KOLLU SAI KUMAR

✉ saikumarkollu855@gmail.com | 📞 +91 7675997126 | 💼 sai kumar kollu | ⊙ sai019

SKILLS

DATA ENGINEER

PROGRAMMING LANGUAGES :
• Python

DATABASES :
• MySQL • Cassandra

BIG DATA FRAMEWORKS :
• Hadoop • Hive • Spark • Kafka
• Airflow

AZURE CLOUD SERVICES :
• Data Factory • Azure Storage
• Azure DataBricks • Data Lake Gen2

OTHERS :
• Git

EDUCATION

K L UNIVERSITY

B.TECH IN COMPUTER SCIENCE
MAY 2019 | Vijayawada

SRI CHAITANYA JUNIOR COLLEGE

STATE BOARD OF INTERMEDIATE
12TH (M.P.C)
MAY 2015 | Vijayawada

NEW VISION THE CONCEPT SCHOOL

10TH SSC
MAY 2013 | khammam

COURSEWORK

UNDERGRADUATE

Object-Oriented Programming
R-DBMS
operating systems
Artificial Intelligence

LANGUAGES KNOWN

• English • Telugu

EXPERIENCE

INEURON | DATA ENGINEER INTERN   DECEMBER 2023 - APRIL 2024
| Remote, india

- Established Azure development environment integrating Azure Databricks and Azure Data Lake Storage Gen2.
- Implemented efficient data ingestion to Azure Data Lake Storage Gen2 using Azure SDK, transferring from web sources to designated folders for subsequent data processing.
- Performed data transformation and cleansing operations using PySpark within Azure Databricks to maintain data integrity and schema standards.
- Configured Azure Key Vaults for secure integration with Data Lake, managing access permissions and authentication for enhanced data management and analysis.

CYIENT | EMBEDDED SOFTWARE ENGINEER   JULY 2022 – JULY 2023
| Bengaluru, india

- Validated incoming project requests against source documentation for various aircraft modules (rudders, spoilers, ailerons, etc.) to ensure accuracy and compliance using IBM DOORS.
- Streamlined the documentation and validation process, improving efficiency and reducing errors in project request handling for Boeing initiatives.

PROJECT(S)

STOCK-MARKET-DATA-ENGINEERING-PROJECT | END – END

- Deployed Confluent Kafka for real-time data streams with secure API access, creating topics for efficient ingestion.
- Developed Python script to produce stock market data to Kafka topic in JSON format.
- Utilized Python consumer script to ingest Kafka JSON data into Azure Data Lake Storage Gen2 via Azure SDK.
- Integrated Azure Synapse Analytics with Azure Data Lake Storage and utilized a serverless SQL pool to run SQL queries on the JSON data.
- Github-Link// https://github.com/sai019/Stock-Market-Data-Engineering-Project

COURSE CERTIFICATIONS

BIG DATA BOOTCAMP : BY INEURON