MACHINE LEARNING PROJECT REPORT

House Price Prediction using ML Algorithms

1. Problem Statement & Objective

Predicting house prices helps buyers and sellers understand the value of properties.

Objective: Build and compare ML models to predict house prices accurately.

2. Dataset Details

Dataset contains 21613 rows and 21 columns with features such as bedrooms, bathrooms, sqft_living, floors, condition, grade, location (lat/long), etc.

3. Algorithms Used and Justification

- Linear Regression: Simple baseline model.

- Random Forest Regressor: Handles non-linear data, more accurate.

- XGBoost Regressor: Boosting-based, high accuracy.

4. Model Architecture

Input Layer $\rightarrow$ Preprocessing $\rightarrow$ Model Training $\rightarrow$ Prediction $\rightarrow$ Evaluation

5. Preprocessing Steps

- Removed unused columns

- Extracted year/month

- Scaled numerical data using StandardScaler

- Train-test split (80–20)

6. Important Code Snippets

Includes imports, preprocessing, training, and evaluation (as discussed).

7. Results & Evaluation

Random Forest and XGBoost performed best.

Used MAE, MSE, RMSE, R^2 Score.

8. Conclusion

The model predicts house prices effectively.

Future work: Hyperparameter tuning, deep learning, and deployment.