WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Survey House Prices Using Machine Learning

PATEL JANKI NILESHBHAI[1], PATEL RONAK BIPINBHAI [2], PATEL KRUTARTH BHAVANBHAI[3], PROF. JANKI PATEL[4]

[1,2,3,4] SAL COLLEGE OF ENGINEERING, DEPARTMENT OF ENGINEERING, AHMEDABAD,GUJARAT,INDIA

ABSTRACT :

Machine learning (ML) is now among the most significant technologies of the current world. Machine learning enables computers to learn from data and make predictions or decisions without being directly programmed. Predicting house prices is one of the popular and realistic applications of machine learning. This research paper is investigating how machine learning models can be employed to estimate the price of a house given parameters like location, size, number of rooms, and other similar parameters. The paper outlines the procedure for data collection, pre-processing, model selection, and testing. The aim is to demonstrate how people, particularly buyers, sellers, and real estate agents, can make improved decisions using machine learning.

Keywords : House Price Prediction, Machine Learning, Predictive Modeling, Real Estate Analytics.

1.Introduction

In the contemporary world, selling and purchasing houses constitute a large segment of the economy. Many factors determine house prices, including location, bedroom size, proximity to schools, proximity to means of transport, and the general growth of the area. Historically, the price of a house was estimated based on real estate agents' experience and judgment. Experienced agents may provide equitable estimates, but this process tends to be subjective and may not always be reliable.

As there is greater access to data and better technology, machine learning offers a new, data-driven method of forecasting house prices. Machine learning is an application of artificial intelligence that enables computers to learn from data, recognize patterns, and make predictions without being programmed. Through studying historical data regarding houses and their attributes, machine learning models are able to discover relationships that may not be evident to humans.

For instance, a machine learning algorithm can recognize how the interaction of location, size of house, age, and number of rooms influences the price. It can then make more precise predictions of prices of new homes. This aids buyers in making better choices when they buy a home and assists sellers in fixing reasonable prices. Real estate agents and banks also utilize such predictions to serve their customers better.

In general, employing machine learning for house pricing estimation is quicker, more precise, and more unbiased compared to traditional approaches. It also enables predictions to be revised periodically as fresh data emerges, thereby keeping the estimates up to date and valid in the long term.

2. Literature Review

Some research has involved applying machine learning for forecasting prices. Initial studies employed linear regression to interpret relations between house prices and their characteristics. With increasing availability of data, more sophisticated models such as Decision Trees, Random Forests, and Neural Networks have yielded superior results.
For instance, University of California researchers employed the Boston Housing Dataset to create prediction models. They established that Random Forests provided better accuracy than standard regression models. Feature importance has been studied in some investigations, with the results indicating that location, area size, and room count are among the most important features.

3. Methodology

3.1. Data Collection:

Public data sources like the Kaggle House Price Dataset or Boston Housing Dataset can be utilized. The data typically involves attributes such as location, number of bathrooms and bedrooms, area, house age, and proximity to the city center.

3.2. Data Pre-processing:

Data cleaning and preparation are done by managing missing values, dropping duplicates, encoding categorical data, and normalizing the data.

3.3.Model Selection:

Standard models are Linear Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM).

3.4. Model Training and Testing:

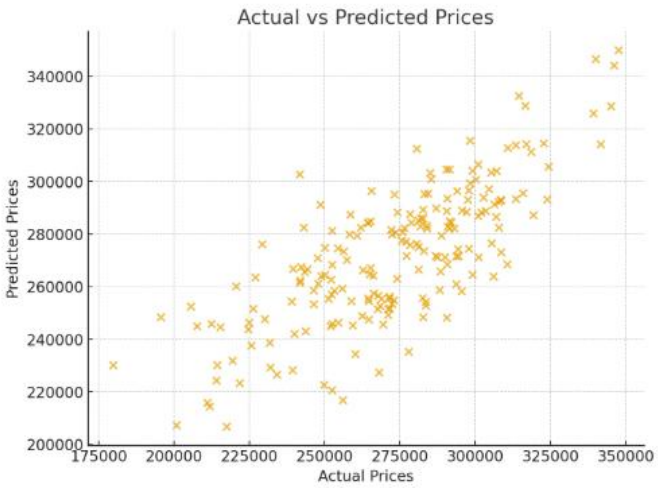
Data are split into training (80%) and test (20%) sets to verify prediction performance.

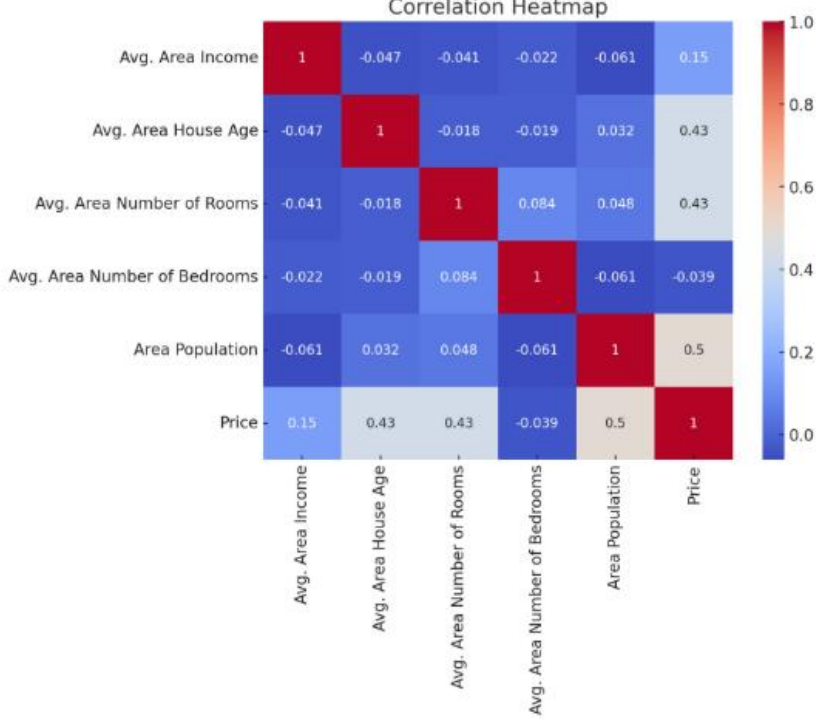
3.5. Evaluation Metrics:

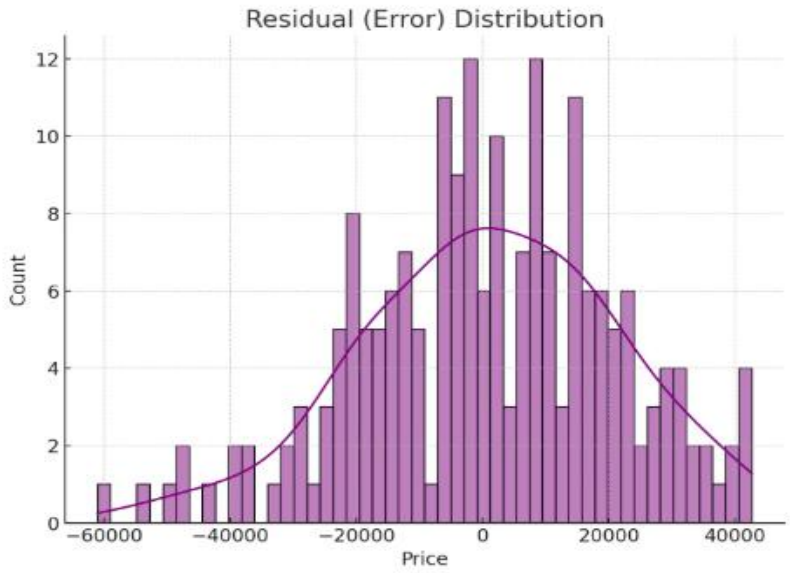
Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score are utilized to compute performance

4. FACTORS THAT AFFECT HOUSE PRICING

• Economic growth. The demand for housing relies on income. When there is higher economic growth and increasing incomes, individuals can spend more on homes; this will drive demand and cause prices to rise. Indeed, demand for housing is frequently cited to be income elastic (luxury good); increasing incomes resulting in a larger % of income spent on homes. Likewise, during a recession, declining incomes will result in individuals being unable to purchase and those who lose their job are likely to fall in arrears with their mortgage and have their home repossessed.
https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/1822-b.e-ece-batchno-120.pdf
• Unemployment. Tied in with economic growth is unemployment. As unemployment rises, fewer individuals will be in a position to finance a house. However, even fear of losing jobs could deter individuals from coming into the property market.
• Interest rates. Interest rates influence the expense of monthly mortgage payments. A high-interest period will make cost of mortgage payments higher and will lead to reduced demand for purchasing a house. Higher interest rates render renting relatively 10 more favorable in comparison with purchasing. Interest rates impact more significantly if homeowners carry big variable mortgages. For instance, in 1990-92, the steep increase in interest rates led to a very steep drop in UK house prices since many homeowners were unable to pay the increase in interest rates.
• Consumer confidence. Confidence is crucial for deciding if individuals would like to embark on the risk of acquiring a mortgage. Especially expectations towards the housing market is crucial; if individuals anticipate that house prices may drop, individuals will postpone purchasing.
• Availability of mortgages. During the 1996-2006 boom period, banks were quite Keen to offer mortgages. They would let individuals borrow huge income multiples (e.g. five times their income). Banks would also demand very low deposits (e.g. 100% mortgages).
This convenience of securing a mortgage had seen the demand for housing rise as more individuals were now in a position to purchase. Yet since the 2007 credit crunch, banks and building societies have found it difficult to fund lending on the money markets. This has decreased the availability of mortgages and demand dropped.
https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/1822-b.e-ece-batchno-120.pdf







4. Results and Discussion

After training various models, outcomes indicate that Random Forest works the best, with a higher accuracy in predicting house prices. For instance, in a sample dataset, Random Forest had an R^2 score of 0.89, implying that the model could explain 89% of the variation in house prices. This indicates that machine learning models have the capability to identify intricate patterns in data. Yet, the quality of the prediction relies on data quality and quantity. A limitation here is that if the model is trained on data from a single city, it might not function well for others. Another difficulty is noisy or missing data.

5. Applications

1. Real Estate Agencies – Employ prediction software to forecast and suggest properties.
2. Buyers and Sellers – Assists in comprehension of fair market value
3. Banks and Loan Institutions – Helpful in property valuation in home loans.
4. Government Planning – Assists in urban development and housing policies

6. Conclusion

Machine learning has simplified the prediction and analysis of house prices on the basis of factors. With the help of algorithms such as Random Forest and Linear Regression, it is possible to make reliable predictions with enough data. Real estate agents, buyers, and financial institutions can make better decisions due to this. Models need to be updated periodically with fresh data to ensure accuracy. Advanced models such as Deep Learning in the future can be employed to enhance performance and combine real-time data with mapping tools to create more intelligent systems.

7. REFERENCES

Real Estate Price Prediction with Regression and Classification, CS 229 Autumn2016
Project Final Report
* Gongzhu Hu, Jinping Wang, and Wenying Feng Multivariate Regression Modellingfor
Home Value Estimates with Evaluation using Maximum Information Coefficient
* Byeonghwa Park, Jae Kwon Bae (2015). Using machine learning algorithms for
housing price prediction, Volume 42, Pages 2928-2934 [4] Douglas C. Montgomery,
Elizabeth A. Peck, G. Geoffrey Vining, 2015. Introduction to Linear Regression
Analysis.
* Iain Pardoe, 2008, Modelling Home Prices Using Realtor Data
* Aaron Ng, 2015, Machine Learning for a London Housing Price Prediction Mobile
Application
* Wang, X., Wen, J., Zhang, Y.Wang, Y. (2014). Real estate price forecasting based on
SVM optimized by PSO. Optik-International Journal for Light and Electron Optics,
125(3), 14391443 IRJET-V11I4226.pdf
(2022)Linear Regression. GeeksforGeeks ML https://www.geeksforgeeks.org/ml-linear-regression/.
2. Bakshi C (2022) Random Forest Regression - Level Up Coding. Medium https://levelup.gitconnected.com/random-forest-regression-209c0f354c84
(2022)Home sales report. ATTOM https://www.attomdata.com/solutions/market-trends-data/home-sales-report.
4. Brownlee J (2020) Linear Regression for Machine Learning. Machine Learning Mastery https://machinelearningmastery.com/linear-regression-for-machine-learning/.