**HOUSING: PRICE PREDICTION**

Submitted by:

Saunak Mukherjee

**ACKNOWLEDGMENT**

This includes mentioning of all the references, research papers, data sources, professionals and other resources that helped you and guided you in completion of the project.

**INTRODUCTION**

House which is one of the human life's most essential needs, along with other fundamental needs such as food, water. Demand for houses that were grew rapidly over the years as people's living standards improved. When there are lots of

people who make their house as an investment and property, yet most people around all over the world are buying a house as

their shelter which is our livelihood.

According to [1], housing markets have a positive impact to the country's currency, which is an important national

economy scale. Homeowners would be purchase goods such as furniture and household that are equipment for the home, and

The contractors will purchase raw material to build houses to asatisfy house demand, that indication

of the economic wave effect created by the new house supply. Besides that, consumers have capital to make a large

investment, and the construction industry is in good condition can be seen through a country's high level of house

supply.

According to [2], numerous international organizations and human rights have emphasized house importance.

House is profoundly rooted in the economic, financial, and political structure of each country. Nevertheless, [3]

reported to the fluctuation of a house prices has been to the issue of the house owners, buildings and real estate,

besides [4] stated that house has the become unaffordable as there are so many substantial price growth in the several countries in the House Price Prediction using a Machine Learning Model: A Survey of Literature 47

Copyright © 2020 MECS I.J. Modern Education and Computer Science, 2020, 6, 46-54

housing sector. Residents' quality of the life that is well in national economy depends which are potential house price when increase.

Ultimately, when this issue will affect the investors they are making their house as an investment.

An increase of the house demand occurs with year, indirectly causing house the price increases in every year.This the problem arises when there are so many numerous variables which is location and property demand that may be influence the house price, thus

Most of the stakeholders that are including to the buyers and the developers, house builders which is the real estate industry would like to know that the

exact of attributes factors influencing the house price to help investors make decisions and help house

builders set the house price.House price prediction that has to be done by using the multiple prediction models (Machine Learning Model) such as

The support vector regression, artificial neural network, and much more. There are so many benefits of that home buyers, to the property of the

investors, and house builders can be reap of the house-price model. This model will be provide a lot of information or

knowledge to the home buyers such as the valuation of house prices in the market, which will help them to determine the house prices. This model can help out potential buyers decide to the characteristics of the house they want according to the budget [5]. Previous studies focused to the analyzing the attributes of that affect and the house price and predicting of the house price which are based on that model of machine learning separately. However, that

article combines to both predicting house price and the attributes together.

There are in literature review focuses on predicting house price based on the model of the ml as well

as analyzing to the attributes primarily used in previous study that the affect house price. This paper which was arranged as follows: the

first section summarizing overall to this study. Second section which is described the common attributes used in prediction of the

house price to the world. It was followed by the brief discussion of the machine learning model which used in previous study to the predict house price. For the next section, the comprehensive effects to the current house price prediction model are addressed. In this section 5 and section 6 respectively provide the description and conclusion of this

comprehensive literature of the analysis.

**Machine Learning Model:**

[20], the paradigm which is the evaluating of the house demand which is classified into the two classes that are traditional method to the advanced to the valuation method. The traditional valuation scheme and stepwise which is the regression process of the hedonic pricing tool, artificial neural network (ANN) and spatial analysis of the framework that are advances valuation method. The model selection that is predict house price which is quite

critical to the varieties of models that are available. One of the most commonly which is utilized models that are research field which is a Regression Analysis that are is used in many studies, including [3, 10, 21]. Another common model for house price predictions which is Support Vector Regression (SVR) [7, 22, 23].

A. Regression Analysis

i. Hedonic Price Model

The housing market that is slightly different from the normal good consumption. According to [13], housing market which is

unique because it displays that the characteristics of resilience, flexibility and spatial fixity. Therefore, hedonic approach is

preferred to accurately predict market differential. [24] conceived hedonic model back in 1939, but this research was

popularized in the early 1960s with comprehensive uses by Zvi Griliches and Rosen [24]. In the early 1930s, Court used to this model to analyse to the automotive value in pricing and quality characteristics. [25] defined hedonic"the implicit

prices of attributes and are revealed to economic agents from the observed prices of differentiated products and the

specific quantities of their characteristics." Following years of progress, Rosen applied the approach to the residential

home price study and became commonly included in real estate sector research [3].

Rosen's philosophy or model comprises two separate phases. The regression of a product price on its attributes is

performed in the initial stage to calculate the aggregate price of the component. A measure of a goods price will be

determined in the first stage, but the inverse demand function cannot be generated at this stage. Thus, the second stage

of estimation is needed to identify the inverse demand function that can be derived from the first stage implicit price

function.

Locational and structural attributes

-

[22]

Locational attributes

0.0047

[23]

Artificial Neural Network

Locational and structural attributes

0.5155

[7]

Locational attributes

-

[32]

Locational attribute

-

[40]

Locational attributes

0.0581

[23]

XGBoost

Structural attribute

0.1212

[38]

Based on reviewing numerous papers, there are several attributes used by researchers in their work to forecast

house prices. All of these attributes can be divided into 4 main categories which are locational, structural, neighborhood

and economic attributes. The locational attribute consists of variables which described the accessibility to shopping mall,

accessibility to school, accessibility to hospital, restaurants and availability of public transport. Meanwhile, the

structural attribute consists variables of number of bedrooms, number of bathrooms, floor area, garage and patio,

property age housing and lot size. The neighborhood on the other hand described the socioeconomic variables, crime

rates, place of worship, pleasant landscape, and quiet atmosphere which the variables is quite subjective. And lastly the

economic attributes consist of income factor and cost of material factor. The most common used attributes by previous

research regarding house price prediction are locational and structural attributes, as neighborhood and economic

attributes are difficult to define and measure.

The classification of attributes made it easier to analyze the effects of different attributes on different models. The

tabulation of finding based on evaluation of previous study can be seen in Table 3. Looking at the table above, it can be

seen that support vector regression has the smallest RMSE value, 0.0047. Technically, the RMSE value of a model is

highly dependent on the attributes used during the prediction process. Most of the model that are using the same

attributes (locational attributes) will generate a very low RMSE value indicating the best model. However, this cannot

simply show the best model because several previous studies rarely provide the RMSE value to justify their model

being the best model. On the basis of the analysis table, the locational attribute may be assumed to be the main attribute

used by several models such as support vector regression and artificial neural network support. The RMSE value is very

low with the presence of the locational attribute only, however, the RMSE value is quite high when the structural

attribute is combined with the locational attribute for the input to make a prediction.

Next, ANN provides the second lowest of RMSE value which is 0.0581. Findings revealed that locational

attributes are indeed the relevant attributes used in the ANN model to forecast house price. Similar to SVR, the RMSE

value is fairly high when locational and structural attributes are used together. This indicates that the locational attribute

should be used alone to achieve low RMSE values by ANN model. In the meantime, the XGBoost model also yields the

lowest RMSE while only the structural attribute is involved.

However, research on model with structural attributes alone is very limited as previous research focused mainly on

locational attributes or the combination of locational and structural attributes in order to predict house prices. In general,

our analysis suggests that SVR, ANN and XGBoost are the most efficient models compared to other models, whereas

locational attributes are the main attribute in predicting house price.

6. Conclusions

This paper examined and analyzed the current research on the significant attributes of house price and analyzed the

data mining techniques used to predict house price. Technically, houses with a strategic location such as the

accessibility to shopping mall or other facilities tend to be more expensive than houses in rural areas with limited

numbers of facilities.

The accurate prediction model would allow investors or house buyers to determine the realistic price of a house as

well as the house developers to decide the affordable house price. This paper addressed the attributes used by previous

researchers to forecast a house price using various prediction models. Taken together, the results of the survey have

shown the potential of SVR, ANN and XGBoost in predicting house prices. These models were developed based on

several input attributes and they work significantly positive with house price. In conclusion, the impact of this research

was intended to help and assist other researchers in developing a real model which can easily and accurately predict

house prices. Further work on a real model needs to be done with the utilization of our findings to confirm them.

References

[1] A. S. Temür, M. Akgün, and G. Temür, “Predicting Housing Sales in Turkey Using Arima, Lstm and Hybrid Models,” J. Bus.

Econ. Manag., vol. 20, no. 5, pp. 920–938, 2019, doi: 10.3846/jbem.2019.10190.

[2] A. Ebekozien, A. R. Abdul-Aziz, and M. Jaafar, “Housing finance inaccessibility for low-income earners in Malaysia: Factors

and solutions,” Habitat Int., vol. 87, no. April, pp. 27–35, 2019, doi: 10.1016/j.habitatint.2019.03.009.

[3] A. Jafari and R. Akhavian, “Driving forces for the US residential housing price: a predictive analysis,” Built Environ. Proj.

Asset Manag., vol. 9, no. 4, pp. 515–529, 2019, doi: 10.1108/BEPAM-07-2018-0100.

[4] Choong Wei Cheng, “Statistical Analysis of Housing Prices in Petaling,” Universiti Tunku Abdul Rahman, 201