**GOVERNMENT COLLEGE OF ENGINEERING ERODE**

****

B.E Electronics and Communication Engineering

PREDICTING HOUSE PRICES USING MACHINE LEARNING

**Name of the Students: University Register no:**

**Team Leader:**

M.Priyadharshini 731121106038

**Team Members:**

M.Sugi 731121106048

K.V.Hema 731121106017

Under the mentor of

**Dr.M.Sathyakala**

**Department of Information Technology(IT)**

**Department of Electronics and Communication Engineering**

Government College of Engineering

Near Vasavi College(PO),Erode,TamilNadu-638316,

Affiliated to Anna University ,Chennai.

**PROJECT DEFINITION:**

Predicting house prices using machine learning involves developing a model or algorithm that uses historical data and various features of a property to estimate its market value or selling price.

**PROJECT OVERVIEW:**

The overall, predicting house prices using machine learning involves a comprehensive data science pipeline, from data collection and preprocessing to model development, evaluation, deployment and ongoing maintenance.

**OBJECTIVE:**

* To develop a machine learning model that accurately predicts house prices based on various features.
* It is a helping tool for real estate professionals and individuals looking to buy or sell properties.

**PROPOSED SYSTEM PHASES:**

Phase 1: Collection of data

The data set contains attributes like location, carpet area, built-up area, age of the property, zip code, price, no of bedrooms etc which is structured and categorized.

Phase 2: Data preprocessing

Data preprocessing is the process of cleaning the data set. There might be missing values or outliers in the dataset. These can be handled by data cleaning. If there are many missing values in a variable, those values can be dropped or replaced by the average value.

Phase 3: Training the model

Since the data is broken down into two sets: a Training set and Testing set, the model must be trained initially. The training set includes the target variable. The decision tree regressor algorithm is applied to the training data set. The Decision tree builds a regression model in the form of a tree structure.

Phase 4: Testing and Integrating with UI

The trained model is applied to test dataset and house prices are predicted. The trained model is then integrated with the front end using Flask in python.

**The dataset contains 13 features**:

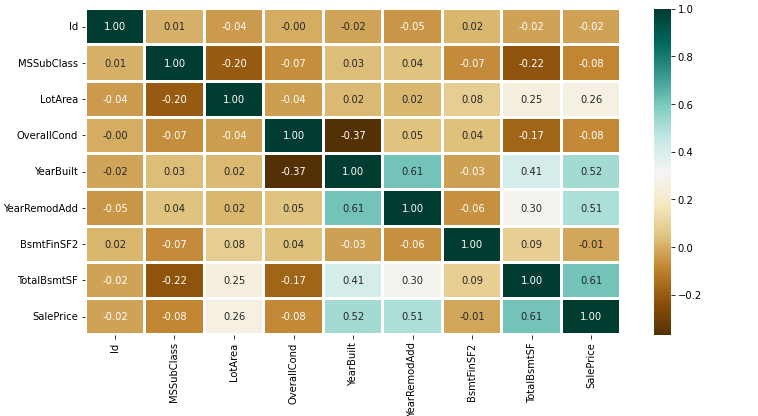
|  |  |  |
| --- | --- | --- |
| 1. | Id | To count the records. |
| 2. | MSS | Identifies the type of dwelling involved in the sale. |
| 3. | MSZoning | Identifies the general zoning classification of the sale. |
| 4. | Lot area | Lot size in square feet. |
| 5. | LotConfig | Configuration of the lot. |
| 6. | BldgType | Type of dwelling. |
| 7. | OverallCond | Rates the overall condition of the house. |
| 8. | YearBuilt | Original construction year. |
| 9. | YearRemodAdd | Construction Date |
| 10. | Exterior1st | Exterior covering on house. |
| 11. | BsmtFinSF2 | Type 2 finished square feet. |
| 12. | TotalBsmtSF | Total square feet of basement area. |
| 13. | SalesPrice | To be predicted. |

**IMPORTING LIBRARIES:**

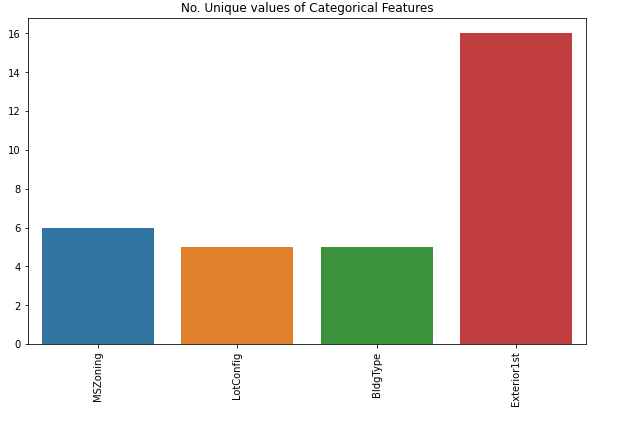
Libraries used are ,

* [**Pandas**](https://www.geeksforgeeks.org/python-pandas-dataframe/)**–** To load the Dataframe
* [**Matplotlib**](https://www.geeksforgeeks.org/matplotlib-tutorial/)**–** To visualize the data features i.e. barplot
* [**Seaborn**](https://www.geeksforgeeks.org/introduction-to-seaborn-python/)**–** To see the correlation between features using heatmap

**HEATMAP:**



**BARPLOT:**



**CONCLUSION:**

This project has shown that machine learning models can be an effective tools for making accurate house price predictions.