**House price prediction**

**Project group 3 Professor:**Dr.Ramakrishna Koganti

Akhil Ade - 1002074926 **Course Number**: CSE-5301

Nikhita Allada - 1002057382 **Date:** 28th Nov 2022

Snehith Gandla – 1002078488

Sai Shiva Deepthi Gopireddy – 1002067864

Tejaswi Gullapalli - 1002087302

**Introduction**

In this ever-growing real estate market, it is essential for the stakeholders and the investors to understand the driving factors of the industry better. For this reason, there is a need of a house prediction model with good accuracy. A correct prediction of the price of the house beforehand can earn huge profits.

House price prediction also plays a vital role in financial aspect of an individual’s growth as it analyzes the different features of a house and its surroundings to generate the most accurate price. It can help people who are planning to buy a house, so they can know the price range in the future and plan their finances well accordingly.

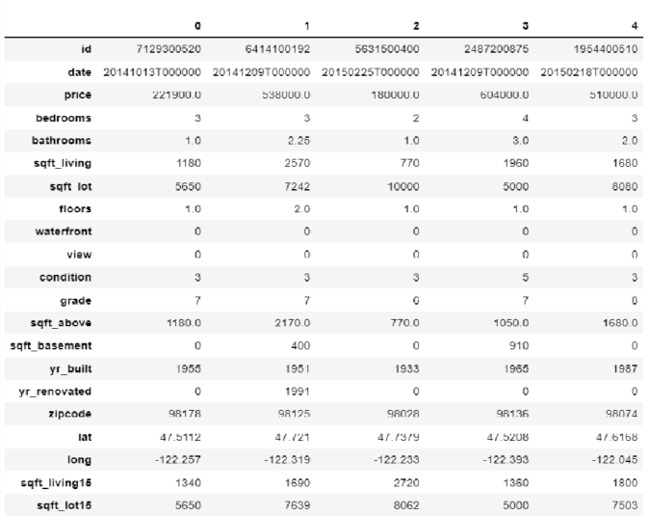
The goal of this project is to take the data of the house prices and consider various factors like size, number of rooms, view etc that majorly impact the price and use one of the various predictive models to predict the price of the house. The outcome from this project will be advantageous for investors, stakeholders and common people planning to buy a house to make an informed decision before purchasing any house.

**Problem Definition**

The proverb "buy low, sell high" is well known to investors, yet it lacks sufficient context to guide wise investing choices. Consider yourself a real estate professional in charge of advertising a just built house. We want to estimate the price since we don’t know the precise amount by comparing it to similar residences. So it is important to have a prediction model that can predict the price of the houses accurately considering various factors like size, view, waterfront, number of bedrooms using machine learning algorithms.

**Data Selection**

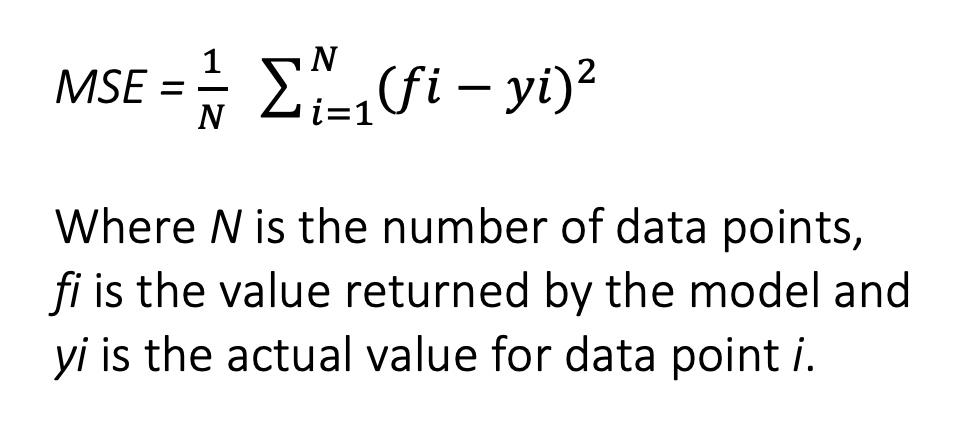
We have picked our dataset from Kaggle. Dataset contains columns like size, bedrooms, bathrooms, condition, view. We have dropped certain characteristics like date, id, zip code which do not affect the price of the house. We have cleaned the data for any null values.



**Predictive Model Selection**

**Model Selected: Random Forest**

The Random Forest model is based on the idea that several uncorrelated models work significantly better together than they do separately. We have chosen random forest and we will do the model fitting next. The fact that there is little to no correlation between the modeling, between the decision trees that make up the larger Random Forest model.



**Predictive Model Analysis**

Perfect predictions are shown by the red line. All points would be on the red line if the prediction and actual value were equal. As you can see, there are a few outliers and deviations in this data, but this is primarily true for prices that are quite high. There are a few anomalies in the low range, and it would be intriguing to see what these anomalies' causes are.

**Chart, scatter chart

Description automatically generated**

**Result:**

It is possible to make a decent prediction of property prices with only a Random Forest. The best accuracy when compared to other models is provided by random forest, which is based on three strategies.

Table

Description automatically generated

**Conclusions**

Since home prices increase annually, a system to predict future home prices was needed. The estimation of a home's value and an acceptable sale price can be done via house price prediction by landowners, estate valuers, and policymakers. This can help prospective buyers choose the ideal period of time for a house purchase. Although the three primary elements that affect a house's price are its location, style, and physical characteristics, the specific elements that affect a house's price differ. A flawless prediction model must therefore take into account the specific factors that affect the cost of a home in the area under consideration. This study has confirmed the ability of the random forest machine learning technique to forecast the values of a house using kc\_house data set.

**Recommendations**

**Appendix**

**Python code:**

**Graphical user interface, text, application

Description automatically generated**

**Text

Description automatically generated with medium confidence**

**Graphical user interface, text

Description automatically generated**

**Graphical user interface, text

Description automatically generated**

**Chart

Description automatically generated with low confidence**

**Text

Description automatically generated with medium confidence**

**Graphical user interface, text

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated**

**Text

Description automatically generated**

**Text

Description automatically generated**

**Additional pictures:**

1. Below screenshot shows graph of the columns we have in our data set.

Chart

Description automatically generated

Chart, histogram

Description automatically generated

Chart, scatter chart

Description automatically generated