

Lokesh Bisen

Data Engineer

✉ lokeshbisen989@gmail.com ☎ 9717302074 📍 New Delhi, India in LinkedIn 🐙 GitHub

Profile

Experienced in field of technology and leveraging data to drive business insights and enhance operational efficiency. looking for transition to the Data engineering Role with a solid foundation in data processing, data modeling, and ETL processes, proficient in a Big Data technologies, I am committed to delivering high-quality, actionable insights that empower decision-making and fuel organizational growth.

Skills

- | | | | |
|--------------|-----------------|--------------------|-----------|
| • Big Data | • Spark | • Hadoop | • Scala |
| • Python | • Tableau | • Hive | • AWS |
| • SQL | • Agile | • Docker | • Kafka |
| • Databricks | • Azure | • HBASE | • Airflow |
| • NLP | • Deep Learning | • Machine Learning | • Sqoop |
| • MS Office | • Redshift | • Sabre | • Amadeus |

Projects

Retail Company ETL Pipeline

- **Problem Statement-** The retail company faced challenges due to scattered and disconnected data sources, causing inefficiencies, data silos, and inconsistent reporting. The lack of real-time insights hindered decision-making and operational efficiency.
- **Approach-** 4000+ worldwide retail store data extracted from S3 and added transformation layer of spark. then enriched data sent to DataMart (MYSQL). Implemented ETL for the retail company enables streamlined data processing from point of sale systems to comprehensive analysis. By utilizing Python, AWS, SQL and Power BI to enhance operational efficiency of worldwide stores chain.
- **Conclusion-** After creation of pipeline, it helped company to make decisions faster and also in creation of real time analysis.

Heath Care Domain Analysis ETL Pipeline

- **Problem Statement:** The healthcare company struggled with managing data from multiple sources, including patient records, medical devices, and billing systems. The lack of data integration led to inefficiencies, delayed reporting, and challenges in providing timely patient care.
- **Approach :** Used spark for transformation, airflow for scheduling and enriched data sent to AWS s3 and s3 connected with Quicksight for analysis. Implemented an infrastructure that handles diverse datasets by using big data technology and cloud service to enhancing service delivery in health sector.
- **Conclusion:** After the creation of pipeline, process of reporting of each patient became fast which helped in managing patient care.

Certificates

Data Engineering Certification

CloudyML

Awards

SID KPI Award 2023

Gold badge in SQL & python on Hacker rank

Professional Experience

Business System Analyst, American Express GBT

02/2022 – present

- worked on data ingestion pipeline to ingest the flat-file in the data lake called CPR tool. Analyzed complex data sets, identifying key areas for improvement. developed data-driven solutions that increased customersatisfaction.
- Provide technological solutions, product deployment, and end-to-end technology solutions implementation, integration, and maintenance.

Data Science Intern (Internship), Flip Robo Technologies

07/2021 – 01/2022

- Worked on data extraction from API and created ETL pipeline, stored data in hive and further connected with powerBI for analysis.
- Worked on Fake news Detection Project using NLP/ML and also worked in telecom industry ML project in providing solution for micro credit loan .

Consultanting Analyst, Travel Boutique Online (TBO)

07/2018 – 01/2020

- Part of the Technology Development Team. The work is to extract data from API and update information in database.
- Maintaining statistical and financial records. Providing API information support.

Languages

English | Hindi

Education

PG diploma in Data Science ML/DL and NLP,

Data trained Education

2021 – 2022

Bachelor of Computer Application,

Veer Bahadur Singh Purvanchal University

2014 – 2017

Interests

Reading | Gardening