

## CSE 3505 FDA- Digital Assignment 2

Team 13 Prediction of Sales Prices of Houses

Doddamani Apurv Sangmesh 19BPS1012

Anushka Pandey 19BPS1048

Mithun P 19BPS1092

---

Sr N o	Title	Journal/ year of publicatio n	Data set used	Methodolo gies used	Metrics used	Interpretation of Results
1	Machine Learning based Predicting House Prices using Regression Techniques	Institute of Electrical and Electronics Engineers (IEEE),2020	House Prices in Bengaluru	Multiple Linear Regression, Ridge Regression, Lasso Regression, Support Vector Regression	R-Squared, RMSE, Root Mean Squared Logarithmic Error (RMSLE)	Multi linear Regression had the lowest RMSE value compared to others meaning it had low standard deviation between error and Actual value but had low R-Squared value which means certain values can't be explained. The SVR had the highest R-Squared explaining that dependent variable is explained collectively by the independent variables and comparatively good RMSLE hence would be the best model here.

2	Real Estate Price Prediction	INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT),2021	Boston Housing Dataset	XGB,Random Forest,Decision Tree, Linear Regression	Root Mean Square Error (RMSE) and Cross-Validation Score	Heat Map to find the Correlation helped to find the highly correlated values. The best accuracy was provided by the XGB as it had the lowest RMSE and lowest standard deviation for the error value then came the Random Forest followed by the Decision Tree.
3	Housing Price Prediction using Machine Learning Algorithms : The Case of Melbourne City, Australia	2018 International Conference on Machine Learning and Data Engineering (iCMLDE)	Melbourne City Housing Dataset	Regression Trees, Neural Network, Support Vector Machine, Stepwise, Principal component analysis	Mean Square Error(MSE) and Run Time	Linear regression will be used as a baseline for model evaluation, which based on Mean Squared Error (MSE) measured on an evaluation dataset. In other words, other dataset will be compared with linear regression for evaluation. The smaller evaluation ratio, the higher accuracy of the model's prediction. Regression tree delivers a prediction result as good as linear regression. The combination of Stepwise and tuned SVM, which produces the lowest error on this dataset, is the most competitive models.

4.	HOUSE PRICE PREDICTION USING VARIOUS REGRESSION: A COMPARATIVE STUDY	April 2020	They have web scrapped the Data from 99acres.com website which is one of the leading real estate websites operating in INDIA. And Data contains Bombay Houses only.	Regression Model, Random Forest Regression Model, XG Boost Regressor Model	Root Mean Squared Value (RMSE)	Study showed that Linear Regression displayed the best performance for this Dataset and can be used for deploying purposes.
5.	VALUATION OF HOUSE PRICES USING PREDICTIVE TECHNIQUES	Jun.-2018 International Journal of Advances in Electronics and Computer Science	Dataset used in this project was an open-source dataset from KaggleInc . It consists of 3000 records with 80 parameters that have the possibility of affecting the property prices	Logistic Regression, Support Vector Regression, Lasso Regression, Decision Tree	Performance metrics such as Accuracy, R-squared value, Root Mean Squared Value (RMSE), Mean Absolute Value (MAE) and Mean Squared Value (MSE)	Decision Tree gives a higher accuracy and R-squared value and low error values. On comparing the various models, they find that decision tree works the best with highest accuracy of 84.64% and Lasso performs least with an accuracy of 60.32%. Decision Tree produces hardly any error with RMSE value of 217 and Lasso performs worst with its RMSE value as 34245

6.	House Price Prediction	July 2020 International Journal of Computer Sciences and Engineering Open Access	They have collected house sales related data to estimate the house prices based on real world dataset IA. It is a public output dataset of that specified region in USA	Random Forest Regression, Ridge Regression, LASSO Regression, Decision Tree Regression, XGBoost Regression and they have used Ada-Boost algorithm for boosting up the weak learners to strong learners	RMSE, K-fold Cross-validation	From the analysis they set the threshold value of RMSE as 0.12 and integrate those algorithms (Ridge regression, Lasso regression, XGBoost regression) with RMSE value less than 0.12. This definitely increases the accuracy. In future this paper may help in the upcoming development of these areas.
7	A hybrid regression technique for house prices prediction	2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)	The dataset contains 79 explanatory variables for part of residential home transactions in Ames, Iowa, and opens to all to predict price of each covered home transaction SalePrice.	The paper also proposes a hybrid Lasso and Gradient boosting regression model to predict individual house price.	RMSE value is used to test the different hybrid combination	The combination of 65% Lasso + 35% XGB had the best RMSE value of 0.11539

8	Prediction of House Price Using XGBoost Regression Algorithm	January 2021 Turkish Journal of Computer and Mathematics Education (TURCOMAT)	The datasets are taken from Kaggle competition on house price prediction. Dataset consists of 80 data attributes and nearly 1500 records.	The dataset was divided into training, validation and testing and xgb regression was applied.	Mean square error (MSE)	When 90% of the dataset is set for training the model has the least MSE at below 5%.
9	Predicting property prices with machine learning algorithms	Journal of Property Research	Housing district in Hong Kong	Support vector machine, Random forest and gradient boosting methodologies were studied	Performance was evaluated using MSE, RMSE and MAPE (mean absolute percentage error)	It is found that advanced ML algorithms can be effectively used to accurately predict housing prices and gradient boosting machine algorithm had the best performance with the MSE, RMSE and MAPE are estimated to be 0.00793, 0.08903 and 0.32251%. But the paper concludes by saying that different algorithms need to be selected according to the dataset used

## REFERENCES:

1. J. Manasa, R. Gupta and N. S. Narahari, "Machine Learning based Predicting House Prices using Regression Techniques," 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2020, pp. 624-630, doi: 10.1109/ICIMIA48430.2020.9074952.

Link: <https://ieeexplore.ieee.org/document/9074952>

2. Smith Dabreo , Shaleel Rodrigues , Valiant Rodrigues , Parshvi Shah, 2021, Real Estate Price Prediction, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 04 (April 2021),

Link: <https://www.ijert.org/real-estate-price-prediction>

3. T. D. Phan, "Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia," 2018 International Conference on Machine Learning and Data Engineering (iCMLDE), 2018, pp. 35-42, doi: 10.1109/iCMLDE.2018.00017.

Link: <https://ieeexplore.ieee.org/document/8614000>

4. C. R. Madhuri, G. Anuradha and M. V. Pujitha, "House Price Prediction Using Regression Techniques: A Comparative Study," 2019 International Conference on Smart Structures and Systems (ICSSS), 2019, pp. 1-5, doi: 10.1109/ICSSS.2019.8882834.

Link: <https://ieeexplore.ieee.org/document/8882834>

5. [http://www.ijer.in/journal/journal\\_file/journal\\_pdf/12-477-153396274234-40.pdf](http://www.ijer.in/journal/journal_file/journal_pdf/12-477-153396274234-40.pdf)

6. [https://ijcseonline.org/pub\\_paper/16-IJCSE-08250-47.pdf](https://ijcseonline.org/pub_paper/16-IJCSE-08250-47.pdf)

7. Lu, Sifei & Li, Zengxiang & Qin, Zheng & Yang, Xulei & Goh, Rick. (2017). A hybrid regression technique for house prices prediction. 319-323. 10.1109/IEEM.2017.8289904.

Link: <https://ieeexplore.ieee.org/document/8289904>

8. Jangaraj, Avanija & Sunitha, Gurram & Madhavi, Reddy & Kora, Padmavathi & Hitesh, R & Associate, Sai. (2021). Prediction of House Price Using XGBoost Regression Algorithm. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 12. 2151-2155.

Link: [https://www.researchgate.net/publication/350810698\\_Prediction\\_of\\_House\\_Price\\_Using\\_XGBoost\\_Regression\\_Algorithm](https://www.researchgate.net/publication/350810698_Prediction_of_House_Price_Using_XGBoost_Regression_Algorithm)

9. Winky K.O. Ho, Bo-Sin Tang & Siu Wai Wong (2021) Predicting property prices with machine learning algorithms, Journal of Property Research, 38:1, 48-70, DOI: 10.1080/09599916.2020.1832558 Link: <https://www.tandfonline.com/doi/full/10.1080/09599916.2020.1832558>