Deccan AI Machine Learning Engineer Project (California Housing Dataset)

1. Project Overview

This project involves predictive modeling using Kaggle datasets and consists of two notebooks:

- **Final_Projects_Kaggle.ipynb:** Handles data preprocessing, feature engineering, model selection, optimization, and validation.
- **HandleTestData.ipynb**: Applies the preprocessing steps from training to the test dataset and generates final predictions for submission.

2. Data Preprocessing & Feature Engineering

- **Data Cleaning**: Addressed missing values by imputation and handled outliers to enhance data quality.
- **Feature Engineering**: Generated additional meaningful features using transformations and encoding of categorical variables.
- **Data Splitting**: Created training and validation datasets to reliably evaluate model performance.

3. Model Selection & Optimization

- **Algorithms Evaluated**: Tested multiple models including Random Forest, XGBoost, and Gradient Boosting.
- **Hyperparameter Tuning**: Optimized model parameters using Grid Search for improved accuracy.
- **Validation Method**: Performed cross-validation to ensure robustness and consistent model performance.

4. Model Deployment Strategy

- **Deployment Method**: Recommended deployment as a REST API using Flask or FastAPI frameworks.
- **Containerization**: Suggested Docker container usage for consistent, reproducible deployments.

5. API Usage Guide

- **Endpoint**: `/predict`
- **HTTP Method**: POST (JSON format)
- **Input**: Feature data in JSON
- **Output**: Predicted outcomes in JSON format

6. Installation & Usage

1. **Clone Repository**: git clone https://github.com/ashish21jan/DeccanAI_Assignement_MLEngineer

   Environment Setup:

2. python -m venv venv
3. source venv/bin/activate # Linux/Mac
   .\venv\Scripts\activate   # Windows

4. **Install Dependencies**: pip install -r requirements.txt
5. **Run Notebooks**: Execute Final_Projects_Kaggle.ipynb followed by HandleTestData.ipynb.

7. Repository Structure

```
project/
├── Final_Projects_Kaggle.ipynb
├── HandleTestData.ipynb
├── requirements.txt
└── README.md
```

8. Contact

For any queries or further details:

- **Name: Aashish Ranjan**
- **Email**: abhirajashish120@gmail.com