Real Estate Price & Investment Prediction — Final Project Report

1. Project Overview

This machine learning project predicts:

• Whether a property is a **Good Investment**.

• The **Estimated Property Price after 5 Years** based on historical and structural features.

The system includes:

• A complete ML pipeline with preprocessing

• Classification and regression models

• MLflow experiment tracking

• A Streamlit deployment-ready application

2. Methodology

• Data Cleaning & Preprocessing

• Feature Engineering

• Exploratory Data Analysis

• Model Training: LightGBM (best), Random Forest, XGBoost, Logistic Regression

• Hyperparameter tuning

• Evaluation using RMSE, R^2, AUC, Precision, Recall

• MLflow Logging & Registry Setup

• Streamlit App Integration

3. Key Findings

• Locality & Amenities strongly influence investment value.

• Size_in_SqFt, Floors, and Age_of_Property drive pricing trends.

• Transport Accessibility improves likelihood of being a Good Investment.

• LightGBM consistently outperformed other algorithms.

4. Use Cases

• Real estate platforms for price forecasting

• Investors evaluating appreciation potential

• Banks assessing loan risk

• Builders performing project feasibility studies

5. Conclusion

The project delivers a robust ML framework capable of real-world deployment.

The final models generalize well, and the Streamlit app enables easy user interaction.