MOUSAM KUMAR

Male , 24 years | +918340786513
Email id: mousamkumar312001@gmail.com | linkedin.com/in/mousam-kumar-6289651a3/

SUMMARY

- Data oriented professional with 2.5 years of work experience in data integration and analysis with proficiency in **SQL, Python , Spark , GIT and Pyspark Plugins** .
- Skilled in handling large-scale datasets, designing efficient data integration batch, and automating workflows.
- Strive to leverage interpersonal skills for team-oriented task.

ACADEMIC QUALIFICATION

Education	Institute	%/CGPA	Year
B.Tech , Civil Engineering	NIT Jamshedpur	8.24	2023
XII (JAC)	St Xavier's College, Ranchi	80.6	2018
X (JAC)	Bhaskar Saraswati Vidya Mandir	88.6	2016

SKILLS

TECHNICAL SKILLS	Soft Skills
O9 Platform : Integration 3.0 , Config 2.0 , Python/Pyspark Plugins **Spark:** Spark Sql , Pyspark **Python :** Numpy , Pandas **Data Engineering :** E2E Batch Architecture Designing , Data Integration, Data Validations SQL/Spark SQL, GIT, MS Excel	- Problem Solving - Interpersonal skills - Team Work - Leadership - Time Management

WORK EXPERIENCE

O9 Solutions Pvt Ltd	Senior Business Operations Specialist

- Designed and developed **data integration pipelines** in the o9 platform for **Pandora** project, enabling seamless data ingestion with **validation checks** and **data pre-processing** through **o9 latest Int3.0 technology.**
- Collaborated with the functional team to conduct exploratory data analysis (EDA) on large datasets using SQL and Python.
- Performed **forecast accuracy checks**, including MAPE , WAPE calculations, across multiple levels on extensive datasets.
- Utilized advanced SQL techniques, including CTEs , window functions,regular expressions and complex joins, to derive actionable insights and reports.
- Responsible for **version control** in **Git** for **data integration pipelines** for each sprint .
- Provided knowledge transfer sessions to partner resources and explained them about integration 3.0 , its capabilities along with EKG Configurator and EKG Deployment.
- Developed **pyspark plugins** to simplify complex ibpl procedures and to improve performance.
- Improved the **batch performance** with collaboration with functional team by implementation of parallel processes and via tuning parameters.
- Developed many innovative utilities not available in the platform to reduce the manual works and shared with the product team to receive the feedback.

INNOVATIVE TOOLS & UTILITIES DEVELOPED

Easy way to move delta-lake backup Utility

- Developed a utility pipeline , which takes backup of the delta table schema from source environment and restore it into the destination environment.
- Created a detailed documentation and package of the utility so that other projects can utilize this.

Automation of Unit Testing File Creation

- Developed a python code which automated the manual and time taking process of unit testing file creation which generates file based on the parameter set in the EKG Configurator.
- Gave KSS to product team and shared the code with them for the further enhancement on this and to make available this for others.

Azure Data Lake Archive PIpeline

Developed a python based utility pipeline , which can archive the files present in ADLS datastore. This solution streamlines the process of archiving client data, improving data management.

COURSES & CERTIFICATIONS

Integrations using **Int 3.0** and **EKG Configurator**:**o9 Academy**	Advance Microsoft Excel Certification - **Tata Steel**
Configuring Solutions using Config 2.0 : **o9 Academy**	**I**ntroduction to Spark SQL and DataFrames – **Linkedin Learning**
o9 Platform Architecture : **o9 Academy**	Learning Apache Airflow **– Linkedin Learning**
o9 Knowledge Masters in Supply Chain Management:**o9 Academy**	Business Analytics For Beginners - **Upgrade**
Staging API : **o9 Academy**	Apache Pyspark by Example **- Linkedin Learning**

HOBBIES AND INTERESTS

Chess,Table Tennis, Foosball