Suraj Raghuvanshi

Data Analyst | Data Engineer

surajraghuwanshi05@gmail.com 8120742013 Bhopal LinkedIn GitHub

SUMMARY

Data Analyst and Data Engineer with a strong foundation in data architecture, data transformation, and machine learning. Skilled in building efficient data pipelines, processing complex datasets, and creating predictive models that drive data-driven decision-making. Proficient in both data analytics and machine learning,Proven record in both standalone and collaborative projects, driving solutions that enhance efficiency and data integrity.

EXPERIENCE

iDataMinds
Data Analytics and Engineering Intern | 07/2024 – Present
- Improved data processing by **20%** through a multi-layer data architecture, organizing data flows and enhancing data accessibility for complex analytics.
- Achieved a **15% reduction** in query runtime by refining SQL performance and leveraging indexing and partitioning for optimized data retrieval.
- Built a dynamic Streamlit app for real-time data visualization, accelerating decision-making by **25%** through clear, interactive dashboards.

Zidio Development
Data Science and Analyst Intern | 04/2024 – 06/2024
- Increased model precision by **10%** with advanced statistical analysis and feature engineering, improving data quality for predictive models.
- Enhanced data handling efficiency by **20%** through engineering solutions in collaboration with cross-functional teams, supporting faster data access for stakeholders.

Nexus Info
Data Science and Analytics Intern | 01/2024 – 02/2024
- Created a **Support Vector Machine**-based machine learning model for breast cancer prediction with an **85% accuracy rate**.
- Conducted Twitter sentiment analysis using advanced text processing, extracting key insights from large-scale datasets.

PROJECTS

Multi-Layer SQL Data Architecture
- Designed and implemented a **Bronze, Silver, Gold** data architecture, increasing data processing efficiency by **25%** with clean, consistent data structures.
- Utilized SQL triggers, window functions, and procedures to streamline data operations, significantly improving data transformation workflows.
- Enhanced query-based insights for users with a Streamlit app, offering flexible data querying and filtering that increased insight generation by **30%**.

Twitter Sentiment Analysis
- Conducted sentiment analysis on Twitter data, applying data preprocessing techniques like tokenization, lemmatization, and stop word removal to normalize large text datasets.
- Engineered a Support Vector Machine (SVM) model for sentiment classification, achieving high accuracy in identifying positive, negative, and neutral sentiments.
- Developed a simple web application using Flask to showcase real-time sentiment analysis, enabling accessible sentiment-based insights for users.

Fraud Detection system
- Improved fraud detection capabilities with a machine learning model, using data engineering techniques to identify patterns in transaction behavior.
- Conducted exploratory data analysis on **10,000+ transactions**, employing pandas and visualization libraries to uncover transaction patterns and fraud signals.

EDUCATION

B.TECH (Agricultural Engineering) 2019 – 2023
Jawaharlal Nehru Agricultural university Jabalpur
CGPA -7.69

SKILLS

Data Analysis — Python, EDA, Feature Engineering, Predictive Modeling, pandas

Data Engineering — SQL, ETL, Data Architecture (Bronze/Silver/Gold), Data Cleaning, Data Pipelines,Numpy, Spark, databricks

Data Visualization — Streamlit, Matplotlib, Seaborn, Dashboards

Machine Learning — Scikit-Learn, Predictive Modeling, Neural Networks, SVM, NLP

Additional Tools and Technologies — Git and GitHub, MongoDB, Flask, HTML, CSS, JavaScript,