **Databases** - MySQL (3+ years)
- **Containers and Orchestration** - Kubernetes, Docker, Linux & Ubuntu
- **Development Tools & Methodologies** - IntelliJ, Eclipse, VS Code, Git, Jenkins, Maven, Agile (Scrum), CI/CD with Jenkins

WORK EXPERIENCE

Software Engineer | Rakuten Symphony India Pvt. Ltd., Indore, MP | July 2021 – Present

- **Developed and enhanced Big Data ETL tool** using Java, NIFI, Spark, Spring Boot Microservices, and Kafka, allowing users to design and customize their data pipelines effectively.
- **Built and maintained a streaming job** for Spark job audit log monitoring, ensuring real-time tracking and auditing of data processing jobs.
- **Developed data engineering pipeline** for Mobile Number Portability (MNP) in telecom, processing massive amounts of data using Spark and NIFI.
- **Ingested largescale data using NIFI pipelines** via SFTP and HTTP protocols as part of the MNP big data project, ensuring seamless and secure data flow.
- **Processed and transformed data** using Spark data pipelines for MNP, enabling efficient data handling and transformation at scale.
- **Developed an ETL tool from scratch with Airflow**, Spring Boot Microservices, Spark, and Trino, ensuring robust, scalable, and reliable data integration across diverse environments.
- **Utilized Airflow to schedule Spark jobs**, automating complex data workflows and enabling efficient orchestration and monitoring of big data processing tasks.
- **Troubleshoot and debugged Spark job** big data pipelines, resolving user reported issues to maintain the stability and performance of the data processing framework.
- **Worked on Spark Streaming job module for ETL**, implementing solutions to automatically recover and restart hanging or failed Spark streaming jobs.
- **Developed new data source connectors** and enhanced existing connectors for the big data pipeline, facilitating the integration of various data sources into the ETL workflow.

- **Created custom NIFI processors** and modified existing processors, enhancing the functionality of the NIFI pipeline and adapting it to new data processing requirements.
- **Leveraged Kubernetes, Docker, Git, Maven, and Jenkins** while following Agile practices for efficient CI/CD pipelines and deployment in big data environments.

PROJECTS

Product - Process-Runner Big Data ETL Tool

- **Developed and enhanced a Big Data ETL tool** using Apache Spark, NIFI & Kafka for designing, scheduling, and monitoring data pipelines.
- **Created a user-friendly Spark pipeline designer** with pre-built components, integrated with Apache NIFI for data processing like Kafka connectors, DB Connectors, S3 connectors, etc.
- **Created a streaming data pipeline** for job monitoring and customized NIFI, including the development and deletion of new processors and connectors to streamline Big Data operations.

Data Engineering Project – MNP Telecom (Client: 1&1, Germany)

- **Designed and implemented a robust Big Data ETL pipeline** for Mobile Number Portability (MNP) using Apache NIFI and Apache Spark, facilitating seamless data integration and processing across telecom systems.
- **Used NIFI for Big Data ingestion** from SFTP
- **Used Apache Spark** to processed large volumes of data to ensure efficient data transformation and loading.
- **Configured mail alerts for the client**, enabling real-time notifications and status updates for critical ETL data processing events.

Product – Data-Transformation Big Data ETL Tool

- **Engineered Scala-based Spark processing engine**, designed data catalogs using Trino, and processors for easy management of Big Data.
- **Utilized Airflow for scheduling Big Data ETL pipelines**, automating complex workflows and ensuring reliable execution of data jobs.
- **Developed new connectors** for additional data sources, enhancing the flexibility and scalability of the ETL pipeline.
- **Designed and managed data catalogs using Trino**, improving Big Data discoverability and governance across the system.

EDUCATION

Bachelor of Technology in Computer Science and Engineering (August 2017 – May 2021)

- BTIRT, Sagar, RGPV, Bhopal, MP
- CGPA: 8.81/10