Forecasting house price using smart regression techniques using data science

Student name: S.Anusiya

Register number: 620123106005

Institute : AVS Engineering college

Department : ECE

Date of submission : 30.04.2025

Problem statement

Accurately predicting house prices is crucial for buyers, sellers, investors, and policymakers. Traditional models often fall short due to lack of adaptability and feature selection.

Objective

To build a robust, data-driven regression model that accurately forecasts house prices using advanced machine learning techniques.

Scope of the Project

- - Analyze historical housing data
- - Identify key factors influencing house prices
- - Develop and evaluate multiple regression models
- - Deliver a predictie model with high accuracy and interpretability

Data Source

- - Kaggle's House Prices Dataset

- - Real estate APIs or government property databases

High-Level Methodology
- - Data collection and preprocessing
- - Exploratory Data Analysis (EDA)
- - Feature engineering and selection
- - Model building (Linear Regression, Lasso, Random Forest, XGBoost, etc.)
- - Model evaluation and tuning
- - Deployment (optional)

Tools and Techniques
- - Python (Pandas, NumPy, Scikit-learn, XGBoost)
- - Jupyter Notebook
- - Tableau or Matplotlib/Seaborn for visualization
- - Git for version control

Team Members and Roles

1) J.Ayishabanu Project Manager

J.Ayishabanu, she was responsible for planning and coordinating the entire project workflow. She ensured that deadlines were met, tasks were distributed efficiently, and communication among team members remained clear. She also supervised model selection and guided the team through problem-solving processes.

2) S.Anusiya Data Engineer

S.Anusiya, she handled the data collection and preprocessing phase. She role involved cleaning the raw housing data, managing missing values, encoding categorical features, and preparing the dataset for modeling. She also worked on feature engineering to improve

model performance.

3) M.Dharani Data scientist

M.Dharani, she responsibility was to develop and fine-tune machine learning regression models to predict house prices accurately. She tested different models like Linear Regression, Random Forest, and XGBoost, evaluated their performance using metrics like RMSE and R^2, and selected the best one based on results.

4) M.Kaviya Data Analyst

M.Kaviya, she conducted exploratory data analysis to understand the dataset and uncover key insights. She created visualizations to show trends, correlations, and patterns that informed our feature selection and model development. She work helped the team interpret complex data more intuitively.