**Midterm Project**

**Team** - Technosapiens

**Members -**

Soumyendra Shrivastava (016670121)

Tanuja Reddy Maligireddy (016715348)

Siddhant Sancheti (016710421)

# **Objective**

Building a model that can precisely forecast house prices based on a collection of input data is the goal of a house price prediction project. This can assist developers, investors, and real estate agents in making educated decisions on the value of real estate, as well as house buyers and sellers.

**Goals**

1. The most typical application of house price prediction is to determine a home's current market value. It can be helpful for both homeowners who want to estimate the value of their home and prospective buyers who want to assess whether a home is priced reasonably.
2. Assessing the possible return on investment (ROI) of a real estate investment: Investors can use home price forecasts to assess the possible ROI of a property and decide whether to buy or sell in accordance with the results.
3. Increasing the accuracy of real estate evaluations: Real estate appraisals frequently depend on judgments about how much a property is worth. By supplying unbiased data, house price prediction algorithms can increase the accuracy of appraisals.
4. Finding locations with the greatest potential for price growth: By forecasting house values in various communities or cities, developers, investors, and governmental bodies can choose locations with the greatest potential for price growth and focus their investments there.
5. Real estate market efficiency can be increased by minimizing the time and effort needed for buyers and sellers to come to an agreement on a price. Accurate house price predictions can do this.
6. A house price prediction project's overall objective is to offer knowledge and data that can assist people and businesses in making wise real estate investment and transaction decisions.

**Metrics**

1. Mean Absolute Error (MAE)
2. Mean Squared Error (MSE)
3. Root Mean Squared Error (RMSE)
4. R-squared (R2)
5. Coefficient of Determination (R2)
6. Mean Absolute Percentage Error (MAPE)
7. Receiver Operating Characteristic (ROC) curve
8. Confusion Matrix

**Data**

1. <https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx>
2. <https://www.zillow.com/research/data/>

**Applications**

1. Real estate investment: Real estate investors may choose the right kinds of properties to invest in and when to buy or sell properties by accurately anticipating future house prices.
2. Home Buyers and Sellers: Home buyers and sellers can utilize house price prediction to better understand the state of the market and decide whether to buy or sell a home.
3. Lenders: Home price forecasting can assist lenders in making wise lending and risk management decisions.
4. Real estate agents: By offering market insights, pricing comparisons, and other pertinent data, real estate agents can use house price prediction to assist their clients in buying or selling a home.
5. Government: In order to monitor and control the housing market, evaluate home affordability, and make educated judgments about housing laws and initiatives, government organizations can employ house price prediction.
6. Academic Research: To better understand the dynamics of the housing market, analyze the variables that affect housing prices, and develop new models and techniques for predicting home prices, employ house price prediction in academic research.

**Business Logic**

In order to increase the efficacy and efficiency of real estate transactions and assist informed decision-making in the housing market, the business logic behind home price prediction is founded on the notion that reliable and timely information on future house prices may be provided. All parties involved may benefit from this, and the housing market may expand and develop as a result.

**Models**

* Linear Regression
* Decision Trees
* Random Forest
* Gradient Boosting
* Neural Networks
* Time Series Models

# **Data Narrative**

Our Datasets contains the following information -

1. **FHFA Dataset**

The Federal Housing Finance Agency (FHFA) serves as a source of data and research on the U.S. housing finance system, and provides information and analysis on topics such as housing prices, mortgage rates, and market trends. This information can be used by a variety of stakeholders, including policymakers, industry professionals, and consumers, to make informed decisions about the housing market and the future of mortgage finance in the United States.

A House Price Index (HPI) is a statistical measure that tracks changes in the prices of residential properties over time.It can be used to measure changes in the overall market, as well as changes in specific regions, cities, or neighborhoods. The index is calculated based on data collected on sales of individual homes, and it typically reflects the change in prices for a representative basket of properties in the market.

In a house price prediction project, the HPI can be used as a benchmark for evaluating the performance of the prediction model.

1. **Zillow House Price Dataset**

The Zillow House Price dataset includes a substantial amount of data on a variety of housing market variables, such as median home values, median rent costs, and the availability of homes for sale in a certain area. The information is gathered at the zip code, city, and state levels and includes a sizable number of US urban regions.

The Zillow House Price dataset can be utilized as a source of training data for a machine learning model in a project to predict home prices. The model may be trained using past data to discover the connections between different variables and house prices, after which it can be applied to forecast future prices.

**Fractal Clustering**

We performed fractal clustering on the following objective functions:

**Maximize ROI rate** : Performing k-means fractal clustering on the monthly percentage change of property prices and identifying the optimal cluster with the highest rate of increase is a data-driven approach to maximizing ROI rate. This approach enables businesses to group similar data points and identify patterns and trends in property prices, which can inform decision-making. By iterating through the k-means fractal clustering process and identifying the optimal cluster, businesses can make informed investment decisions that have the highest potential for ROI.

**Finding greatest potential locations**: Fractal clustering is used to group similar neighborhoods together based on their spatial properties and other features, such as locality, access to amenities etc. Then, a Gaussian mixture model is used to model the distribution of house prices within each cluster, based on factors such as the size of the house, the number of rooms, and the age of the property.

# **Articles for Reference**

* <https://towardsdatascience.com/house-price-prediction-with-zillow-economics-dataset-18709abff896>
* <https://towardsdatascience.com/predicting-house-prices-with-linear-regression-machine-learning-from-scratch-part-ii-47a0238aeac1>