+447990932599
London, United Kingdom
shrayansh1003@gmail.com

# Shrayansh Bhardwaj

Data Analyst / Data Scientist

github.com/shrayansh0310
linkedin.com/in/shrayansh0310/

As a highly motivated Data Analyst and Data Scientist with a Master's in Big Data Science, I have over 2 years of experience as a Data Analyst and strong skills in Python, SQL, Power BI, and data libraries like Pandas, Matplotlib, and NumPy. I specialize in designing data science projects that turn complex datasets into actionable insights. My analytical mindset, attention to detail, and problem-solving abilities drive data-driven decision-making. I excel in cross-functional teams and am committed to advancing my expertise to contribute to innovative projects.

## Skills

| | |
|---|---|
| **Tools and Languages** | Python, Python Pandas, Python Numpy, Python Matplotlib, Python Seaborn, SQL, DB2, Pyspark |
| **Data Visualization** | Power BI, MySQL, SQL Server, DB2 servers, Power Bi Dashboards, Data Storytelling, DAX |
| **Machine Learning** | Sampling Data, Data Preprocessing, Data cleaning, Regression, Classification, Clustering Analysis, Anomaly Detection, SVM and KNN classifiers, underfitting, and over fitting analyses |
| **People Skills** | Effective Decision-Making, Effective Leadership, Proficiency in Overcoming Challenges |
| **Languages** | English, Hindi |

## Technical Experience

**Data Analyst** **Nov 2020 — Aug 2023**
*Tata Consulting Services* *New Delhi, India*

- Developed and maintained complex SQL queries for extracting, analyzing, and processing large datasets, improving data retrieval efficiency by 20
- Automated data cleaning and transformation processes using Python libraries such as Pandas and NumPy, reducing data preparation time by 30
- Implemented data visualization dashboards in Power BI to present actionable insights to stakeholders, leading to a 15% increase in decision-making speed.
- Conducted in-depth analysis on customer data, uncovering key trends and patterns that drove a 10% increase in customer retention.
- Optimized database performance by restructuring tables and indexing, resulting in a 25% reduction in query processing time.

## Education

**Master's in Big Data Science**, *Queen Mary University of London, United Kingdom* Sep 2023 — Sep 2024
**Bachelor of Technology in Electrical Engineering**, *Swami Keshvanand Institute of Technology, Jaipur, India* Aug 2016 — July 2020

## Certifications

Microsoft Power BI Desktop for Business Intelligence
SQL - MySQL for Data Analytics and Business Intelligence

## Projects

**Advanced Time Series Forecasting for Energy Consumption** **June 2024 — Aug 2024**

- Implemented and evaluated multiple time series forecasting models (ARIMA, SARIMA, Auto ARIMA, Prophet, XGBoost) on PJM energy consumption data, achieving a 98.9% accuracy with XGBoost.
- Conducted extensive feature engineering and cross-validation techniques, enhancing model robustness and reducing forecast errors by over 50% compared to baseline models.
- Developed advanced predictive analytics to inform energy grid management decisions, reducing RMSE to as low as 431.27 and MAPE to 1.011%, optimizing energy demand forecasts.

**Deep Learning Model Development for CIFAR-10 Classification** **Feb 2024 — May 2024**

- Designed and implemented a custom Convolutional Neural Network (CNN) using PyTorch for the CIFAR-10 dataset, progressively improving test accuracy to 82.02% over 60 epochs.
- Applied advanced feature extraction through intermediate blocks, enhancing model performance and stability, with a peak training accuracy of 86.77%.
- Utilized cross-entropy loss and Adam optimizer to efficiently train the model, demonstrating significant improvement in training and testing accuracy over the course of training.

**Machine Learning Pipeline for Cuisine Classification** **Oct 2023 — Dec 2024**

- Developed a machine learning pipeline to classify images of dishes as either Indian or Chinese cuisine using a Support Vector Machine (SVM) and other models.
- Applied advanced feature extraction techniques, including Yellow Component and Gray-Level Co-occurrence Matrix (GLCM), to enhance model interpretability and performance.
- Conducted model evaluation and comparison, demonstrating the effectiveness of SVM over Random Forest and K-Nearest Neighbors for this classification task.