SAIF IBRAHIM

+91 9170055886
saifibrahim.up@gmail.com
LinkedIn
Hyderabad, India

Data Scientist: Portfolio

Professional Summary

Associate Data Scientist combining strong **Python, SQL, and ML** skills with **backend engineering** experience. Expert in the full data lifecycle—from **ETL pipelines** to **predictive modeling**. Dedicated to transforming raw data into insights that drive **operational efficiency** and innovation at **Swiggy**.

Skills

- **Languages & Databases:** Python, SQL (MySQL/PostgreSQL), PHP (Laravel), JavaScript.
- **Python Libraries:** Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Beautiful Soup.
- **Data Science & Stats:** Exploratory Data Analysis (EDA), Statistical Analysis, Hypothesis Testing, Predictive Modeling, Regression & Classification.
- **Tools & Version Control:** Git, GitHub, VS Code, Jupyter Notebooks.

Experience

Project: Inventory Optimization in Pharmaceutical Supply Chain (SQL + Python)

- **Situation:** Managed pharmaceutical inventory optimization addressing **₹528M financial losses** from expiry and overstock
- **Action:** Implemented FEFO system and JIT procurement to optimize inventory management
- **Result:** Achieved **15% overstock reduction**, identified **20-30% profit improvement** potential while maintaining **99.8% fulfilment rate**

Project: Content-Based Restaurant Recommendation Engine (Python)

- **Situation:** Users often struggle to discover new restaurants that match their specific taste preferences among thousands of options.
- **Action:** Built a **Content-Based Recommendation System** using **Natural Language Processing (NLP)**. Utilized **TF-IDF Vectorization** to analyze restaurant descriptions and reviews, and applied **Cosine Similarity** to measure the closeness between user preferences and restaurant profiles.
- **Result:** The model successfully generates the "Top 5 Similar Restaurants" for any given venue, mimicking the "You might also like" feature found in food delivery apps.

Project: Delivery Time Estimation Model (Python & Scikit-Learn)

- **Situation:** Inaccurate delivery time estimates lead to customer frustration and lower app ratings.
- **Action:** Developed a Machine Learning Regression Model (Random Forest) to predict delivery duration based on distance, traffic density, and weather conditions. performed Feature Engineering to convert categorical data (Traffic/Weather) into numerical values.
- **Result:** Achieved an **R2 Score of 0.82** (82% accuracy), proving that traffic density was the most critical factor affecting delays.

EDUCATION

B. Tech (Computer Science Engineering) 2021-2025
KMC Language University, Lucknow **(CGPA: 7.04)**

Certifications

Certifications: Certification in Software Backend Developer by **Routa Digital India Pvt Ltd**.