**House price prediction**

**Preprocessing and feature engineering:**

* First, I checked the data if it has null values or not
* Then I checked if there is duplication in data
* After that I started to see if there are outliers in the data
* Then I print the correlation matrix to know which will be the important features

A screenshot of a computer generated graph

Description automatically generated

* From this picture I discovered that the Square Feet is the most important feature, and the Location rating is not affecting the output price, so I dropped it
* I also try to create new feature (price per square feet) it gave me excellent loss, but I forgot that I must predict the price, and the price square is calculated using formula of price / square feet, so I removed it after that
* When I trained the model, the loss was very large because of the range of prices so I use log transformation to reduce the range of it
* Then I split the data into train and test, validation
* After that I transform the data into Numpy arrays as expected type for neural networks
* Then I start to build my model with linear activation function in the output layer as it is regression task
* I used Adam optimizer with loss function of Mean squared error
* After training the model I saved the model and the scaler to use it in the Api
* Then I make flask Api and use the saved model to predict the price
* I then applied the exponential function to the output price to convert it from its logarithmic form
* When you are testing the Api, you will pass input of the three features as provided in this pictureA screenshot of a computer

  Description automatically generated
* I made Docker file for the task