

# Application for House Price Prediction Using ML

<sup>1</sup>Ms. Soniya Xavier, <sup>2</sup>Ms. Iqra Ansari, <sup>3</sup>Ms. Kajal Kokane, <sup>4</sup>Mr. Kelhe Priti, <sup>5</sup>Prof. Gursal P. S.

Department of Computer Engineering  
SND Polytechnic, Yeola

**Abstract:** Now a days house resale is majorly seen in metro cities. The market demand for housing is always increasing every year due to increase in population and migrating to other cities for their financial purpose. Prediction of house resale price for long-term temporary basis is important especially for the people who stays who will stay the long time period but not permanent and the people who do not want to take any risk during the house construction. In this paper, the resale price prediction of the house is done using different classification algorithms like Logistic regression, Decision tree, Naive Bayes and Random forest is used and we use AdaBoost algorithm for boosting up the weak learners to strong learners. Several factors that are affecting the house resale price includes the physical attributes, location as well as several economic factors persuading at that time. Here we consider accuracy as the performance metrics for different datasets and these algorithms are applied and compared to discover the most appropriate method that can be used the reference for determining the resale price by the sellers

**Keywords:** Processing, Price, Prediction, Machine learning

## INTRODUCTION

A House/Home are a basic necessity for a person and their prices vary from location to location based on the facilities available like parking space, locality, etc. The house pricing is a point that worries a ton of residents whether rich or white collar class as one can never judge or gauge the valuing of a house based on area or offices accessible. Buying of a house is one of the greatest and significant choice of a family as it expends the entirety of their investment funds and now and again covers them under loans. It is the difficult task to predict the accurate values of house pricing. Our proposed model would make it possible to predict the make it possible to predict the exact prices of houses.

## MOTIVATION

This project is proposed to predict house prices and to get better and accurate results. The stacking algorithm is applied on various regression algorithms to see which algorithm has the most accurate and precise results. This would be of great help to the people because the house pricing is a topic that concerns a lot of citizens whether rich or middle class as one can never judge or estimate the pricing of a house on the basis of locality or facilities available

## LITRATURE SURVEY

1. Mohammadhossein Kiyaei; Farkhondeh Kiaee [1], The housing market is increasing huge, predicting housing prices is not only important for a business issue, but also for people. However, housing price fluctuations have a lot of influencing factors. Also, there is a non-linear relationship between housing prices and housing factors. Most econometric or statistical models cannot capture non-linear relationships yet. Therefore, we propose housing price prediction models based on deep learning methods, which can capture non-linear relationships. In this work, we construct a dataset, including the housing attributes data and macroeconomic data in Taiwan from January 2013 to December 2018. The housing attributes data includes two types of housing transactions, which are "land + building" (Type1) and "land + building + park" (Type2). Macroeconomic data includes housing investment demand ratio, owneroccupier housing ratio, housing price to income ratio, housing loan burden ratio, and housing bargaining space ratio. Then, this dataset is utilized to evaluate the prediction methods based on deep learning algorithms BPNN and CNN to predict housing prices. Experimental results show that CNN with housing features has the best prediction effect. This study can be used to develop targeted interventions aimed at the housing market.

2. Dhanabal Thirumorthy; Umang Rastogi; B.Barani Sundaram [2], Nowadays, house price is extremely essential for human living and plays an important part in the financial market, especially in the real estate industry. In this paper, we proposed a deep learning based network for house price prediction. Firstly, a brief introduction and related work are discussed. Then, Linear Regression, Random Forest Regression, XGBoost, and SVM are all detailed explained, developed and tested. In addition, several types of densely connected based neural network are proposed and developed. Finally, all methods are evaluated on a publicly available dataset, Boston house price dataset. Our method performs competitively performance compared with other classical methods including Linear Regression, Random Forest Regression, XGBoost, and SVM, and evaluated with a variety of evaluation methods including R2, Adjustable R2, MAE, MSE, RMSE..

3. RPrateek Singh; Mukesh Kumar Lohani; Riya Nagarkoti [3], This paper provides an overview about how to predict house costs utilizing different regression methods with the assistance of python libraries. The proposed technique considered the more refined aspects used for the calculation of house price and provide the more accurate prediction. It also provides a brief about various

graphical and numerical techniques which will be required to predict the price of a house. This paper contains what and how the house pricing model works with the help of machine learning and which dataset is used in our proposed model.

4. P. Raja; A. Thivya; C. Janani [4] ,House price prediction is a popular topic, and research teams are increasingly performing related studies by using deep learning or machine learning models. However, because some studies have not considered comprehensive information that affects house prices, prediction results are not always sufficiently precise. Therefore, we propose an end to end joint self-attention model for house prediction. In this model, we import data on public facilities such as parks, schools, and mass rapid transit stations to represent the availability of amenities, and we use satellite maps to analyze the environment surrounding houses. We adopt attention mechanisms, which are widely used in image, speech, and translation tasks, to identify crucial features that are considered by prospective house buyers. The model can automatically assign weights when given transaction data. Our proposed model differs from self-attention models because it considers the interaction between two different features to learn the complicated relationship between features in order to increase prediction precision. We conduct experiments to demonstrate the performance of the model. Experimental data include actual selling prices in real estate transaction data for the period from 2017 to 2018, public facility data acquired from the Taipei and New Taipei governments, and satellite maps crawled using the Google Maps application programming interface. We utilize these datasets to train our proposed and compare its performance with that of other machine learning-based models such as Extreme Gradient Boosting and Light Gradient Boosted Machine, deep learning, and several attention models. The experimental results indicate that the proposed model achieves a low prediction error and outperforms the other models. To the best of our knowledge, we are the first research to incorporate attention mechanism and STN network to conduct house price prediction

### LIMITATION OF EXISTING SYSTEM

- Costing: The Existing system is high cost and this is main reason most of the system is failed.
- Technology Complexity: Most of system is the complex to understand, Not user friendly as compare to our proposed system
- Time Consuming Feature: In existing system, the performance is low and most of the time system gets hanged due to load.
- Not Easy to Understand: Systems are complex to understand and they were not user friendly

### EXPERIMENTAL SETUP

This section describes the various features of the system and also describes the implementation methods. Following are some of the features explained with their implementation details:

- Pattern Recognition: Pattern recognition is the automated recognition of patterns and regularities in data. It has applications in statistical data analysis, signal processing, image analysis, information retrieval, bio informatics, data compression, computer graphics and machine learning.
- Authentication: Authentication is the act of proving an assertion, such as the identity of a computer system user. In contrast with identification, the act of indicating a person or thing's identity, authentication is the process of verifying that identity.

### Hardware and Software Requirements

#### Hardware Requirements

- i3 Processor
- RAM 4GB
- Mobile Device
- Hard Disk 150 GB or above

#### Software Requirement

- Language : Python 3.3 and Andriod
- Operating system : Windows XP/Vista/7/8/8.1
- IDE : Pycharm

Python PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company JetBrains.[6] It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as Data Science with Anaconda. MySQL 5.1 MySQL provides our small, medium and large enterprise customers with affordable, open access to their web data warehouses.

MySQL 5.1 or Firebase MySQL provides our small, medium and large enterprise customers with affordable, open access to their web data warehouses. MySQL allows us to offer our System Administrator low cost, low maintenance database solution for applications without sacrificing power, performance or scalability. Benefits of MySQL are as follows:

- Easy to maintain & upgrade, does not have a slew of administrative tasks to put up.
- Its table format does not vary between releases
- It has cleanly separated table handler modules and can mix access to different type of tables.
- It seems to be developed iteratively, and the features are very stable when they ship them.
- It is a relational database. Over the past several years, this relational database management systems have become the most widely accepted way to manage data.
- It offers benefits such as:
  - Easy to access data
  - Flexibility in data modeling
  - Reduced data storage and redundancy
  - Independence of physical storage and logical data design
  - A high level data manipulation language

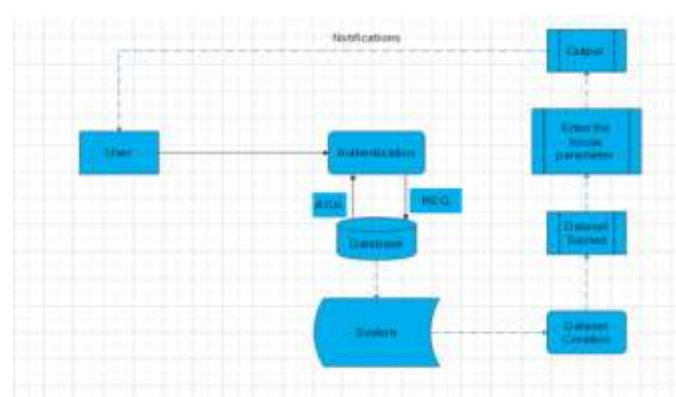
#### SCOPE:

We propose an end to end joint self-attention model for house prediction. In this model, we import data on public facilities such as parks, schools, and mass rapid transit stations to represent the availability of amenities, and we use satellite maps to analyze the environment surrounding houses

#### PROBLEM STATEMENT:

There are a couple of components that impact house costs. In this exploration, partition these components into three essential get-togethers, there are state of being, thought and territory [2]. States of being are properties constrained by a house that can be seen by human recognizes, including the range of the house, the amount of rooms, the availability of kitchen and parking space, the openness of the yard nursery, the zone of land and structures, and the age of the house [3], while the thought is an idea offered by architects who can pull in potential buyers, for instance, the possibility of a moderate home, strong and green condition, and world class condition. Zone is a critical factor in shaping the expense of a house.

#### SYSTEM ARCHITECTURE



**Fig -1:** System Architecture Diagram

The proposed technique considered the more refined aspects used for the calculation of house price and provide the more accurate prediction. It also provides a brief about various graphical and numerical techniques which will be required to predict the price of a house. This paper contains what and how the house pricing model works with the help of machine learning and which dataset is used in our proposed model.

#### ADVANTAGES

1. Easy to use

2. High Performace
3. Scalable

## METHODOLOGY

The single problem can be solved by different solutions. This considers the performance parameters for each approach. Thus considers the efficiency issues.:

- Problem Solving Methods are concerned with efficient realization of functionality. This is an important characteristics of Problem Solving Methods and should be deal with it explicitly.
- Problem Solving Methods achieve this efficiency by making assumptions about resources provided by their context (such as domain knowledge) and by assumptions about the precise definition of the task. It is important to make these assumptions explicit as it give the reason about Problem Solving Methods.
- The process of constructing Problem Solving Methods is assumptionbased. During this process assumptions are added that facilitate efficient operationalization of the desired functionality

## CONCLUSION

The sales price for the houses are calculated using different algorithms. The sales prices have been calculated with better accuracy and precision. This would be of great help for the people. To achieve these results, various data mining techniques are utilized in python language. The various factors which affect the house pricing should be considered and work upon them. Machine learning has assisted to complete out task. Firstly, the data collection is performed. Then data cleaning is carried out to remove all the errors from the data and make it clean. Then the data pre-processing is done. Then with help of data visualization, different plots are created. This has depicted the distribution of data in different forms. Further, the preparation and testing of the model are performed.

## REFERENCES

- [1] Jain, N., Kalra, P., Mehrotra, D. (2019). Analysis of Factors Affecting Infant Mortality Rate Using Decision Tree in R Language. In *Soft Computing: Theories and Applications* (pp. 639-646). Springer, Singapore.
- [2] R. A. Rahadi, S. K. Wiryono, D. P. Koesrindartotoor, and I.B. Syamwil, —Factors influencing the price of housing in Indonesia, *Int. J. Hous. Mark. Anal.*, vol. 8, no. 2, pp. 169–188, 2015
- [3] Limsombunchai, —House price prediction: Hedonic price model vs. artificial neural network, *Am. J. ...*, 2004
- [4] Kadir, T., Gleeson, F. (2018). Lung cancer prediction using machine learning and advanced imaging techniques. *Translational Lung Cancer Research*, 7(3), 304-312
- [5] Liu, J., Ye, Y., Shen, C., Wang, Y., Erd'elyi, R. (2018). A New Tool for CME Arrival Time Prediction using Machine Learning Algorithms: CATPUMA. *The Astrophysical Journal*, 855(2), 109.