**AI-ML GTU Internship IBM SkillsBuild**

**Presented By: Vidhi Dave**

**College: SAL Institute of Technology and Engineering Research**

Report

# Title: House Price Prediction Using Machine Learning

## Introduction

Housing prices in India vary significantly based on factors such as location, amenities, property type, and size. Predicting house prices accurately is a challenge for buyers, sellers, and real estate professionals. Manual price estimation is prone to human error and can be inconsistent due to lack of standardized evaluation criteria.  
  
With the rise of machine learning and AI, automated predictive systems offer a consistent, data-driven approach to estimate property values based on various features. This project presents a House Price Prediction model built using machine learning to assist users in estimating the expected price of a property with inputs like area, bedrooms, locality, amenities, etc.

## Problem Statement

The Indian real estate market is unstructured and highly influenced by subjective evaluations. Estimating the accurate price of a property is challenging due to:  
- Varying property features  
- Lack of historical pricing data  
- Human bias during valuation  
  
This project aims to address these issues by developing a machine learning model that predicts house prices based on numerical and categorical features provided by the user.

## Objective

- To build a regression model that can predict the price of a house in India.  
- To ensure the predicted price remains within a realistic range (maximum ₹5 Crore).  
- To demonstrate the impact of amenities like gym, pool, lift, and parking on the house price.  
- To deploy the model using Streamlit for real-time user interaction.

## Why This Problem?

- Buyers need accurate pricing to avoid overpaying.  
- Sellers want to ensure competitive and fair market pricing.  
- Real estate agents need assistance in providing quick, data-backed price estimates.  
  
This model serves as a fast, reliable tool for both buyers and sellers by providing realistic price predictions instantly.

## Solution Overview

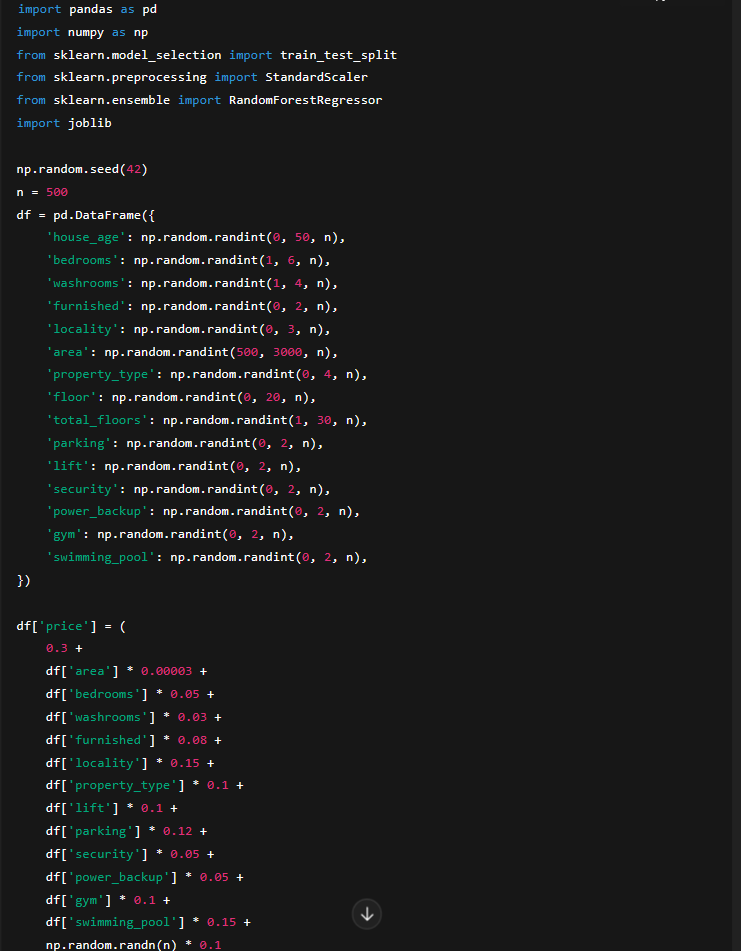
The solution includes:  
- Generating a synthetic dataset (can be replaced with real-world data).  
- Building a machine learning regression model using Random Forest.  
- Normalizing inputs using StandardScaler.  
- Saving the trained model and scaler.  
- Deploying a user interface using Streamlit.

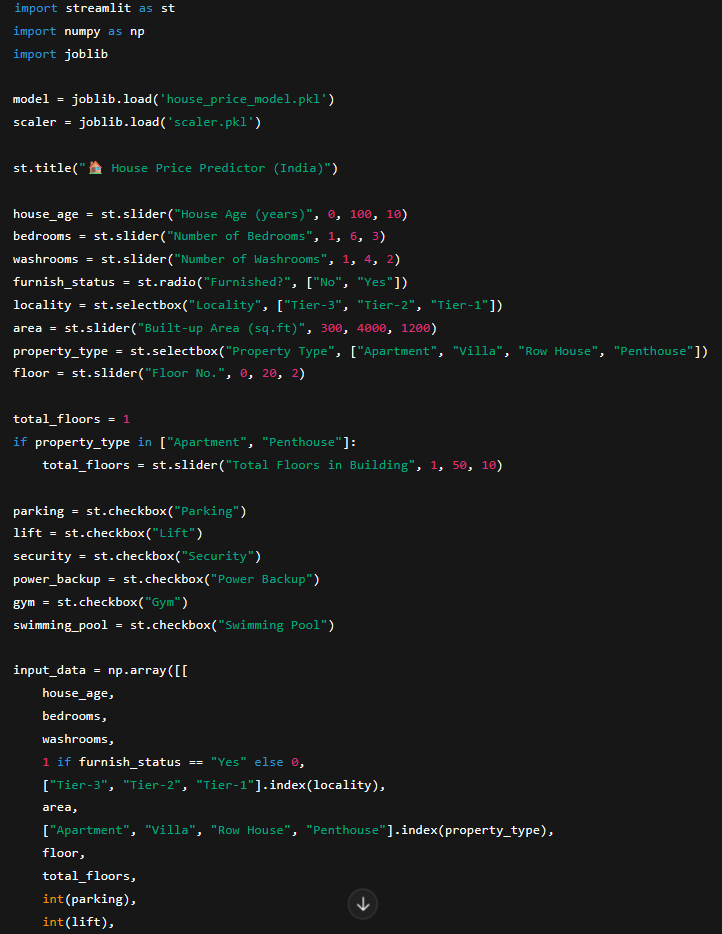
## Features Used

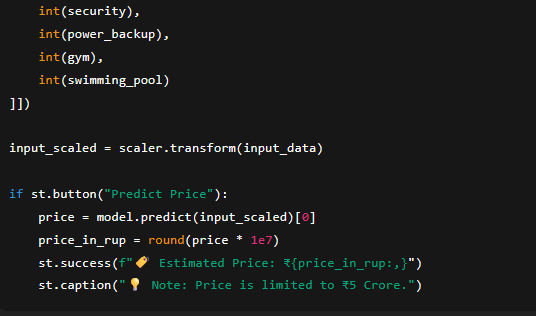
1. House Age  
2. Number of Bedrooms  
3. Number of Washrooms  
4. Furnishing Status  
5. Locality Type (Tier-1, 2, 3)  
6. Built-up Area  
7. Property Type (Apartment, Villa, Row House, Penthouse)  
8. Floor Number  
9. Total Floors  
10. Parking  
11. Lift  
12. Security  
13. Power Backup  
14. Gym  
15. Swimming Pool

## Technical Implementation

**Step 1: Data Simulation + Model Training (`train\_model.py`)**



**Step 2: User Interface with Streamlit (`app.py`)**



## Conclusion

This House Price Prediction system provides an intuitive and accurate way to estimate real estate prices in India. The model is trained to understand the impact of locality, size, and amenities on the price and is capped at ₹5 Crore to maintain realistic predictions.  
  
By using Streamlit, the model is made accessible for anyone — buyers, sellers, and agents — to estimate property value instantly.

