Rajdeep Sodha

rajdeep.bsodha444@hotmail.com | +91-7208213422 | Mumbai, Maharashtra

EDUCATION

MUMBAI UNIVERSITY
Bachelor of Information Technology
2020 | Score: 7.83 CGPA

MSBTE
Diploma in Information Technology
2017 | Score: 84.76% (with Distinction)

SKILLS

Programming language: Python, SQL.
Linux Shell Scripting
Libraries: Pandas, NumPy
Data Visualization: Qlik, PowerBI
Database: Oracle, MySQL, SAP HANA
Big Data Tech: PySpark, Hadoop
Cloud: GCP
CI Tools: Jenkins
SCM Tools: GIT

PUBLICATION

Bird Species Identifier using Convolutional Neural Network, IJERT, Feb 2021

LINKS

Github:// **rajdeep999**
LinkedIn:// **rajdeepsodha**
Leetcode:// **rajdeep_999**
Hackerrank:// **rajdeep_bsodha91**
GFG:// **rajdeepbsodha999**

CERTIFICATION

Google Cloud Certified - Professional Data Engineer by Google
MTA: Security Fundamentals - Certified 2017 by Microsoft Qlik Sense Qualification Exam – Business Analyst 2022

EXPERIENCE

TATA CONSULTANCY SERVICES | Data Engineer
Aug 2020 - Present | Mumbai

Lloyds Banking Group | Oct 2022 - Present

- Leveraged Python, BeautifulSoup, Selenium, Pandas, NumPy, and SQLAlchemy to efficiently extract, clean, and manipulate data from various UK public registers. Ensured seamless data processing for analysis and storage.
- Developed and implemented algorithms to process the data from 30 public registers, enhancing data quality and accuracy.
- Expanded data collection capabilities by writing Python scripts to fetch data from external APIs, enabling access to a wider range of data sources. Enhanced data collection efficiency and enriched the available datasets.
- Successfully scraped and processed a substantial volume of data, approximately 30 million rows, from multiple public registers on a monthly basis. Provided a comprehensive dataset for in-depth analysis and reporting purposes.
- Implemented simultaneous data cleaning during Python script execution, reducing the overall data cleaning and extraction process from approximately 2 working days to just 5 hours. Increased efficiency and streamlined data preparation.
- Developed and delivered customized Power BI dashboards for the sales and marketing teams, enabling them to personalize offers and improve data quality. Strengthened decision-making capabilities and fostered data-driven practices.
- Created Power BI dashboards for the executive team, facilitating data analysis and understanding of data quality. Achieved a significant reduction in error rates by 95%, supporting strategic decision-making at the organizational level.

TCS Ultimatix | Aug 2020 - Sep 2022

- Managed a high volume of raw data, handling around 50 million records with 110 columns.
- Implemented an end-to-end ETL pipeline in Python and SAP HANA, improving customer billing compliance by 80%.
- Optimized ETL script to reduce run time by 60% and resolve major bottlenecks independently.
- Proficient in GCP's BigQuery for efficient data processing and familiar with other big data technologies.
- Executed end-to-end data lifecycle activities with a focus on data gathering, pre-processing, cleaning, analysis, modeling, and visualization.
- Deployed QlikSense Dashboards using CI/CD pipelines via Git/Github, ensuring accurate and streamlined deployment.
- Developed 50+ risk scenarios using Python and SAP HANA, aligning with use cases provided by the finance team.
- Redesigned Jenkins UI to align with user requirements and adhere to TCS standards.
- Actively involved in AGILE/SCRUM methodologies and contributed to the onboarding of new team members.

ACHIEVEMENTS

- Solve more than 100 SQL coding question at site like hackerrank, Leetcode.
- Awarded star of the month twice at TCS.
- Rank 3705 (Pre-Qualifier Zonal Round) in TCS Codevita Season 8.