Viraj Pitale

"Self Taught Data Scientist"

| virajpitale.github.io | linkedin.com/in/viraj-pitale/ |
| virajpitaleworkmail@gmail.com |

SKILLS

Programming Languages: Python, SQL, Java
Python Libraries & Tools: Numpy, Pandas, Matplotlib, Seaborn, scikit-learn, FastAPI, Flask
Machine Learning: Supervised and Unsupervised Learning, Regression, Classification, Decision Trees, Random Forest, XGBoost, K-means, PCA
Deep Learning: Neural Networks, CNN, RNN, LSTM
Data Analysis & Visualization: Exploratory Data Analysis (EDA), Data Cleaning, Data Manipulation, Data Normalization, Feature Engineering
Statistics & Mathematics: Descriptive and Inferential Statistics, Probability, Hypothesis Testing, Correlation, Central Limit Theorem
Database Management: Relational Databases, Basic and Advanced SQL Queries, Joins, CTE, Subqueries, Window Functions
Web Development: HTML, CSS, JavaScript, Flask, FastAPI
Deployment: AWS, Azure
Tools & Platforms: Excel, ChatGPT, PowerPoint, JIRA, Notion, GitHub

EDUCATION

Indian Institute Of Technology (IIT) Madras
Bs In Data Science And Applications, Expected Graduation 2027

PROJECTS

Mumbai House Price Prediction

- **Tools & Technologies:** Python, Pandas, scikit-learn, FastAPI, HTML, CSS, JavaScript
- **Situation:** Tasked with predicting house prices in Mumbai using a dataset.
- **Task:** Cleaned data, engineered features, and developed a machine learning model to predict house prices.
- **Action:** Explored various machine learning models, including Decision Tree Regressor, fine-tuning parameters to achieve optimal accuracy. Implemented FastAPI for backend deployment and used HTML, CSS, and JavaScript for frontend development.
- **Result:** Achieved an accuracy of 85%-90% on test data, demonstrating a significant understanding of predictive analytics and web deployment. The project provided valuable insights for real estate firms to enhance their pricing strategies.
- **Explore the Project:** Check out the LinkedIn Post for an overview and view the full code on GitHub.

Loan Approval Prediction Model

- **Situation:** Faced with the challenge of predicting loan approvals accurately, which is crucial for financial institutions to streamline their lending processes and minimize risk.
- **Task:** Develop a machine learning model to predict whether a loan application will be approved or rejected based on various applicant features.
- **Action:** Cleaned data, engineered features, trained models achieving 98%-99% accuracy with Random Forest and Gradient Boosting, deployed with Flask, and created a user-friendly website interface.
- **Result:** Successfully developed a prediction model that provides accurate loan approval outcomes, enhancing decision-making processes for lenders and improving user experience through a seamless web interface.
- **Explore the Project:** Check out the LinkedIn Post for an overview and view the full code on GitHub.

SQL Challenge at Codebasics

- **Tools & Technologies:** SQL, Excel, PowerPoint
- **Situation:** Participated in a hypothetical scenario at Atliq Hardwares, a leading computer hardware company in India.
- **Task:** Transform raw data into actionable business strategies using SQL queries.
- **Action:** Employed SQL for data extraction and manipulation, used Excel for data cleaning and visualizations, and PowerPoint for presenting findings.
- **Result:** Developed a comprehensive business strategy that showcased the power of SQL analytics in driving informed decisions. The project deepened my understanding of market dynamics and data-driven strategic planning.
- **Explore the Project:** Check out the LinkedIn Post for an overview and view the full code on GitHub.

Grocery Purchase Patterns Analysis

- **Tools & Technologies:** Python, Pandas, Matplotlib, Seaborn
- **Situation:** Conducted an exploratory data analysis on a grocery purchase dataset from Kaggle.
- **Task:** Analyzed purchasing trends to provide actionable insights for the grocery retail sector.
- **Action:** Used Pandas for data cleaning and manipulation, and Matplotlib and Seaborn for data visualization. Analyzed seasonal trends, customer segmentation, and top-selling items.
- **Result:** Identified peak purchase periods and top items, enabling targeted promotions and better inventory management. Successfully completed the project within 3-4 hours, demonstrating efficiency in data analysis.
- **Explore the Project:** Check out the LinkedIn Post for an overview and view the full code on Kaggle.

Hotel Booking Demand Analysis

- **Tools & Technologies:** Python, Pandas, Matplotlib, Seaborn
- **Situation:** Conducted an exploratory data analysis on the Hotel Booking Demand dataset from Kaggle.
- **Task:** Investigate booking patterns and customer behavior to inform business decisions.
- **Action:** Cleaned and prepared the data, analyzed lead times, cancellation rates, average daily rates, and booking changes using Pandas, Matplotlib, and Seaborn.
- **Result:** Provided insights such as last-minute booking trends and high cancellation rates among transient customers, aiding in strategic planning for hotel management.
- **Explore the Project:** Check out the LinkedIn Post for an overview and view the full code on Kaggle.

CERTIFICATIONS

Great Learning
Basics of Exploratory Data Analysis

IIT Madras
Programming and Data Science Foundation Level

Cognitive Class
Python 101 for Data Science

IBM
Python For Data Science

GUVI Geek Networks
Python Zero To Hero