Rutuja Deepak Jadhav

Pune, MH rutujadjadhav0846@gmail.com +91 8080118106 rutuja-jadhav-677aa4231

RutujaJadhav80

Professional Summary

Results-driven Data Analyst with a strong foundation in data processing, statistical analysis, machine learning, and cloud computing. Proficient in Python, SQL, Hadoop, Spark, and data visualization tools like tableau, PowerBI. Skilled in extracting insights from large datasets to drive business decisions. Adept at building predictive models, performing ETL processes, and working with both structured and unstructured data. Passionate about leveraging analytical skills to solve complex data challenges. Strong problem-solving abilities and a keen interest in continuous learning.

Technical Skills

- Programming Languages: Python, SQL, R, Java, C, C++
- Data Visualization Tools : Tableau, Power BI
- Big data Technologies : Apache spark, hadoop, Airflow etc.
- Cloud Services : AWS S3, GCP, Glue, Lambda.
- Machine Learning
- Database management : MySQL, MongoDB

Education

KNOW-IT, C-DAC ACTS (ATC), Pune *Aug 2024 – Feb 2025*
Postgraduate Diploma in Big data analytics (72.13 %)

JSPM's Rajarshi Shahu college of Engineering, Tathawade, Pune *Aug 2019 – Aug 2023*
Bachelor of Technology (Civil Engineering) (8.57 CGPA)

Ligade-Patil junior college of science, Karad *2019*
12th HSC (61.69 %)

S.R.K. Lahoti kanya prashala, Karad *2017*
10th SSC (95.20 %)

Specialization

KNOW-IT, C-DAC ACTS (ATC), Pune
PG- DBDA

- Specialized in Big Data technologies (Hadoop, Apache Spark) with expertise in Python, R, Java, and SQL for data processing using PySpark, Pandas, and NumPy. Worked with MySQL, MongoDB, and gained expertise in machine learning Algorithms using Scikit-learn, TensorFlow. Learned AWS and gained experience in real-time data processing (Kafka, Apache Airflow) and data visualization (Tableau, Matplotlib).

Projects

Diabetes Prediction *gRutujaJadhav80/RutujaDBDA*

- This project develops a robust classification model to predict diabetes risk using the Diabetes Health Indicators Dataset (BMI, glucose levels, blood pressure). Utilizing PySpark, Python (Pandas, Sklearn, Matplotlib), Tableau, and Streamlit, we apply advanced data processing and machine learning techniques to analyze health data. The goal is to enhance healthcare outcomes, promote preventive care, and improve diabetes management through data-driven insights.
- Technologies Used: PySpark, Python (Pandas, Sklearn), Meachine Learning algorithms, Tableau, AWS EC2, Stramlit.

You Tube analysis

- The project aimed to extract YouTube data, transform it for meaningful insights, and load it into a database for analysis. The goal was to analyze video trends, engagement metrics, and user behavior. Key Insights: 1.Identified trending topics based on views, likes, and comments. 2.Analyzed audience engagement metrics over time. 3.Compared content performance across different categories.
- Technologies Used: Python (Pandas), PySpark, MySQL, Tableau.