

# House Price Prediction Using Machine Learning Techniques

Anirudha Chavan<sup>1</sup>, Shubham Bobade<sup>2</sup>, Madhuri Bhosale<sup>3</sup>, Prof. S. H. Thengil<sup>4</sup>

Students, Department of Computer Engineering<sup>1,2,3</sup>

Professor, Department of Computer Engineering<sup>4</sup>

Navsahyadri Education Society's Group of Institutions, Pune, Maharashtra, India

**Abstract:** *The real property retail business is one of the main area where machine learning can be used to determine and predict the amount with high rated value. Housing price is one of the most anxious issues to the user in all over the world. The imprudent increasing of housing price will impact not purely the grade of life, but also the business cycle spirited. Although, the element sway residential real estate amount are not easy and the chosen of effective parameter is indistinct, which leads to a lower precision in many of the conventional housing price detection approaches. Accordingly, a original detection model situated on CNN is suggest for detection of housing price as well as the process of properties selection. Differentiate with other newest methods, our work can acquire a better performance via test using training data of property transaction. Influence housing price is crucial model for resolution making for user in which number of element can be used to detect amount of regret house. The parameter that are intricate in the process are not acquire of the many analytical system available to examine the house price considering various element relating to atmosphere, environment and other entities etc. The system will help people to invest in a possessions without approaching a third party user. It also overcome the risk present in the transaction. Use of CNN algorithm is done as model because of its adaptable and probability techniques on model selection. The output displays that the approach of the problem needs to be successful, and has the capacity to operate predictions that would be relative with other house price detection models.*

**Keywords:** CNN algorithm, House Price, Prediction, Machine Learning, Feature Selection

## I. INTRODUCTION

Urban real estate has consistently been not only a key part of household issue, but also a great collision or on copious economy. As such, it is of best moment to study the action of housing price, point to provide discernment in the real estate industry and help prepare genuine policy advice for the collaborator. Associated works on real estate work around the 17th century and a lot of reaching have been fulfil. Cong Wang design a logistic model to detect housing price for different cities in China and encourage that their model was better than newest MLR model, while Ruina Shen developed the SVM to determine the price for Shanghai city, in which PCA was used to overcome the size of data and increment the prediction property, but it was trust more fitted for a small testing scale of data. Guanglan Wei design the Markov chain to detect housing price prediction for the city of Kunming, but with some dominant factors disregard. In addition, Shengping Jin design a system based on Random Forest with a great output that satisfied outcome than ARMA model. As same works also have BP neural network and GM model etc.

In this proposed system, differentiating that many factors may have unpredicted convolute reach on housing amount, we design a CNN-based model to deal with the intricacy. The work most connected to ours is from Yanjun Liu, in which Convolutional nural network also played and main role. However, they attentive mainly on fining text data to predict housing amount, while our system put the highlight on detecting rational attribute and modify the network structure of CNN so as to detect the amount correctly. As introduce, we believe that the attribute influence housing amount are wide reach and compound, and different attribute at many times may even have different parameter in predicting process. In view of this state, we initially delve into the characteristic selection and combine indicators like GDP, real estate evolution investment, per capita area and other important element into the prediction model. Secondly, we developed an effective CNN model based on a training dataset of property transaction. Finally, we said that the system based on CNN is best to other traditional experiments.

The quarry characteristic in this proposed model is the amount of the house price estate property and the own attribute are: available bedrooms, bathrooms, carpet area, near about area, built-up area, the floor, , zip code, latitude and longitude of the property, age of the property. Other than those of the introduce features, which are mainly essential for predicting the house amount, we design two other characteristic - air quality and crime rate. These attribute provide a main parts towards detecting property prices since the higher values of these attribute will lead to a reduction in house amount.

## **II. RELATED WORK**

The amount of a specific property depends on the architecture amenities nearby the property. Lately, a few authors' confines for detect the best properties for the users came along with various system. Raghunandhan discuss the basic data mining abstraction of how it works and underpin algorithms for the purpose of detection. The main important element is which machine learning system is best benefit for predicting the house amount. Frequently the location's situation conditions determine what kind of amount we can await for different types of home, Manjula describe many important attribute to use when predict property amount with good accuracy using a regression model. A. Varma developed a techniques that used real-time locality data to get amount real-world evaluation using Google maps. Author's also showed that there subsist connection between the optical appearance and non-visual element such as crime circumstance, housing amount, native's density, etc. of a city. Hujia Yu, Jiafu Wu used detection and regression algorithms. According to survey, living area square feet, neighborhood have the greatest statistical importance in estimating the selling price for a home and roof content. And detection analysis can be upgrade by the PCA method. Li Li and kai-Hsuan Chu studied different algorithms such as Radial basis functional (RBF) neural networks and Back propagation neural network (BPN). The use of RBF and BPN models is initiate to determine the difference between the house amount index such as Cathy and sinny amount index and complex connection function to examine the macroeconomic analysis. Nihar Bhagat, Ankit Mohokar, Shreyash Mane careful linear regression algorithms for prediction of the houses. The goal of the paper is to divine the efficient price of real estate for users with respect to their forecast and priorities. Analysis of past market drift and price ranges will predict attribute house amount.

## **III. PROPOSED SYSTEM MODULE**

The proposed system have main four module as discuss below:

### **Module1: Dataset**

The prediction issue lies in the kaggal dataset from the different types of resources and housing data. The dataset contain many type of attribute as near area, carpet area, bedroom, kitchen, hall and other required area. There are some feature, the data put into which result in final outcome experimental result. Each attribute have its own value, which are listed in the common element. Dataset is required before any type of machine learning project is carried out. Dataset is analyzing the train and test dataset.

### **Module2: Feature Extraction**

In order to pull out some significant particular of the object quality extraction is used. Characteristic is the function of extent measurable things of an object. Feature extraction is connected to dimensionality pruning. All the attribute of data can be differentiate into high-level aspect and low-level aspect production. Feature extraction may include area, rooms, carpet area and location. In our system, we are using convolutional neural network for feature extraction.

### **Module3: Training Dataset**

The proposed system uses two dataset as training dataset and testing dataset. The training dataset must have target value. System train the complete dataset as 70% data used for training and 30% used for testing dataset. In machine learning training module is used for predict the house amount.

### **Module4: Testing and result**

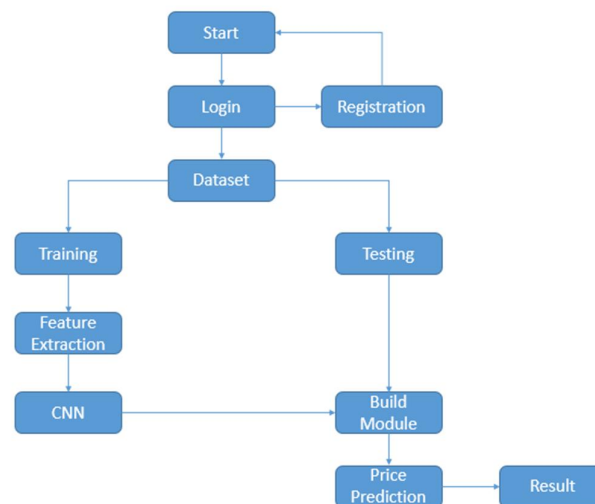
In proposed system these is final module where we trained complete module that applied to test dataset and predict the house price. The final outcome is trained with CNN algorithm. The proposed system uses spyder IDE for develop the code in python.

#### IV. SYSTEM ARCHITECTURE

The proposed system architecture is design as below:

The workflow of our prosed system is as follows:

1. Step1: User should login itself
2. Step2: If user don't have account they register first and then login to our system
3. Step3: We have already training dataset and testing dataset. With the help of CNN we developed appropriate module for extract the feature of system.
4. Step4: User can check the attribute of the required house.
5. Step5: On the basis of trained module system predict the price.
6. Step6: System shows the predicted result.



**Figure:** System Architecture

#### V. CONCLUSION

The proposed system is mainly focused on detection of the classified the features from the training housing dataset, for get better result we use CNN model to predict the housing price with better outcome as compared with other module. It is of great consequence to provide stakeholders valuable insight in the real estate market and help supply exact policy. As we compared with other newest methods, our model gives a more pleasing result, which shows that the data element selected as prediction attribute are explicit parameter and the CNN is suitable for housing price detection. HPP model just about try to find the same one. Proposed system attentive on predict the house amount according to the area. As final when we design these system that give accurate prediction of house amount.

#### 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, Suchibrota Dutta, "Predicting the Housing Price Direction using Machine Learning Techniques" 2018.
- [4]. F. J. López-Iturriaga and I. P. Sanz, "Predicting Public Corruption with Neural Networks: An Analysis of Spanish Provinces," Soc. Indic. Res., vol. 140, no. 3, pp. 975–998, 2018, doi: 10.1007/s11205-017-1802-2.
- [5]. D. Miao, W. Chen, W. Zhao, and T. Demsas, "Parameter estimation of PEM fuel cells employing the hybrid grey wolf optimization method," Energy, vol. 193, p. 116616, 2020, doi: 10.1016/j.energy.2019.116616.

- [6]. S. Dhargupta, M. Ghosh, S. Mirjalili, and R. Sarkar, "Selective Opposition based Grey Wolf Optimization," *Expert Syst. Appl.*, vol. 151, p. 113389, 2020, doi: 10.1016/j.eswa.2020.113389.
- [7]. E. Ileri, A. D. Karaoglan, and S. Akpınar, "Optimizing cetane improver concentration in biodiesel-diesel blend via grey wolf optimizer algorithm," *Fuel*, vol. 273, no. March, p. 117784, 2020, doi: 10.1016/j.fuel.2020.117784.
- [8]. J. I. Pérez-Rave, J. C. Correa-Morales, and F. González Echavarría, "A machine learning approach to big data regression analysis of real estate prices for inferential and predictive purposes," *J. Prop. Res.*, vol. 36, no. 1, pp. 59–96, 2019.