RUSHIKESH YADAV

DATA SCIENTIST

📞 +91 7448069149

✉️ rushikeshyadav301220@gmail.com

in linkedin.com/in/rushikeshyadav149/

💻 Portfolio Website ↗

📍 Pune, Maharashtra-411044

EDUCATION

Bachelor of Technology
SKN Sinhgad College of Engineering
2019 - 2022
79.48 %

Diploma
Shivaji Polytechnic, Atpadi
2016 - 2019
66.41 %

SSC
Dighanchi High-School, Dighanchi
2015 - 2016
76 %

TECHNICAL SKILLS

- Python
- MySQL
- Tableau
- Machine Learning
- Statistics
- Numpy
- Pandas
- Scikit-Learn
- Matplotlib
- OOPs
- DSA

SOFT SKILLS

- Communication Skills
- Teamwork
- Critical thinking
- Problem Solving
- Adaptability

PROFILE

- An aspiring Data scientist with 2 months internship experience at Pantech Prolabs India Pvt Ltd.
- Enthusiastic about data science and machine learning, seeking opportunities to apply my skills and knowledge in a professional setting to contribute to the field, further develop my expertise, and drive meaningful insights from data

Project Works

Data Science Project: Flight Fare Prediction ↗

Project Overview:
- Developed a machine learning model to predict flight fares using Python and popular libraries such as Pandas, Matplotlib, Seaborn, and Scikit-Learn.
- Conducted exploratory data analysis (EDA) to understand the dataset and visualize the relationship between different features and flight prices.

Data Cleaning and Preprocessing:
- Performed data inspection and identified missing values.
- Cleaned the dataset by handling missing values, converting data types, and encoding categorical variables.
- Extracted meaningful features from the "Date_of_Journey" column and converted time-related columns for analysis.

Data Visualization:
- Utilized Matplotlib, Seaborn, and Plotly libraries to create insightful visualizations.
- Plotted bar graphs to showcase the relationship between different airlines and their corresponding ticket prices.
- Visualized the maximum flight prices for each airline using various Python data visualization techniques.

Machine Learning Models:
- Implemented three machine learning algorithms: K-Nearest Neighbors, Decision Tree, and Random Forest.
- Split the dataset into training and testing sets and evaluated the model performance using metrics like R-squared, Mean Absolute Error, and Mean Squared Error.

Hyperparameter Tuning:
- Applied RandomizedSearchCV for hyperparameter tuning on the Random Forest model to optimize its performance.

Skills Demonstrate
- Data cleaning and preprocessing techniques.
- Exploratory Data Analysis (EDA) using Python data visualization libraries.
- Implementation and evaluation of machine learning algorithms.
- Hyperparameter tuning for model optimization.
- Effective communication of complex technical concepts.

Course Works

FULL STACK DATA SCIENCE ↗
- Comprehensive training in Python programming, including data manipulation and scripting.
- Familiarity with data visualization tools like Tableau and Power BI to translate complex data into actionable insights.
- Practical experience with database management using Python and MySQL.
- In-depth exploration of Data Science techniques, including implementation of machine learning algorithms such as supervised and un-supervised using libraries like Pandas, Matplotlib, and Scikit-Learn

Google Advanced Data Analytics Professional Certificate ↗
- Completed Google's Advanced Data Analytics Professional Certificate, mastering a 7-course series
- Applying over 200 hours to gain proficiency in Jupyter Notebook, Python, and Tableau
- Developing expertise in data analysis, statistical modeling, machine learning,