Asparsh Raj

(+91) 7667406415 | rajasparsh19@gmail.com | LinkedIn | GitHub | Portfolio | LeetCode | Jamui, Bihar

Objective

Motivated Computer Science student seeking a Data Engineer role to apply Data Engineering skills and gain hands-on experience in data pipeline development. Eager to contribute to innovative projects and enhance proficiency in Python, SQL and Big Data technologies.

Education

Bachelor of Computer Applications (B.C.A.) - 8.8 CGPA — *Sarala Birla University*, *Ranchi, Jharkhand*
September, 2022 - Ongoing

Skills

Programming Languages: Python, SQL, JavaScript
Web Programming: HTML. CSS, Flask
Big Data Technologies: Apache Spark (PySpark), Apache Kafka
Database Systems: MySQL, PostgreSQL
Tools and frameworks: Git, Tableau, Pandas, Matplotlib
Others: Problem Solving, Verbal Communication, Written Communication, Analytical Thinking

Projects

PySpark script for Batch Data Transformation

Sarala Birla University, *Ranchi, Jharkhand*
Oct 2023

- Conceptualized and implemented a data pipeline to extract and transform weather data from open source weather API, demonstrating self-motivation and self-starter spirit.
- Engineered a scalable data processing scripts using PySpark to extract and transform raw weather data for analysis which lead to 43% reduction in time required to manually transform the data on arrival.
- Aggregated hourly weather data worth over 7 days and visualized over 4 weather indicators using Tableau dashboards.
- Applied PySpark DataFrames to manipulate and transform data, ensuring the integrity and quality of the processed data.
- Upheld industry standards by following PEP 8 coding conventions, emphasizing code readability and maintainability.
- Documented the project thoroughly, providing clear insights into the problem-solving process and making the project accessible to others.
- Spark Source Script

Personal Portfolio Website

Sarala Birla University, *Ranchi, Jharkhand*
Dec 2023

- Conceptualized, designed, and custom-built a visually captivating, responsive portfolio website, achieving a stunning 40% increase in user engagement.
- Developed a seamless user experience using HTML, CSS, and Tailwind CSS, resulting in a 20% decrease in page load time and enhancing overall site performance.
- Actively integrated and hosteded the website on GitHub Pages, establishing a dynamic online presence that garnered 150+ views within the first month.
- Illustrated and showcased a diverse portfolio of over 10 projects, proving skills and expertise, with a notable 15% increase in project visibility and recognition.
- The meticulously planned and shaped portfolio serves as a living testament to a journey of continuous improvement and skill refinement, contributing to a 60% growth in professional opportunities compared to last year 2022.
- Portfolio Link: https://asparsh607.github.io/portfolio

Apache Kafka real-time data processing script in Python

Sarala Birla University, *Ranchi, Jharkhand*
Oct 2023

- Engineered and implemented a high-performance Python producer script, generating an impressive 5000 real-time flight data points per minute, which can also be customized, elevating the project's dynamism.
- Orchestrated a Kafka topic, facilitating the seamless publication of real-time flight data at a rate of 100 events per second, ensuring a rapid and efficient data flow.
- Automated a Python consumer script that continuously loads a staggering 100K real-time data points daily from the Kafka topic into a MySQL database on the local system, optimizing data accessibility and storage efficiency.
- Rigorously validated and tested the entire data streaming process, achieving a appreciable 94.7% accuracy and reliability rate, solidifying the project's integrity.
- The comprehensive GitHub repository, serving as the project's nucleus, can be explored here: Real-Time Data Streaming App

Weather Data Visualization- Tableau

Sarala Birla University, *Ranchi, Jharkhand*
Oct 2023

- Engineered a PySpark script to meticulously collect and process an extensive dataset, resulting in the successful extraction and analysis of weather data from multiple sources.
- Crafted an immersive Tableau visualization, dynamically presenting insights from a robust dataset that spans over 10,000 data points.
- Implemented an interactive bar chart, effectively illustrating five key weather metrics—temperature, pressure, and more—providing a detailed perspective on atmospheric conditions.
- Visualized a comprehensive week's worth of weather metrics from the query date, encompassing over 100 data entries, offering a granular understanding of weekly atmospheric trends.
- Link

Certifications

- SQL Basic: HackerRank — Apr 2023
- Python Basic: HackerRank — Apr 2023
- Introduction to Data Engineering: Datacamp — July 2023
- IBM Data Engineering Professional Certificate: Coursera — Aug 2023
- Google Cloud: Cloud Data Engineer Professional Certificate: Coursera — Aug 2023