

# Implementation of House Price Prediction Model Using Image Processing and Machine Learning

Gaikwad Purva Chandrakant<sup>1</sup>, Ganjave Pratiksha Namdev<sup>2</sup>, Gorade Pooja Subhash<sup>3</sup>, S. S. Gore<sup>4</sup>

<sup>1,2,3</sup>Student, Department of Computer Engineering, Jaihind College of Engineering, Pune, India

<sup>4</sup>Professor, Department of Computer Engineering, Jaihind College of Engineering, Pune, India

**Abstract:** While purchasing the house, the price of house is the main factor which is considered by people. The pricing of house not only depends on the size of the property and no. of rooms, but also on the neighborhoods like transport facility, banks, schools or colleges, shops etc. When a person buys a home, they consider structural features, working accessibility, neighborhood services. Hence, a house price prediction system is invented to improve estimation of house prices. This system presents a House Price Prediction using Image Processing and Machine Learning. The system would give comparison of prices of house at particular location for customers. It would also give comparative pricing rates to builder so that he can estimate his construction budget to compete with other builders at that area. The satellite images have been used to visualize impression of neighborhood. The Image Processing will have applied to satellite images and Machine Learning Algorithm Convolutional Neural Network (CNN) and Linear Regression is used for estimation of house pricing. The project is purpose to predict price of houses at particular area to people and builders

**Keywords:** Convolutional Neural Network, Feature Extraction, House price prediction, Image Processing, Linear Regression, Machine Learning, Preprocessing.

## 1. Introduction

The fundamental need of shelter is can be fulfill by housing. Housing will also be a part of investment. As civilization is developing day by day the need of housing is increasing. For buyers, sellers and bankers the accurate price prediction is always greed. There are many influences while deciding the price of house. Many researchers had been proposed their work for predicting the house price accurately. There are many machine learning regression algorithms to use. Machine Learning and neural networks were used by many researchers for price prediction model. In our system, we are going to use machine learning with Image processing. Our results are not determined by one technique solely but it is combination of image processing and machine learning.

For our research work, we select Pune is as primary location for predicting the house prices. The nearby infrastructure facilities like school, colleges, banks, hospitals, company's etc. will be considered while predicting the prices of houses. This

system can be used by common people easily. People don't want to go anywhere for knowing the prices of house. They can get information about house prices in any area from their home also. In construction estimation plan, builder can get comparative prices of house offered by different builders at particular area. So, builder can easily have made their plans that will affordable to customers.

### A. Motivation

The family purchase the house in city. They invest whole profit into house they don't have idea about that price of house at the certain area. So, sometimes they give extra money to builder for same construction and some distance. But they don't know about the two different prices of the model. In some cases, builder can't predict the price of the current area where his construction is working so that he can loss his business. So, proposed system is developing for price prediction.

### B. Objectives

- To predict the house price according to the area.
- To calculate house price depending upon surrounding environment like railway station, hospital area, ATM, college, banks so customer can purchase flat with full facilities.
- To suggest builder price prediction for the new constructions.
- To provide comparison of house pricing to customers.

This paper consists of V section, section I represents introduction of proposed system and techniques used. Reviews related previous work in part II. Section III consists of the system that we proposed to predict house prices. Section IV is about conclusion and the last section V consists of the future work where, with reference to this system new modules can be formed.

## 2. Literature analysis

Literature survey is the most important step in any kind of research. Before start developing we need to study the previous papers of our domain which we are working and on the basis of study we can predict or generate the drawback and start working

with the reference of previous papers. In this section, we briefly review the related work on house price prediction and the techniques used.

In this paper, Machine learning techniques are applied to analyze historical property transactions in Australia to discover useful models for house buyers and sellers. There is the high inconsistency between house prices in the most expensive and most affordable suburbs in the city of Melbourne is showed in this paper. Moreover, experiments demonstrate that the combination of Stepwise and Support Vector Machine that is based on mean squared error measurement is a competitive approach. This paper seeks useful models for house price prediction. It also provides insights into the Melbourne Housing Market. Firstly, the original data is prepared and transformed into a cleaned dataset ready for analysis. Data reduction and transformation are then applied by using Stepwise and PCA techniques. Different methods are then implemented and evaluated to achieve an optimal solution. The evaluation phase indicates that the combination of Step-wise and SVM model is a competitive approach. Therefore, it could be used for further deployment [1].

With limited dataset and data features, a practical and composite data preprocessing, creative feature engineering method is examined in this paper. To predict individual house price hybrid Lasso and Gradient boosting regression model were used in this paper. Regression algorithms with parameter more than 10000 iterations are applied by considering the homoscedasticity verification. But the result is determined by the homoscedasticity between training data and test data. Linearity of each feature is the statistic fundamental of regression algorithm, therefore, many transformations are applied to enhance the linearity of input features [2].

This paper considers the issue of changing house price as a classification problem and applies machine learning techniques to predict whether house prices will rise or fall. This work applies various feature selection techniques such as variance influence factor, Information value, principle component analysis and data transformation techniques such as outlier and missing value treatment as well as box-cox transformation techniques. Random Forest provides more accuracy however at the same time this particular type of classifier also prone to over fitting therefore the performance of Support Vector Machine classifier can have said to be reliable and consistent over the rest of the two classifiers [3].

To predict Singapore housing market and to compare the predictive performance of artificial neural network (ANN) model, i.e., the multilayer perceptron, with auto regressive integrated moving average (ARIMA) model, this paper proposes two algorithms. To predict the future condominium price index (CPI) the more superior model is used. The lower mean square error (MSE) of the ANN models showed the superiority of ANN over other predictive tools. The forecasts were based on time series data of variables that are believed to influence the condominium prices in Singapore. These

variables and the CPI were the inputs and output to the models, respectively [4].

In this paper, the study of real estate value is felt critical to help the choices in urban arranging. The land framework is a precarious stochastic process. Financial specialist's choices depend on available patterns to procure most extreme returns. Designers are intrigued in knowing the future patterns for their basic leadership. To precisely gauge real estate costs what's more, future patterns, vast measure of information that impacts arrive cost is required for examination, demonstrating and determining [5].

This paper presents how to make optimal use of Linear Regression, Forest regression, Boosted regression. The efficiency of the algorithm has been further increased with use of Neural networks. The system will satisfy customers by providing accurate output and preventing the risk of investing in the wrong house. Additional features for the customer's benefit can also be added to the system without disturbing its core functionality [6].

In this paper, there is discussion on which feature extraction technique is best by considering various types of features and feature extraction techniques. In case of character recognition application, we are going to refer features and feature extraction methods in this paper. Using this paper, a quick idea of feature extraction techniques may be got and it can be decided that which feature extraction technique will be better for the work to be done based on type of image [7].

The variety of approaches either consider the entire data for modeling, or split the entire house data and model each partition independently for the choice of prediction model. There were may not sufficient training samples per partition for the latter approach. So, such modeling ignores the relatedness between partitions, and for all prediction scenarios. In this paper, this problem is exploited by conducting a careful study of the Multi-Task Learning (MTL) model. Specifically, the strategies for splitting the entire house data to the ways the tasks are defined in MTL, and each partition obtained is aligned with a task. The specific MTL-based methods with different regularization terms are selected to capture and exploit the relatedness between tasks.

### 3. Proposed system

The architectural design of the system is as follow which contain some important points such as algorithm etc.

#### A. Algorithms used

##### 1) Convolutional Neural Network

For the purpose of image detection or image and video recognition, image classifications, recommender systems, natural language processing, object detection and face recognition etc., convolutional neural networks are widely used. The image classification under CNN takes an input image and classify it under certain categories. The CNN has required little pre-processing as compared to other classification

algorithms. Hence, in our system we are using CNN for detection of signs and text present on the satellite image.

## 2) Linear Regression

The simplest method for prediction is Linear regression. On the available data that we got from processing on image the linear regression algorithm will be applied to get predicted price of house. The equation of the regression equation with one dependent and one independent variable is defined by the formula,

$$a = s + r*b$$

where, a = estimated dependent variable score, s = constant, r = regression coefficient, and b = score on the independent variable.

## 4. System architecture

### A. Image Processing

In order to get an enhanced image or to extract some useful information from it, some operations are required to perform on image. Image processing is the technique to extract useful information from image. In an image processing, input is an image and output may be image or characteristics or features associated with that image. We are using snapshots of satellite images to gain text based information from image of particular area.

### B. Feature Extraction

In order to extract some significant characteristics of the object feature extraction is used. Feature is the function of measuring measurable property of an object. Feature extraction is related to dimensionality reduction. All the features of image can be classified into high-level feature and low-level feature extraction. Feature extraction may include color feature, shape feature and texture feature. Due to powerful parallel mechanism of computation neural networks are mainly used for feature extraction. In our system, we are using convolutional neural network for feature extraction.

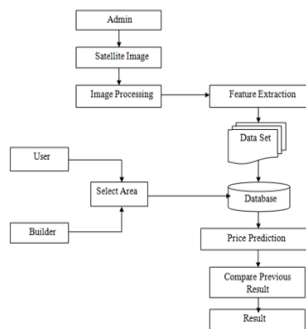


Fig. 1. System Architecture

### C. Price Dataset

In Machine Learning, the training data set is the actual dataset used to train the model for predicting house price.

## 5. Conclusion

The goal is to achieve the system which will reduce the human effort to find a house having reasonable price. The proposed system. House Price Prediction model approximately try to achieve the same one. Proposed system focused on predict the house price according to the area for that image processing and machine learning methods are used. The experimental results showed that this technique that are used while developing system will give accurate prediction of house price.

## 6. Future work

The usability of system will be further increased by making real time use of google maps instead of using satellite images of google map. The use of online database will increase the capacity of system to store any location available on google map.

## Acknowledgment

It gives us great pleasure in presenting the preliminary project report on House Price Prediction Model Using Image Processing & Machine Learning. I would like to take this opportunity to thank my guide Prof. Prof. S. S. Gore. And Project co-ordinator Prof. Gunjal S. D. for giving me all the help and guidance we needed. we really grateful to the for their kind support. Her valuable suggestions were very helpful. We also grateful I to Prof. Wavhal D. N. Head of Computer Engineering Department, Jaihind college of engineering, Kuran for his in dispensable support, suggestions. In the end our special thanks to Prof. Khatri A. A. for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Project.

## References

- [1] The Danh Phan, "Housing Price Prediction using Machine Learning Algorithms: The Case of Melbourne City, Australia", 2018.
- [2] Sifei Lu, Zengxiang Li, Zheng Qin, Xulei Yang, Rick Siow Mong Goh Institute of High Performance Computing (IHPC), Agency for Science Technology and Research, Singapore, "A Hybrid Regression Technique for House Prices Prediction", 2018.
- [3] Debanjan Banerjee, Suchibrotta Dutta, "Predicting the Housing Price Direction using Machine Learning Techniques" 2018.
- [4] Wan Teng Lim, Lipo Wang, Yaoli Wang, and Qing Chang, "Housing Price Prediction Using Neural Networks", 2016.
- [5] Nehal N Ghosalkar, "Real Estate Value Prediction Using Linear Regression", 2018.
- [6] Ayush Varma, Abhijit Sharma, Sagar Doshi, Rohini Niar, "House Price Prediction Using Machine Learning and Neural Networks", 2018.
- [7] Gaurav Kumar, Pradip Kumar Bhatiya, Fourth International Conference on Advanced Computing & Communication Technologies, "A Detailed Review of Feature Extraction in Image Processing Systems", 2014.
- [8] Guangliang Gao, Zhifeng Bao, Jie Cao, A. K. Qin, Timos Sellis, Zhiang Wu, "Location-Centered House Price Prediction: A Multi-Task Learning Approach", 2019.