

PROJECT REPORT

Topic: House Price Prediction System

Submitted To:

Kochumol Abraham

Assistant Professor

Submitted By:

Asim Thaha Azeez

22PMC118

ABSTRACT OF HOUSE PRICE PREDICTION SYSTEM

The aim of this project is to develop a machine learning model using Python and Multi-Linear Regression for accurate prediction of house prices. The model will be trained on a comprehensive dataset containing various variables such as area, land slope, water front, alley, street, lot Shape, and several others. Python, with its rich ecosystem of libraries such as NumPy, Pandas, and Scikit-learn, will be the primary language used for implementing the model.

The dataset used for training and evaluation consists of a wide range of features that have the potential to influence house prices. These features include quantitative variables like the area of the property, as well as categorical variables such as land slope (e.g., flat, gentle slope, or steep slope), water front (e.g., whether the property has a water view or not), alley (e.g., presence of an alley access), street (e.g., type of road access to the property), and lot Shape (e.g., regular, irregular, or other shapes).

The project will follow a systematic approach starting with data preprocessing. This phase involves handling missing values, addressing outliers, and transforming variables if needed. Feature engineering techniques will be applied to extract additional meaningful information from the existing dataset, ensuring that the features are suitable for Multi-Linear Regression.

The next step involves model development using Multi-Linear Regression, which is well-suited for predicting house prices based on multiple independent variables. The algorithm will be implemented using Python's Scikit-learn library. The model will be trained on the preprocessed dataset, utilizing techniques such as gradient descent to optimize the model's coefficients and minimize the error between predicted and actual house prices.

Evaluation of the model's performance will be carried out using appropriate metrics such as mean squared error (MSE), root mean squared error (RMSE), and R-squared. These metrics will provide insights into the accuracy and goodness-of-fit of the Multi-Linear Regression model for house price prediction.

Furthermore, the project will include a comprehensive analysis and interpretation of the model's coefficients to understand the impact of each variable on the predicted house prices. This analysis will provide valuable insights into the factors driving property prices and aid in making informed decisions.

To summarize, this project aims to develop a machine learning model using Python and Multi-Linear Regression for accurate house price prediction. The project will leverage a dataset containing variables such as area, land slope, water front, alley, street, lot Shape, and others. The resulting model will provide valuable insights into the housing market, assisting real estate professionals, homeowners, and potential buyers in making informed decisions based on accurate price estimates.