HOUSE PRICE PREDICTION

Basic Details of the Team and Problem Statement

Project Name :   **House Price Prediction Project**

Institute Name: ASANSOL ENGINEERING

   COLLEGE

Institute Code: 108

Theme Name :   Predict the house rent

Team Member Details

Team Member 1 Name : SHRUTI KIRAN

Branch : Btech   **Stream :** IT(6TH SEM)

Team Member 2 Name: PRIYA JHA

Branch : Btech   **Stream :** IT(6TH SEM)

Team Member 3 Name: ANKITA

Branch : Btech   **Stream :** IT(6TH SEM)

Team Member 4 Name: ISHIKA PRABHAKAR

Branch : Btech   **Stream :** IT(6TH SEM)

- **Problem Statement**:

  We are given dataset of house price with some feature like number of bedroom,crime rate in area,etc. our task is to create a model which will predict the price for any new house by looking at the features.

  Prediction of property prices is becoming increasingly important and beneficial. Property prices are a good indicator of both the overall market condition and the economic health of a country. The buyers are just not concerned only about the size(square feet) of the house but there are various other factors that play a key role to decide the price of a house/property. Considering the data provided, we are wrangling a large set of property sales records with unknown data quality issues.

Introduction:

Problems faced during buying a house. Buying a house is a stressful thing. Buyers are generally not aware of factors that influence the house prices. Many problems are faced during buying a house. Hence real estate agents are trusted with the communication between buyers and sellers as well as laying down a legal contract for the transfer. This just creates a middle man and increases the cost of house.

Technology Used

Machine Learning →In general, any machine learning problem can be assigned to one d classifications: Supervised learning and Unsupervised learning.

1) Supervised Learning: In supervised learning, we are given a dat already know what our correct output should look like, having the that there is a relationship between the input and the output. Supervised learning problems are categorized into "regression" and "classification problems.

2) 2) In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function.In a classification problem, we are instead trying to predict ht discrete output. In other words, we are trying to map input va discrete categories.

3) Unsupervised Learning:-Unsupervised learning allows us to approach problems with little or no idea what our results should look like. We ca derive structure from data where we don't necessarily know the effect of the variables. We can derive this structure by clustering the data based an relationships among the variables in the data. With unsupervised learo there is no feedback based on the prediction results.

Idea/Approach Details

Prototype

In recent times, finding the ideal housing option according to budget and preference is such a hussle. The cost of house rent depends on many factors such as; the house size,number of bedrooms ,locality,number of bathrooms,halls,and kitchen,furnishing status, and a lot more.
With the use of appropriate machine learning algorithms,real estate owners can find the ideal house according to customers budgets and preferences with ease.

Describe Technology stack here:
➢ Python
➢ Numpy
➢ Pandas
➢ seaborn
➢ Matplotlib
➢ Sklearn
➢ t-Distributed Stochastic Neighbor Embedding (t-SNE) (for graph)
➢ Nltk (Stopwords)
➢ Xgbclassifier in Xgboost

• **Architecture Diagram**
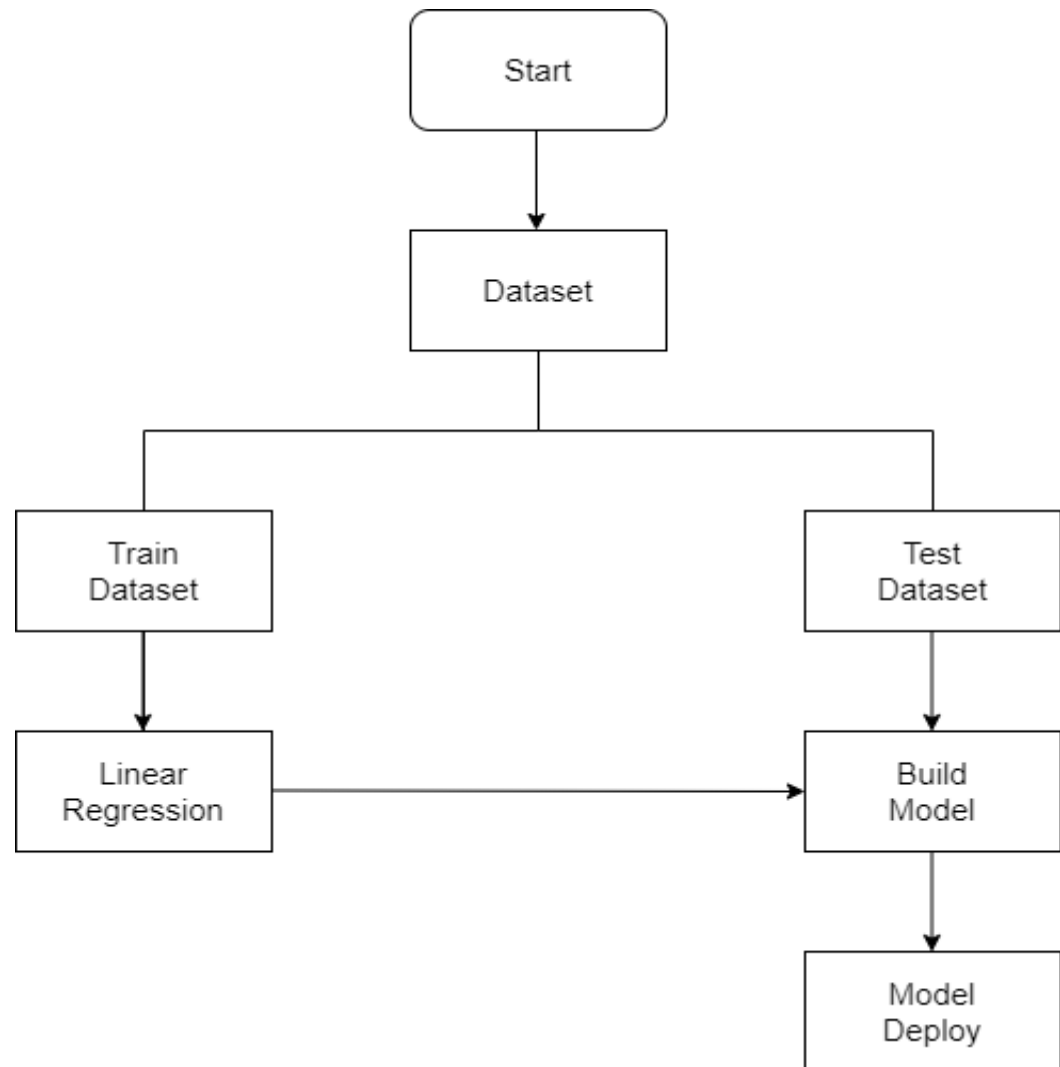
Use Cases

1. Implemented a machine learning model capable of predicting the house rent for homes.

2. Determine the essential features significantly needed to predict the house.

The basic procedures implemented to achieve the goals of this research are:

1. Data collection
2. Data cleaning and exploration
3. Feature encoding
4. Train-test split validation
5. Feature scaling
6. Modeling
7. Model evaluation

HOW IT WORKS?

➢ COLLECTING DATA: FIRST STEP WAS TO COLLECT DATA WE COLLECTED DATA FROM DIFFERENT SOURCES & MERGED THEM TOGETHER TO FORM OUR TRAINING DATA SET.

➢ THEN WE TRAINED THE MODEL USING MACHINE LEARNIG ALGORITHM WHICH IN THIS CASE IS MULTIPLE LINEAR REGRESSION.

➢ BASED ON THE GENERATED GRAPHS WE PREDICT THE COST OF THE HOUSE

Future Scope

Our model had a low root mean square residual correlations score, but there is still room for improvement. In a real world scenario, we can use such a model to predict house prices. This model should check for new data, once in a month, and incorporate them to expand the dataset and produce better results.

We can try out other dimensionality reduction techniques like Univariate Feature Selection and Recursive feature elimination in the initial stages.

We can try out other advanced regression techniques, like Random Forest and Bayesian Ridge Algorithm, for prediction. Since the data is highly correlated, we should also try Elastic Net regression technique.

The aim is to predict the efficient house pricing for real estate customers with respect to their budgets and priorities. By analyzing previous market trends and price ranges, and also upcoming developments future prices will be predicted.

CONCLUSION

Thus, we studied and applied the concept of Linear Regression in real time implementation so as to ease the life of human.

Determining the price of property without complete knowledge about the surrounding is quite riskier for both customer and the seller. In order to overcome this problem we have tried to develop application which determines the price of the property based on various parameters of the surrounding.

Data provide us with the complete data about the surrounding in the form of dataset. Dataset helps to get the insight of the surrounding and machine learning model helps to predict the price of the property based on the training provided by the dataset. We successfully implemented linear regression model to predict the price of the houses.

THANK YOU…..