Comparative Assessment of Radom Forest, SVC and Cat Boost Performances as Property Price Forecasting Models

Oyedeji, Joseph Oyewale Estate Management Department, Bells University of Technology, Ota[1]

Oyediran, Mayowa Oyedepo Computer Engineering Department Ajayi Crowther University[2]

Majekodunmi, Olumide Samson, Estate Management Dept, Bells University of Technology

Article Info
Page Number: 1283-1289
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: 25 March 2022
Revised: 30 April 2022
Accepted: 15 June 2022
Publication: 19 August 2022

Abstract
Associated risk and uncertainties that characterized property investment necessitate adoption of property price forecasting models with high precision rate. This study aimed at comparing the predictive performances of the random forest, support vector machine, cat boost as residential property price forecasting models. Data for the study were gathered from the records of recent lettings as provided by residential property managers in the study area. For the purpose of precision, this study adopted random forest, SVM and Cat Boost as model of classifying rental value of residential properties in the study area. Important property value determining factors like; distant to cultural site, age of building, house type, road network, availability of water, state of exterior, state of interior and security were considered as input variables. It was revealed that the three adopted forecasting models achieved over 80% of precision and accuracy in the classification of residential properties in the study area. Also, the study established that random forest as the best rental value forecasting model among the three considered rental value forecasting models.
Keywords: Property Price; Forecasting: Radom Forest; SVC; Cat Boost

Introduction

Real estate investment is a viable class of investment in an investment portfolio (Klimczak, 2010). Among the available classes of real estate, investment in residential properties is attractive to investors (Oyedeji et. al, 2018). Proceeds from investment in residential real estate can be in the form of capital value appreciation and rental income. These two categories of real estate investment are characterized by different risk factors and uncertainties. These risk factors and uncertainties necessitate the adoption of different forecasting models with established high precision rate. Abidoye and Chan (2017) adopted Artificial Neural Network in forecasting property value, Oyedeji (2018) adopted logistic regression, Artificial Neural Network and Support Vector Machine as a prediction model, NunezTabales, Caridad and Rey (2013) adopted Artificial Neural Network as property forecast model, Jadevicius and Houston (2015) employed auto-regression as an property price forecasting model and Adetunji et al (2021) predicted house price using random forest machine learning technique.

Previous studies on real estate rental and price forecasting have attributed different precision rate to the various forecasting models adopted. In addition to the adoption of artificial intelligence property forecasting models, previous studies also employed non-artificial intelligence prediction models like hedonic model Li and Li (1996), hedonic and regression forecasting models (Babawale et al, 2012). In the same vein, there are studies that compared the adoption of artificial intelligence and non-artificial (Abidoye and Chan, 2016; Abidoye and Chan, 2017; Oshodi et. al, 2020). Hepsen and Vatansever (2011) classified the different property rental and price forecasting models into multivariate and univariate forecasting models.

There are different variables considered in the various models employed for residential property rental and price forecasting. Abidoye and Chan (2016) classified these variables broadly into three namely; structural, neighborhood and location variables. Ajibola et. al (2011) and Olajide and Lizam (2016) examined the impact of different neighborhood variables on residential property value. Abidoye and Chan (2017) examined the impact of structural variables on residential property value and Popoola et. al (2015) examined the impact of location variable on residential property value. These various variables have different impact on residential property value. This makes it necessary to consider the three broad variables in the development of model for property rental and price forecasting (Oyedeji, 2018; Oshodi et. al, 2020). This study aimed at comparing the forecasting performances of random forest, Cat boost and support vector machine models of residential property rental value in proximity to a UNESCO site at Osogbo, Nigeria.

Determinants of Residential Properties Value

There are different factors that influence residential property value. These factors can be categorized broadly as: structural, neighborhood and location factors (Chin and Chau, 2002). Other scholars Pozo (2009), Ajide and Alabi (2010), Babawale et. al (2012) gave credence to this assertion. However, Jekins (2000) and Olayiwola (2005) asserted that the influence of these factors on property value depends on cultural, economic, financial and legal structures of different countries. The various sub-classifications of the three broad classifications (structural, neighborhood and location factors) are influenced differently by cultural, economic, financial and legal structures of different countries and consequently influence the property value.

Abidoye and Chan (2016) in a study conducted in Lagos, Nigeria, classified structural factors determining rental value as; state of property improvement, living area size, stage in estate life cycle, number of the bedrooms, number of the bathrooms/toilets, building attributes, presence of security fence and size of bedrooms. The scholars classified neighborhood factors as; neighborhood attributes, presence of neighborhood security, presence of electricity, presence of pipe-borne water supply and presence of waste disposal. In the same vein, location factors are: easy access to place of work, easy access to CBD, easy access to public transport facility, closeness to highway, easy access to school, and easy access to shopping mall.

This study factored in the uniqueness of the study area in the identification and adoption of residential properties rental value determinants. This is important so as to ensure high precision of the predictive models.

Empirical Review

The three machine learning models; random forest, cat boost and support vector machine have been adopted in previous studies as property rental or market value predictive models. These studies established that the precision rates of the predictive abilities of these models are high. Zeffora and Shobarani (2019) established the high precision rate of random forest as a property price forecasting model. The study established that random forest model predicted housing prices more accurately than multiple regression models. Also, Hong et al (2020) compared property price prediction ability of random forest model and hedonic pricing model. The study adopted a data set covering 40percent of all transaction in the study area. Finding from the study revealed that the deviation between the predicted and actual market value is 5.5% in random forest predictor and 20% in hedonic pricing model. Wu (2017) adopted Support Vector Machine in predicting house price in King County, USA. The study established high precision in the predictive performance of the model. It can be inferred from the various empirical studies that the predictive ability of machine learning models is high.

Research Method

Data

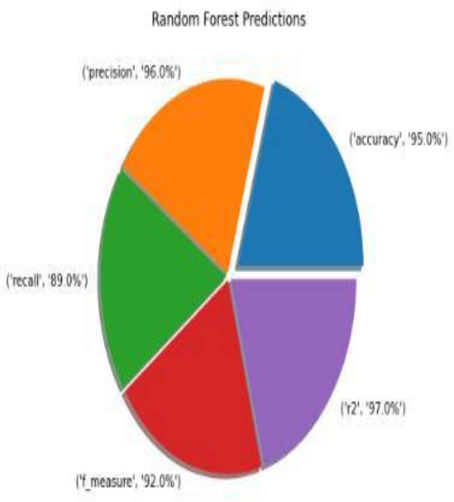
The study area for the study is Osogbo, Osun State Nigeria. The study area is a good test bed for the study because a major input variable is distance from cultural site. The study adopted distance from UNESCO site (Osun Osogbo Groove) as a reference point as established in previous studies (Oyedeji, 2018; Oshodi et al, 2017). Data on rents of recently let residential properties were gathered from 168 property owners and property managers in the study areas. Data on the input variables like: distance from cultural site, house types, age of building, state of exterior, state of interior, availability of water, road network and neighborhood security were also gathered from the property owners and managers due to absence of secondary data in this regard. The analysis of the gathered data was done using Python 6.0.

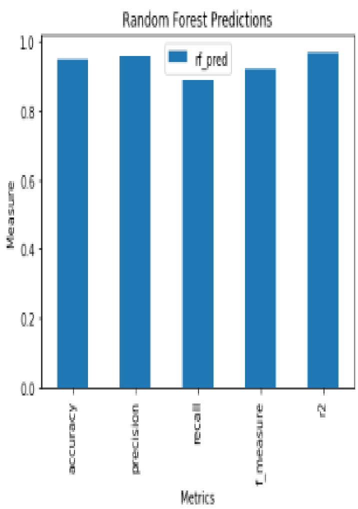
Results and Description

Results

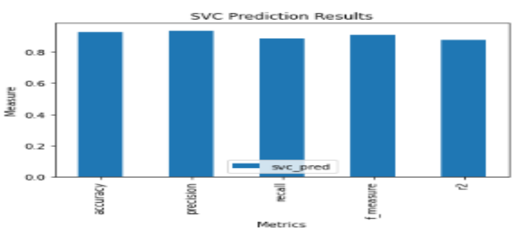
Machine learning models are used to predict numerical values and classifications. Rental value of residential properties is classified as follows in the study: (A.) less than #50,000 (B.) #50,001 - #100,000 (C.) #100,001 - #150,000 and (D.) over #150,001. The predictive performances of the three machine learning models (Radom Forest, SVC and Cat Boost Performances) were established by precision and accuracy rate. The predictive performance of the machine learning models ranges from 0 – 100% where a value close to 100% shows perfect performance. The predictive performances of the machine learning models are as stated below.

Random Forest Predictions

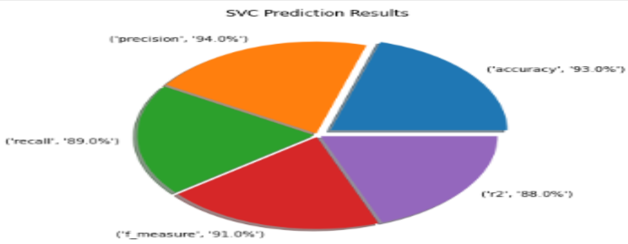




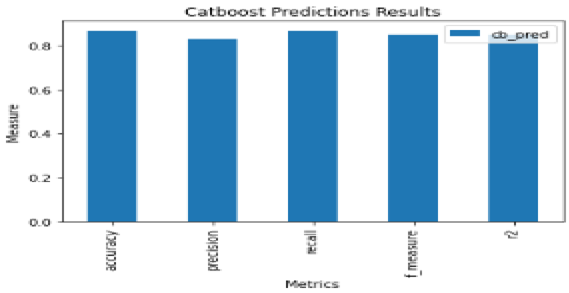
Support Vector Classification Prediction



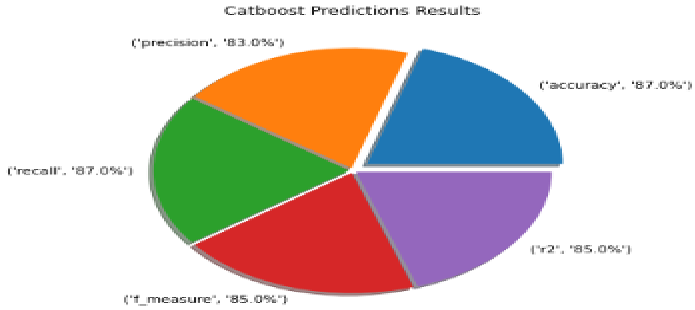

```
plot_pie(svc_dico_results, 'SVC Prediction Results')
```



Catboost Prediction




```
plot_pie(cb_dico_results, 'Catboost Predictions Results')
```



Discussion of Findings

Accuracy is the degree of closeness to the true value while precision is the degree to which an instrument or process will repeat the same value if the same study is repeated. Random forest model reveals that the predicted rental value is 95% close to the true value and there is 96% possibility that the predicted rental value will be returned if the study is repeated. SVC reveals that 93% of the predicted rental value is close to the true value and there is 94% possibility that the predicted rental value will be returned if the study is repeated. Cat boost model reveals that 87% of the predicted rental value is close to the true value and there is 83% possibility that the predicted rental value will be returned if the study is repeated. This study corroborates findings from previous studies that established high prediction performances of machine learning models (Wu, 2017; Hong et. al, 2020; Adetunji et.al, 2021).

Conclusion

It can be concluded that machine learning models are good rental value predictive models. The precision and accuracy rate of the prediction potentials of the three machine learning models adopted is above 80%. The finding corroborates findings from previous studies. Also, the study established that the determinants of rental value are reliable predictors of rental value. Machine learning should be adopted by practitioners and academics as property value predictive models.

References

1. Adetunji, A. B, Akande, O. N, Ajala, F.A, Oyewo, O, Akande, Y. F and Oluwadara, G (2021) House Price Prediction Using Radom Forest Machine Learning Technique. *Procedia Computer Science 199* (2022), 806 – 813.

2. Abidoye, R. B., and Chan, A. P. C. (2016). Critical determinants of residential property value: professionals' perspective. Journal of facilities management, 14(3), 283-300
3. Abidoye, R. B., and Chan, A. P. C. (2017).Modeling property values in Nigeria using artificial neural network. Journal of Property Research, 1-18.
4. Ajibola, M.O., Oloke, O.C., and Ogungbemi, A.O. (2011) Impact of Gated Communities on Residential Property Values: A Comparison of Onipetesi Estate and its Neighbourhoods in Ikeja, Lagos State, Nigeria. Journal of Sustainable Development, 4(2), 063-071.
5. Ajide, K. and Alabi, M. (2010), "Does the functional form matter in the estimate of hedonic price model for housing market", *The Social Science*, Vol. 5 (6), 559-564.
6. Babawale, G., Koleoso, H. and Otegbulu, C. (2012), "A hedonic model for apartment rentals in Ikeja area of Lagos metropolis", *Mediterranean Journal of Social Sciences*, V (3), 109-120.
7. Chin, T.L. and Chau, K.W. (2002), "A critical review of literature on the hedonic price model", *International Journal for Housing Science and its Applications*, Vol. 27 No. 2, pp. 145-165.
8. Hepen, A. and Vatanserer, M. (2011) Forecasting Future Trends in Dubai Housing Market by Using Box-Jekins autoregressive integrated moving average. International Journal of Housing Market and Analysis4 (3):210- 223
9. Hong, J, Choi, H and Kim, W-S (2020) A House Price Valuation Based on the Random Forest Approach: The Mass Appraisal of Residential Property in South Korea. *International Journal of Strategic Property Management*, 24(3), 140 – 152.
10. Jenkins, D. (2000), *Residential Valuation Theory and Practice*, Chandos Publishing, Oxford.
11. Li, H., & Li, V. (1996). Forecasting house rental levels: Analytical rent model versus neural network. *Journal of urban planning and development*, *122*(4), 118-127.
12. Klimczak, K (2010) Determinants of Real Estate Investment, *Economics and Sociology*, 3(2), 58 – 66
13. NunezzTabales J.M, Caridad YOJ, Rey C F J (2013) Artificial Neural Network for Predicting Real Estate Prices, Revista de Metodos Cuantitativosparala economia y laempresa (15), 29 – 44
14. Olajide, S.E and Lizam, M (2016) Determining the Impact of Residential Neighborhood Crime on Housing Investment Using Logistic Regression. International Electronic Scientific Journal 2(12), 9 – 18
15. Oshodi, O, Oyedeji, O and Aigbovboa, C (2020) Does proximity to tourist site affect rental value of residential properties? Empirical evidence from Nigeria. Matec Web of Conferences 312 04003 https://doi.org/10.1051/matecconf/202031204003
16. Oyedeji, J.O (2018) Classification Modeling for the Impact of Historical Site on Residential Property Rental Value in Osogbo, Nigeria, *International Journal of Property Science*, 18(1), 13 – 26
17. Oyedeji, J.O, Oshodi, O.S and Oloke, O.C (2018) Property rental classification model: A case of Osogbo, Osun State, Nigeria, *Covenant Journal of Research in Built Environment*, 6(1), 52-64

18. Popoola, N.I, Jinadu, M.A, Liman, H.S &Abd'Razack, N.T.A (2015) Effectof Environmental Quality on Rental Values in Peri-urban Neighborhoods of Minna, Nigeria. *ATBU Journal of Environmental Technology*, 8(2), 42 - 52.
19. Wu, J (2017) House Price Prediction Using Support Vector Regression. Masters Project. 540 https://scholarworks.sjsu.edu/etd_projects/540
20. Zeffora, J and Shobarani, A.R (2019) Statistical Analysis of Random Forest on Real Estate Prediction. International Journal of Innovative Technology and Exploring Engineering, 8(8S), ISSN: 2278-3075