International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

HOUSE PRICE PREDICTION USING MACHINE LEARNING

Sonal Naidu, Sagar Kendre, Megha Dumbre, Ankit Pandey
Prof.S.R.Dhotre (Guide)
DEPT. OF INFORMATION TECHNOLOGY
JAYAWNATRAO SAWANT COLLEGE OF ENGINEERING, PUNE
INDIA

Abstract—

parcels is told by multiple factors similar as position, square footage, amenities, and structure. The proposed model aims to prognosticate house prices with a high degree of delicacy, taking into account all applicable factors that impact real estate prices.

The model is grounded on machine literacy algorithms that learn from literal data and use it to give new perceptivity and prognostications. This approach has multitudinous profitable operations, including in banking, business,e-commerce, healthcare, and entertainment diligence. In metropolises like Bangalore, where real estate prices are largely dynamic and told by a range of factors, the proposed model can be particularly useful.

To prognosticate house prices, the proposed model uses colorful retrogression ways, enforced using the Python programming language. The model takes into account sophisticated aspects of house price computation, including position, land size, distance from amenities, and other applicable factors. By using machine literacy, the model can give a more accurate cast of house prices, helping homeowners and buyers make informed opinions.

Real estate is known to be one of the least transparent diligence in our ecosystem, with casing prices changing constantly and frequently grounded on hype rather than factual valuation. This makes

prognosticating casing prices a complex and grueling task. still, the proposed model uses retrogression ways and machine literacy to give further robust and accurate prognostications of real estate prices. To epitomize, the proposed model for prognosticating house prices grounded on machine literacy algorithms and retrogression ways is a precious tool for homeowners, buyers, and real estate professionals. The model takes into account multiple factors that impact real estate prices, furnishing a more accurate cast of house prices. With this tool, people can make further informed opinions when buying or dealing a home, and the real estate assiduity can come more transparent and effective.

• INTRODUCTION

Every single association in moment's real estate business is operating fruitfully to achieve a competitive edge over indispensable challengers. There's a need to simplify the process for a normal human being while furnishing the stylish results. This paper proposes a system that predicts house prices using a retrogression machine literacy algorithm. In case you are going to vend a house, you have to fete what sticker price to put on it. What is further, a PC computation can give you a precise hand!. This retrogression model is erected not only for prognosticating the price of the house which is ready for trade but also for houses that are under construction.

Retrogression is a machine literacy outfit that encourages you to make prospects by taking in – from the current measurable information – the connections between your target parameter and a lot of different independent parameters. As per this description, a house's cost relies upon parameters, for illustration, the number of apartments, living region, area, and so forth. On the off chance that we apply fake figuring out how to these parameters, we can cipher house valuations in a given land region. The target point in this proposed model is the price of the real estate property and the independent features are no. of bedrooms,no.of bathrooms, carpet area, erected- up area, the bottom, age of the property, zip law, latitude and longitude of the property. Other than those of the mentioned features, which are generally needed for prognosticating the house prices, we've included two other features- air quality and crime rate. These features give a precious donation towards prognosticating property prices since the advanced values of these features will lead to a reduction in house prices.

The whole perpetration is done using the python programming language. For the construction of the prophetic model, a Decision tree regressor is used from the " Scikit- learn " machine literacy library. Grid Search CV helps to find the stylish maximum- depth value for constructing the decision tree. After the trained model is ready, it's integrated with the stoner interface using Beaker(a python frame).

• SYSTEM DESIGN AND ARCHITECTURE

Phase 1: Collection of data

Data processing ways and processes are multitudinous. We collected data for Mumbai's real estate parcels from colorful real estate websites. The data would be having attributes similar as position, carpet area, erected- up area, age of the property, zip law, etc. We must collect the quantitative data which is structured and distributed. Data collection is demanded before any kind of machine literacy exploration is carried out. Dataset validity is a must else there's no point in assaying the data.

Phase 2: Data preprocessing

Data preprocessing is the process of drawing our data set. There might be missing values or outliers in the dataset. These can be handled by data cleaning. If there are numerous missing values in a variable we will drop those values or substitute it with the average value.

Phase 3: Training the model

Since the data is broken down into two modules a Training set and Test set, we must originally train the model. The training set includes the target variable. The decision tree regressor algorithm is applied to the training data set. The Decision tree builds a retrogression 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 prognosticated. The trained model is also integrated with the frontal end using Beaker in python.

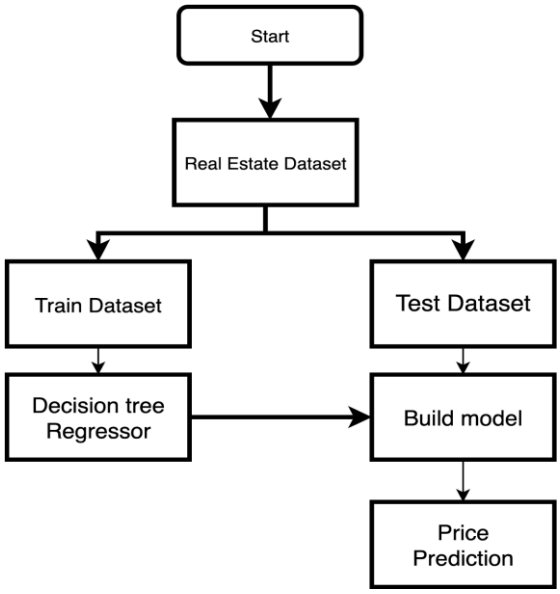


Fig 1. The generic flow of development

- **METHODOLOGY**

- *STUDIED ALGORITHMS:*

In the process of developing this model, colorful retrogression algorithms were studied. SVM, Random Forest, Linear retrogression, Multiple direct retrogression, Decision Tree Regressor, KNN, all were tested upon the training dataset. still, the decision tree regressor handed the loftiest delicacy in terms of prognosticating the house prices. The decision to choose the algorithm largely depends upon the confines and type of data in the data used. The decision tree algorithm suited stylish for our dataset.

- *DECISION TREE REGRESSOR:*

The decision tree regressor observes features of an trait and trains a model in the form of a tree to prognosticate data in the future to produce meaningful affair. Decision tree regressor learns from the maximum depth, min depth of a graph and according to system analyzes the data. Grid Search CV is a way to deal with parameter tuning that will efficiently manufacture and assess a model

for every blend of computation parameters indicated in a grid. Grid Search CV in this algorithm is used to assess the stylish value for maximum- depth, using which the decision tree is constructed.

- *FLASK INTEGRATION*

After erecting the model and successfully giving the result, the coming step is to do the integration with the UI, for this purpose beaker is used. Beaker is a web frame. This means beaker provides you with tools, libraries, and technologies that allow you to make a web operation. Beaker is easy to put away routes together and this frame is substantially used for integrating python models.

- **IMPLEMENTATION**

- *Data preprocessing:*

Age and bottom parameters were handled for their missing values. the target trait is also dropped off from the training dataset. Pandas library is used for this purpose. For statistical visualization of the dataset, the min, maximum, standard divagation, mean of the target trait were set up out. We resolve the dataset into a training set(80%) and a test set(20%).

Max-depth:

As mentioned earlier grid hunt cv helps to find maximum depth for the tree. We've used Matplotlib to fantasize the different maximum- depths and complexity performance.

Following are the visualizations:

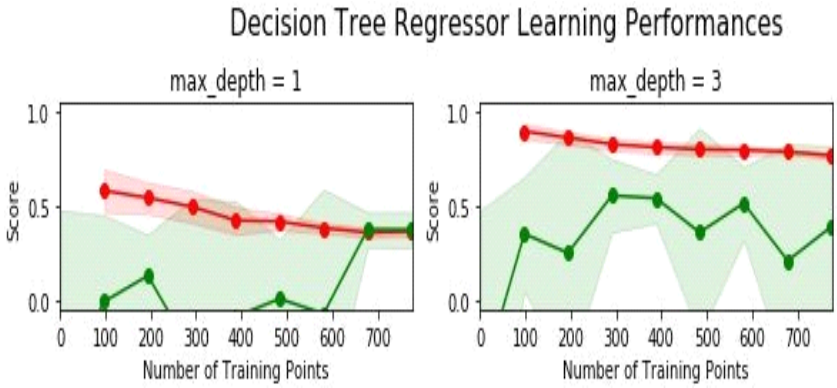


Fig 2.Testing max-depth values(1)

(On axis: Number of training pts. vs Score)

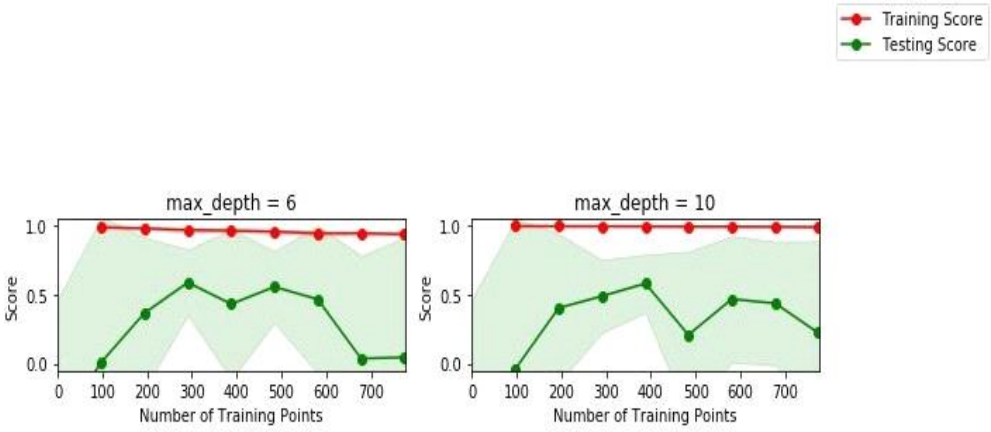


Fig 3.Testing max-depth values(2)

(On axis: Number of training pts. vs Score)

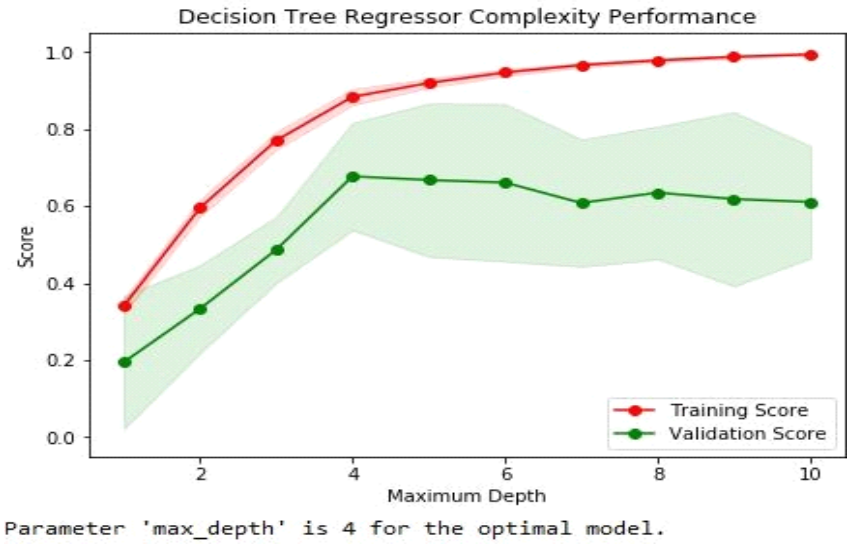
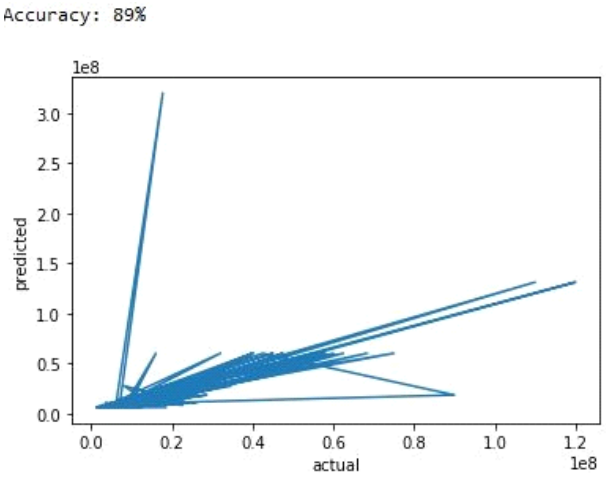


Fig 4. Max-depth value for optimal model

- *Fitting the model:*

The model is trained using a Decision tree regressor from the Scikit-learn toolkit.. The prognosticate function is used to prognosticate the test set results.

- **RESULTS**



The following shows the plot of predicted vs actual priceswith the accuracy of prediction:

Fig 5. Actual vs predicted price graph based on thedataset

Accuracy is nothing but the r2 score of the regression model.

• FUTURE SCOP

House Price Prediction model helps buyer to choose the house grounded on their fiscal status and also saves the time of searching the house. This helps the dealer to know more value of their property and avoids thebrokers.We're adding carriers and packers in the website so that if a person who buy the property can get installation of shifting individualluggage.If an existent who agreed with the price can directly communicate the separate property proprietor and purchases it and if and existent who wants to vend their property can put an announce on the web gate.

• CONCLUSION

In this paper, the Decision tree machine literacy algorithm is used to construct a vaticination model to prognosticate implicit selling prices for any real estate property. fresh features like air quality and crime rate were included in the dataset to help prognosticate the prices indeed more. These features aren't substantially included in the datasets of other vaticination systems, which makes this system different. These features impact people's decision while copping a property, so why not include it in prognosticating house prices. The trained model is integrated with the stoner Interface using the Flask Framework. The system provides 89 delicacy while prognosticating the prices for the real estate prices.