International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

Forecast Home Price Using Different types of Machine Learning Algorithms

Bramhanand Dubey, LokendrasinghSongare

PG Scholar, CSD, Dr. APJ Abdul Kalam University Indore, M.P., India E mail -

Assistant Professor, CSD, Dr. APJ Abdul Kalam University Indore, M.P., India E mail - lokendrasinghsongare@aku.co.in

ABSTRACT

Home prices increase every year and it indicates the current economic situation so there is a need for a system to predict Home sales in the future for both buyer and the seller. Here we use dataset of India across different cities and having more then 68,613 entries of test data and 28000 train data of housing sales in whole, India. This analysis includes the effect of markdowns on the sales and the extent of effects on the sales by size, price, area etc. has been analysed using different machine learning algorithms. Home sales prediction can help the developer determine the selling price of a Home and can help the buyer to arrange the right time to purchase a Home. Algorithms predict the output values based on input features from the data fed in the system and analysis is a set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables. There are three factors that influence the sales price of a Home which include physical conditions, concept and location.

An exact forecast of forthcoming development market interest, particularly the private market, is central imperative to strategy producers, as it could assist with forming procedures to develop/balance out the economy and fulfill the social requirements (at full scale level). In spite of that, a sensible forecast of future private interest is never a simple assignment, as it is represented by various social and monetary elements. In this paper, four proactive factor models are created and thought about for straightforwardly anticipating India private area a wide range of interest

Key Words: Forecast, Machine Learning Algorithms, Home Price, economy

INTRODUCTION

Investment is a business activity that most people are interested in this globalization

era. There are several objects that are often used for investment, for example, gold, stocks and property. In particular, property investment has increased significantly since 2011, both on demand and property selling. One of the increasing of property demand is because of high population. As of the 2010 India Census, Of the 1789,232 Home holds, 29.2% had children under the age of 18 living with them and 31.0% of all Home holds were made up of individuals. The result of this census indicates that the younger generation will need a Home or buy a Home in the future.

There are several approaches that can be used to determine the price and sales of the Home, one of them is the sales prediction analysis. This research aims to create a Home sales prediction model using regression analysis. Prediction Home sales are expected to help people who plan to buy a Home so they can know the price range in the future, then they can plan their finance well. In addition, Home sales predictions are also beneficial for property investors to know the trend of housing sales price in a certain location.

OBJECTIVE OF THEWORK

In most recent twenty years anticipating the property estimation has turned into a significant field. Valuation of Home Prices utilizing Predictive Techniques 35 property and eccentric conduct of economy force analysts to discover a way that foresee the land costs with no predispositions. In this manner, it is a test for scientists to discover every one of the moment factors that can influence the expense of property and make a prescient model by thinking about every one of the variables. Building a prescient model for land value valuation requires intensive information regarding the matter. Numerous analysts have dealt with this issue and imparted their examination work. A large portion of this exploration work is propelled from

The creator has scratched the lodging informational collection from Centris.ca and duProprio.com. Their dataset comprises of around 25,000 models and 130 variables. Around 70 highlights were scratched from the above sites and land organizations, for example, RE/MAX, Century 21, and Sutton, and so on Other 60 highlights were sociodemographic dependent on where the property is found. Afterward, creator

executed Principal Component Analysis to lessen the dimensionality. The creator utilized four relapse strategies to foresee the value worth of the property. The four strategies are Linear Regression, Support Vector Machine, K-Nearest Neighbours (KNN) and Random Forest Regression and a troupe approach by joining KNN and Random Forest Technique. The outfit approach anticipated the costs with least mistake of 0.0985. Nonetheless, applying PCA didn't further develop the forecast mistake. A ton of investigates have been done on Artificial Neural Networks. This has helped numerous analysts zeroing in on land issue to settle utilizing neural organizations..

Result:-Based on the analyzed studies and practical applications for hedonic pricing models, a list of characteristics is identified and divided into three categories: market characteristics, location characteristics and intrinsic characteristics of the Home. The two most important categories are the intrinsic and location characteristics of the Home, since the market characteristics are global influences impacting all Homes. Nevertheless, the market characteristics have been included for the sake of completeness. This overview is based on the overview of hedonic model variables of Zhou et al. [34]. This overview however focuses mainly on variables that have also been included in geographically weighted regression models.

The market characteristics are identified as global influences on the entire housing market. One large market influence are national policies, such as the recent abolition (April 2021) of transfer tax for starters in the Dutch housing market. These national policies often have an equal impact on the price for all housing [35]. Another global influence is the mortgage interest rate. A lower interest rate leaves the home buyer with more money to spend. As a result, this often drives up Home prices. Since market characteristics are global influences, it does not explain the spatial variance in Home prices. As such, these variables do not belong in a geographically weighted regression model. Nevertheless, they play a crucial role in explaining the temporal difference in Homes prices, as they do play a role when looking at the growth of Home prices on a yearly basis.

In contrast, intrinsic characteristics are the biggest differentiating factors for Home prices. As such, they are also by far the most used variables for hedonic pricing models. Not only in literature, but also in the hedonic price models such as the from practice these were stated as the heaviest influences for Home prices. The largest influences are naturally the living area and volume, commonly followed by the amount of garden space. Amenities such as garages and multiple bathrooms also contribute to higher Home prices. The build year can serve as a moderate indicator of energy efficiency and state of maintenance; however it does not always depict the true condition of the Home. Old Homes are likely renovated once in their life span, so other features such as an energy label are needed. Furthermore, older buildings can also be cultural heritage, which can result in higher prices for older buildings due to their significant historic value as stated in.

Conclusions are carried out:

When it comes to data, a set or collection of data will be called a dataset, Theobjective of this thesis is to fit models to predict the housing sale price and find someimportant aspects of the Home. In order to achieve my goal, I fit four models to thedataset: linear regression, Logistic regression, Naive Bayesian and knn. Here Iimplements three models in dataset linear regression, Naive Bayesian and k NearestNeighbour. As for the first model - linear regression, it doesn't meet the assumptionof equality of the variances. Therefore we can't use the linear model as the candidateof our final model. In order to deal with this problem, I try the Second model –Naïve Bayesian, Naive bayes works well with small datasets Naive bayes is much fasterthan KNN due to KNN's real-time execution. The third model is K NearestNeighbour but the neural networks need large training data compared to KNN toachieve sufficient accuracy. KNN is better than linear regression when the data havehigh Signal-to-noise ratio (SNR). Neural networks need lot of hyperparameter tuningcompared to KNN.we can know that the total square feet, the overall quality, and thetotal number of bathrooms are the three main aspects which influence the housingsale price.

Finally, a model is a trained function that defines the relationships within adetermined set of data. A model is capable of predicting the class or value ofan input, which in this case is new information that isn't present in the data used totrain the model.

REFERENCES

1. D. Belsley, E. Kuh, and R. Welsch, Regression Diagnostics:

IdentifyingInfluential Data and Source of Collinearity. New York: John Wiley, 1980.

2. J. R. Quinlan, "Combining instance-based and model-based learning," MorganKaufmann, 1993, pp. 236–243.
3. S. C. Bourassa, E. Cantoni, and M. Hoesli, "Predicting Home prices with spatialdependence: a comparison of alternative methods," Journal of Real Estate Research,vol. 32, no. 2, pp. 139–160, 2010. [Online]Available:http://EconPapers.repec.org/RePEc:jre:issued:v:32:n:2:2010:p:139-160.
4. S. C. Bourassa, E. Cantoni, and M. E. Hoesli, "Spatial dependence, housingsubmarkets and Home price prediction," eng, 330; 332/658, 2007, ID:unige:5737.[Online]. Available: http:// archive - ouverte. unige. ch/unige:5737
5. P. Falinouss, "Stock trend prediction using news articles: a text miningapproach," 2007Pow, Nissan, Emil Janulewicz, and L. Liu. "Applied Machine Learning Project 4Prediction of real estate property prices in Montréal." (2014).
6. Limsombunchai, Visit. "Home price prediction: hedonic price model vs. artificialneural network."New Zealand Agricultural and Resource Economics SocietyConference. 2004.
7. Park, Byeonghwa, and Jae Kwon Bae. "Using machine learning algorithms forhousing price prediction: The case of Fairfax County, Virginia housing data."ExpertSystems with Applications 42.6 (2015): 2928-2934.
8. Bhuriya, Dinesh, et al. "Stock market predication using a linear regression."Electronics, Communication and Aerospace Technology (ICECA), 2017International conference of.Vol.2.IEEE, 2017.
9. Li, Li, and Kai-Hsuan Chu. "Prediction of real estate price variation based oneconomic parameters." Applied System Innovation (ICASI), 2017 InternationalConference on.IEEE, 2017.

10. Hromada, Eduard. "Mapping of real estate prices using data mining techniques."Procedia Engineering 123 (2015): 233- 240.
11. Wu, Jiao Yang. "Housing Price prediction Using Support Vector Regression."(2017).
12. Hujia Yu, Jiafu Wu hujiay, jiafuwu, “Real Estate Price Prediction withRegression and Classification”, CS 229 Autumn 2016 Project Final Report.
13. Hong Zhao, Rong-Qiu Chen, Wei Xu, Da-Ying Li, “A SVR based forecastingapproach for real estate price prediction”, International Conference on MachineLearning and Cybernetics, 2009.
14. AyushVarma, AbhijitSarma, SagarDoshi, Rohini Nair, “Home Price PredictionUsing Machine Learning and Neural Networks”, ICICCT, KJ Somaiya College OfEngineering, Mumbai, 2018.
15. NeelamShinde, KiranGawande,“Valuation Of Home Prices Using PredictiveTechniques”, International Journal of Advances in Electronics and ComputerScience, ISSN: 2393-2835 Volume-5, Issue-6, Jun.-2018.
16. Raouf Boutaba1*, Mohammad A. Salahuddinet all, “A comprehensive surveyon machine learning for networking: evolution, applications and researchopportunities”, Journal of Internet Services and Applications, 2018