

	Cite	Abstract
1	@article{Zong2023HousePP, title={House Prices Prediction Advanced Regression Techniques}, author={Yue Zong}, journal={Advances in Economics, Management and Political Sciences}, year={2023}, url={https://api.semanticscholar.org/CorpusID:265629120} }	<p>House Prices Prediction Advanced Regression Techniques</p> <p>A neural network model was built, parameters were adjusted, and the trained network was used to predict housing prices, positioning the model among the top performers among all machine learning methods.</p> <p>Abstract</p> <p>In recent years, with the development of the real estate industry, housing prices have continued to rise. The nation, society, and individuals are all concerned about these prices. For commodity housing prices, there are many factors that influence the housing prices. Apart from national regulations, factors such as lighting, layout, and environment of the houses themselves also have a certain impact on the prices, leading to significant flvuctuations in the real estate market. Therefore, researching an accurate model for predicting housing prices has practical significance. It can guide residents in housing consumption and provide policy recommendations for government price regulation. Machine learning methods have become a new type of prediction method in this regard. Based on the theories of data analysis and machine learning, a dataset consisting of 2920 data points with 81 attributes was selected from the publicly available Kaggle housing dataset. The data was normalized and analyzed for feature selection. The ranking of attributes most correlated with housing prices was obtained. Subsequently, a neural network model was built, parameters were adjusted, and the trained network was used to predict housing prices. On the Kaggle leaderboard, the RMSE test result stands at 0.1198, positioning our model among the top performers among all machine learning methods.</p>

2	<pre>@article{KHARB2023ForestallingHP, title={Forestalling House Prices Using Linear Regression in Machine Learning}, author={LATIKA KHARB and Deepak Chahal}, journal={International Research Journal of Modernization in Engineering Technology and Science}, year={2023}, url={https://api.semanticscholar.org/CorpusID:260941791} }</pre>	<p>Forestalling House Prices Using Linear Regression in Machine Learning</p> <p>This paper focuses on using linear regression to predict house prices, discussing the dataset, preprocessing steps, model training, and evaluation and highlights the strengths, limitations, and potential improvements of the linear regression approach in the context of house price prediction.</p> <p>Abstract</p> <p>Predicting house prices is a fundamental problem in the field of real estate and economics. The accurate estimation of house prices is crucial for various stakeholders including homeowners, buyers, sellers, and policymakers. In this paper, we explore the application of linear regression, a simple yet effective machine learning technique, to predict house prices. We discuss the methodology, data preprocessing, feature selection, model training, and evaluation metrics. We also highlight the strengths, limitations, and potential improvements of the linear regression approach in the context of house price prediction. In this paper, we focus on using linear regression to predict house prices, discussing the dataset, preprocessing steps, model training, and evaluation.</p>
	<pre>@article{Nwankwo2023PredictionOH, title={Prediction of House Prices in Lagos-Nigeria Using Machine Learning Models}, author={Mmesoma Peace Nwankwo and Ndukaku Macdonald Onyeizu and Emmanuel Chibuogu Asogwa and Chukwuogo Okwuchukwu Ejike and Okechukwu J. Obulezi}, journal={European Journal of Theoretical and Applied Sciences}, year={2023}, url={https://api.semanticscholar.org/CorpusID:262169298} }</pre>	<p>Prediction of House Prices in Lagos-Nigeria Using Machine Learning Models</p> <p>The model shows that the number of bedrooms, parking spaces, and house types play an important role in determining the price of houses.</p> <p>Abstract</p>

	}	<p>This paper considers the relationship between the price of houses and the features namely the number of bedrooms, parking space, and different house types. In this study, a machine learning approach was used to develop prediction models that predicted house prices in Lagos. Different machine learning techniques were used, train-test split to split the data into training sets for training and building the model and test data to test the accuracy of the model, performance metric mean absolute error to set the baseline for the model, Variance Inflation Factor (VIF) to help remove multicollinearity between features and Streamlit interactive dashboards to communicate with the model. Correlation and regression methods were used to examine the relationship and build the model. It is observed that there is a strong positive correlation between the number of bedrooms and the number of toilets, likewise the number of bedrooms and the number of bathrooms. It also shows that there is a moderate positive correlation between the number of bedrooms and price. The model shows that the number of bedrooms, parking spaces, and house types play an important role in determining the price of houses.</p>
4	<pre>@article{Sundari2023OptimizationHP, title={Optimization house price prediction model using gradient boosted regression trees (GBRT) and xgboost algorithm}, author={Putri Susi Sundari and Mahardika Khafidz Putra}, journal={Journal of Student Research Exploration}, year={2023}, url={https://api.semanticscholar.org/CorpusID:263632752} }</pre>	<p>Optimization house price prediction model using gradient boosted regression trees (GBRT) and xgboost algorithm</p> <p>The researcher utilizes ensemble learning techniques, combining the Gradient Boosted Regression Trees (GBRT) and XGBoost algorithms, to develop a predictive model that can accurately estimate house prices based on relevant features or variables.</p> <p>Abstract</p> <p>In this rapidly advancing technological era, the demand for the real estate industry has also increased, including in the field of house price prediction. House prices fluctuate every</p>

		<p>year due to several factors such as changes in land prices, location, year of construction, infrastructure developments, and other factors. Numerous studies have been conducted on this issue. However, the challenge lies in building a proven accurate and effective model for predicting house prices with the abundance of features present in the dataset. The objective of this research is to develop a predictive model that can accurately estimate house prices based on relevant features or variables. The researcher utilizes ensemble learning techniques, combining the Gradient Boosted Regression Trees (GBRT) and XGBoost algorithms. The dataset used in this article is titled "Ames Housing dataset" obtained from Kaggle. The predictive model is then evaluated using the Root Mean Squared Error (RMSE) method. The RMSE result from a previous study that used the combination of Lasso and XGBoost was 0.11260, while the RMSE result from this research is 0.00480. This indicates a decrease in the RMSE value, indicating a lower level of error in the model. It also means that the combination of GBRT and XGBoost algorithms successfully improves the prediction accuracy of the previous research model.</p>
5	<pre>@article{Jayadharshini2023MachineLT, title={Machine learning techniques for predicting home rental prices in India}, author={P Jayadharshini and S. Santhiya and S Keerthika and N Abinaya and S Priyanka}, journal={Applied and Computational Engineering}, year={2023}, url={https://api.semanticscholar.org/CorpusID:266556584} }</pre>	<p>Machine learning techniques for predicting home rental prices in India</p> <p>Predicting the selling price of houses has become increasingly crucial as land and housing prices rise annually. This task is particularly challenging for metropolitan areas like Chennai and Bangalore. Therefore, there is a growing demand for an easier and more effective approach to forecast house rental prices, ensuring future generations have access to reliable predictions. Several key factors, such</p>

		<p>as the house's location and area, significantly influence rental prices. In this paper, a dataset comprising ten similar crucial features is utilized. The model is developed using a Python library, where the data is preprocessed and prepared to ensure cleanliness for constructing the model. Various machine learning algorithms, including Random Forest, Linear Regression, Decision Tree Regression, and Gradient Boosting, are employed. Through feature extraction, it is determined that area and property type are the most important features that significantly impact rental prices. Among the techniques used, gradient boosting yields the most satisfactory predictive results for rent based on evaluation metrics like Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-Squared Metric (R2).</p>
6	<pre>@article{Yamur2023HousePP, title={House price prediction modeling using machine learning techniques: a comparative study}, author={Ayten Yağmur and Mehmet Kayakuş and Mustafa Ender Terzioğlu}, journal={Aestimum}, year={2023}, url={https://api.semanticscholar.org/CorpusID:257470072} }</pre>	<p>House price prediction modeling using machine learning techniques: a comparative study</p> <p>The artificial neural network (ANN) method made predictions with more meaningful results compared to support vector regression (SVR) and multiple linear regression (MLR) and appears to be a viable model for institutions that supply housing, mediate housing sales, and provide housing financing and valuation.</p> <p>Abstract</p> <p>In the literature, there are two basic approaches regarding the determination of house prices. One of them is the prediction of house price using macroeconomic variables in the country where the house is produced, and another one is the price prediction models, which we can express as micro-variables, by considering the features of the house. In this study, the price of the house was attempted to be predicted using machine learning methods by establishing a model with micro variables that reveal the features of the house. The</p>

		<p>study was conducted in Turkey' Antalya province, where household housing demand of foreigners is also high. The house advertisements in locations belonging to the lower, middle- and upper-income groups were selected as the sample. In the results, it was observed that the artificial neural network (ANN) method made predictions with more meaningful results compared to support vector regression (SVR) and multiple linear regression (MLR). These results appear to be a viable model for institutions that supply housing, mediate housing sales, and provide housing financing and valuation. It is considered that this model, which can be used to predict fluctuating house prices, especially in developing countries, will regulate the housing market.</p>
7	<pre>@article{Basha2023HarnessingDI, title={Harnessing Data-Driven Insights: Predictive Modeling for Diamond Price Forecasting using Regression and Classification Techniques}, author={Md. Shaik Amzad Basha and Peerzadah Mohammad Oveis}, journal={International Journal on Recent and Innovation Trends in Computing and Communication}, year={2023}, url={https://api.semanticscholar.org/CorpusID:265713400} }</pre>	<p>Harnessing Data-Driven Insights: Predictive Modeling for Diamond Price Forecasting using Regression and Classification Techniques</p> <p>This study delves deep into predicting diamond prices in terms of exact monetary values and broader price categories, and highlights the superior performance of the Random Forest regressor in predicting exact prices with an R2 value of approximately 0.975.</p> <p>Abstract</p> <p>In the multi-faceted world of gemology, understanding diamond valuations plays a pivotal role for traders, customers, and researchers alike. This study delves deep into predicting diamond prices in terms of exact monetary values and broader price categories. The purpose was to harness advanced machine learning techniques to achieve precise estimations and categorisations, thereby assisting</p>

		<p>stakeholders in informed decision-making. The research methodology adopted comprised a rigorous data preprocessing phase, ensuring the data's readiness for model training. A range of sophisticated machine learning models were employed, from traditional linear regression to more advanced ensemble methods like Random Forest and Gradient Boosting. The dataset was also transformed to facilitate classification into predefined price tiers, exploring the viability of models like Logistic Regression and Support Vector Machines in this context. The conceptual model encompasses a systematic flow, beginning with data acquisition, transitioning through preprocessing, regression, and classification analyses, and culminating in a comparative study of the performance metrics. This structured approach underscores the originality and value of our research, offering a holistic view of diamond price prediction from both regression and classification lenses. Findings from the analysis highlighted the superior performance of the Random Forest regressor in predicting exact prices with an R^2 value of approximately 0.975. In contrast, for classification into price tiers, both Logistic Regression and Support Vector Machines emerged as frontrunners with an accuracy exceeding 95%. These results provide invaluable insights for stakeholders in the diamond industry, emphasising the potential of machine learning in refining valuation processes.</p>
--	--	--

8	<pre>@article{Chen2022PredictionAA, title={Prediction and Analysis of House Prices in Boston Based on Regression Model}, author={Zuohang Chen}, journal={Proceedings of the 2022 4th International Conference on E-Business and E-Commerce Engineering}, year={2022}, url={https://api.semanticscholar.org/CorpusID:261583319} }</pre>	<h3>Prediction and Analysis of House Prices in Boston Based on Regression Model</h3> <p>The housing price data set of Boston is divided and the regression model linear regression, decision tree regression and support vector machine regression SVR is built so as to obtain the relationship between different data related to Boston house price, and this relationship can finally predict the future house price trend in Boston and display it through visual operation.</p> <p>Abstract</p> <p>In artificial intelligence learning, Boston housing price forecast analysis problem is a classic regression problem. Based on the housing price information collected by the U.S. Census Bureau in Boston, Massachusetts. this paper divides the housing price data set of Boston and builds the regression model linear regression, decision tree regression and support vector machine regression SVR and trains the data set, so as to obtain the relationship between different data related to Boston house price, and use this relationship to connect all data, it can finally predict the future house price trend in Boston and display it through visual operation. Through three regression model prediction value, respectively compared with the actual value, the trend of overall and actual and estimated values of the same, but there is a certain error, especially when spot prices higher or lower, often cannot get accurate forecast, so the data for the selection of the characteristic value still exists space for improvement, future study needs to get more data and the characteristics of abundant data.</p>
---	--	--

9	<pre>@article{Chu2023CovariatedistanceWR, title={Covariate-distance Weighted Regression (CWR): A Case Study for Estimation of House Prices}, author={Hone-Jay Chu and Po-Hung Chen and Sheng-Mao Chang and Muhammad Zeeshan Ali and Sumriti Ranjan Patra}, journal={ArXiv}, year={2023}, volume={abs/2305.08887}, url={https://api.semanticscholar.org/CorpusID:258714640} }</pre>	<p>Covariate-distance Weighted Regression (CWR): A Case Study for Estimation of House Prices</p> <p>Geographically weighted regression (GWR) is a popular tool for modeling spatial heterogeneity in a regression model. However, the current weighting function used in GWR only considers the geographical distance, while the attribute similarity is totally ignored. In this study, we proposed a covariate weighting function that combines the geographical distance and attribute distance. The covariate-distance weighted regression (CWR) is the extension of GWR including geographical distance and attribute distance. House prices are affected by numerous factors, such as house age, floor area, and land use. Prediction model is used to help understand the characteristics of regional house prices. The CWR was used to understand the relationship between the house price and controlling factors. The CWR can consider the geological and attribute distances, and produce accurate estimates of house price that preserve the weight matrix for geological and attribute distance functions. Results show that the house attributes/conditions and the characteristics of the house, such as floor area and house age, might affect the house price. After factor selection, in which only house age and floor area of a building are considered, the RMSE of the CWR model can be improved by 2.9%-26.3% for skyscrapers when compared to the GWR. CWR can effectively reduce estimation errors from traditional spatial regression models and provide novel and feasible models for spatial estimation.</p>
---	--	---

10	<pre>@article{Saeed2023PredictiveMO, title={Predictive Modeling of House Prices; A Regression Approach}, author={Natasha Saeed and Ahsan Saeed and Maria Bashir and Lismer Andres Caceres-Najarro}, journal={2023 18th International Conference on Emerging Technologies (ICET)}, year={2023}, pages={159-164}, url={https://api.semanticscholar.org/CorpusID:266738871} }</pre>	<p>Predictive Modeling of House Prices; A Regression Approach</p> <p>This study compares Random Forest Regressor and the Decision Tree Regressor using an extensive dataset of Bangalore that includes nine property features and shows that the models achieved an optimal accuracy level and provide insights into the variances in model performance.</p> <p>Abstract</p> <p>Housing prices have always been an important issue among the business community. The sudden increase in property prices has an impact on economic cycles as well as our means of living. However, understanding several features that affect residential real estate prices is still challenging. Traditional techniques have been unable to choose essential features for precise predictions, which frequently leads to reduced prediction accuracy. In this research work, we utilized machine learning techniques to increase prediction accuracy and get data-driven insights into property evaluation. This study compares Random Forest Regressor and the Decision Tree Regressor using an extensive dataset of Bangalore that includes nine property features. The findings show that our models achieved an optimal accuracy level and provide insights into the variances in model performance, assisting policymakers in choosing the most suitable approaches to get accurate house prices across various market circumstances.</p>
----	--	---

11	<pre>@article{Li2023ForecastingHR, title={Forecasting House Resale Prices using Ensemble learning}, author={Zhenyu Li and Zhaoji Li}, journal={2023 2nd International Conference on Big Data, Information and Computer Network (BDICN)}, year={2023}, pages={107-111}, url={https://api.semanticscholar.org/CorpusID:258379049} }</pre>	<p>Forecasting House Resale Prices using Ensemble learning</p> <p>This paper proposes a blending model based on Python and its xgboost, DF21, and Geatpy packages to predict resale prices in Singapore that can obtain a mean absolute percentage error (MAPE) of 7.36% in a stationary overall trend of house prices trend.</p> <p>Abstract</p> <p>Forecasting House sales prices is a critical issue for individual investors and government. prediction system has high accuracy and flexibility and plays an important role in decision making. However, previous studies have focused only on limited datasets, as well as nontime series or feature prediction. This paper proposes a blending model based on Python and its xgboost, DF21, and Geatpy packages to predict resale prices in Singapore. First, the high-cardinality categorical attributes are preprocessed by means of mean coding. Then, we propose a linear blending method that includes GA-HL-Xg-Boost, GARandom Forest (GA-RF), deep-Random Forest (DRF) and lightGBM, with Gini impurity to determine the importance of features. Finally, the result shows that it can obtain a mean absolute percentage error (MAPE) of 7.36% in a stationary overall trend of house prices trend. Our research can provide a powerful predictor of house resale prices in different economic environments.</p>
----	---	--

12	<pre>@article{Miao2021SupportVR, title={Support Vector Regression with Gaussian kernel for Housing Prices Prediction}, author={Dingyang Miao and Hongru Tang and Boshen Wang}, journal={Journal of Physics: Conference Series}, year={2021}, volume={1994}, url={https://api.semanticscholar.org/CorpusID:236974103} }</pre>	<p>Support Vector Regression with Gaussian kernel for Housing Prices Prediction</p> <p>The support vector regression (SVR) with Gaussian kernel is developed to predict housing prices and demonstrates superior accuracy and effectiveness compared with the SVR with linear kernel, KNN, and decision tree.</p> <p>Abstract</p> <p>The housing sector is one of the main sources of economic growth in both developing and developed countries. It is reported that nearly half of people buy or sell houses at an inappropriate price. Based on the public data set of Boston housing prices, this essay analyzed the factors affecting house prices and selected the five most important factors based on the decision tree with the ID3 algorithm. Then, this essay developed the support vector regression (SVR) with Gaussian kernel to predict housing prices. Experimental results showed that our method achieves superior accuracy and effectiveness compared with the SVR with linear kernel, KNN, and decision tree. To verify the applicability of our model, this research applied this model in Beijing housing price data, and it also achieved satisfactory fitting results.</p>
13	<pre>@article{Lv2022AnalysisAF, title={Analysis and Forecast of Influencing Factors on House Prices Based on Machine Learning}, author={Chenyue Lv and Yingxin Liu and Lidong Wang}, journal={2022 Global Conference on Robotics, Artificial Intelligence and Information Technology (GCRAIT)}, year={2022}, pages={97-101}, url={https://api.semanticscholar.org/CorpusID:252720096} }</pre>	<p>Analysis and Forecast of Influencing Factors on House Prices Based on Machine Learning</p> <p>The analysis of the influencing factors and trends of Shenzhen's commercial house prices can provide reasonable policy suggestions for the development of the real estate</p>

	<pre> }</pre>	<p>industry and construction of long-term regulation of real estate in the future.</p> <p>Abstract</p> <p>Taking Shenzhen as an example, based on the relevant quantitative data affecting commercial house prices in Shenzhen from 2001 to 2020, the main factors affecting commercial house prices in Shenzhen were screened out by using Lasso regression, and grey prediction model was constructed by using the data passing the grade ratio test to predict the data of each variable in the next 5 years. Afterwards, a Support Vector Regression model was constructed to predict the trend of commercial house prices of Shenzhen based on the values of each variable from 2001 to 2025. The results of the study show that the main influencing factors of Shenzhen commercial house prices are the sales area of residential commercial housing, the purchase and sale prices of second-hand houses, etc. The growth of commercial house prices in Shenzhen from 2021 to 2025 is greater and faster compared to previous years. The analysis of the influencing factors and trends of Shenzhen's commercial house prices can provide reasonable policy suggestions for the development of the real estate industry and construction of long-term regulation of real estate in the future.</p>
14	<pre> @article{Zhang2023ResearchOT, title={Research on the Influencing Factors and Prediction of Housing Prices Based on Regression Analysis --- Taking California as an Example}, author={Kunlong Zhang}, journal={Advances in Economics, Management and Political Sciences}, year={2023}, url={https://api.semanticscholar.org/CorpusID:261784296} }</pre>	<p>Research on the Influencing Factors and Prediction of Housing Prices Based on Regression Analysis --- Taking California as an Example</p> <p>A house is a necessity for everyone's life. But in today's world, with the continuous reduction of land area, the continuous increase of population, and the continuous maturity of the real estate industry, people have to consider more carefully</p>

		<p>what factors are the most important in the house. Based on the background of housing prices in California, this paper uses linear regression, random forest and principal component analysis to determine which variables have the greatest impact on housing prices. The reason for using three methods is to obtain more accurate results. According to the results, linear regression shows that income is the most relevant variable to housing prices. Random Forest shows an R square of 75.7%, meaning the predictions fit the data fairly well. Principal component analysis also shows that income is the most important variable. At the same time, house prices will be predicted based on the data obtained, which is about 116598.48588. Some suggestions for those who want to buy a house - which factors of the house are more important. Second, this article will also analyse how to face the rising housing prices.</p>
15	<pre>@article{Afonso2019HousingPP, title={Housing Prices Prediction with a Deep Learning and Random Forest Ensemble}, author={Bruno Klaus de Aquino Afonso and Luckeciano Carvalho Melo and Willian Oliveira and Samuel Bruno da Silva Sousa and Lilian Berton}, journal={Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2019)}, year={2019}, url={https://api.semanticscholar.org/CorpusID:214699727} }</pre>	<p>Housing Prices Prediction with a Deep Learning and Random Forest Ensemble</p> <p>This study analyzes a dataset composed of 12,223,582 housing advertisements, collected from Brazilian websites from 2015 to 2018, and demonstrates that enriching the dataset and combining different ML approaches can be a better alternative for prediction of housing prices in Brazil.</p> <p>Abstract</p> <p>The development of a housing prices prediction model can assist a house seller or a real estate agent to make better-informed decisions based on house price valuation. Only a few works report the use of machine learning (ML) algorithms to predict the values of properties in Brazil. This study analyzes a dataset composed of 12,223,582 housing advertisements, collected from Brazilian websites from 2015 to 2018. Each instance comprises twenty-four features of five</p>

		<p>different data types: integer, date, string, float, and image. To predict the property prices, we ensemble two different ML architectures, based on Random Forest (RF) and Recurrent Neural Networks (RNN). This study demonstrates that enriching the dataset and combining different ML approaches can be a better alternative for prediction of housing prices in Brazil.</p>
16	<pre>@article{Zong2023HousePP, title={House Prices Prediction Advanced Regression Techniques}, author={Yue Zong}, journal={Advances in Economics, Management and Political Sciences}, year={2023}, url={https://api.semanticscholar.org/CorpusID:265629120} }</pre>	<p>House Prices Prediction Advanced Regression Techniques</p> <p>A neural network model was built, parameters were adjusted, and the trained network was used to predict housing prices, positioning the model among the top performers among all machine learning methods.</p> <p>Abstract</p> <p>In recent years, with the development of the real estate industry, housing prices have continued to rise. The nation, society, and individuals are all concerned about these prices. For commodity housing prices, there are many factors that influence the housing prices. Apart from national regulations, factors such as lighting, layout, and environment of the houses themselves also have a certain impact on the prices, leading to significant fluctuations in the real estate market. Therefore, researching an accurate model for predicting housing prices has practical significance. It can guide residents in housing consumption and provide policy recommendations for government price regulation. Machine learning methods have become a new type of prediction method in this regard. Based on the theories of data analysis and machine learning, a dataset consisting of 2920 data points with 81 attributes was selected from the publicly available Kaggle housing dataset. The data was normalized</p>

		<p>and analyzed for feature selection. The ranking of attributes most correlated with housing prices was obtained. Subsequently, a neural network model was built, parameters were adjusted, and the trained network was used to predict housing prices. On the Kaggle leaderboard, the RMSE test result stands at 0.1198, positioning our model among the top performers among all machine learning methods.</p>
17	<pre> title={Foretelling House Prices Using Linear Regression in Machine Learning}, author={LATIKA KHARB and Deepak Chahal}, journal={International Research Journal of Modernization in Engineering Technology and Science}, year={2023}, url={https://api.semanticscholar.org/CorpusID:260941791} } </pre>	<p>Foretelling House Prices Using Linear Regression in Machine Learning</p> <p>This paper focuses on using linear regression to predict house prices, discussing the dataset, preprocessing steps, model training, and evaluation and highlights the strengths, limitations, and potential improvements of the linear regression approach in the context of house price prediction.</p> <p>Abstract</p> <p>Predicting house prices is a fundamental problem in the field of real estate and economics. The accurate estimation of house prices is crucial for various stakeholders including homeowners, buyers, sellers, and policymakers. In this paper, we explore the application of linear regression, a simple yet effective machine learning technique, to predict house prices. We discuss the methodology, data preprocessing, feature selection, model training, and evaluation metrics. We also highlight the strengths, limitations, and potential improvements of the linear regression approach in the context of house price prediction. In this paper, we focus on using linear regression to predict house prices, discussing the dataset, preprocessing steps, model training, and evaluation.</p>

18	<pre>@article{Mao2022ASO, title={A Study on the Prediction of House Price Index in First-Tier Cities in China Based on Heterogeneous Integrated Learning Model}, author={Yaqi Mao and Yonghui Duan and Yibin Guo and Xiang Wang and Shen Gao}, journal={Journal of Mathematics}, year={2022}, url={https://api.semanticscholar.org/CorpusID:252522835} }</pre>	<p>A Study on the Prediction of House Price Index in First-Tier Cities in China Based on Heterogeneous Integrated Learning Model</p> <p>The state-of-the-art interpretation method SHAP combined with the LightGBM algorithm is used to interpret the model, and the result can be used as a basis for real estate policy planning and adjustment and even guide the demand of home buyers, thus improving the efficiency and effectiveness of government policy making.</p> <p>AbstractTo address the difficulty of low prediction accuracy, insufficient model stability, and certain lag associated with a single machine learning model in the prediction of house price, this paper proposes a multimodel fusion house price prediction model based on stacking integrated learning. Firstly, web search data affecting house prices were collected by web crawler technology, and Spearman correlation analysis was performed on the attribute set to reduce its complexity and establish a prediction index system for four first-tier cities in China. Secondly, with the goal of improving accuracy, diversity, and generalization ability, the types of base learners as well as metalearners are determined, and the parameters of the base learners are optimized using the grey wolf optimization algorithm to produce the GWO-stacking model, and the experimental results from four datasets demonstrate that the model has high prediction accuracy. Finally, to solve the issue of unintelligible black boxes in machine learning models, we have used the state-of-the-art interpretation method SHAP combined with the LightGBM algorithm to interpret the model, and the result can be used as a basis for real estate policy planning and adjustment and even guide the demand of home buyers, thus improving the</p>
----	--	---

		efficiency and effectiveness of government policy making.
19	<pre>@inproceedings{Prakash2020GUIBH, title={GUI Based Housing Price Prediction Using Machine Learning Approach}, author={T. Jaya Prakash and S. Ashok Kumar}, year={2020}, url={https://api.semanticscholar.org/CorpusID:219676309} }</pre>	<p>GUI Based Housing Price Prediction Using Machine Learning Approach</p> <p>The object is to break down gadget reading based absolutely techniques for Housing value Prediction guaging through forecast impacts in top notch precision with a far reaching manual affectability assessment of rendition parameters regarding by and large execution in expectation of air pleasant poisons by utilizing precision count.</p> <p>AbstractLiving arrangement costs development consistently, it's miles a need for a machine to are expecting habitation charges inside what's to come. Habitation rate expectation can help the designer choose the advancing cost of a house and may assist the supporter with arranging the opportune time to purchase a home. To spare you this difficulty in affect the cost of a living arrangement which comprise of physical circumstances, thought and region. The object is to break down gadget reading based absolutely techniques for Housing value Prediction guaging through forecast impacts in top notch precision. The assessment of dataset with the guide of managed Machine Learning acing approach (SMLT) to catch a few measurements resembles, variable ID, uni-variate assessment, bi-variate and multi-variate examination, missing cost medicines and break down the information approval, realities cleaning/making prepared and records representation should be possible on the total given dataset. Our examination shows a far reaching manual affectability assessment of rendition parameters regarding by and large execution in</p>

		<p>expectation of air pleasant poisons by utilizing precision count. To prescribe a gadget picking up information on based technique to precisely expect the house charge forecast outcomes inside the state of palatable exactness from assessing oversee classification framework acing calculations. Besides, to assess and talk about the general execution of assorted Machine Learning calculations GUI based absolutely UI house value forecast by utilizing traits.</p>
20	<pre>@article{Wang2021DeepLM, title={Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism}, author={Pei-Ying Wang and Chiao-Ting Chen and Jain-Wun Su and Tingwei Wang and Szu-Hao Huang}, journal={IEEE Access}, year={2021}, volume={9}, pages={55244-55259}, url={https://api.semanticscholar.org/CorpusID:233262582} }</pre>	<p>Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism</p> <p>This work proposes an end to end joint self-attention model for house price prediction that incorporates attention mechanism and STN network and achieves a low prediction error and outperforms the other models. AbstractHouse price prediction is a popular topic, and research teams are increasingly performing related studies by using deep learning or machine learning models. However, because some studies have not considered comprehensive information that affects house prices, prediction results are not always sufficiently precise. Therefore, we propose an end to end joint self-attention model for house prediction. In this model, we import data on public facilities such as parks, schools, and mass rapid transit stations to represent the availability of amenities, and we use satellite maps to analyze the environment surrounding houses. We adopt attention mechanisms, which are widely used in image, speech, and translation tasks, to identify crucial features that are considered by prospective house buyers. The model can automatically assign weights when given transaction data. Our proposed model differs from self-</p>

		<p>attention models because it considers the interaction between two different features to learn the complicated relationship between features in order to increase prediction precision. We conduct experiments to demonstrate the performance of the model. Experimental data include actual selling prices in real estate transaction data for the period from 2017 to 2018, public facility data acquired from the Taipei and New Taipei governments, and satellite maps crawled using the Google Maps application programming interface. We utilize these datasets to train our proposed and compare its performance with that of other machine learning-based models such as Extreme Gradient Boosting and Light Gradient Boosted Machine, deep learning, and several attention models. The experimental results indicate that the proposed model achieves a low prediction error and outperforms the other models. To the best of our knowledge, we are the first research to incorporate attention mechanism and STN network to conduct house price prediction</p>
21	<pre>@article{Zulkifley2020HousePP, title={House Price Prediction using a Machine Learning Model: A Survey of Literature}, author={Nor Hamizah Zulkifley and Shuzlina Abdul Rahman and Nor Hasbiah Ubaidullah and Ismail Ibrahim}, journal={International Journal of Modern Education and Computer Science}, year={2020}, volume={12}, pages={46-54}, url={https://api.semanticscholar.org/CorpusID:230528609} }</pre>	<p>House Price Prediction using a Machine Learning Model: A Survey of Literature</p> <p>The findings of this analysis verified the use of the Artificial Neural Network, Support Vector Regression and XGBoost as the most efficient models compared to others and suggest that locational attributes and structural attributes are prominent factors in predicting house prices.</p> <p>AbstractData mining is now commonly applied in the real estate market. Data mining's ability to extract relevant knowledge from raw data makes it very useful to predict house prices, key housing attributes, and many more. Research has stated that the fluctuations in</p>

		<p>house prices are often a concern for house owners and the real estate market. A survey of literature is carried out to analyze the relevant attributes and the most efficient models to forecast the house prices. The findings of this analysis verified the use of the Artificial Neural Network, Support Vector Regression and XGBoost as the most efficient models compared to others. Moreover, our findings also suggest that locational attributes and structural attributes are prominent factors in predicting house prices. This study will be of tremendous benefit, especially to housing developers and researchers, to ascertain the most significant attributes to determine house prices and to acknowledge the best machine learning model to be used to conduct a study in this field.</p>
22	<pre>@article{Thamarai2020HousePP, title={House Price Prediction Modeling Using Machine Learning}, author={M. Thamarai and S. Malarvizhi}, journal={International Journal of Information Engineering and Electronic Business}, year={2020}, url={https://api.semanticscholar.org/CorpusID:216654981} }</pre>	<p>House Price Prediction Modeling Using Machine Learning</p> <p>House availability based on desired features of the house and house price prediction are modeled in the proposed work and the model is constructed for a small town in West Godavari district of Andhrapradesh.</p> <p>Abstract—Machine Learning is seeing its growth more rapidly in this decade. Many applications and algorithms evolve in Machine Learning day to day. One such application found in journals is house price prediction. House prices are increasing every year which has necessitated the modeling of house price prediction. These models constructed, help the customers to purchase a house suitable for their need. Proposed work makes use of the attributes or features of the houses such as number of bedrooms available in the house, age of the house, travelling facility from the location, school facility available nearby the houses and Shopping malls</p>

		<p>available nearby the house location. House availability based on desired features of the house and house price prediction are modeled in the proposed work and the model is constructed for a small town in West Godavari district of Andhrapradesh. The work involves decision tree classification, decision tree regression and multiple linear regression and is implemented using Scikit-Learn Machine Learning Tool.</p>
23	<pre>@article{Das2020BoostingHP, title={Boosting house price predictions using geo-spatial network embedding}, author={Sarkar Snigdha Sarathi Das and Mohammed Eunus Ali and Yuan-Fang Li and Yong-Bin Kang and Timos K. Sellis}, journal={Data Mining and Knowledge Discovery}, year={2020}, volume={35}, pages={2221 - 2250}, url={https://api.semanticscholar.org/CorpusID:221397667} }</pre>	<p>Boosting house price predictions using geo-spatial network embedding</p> <p>A novel method is presented, the geo-spatial network embedding (GSNE), that learns the embeddings of houses and various types of points of interest in the form of multipartite networks, where the houses and the POIs are represented as attributed nodes and the relationships between them as edges.</p> <p>Abstract Real estate contributes significantly to all major economies around the world. In particular, house prices have a direct impact on stakeholders, ranging from house buyers to financing companies. Thus, a plethora of techniques have been developed for real estate price prediction. Most of the existing techniques rely on different house features to build a variety of prediction models to predict house prices. Perceiving the effect of spatial dependence on house prices, some later works focused on introducing spatial regression models for improving prediction performance. However, they fail to take into account the geo-spatial context of the neighborhood amenities such as how close a house is to a train station, or a highly-ranked school, or a shopping center. Such contextual information may play a vital role in users' interests in a house and thereby has a direct influence on its price. In this paper, we propose</p>

		<p>to leverage the concept of graph neural networks to capture the geo-spatial context of the neighborhood of a house. In particular, we present a novel method, the geo-spatial network embedding (GSNE), that learns the embeddings of houses and various types of points of interest (POIs) in the form of multipartite networks, where the houses and the POIs are represented as attributed nodes and the relationships between them as edges. Extensive experiments with a large number of regression techniques show that the embeddings produced by our proposed GSNE technique consistently and significantly improve the performance of the house price prediction task regardless of the downstream regression model. Relevant source code for GSNE is available at: https://github.com/sarathismg/gsne.</p>
24	<pre>@article{Ahtesham2020HousePP, title={House Price Prediction using Machine Learning Algorithm - The Case of Karachi City, Pakistan}, author={Maida Ahtesham and Narmeen Zakaria Bawany and Kiran Fatima}, journal={2020 21st International Arab Conference on Information Technology (ACIT)}, year={2020}, pages={1-5}, url={https://api.semanticscholar.org/CorpusID:230999527} }</pre>	<p>House Price Prediction using Machine Learning Algorithm - The Case of Karachi City, Pakistan</p> <p>The proposed house price prediction model is able to predict 98% accuracy and Gradient Boosting Model XGBoost is utilized to predict housing prices in Pakistan. AbstractHouse prices are a significant impression of the economy, and its value ranges are of great concerns for the clients and property dealers. Housing price escalate every year that eventually reinforced the need of strategy or technique that could predict house prices in future. There are certain factors that influence house prices including physical conditions, locations, number of bedrooms and others. Traditionally predictions are made on the basis of these factors. However such prediction methods require an appropriate knowledge and experience regarding this domain. Machine Learning techniques have been a significant source of advanced opportunities to analyze, predict and visualize</p>

		<p>housing prices. In this paper, Gradient Boosting Model XGBoost is utilized to predict housing prices. Publicly available dataset containing 38,961 records of Karachi city is attained from an Open Real Estate Portal of Pakistan. Lot of work has been done in predicting house prices across many countries, however very limited amount of work has been done for predicting house prices in Pakistan. Our proposed house price prediction model is able to predict 98% accuracy.</p>
25	<pre>@article{MoraGarcia2022HousingPP, title={Housing Price Prediction Using Machine Learning Algorithms in COVID-19 Times}, author={Raul Tomas Mora-Garc{\`{a}}a}, journal={Land}, year={2022}, url={https://api.semanticscholar.org/CorpusID:253824372} }</pre>	<p>Housing Price Prediction Using Machine Learning Algorithms in COVID-19 Times</p> <p>This research is one of the first studies to use machine learning and microdata to explore the incidence of the COVID-19 pandemic on house prices, and shows that machine learning algorithms perform better than traditional linear models because they are better adapted to the nonlinearities of complex data such as real estate market data.</p> <p>AbstractMachine learning algorithms are being used for multiple real-life applications and in research. As a consequence of digital technology, large structured and georeferenced datasets are now more widely available, facilitating the use of these algorithms to analyze and identify patterns, as well as to make predictions that help users in decision making. This research aims to identify the best machine learning algorithms to predict house prices, and to quantify the impact of the COVID-19 pandemic on house prices in a Spanish city. The methodology addresses the phases of data preparation, feature engineering, hyperparameter training and optimization, model evaluation and selection, and finally model interpretation. Ensemble learning algorithms based on boosting (Gradient Boosting Regressor,</p>

		<p>Extreme Gradient Boosting, and Light Gradient Boosting Machine) and bagging (random forest and extra-trees regressor) are used and compared with a linear regression model. A case study is developed with georeferenced microdata of the real estate market in Alicante (Spain), before and after the pandemic declaration derived from COVID-19, together with information from other complementary sources such as the cadastre, socio-demographic and economic indicators, and satellite images. The results show that machine learning algorithms perform better than traditional linear models because they are better adapted to the nonlinearities of complex data such as real estate market data. Algorithms based on bagging show overfitting problems (random forest and extra-trees regressor) and those based on boosting have better performance and lower overfitting. This research contributes to the literature on the Spanish real estate market by being one of the first studies to use machine learning and microdata to explore the incidence of the COVID-19 pandemic on house prices.</p>
26	<pre>@article{Guo2020CANML, title={CAN MACHINE LEARNING ALGORITHMS ASSOCIATED WITH TEXT MINING FROM INTERNET DATA IMPROVE HOUSING PRICE PREDICTION PERFORMANCE?}, author={Jian-qiang Guo and Shu-hen Chiang and Min Liu and Chien-Chun Yang and Kai-yi Guo}, journal={International Journal of Strategic Property Management}, year={2020}, url={https://api.semanticscholar.org/CorpusID:225762221} }</pre>	<p>CAN MACHINE LEARNING ALGORITHMS ASSOCIATED WITH TEXT MINING FROM INTERNET DATA IMPROVE HOUSING PRICE PREDICTION PERFORMANCE?</p> <p>This research adopts a broader version of text mining to search for keywords in relation to housing prices and then evaluates the predictive abilities using machine learning algorithms to better understand the trends of house prices in China.</p> <p>AbstractHousing frenzies in China have attracted widespread global attention over the past few years, but the key is how to more accurately forecast housing</p>

		<p>prices in order to establish an effective real estate policy. Based on the ubiquitousness and immediacy of Internet data, this research adopts a broader version of text mining to search for keywords in relation to housing prices and then evaluates the predictive abilities using machine learning algorithms. Our findings indicate that this new method, especially random forest, not only detects turning points, but also offers prediction ability that clearly outperforms traditional regression analysis. Overall, the prediction based on online search data through a machine learning mechanism helps us better understand the trends of house prices in China.</p>
26	<pre>@article{Wang2021PredictionOH, title={Prediction of House Price Index Based on Bagging Integrated WOA-SVR Model}, author={Xiang Wang and Shen Gao and Shiyu Zhou and Yibin Guo and Yonghui Duan and Daqing Wu}, journal={Mathematical Problems in Engineering}, year={2021}, url={https://api.semanticscholar.org/CorpusID:240259377} }</pre>	<p>Prediction of House Price Index Based on Bagging Integrated WOA-SVR Model</p> <p>A whale algorithm optimized support vector regression model based on bagging ensemble learning method is proposed, and the results show that the prediction accuracy of the proposed integrated model is better than the comparison model in all cases.</p> <p>AbstractAiming at the shortcomings of a single machine learning model with low model prediction accuracy and insufficient generalization ability in house price index prediction, a whale algorithm optimized support vector regression model based on bagging ensemble learning method is proposed. Firstly, gray correlation analysis is used to obtain the main influencing factors of house prices, and the segmentation forecasting method is used to divide the data set and forecast the house prices in the coming year using the data of the past ten years. Secondly, the whale optimization algorithm is used to find the optimal parameters of the penalty factor and kernel function in the SVR model, and then, the WOA-SVR model is established. Finally, in order to further improve</p>

		<p>the model generalization capability, a bagging integration strategy is used to further integrate and optimize the WOA-SVR model. The experiments are conducted to forecast the house price indices of four regions, Beijing, Shanghai, Tianjin, and Chongqing, respectively, and the results show that the prediction accuracy of the proposed integrated model is better than the comparison model in all cases.</p>
26	<pre>@inproceedings{Jain2020HOUSINGPP, title={HOUSING PRICE PREDICTION USING MACHINE LEARNING}, author={Prof. N. R. Jain and Prachi S. Zende and Mohini M. Gaikwad and Vibha Bhor}, year={2020}, url={https://api.semanticscholar.org/CorpusID:216491060} }</pre>	<p>HOUSING PRICE PREDICTION USING MACHINE LEARNING</p> <p>The high discrepancy between house prices in the most expensive and most affordable suburbs is revealed and experiments demonstrate that the combination of stepwise and support vector machine that is based on mean squared error measurement is a competitive model.</p> <p>AbstractHouse price forecasting is an important topic of real estate. The literature attempts to drive useful knowledge from historical data of property markets. Machine learning techniques are applied to analyze historical property transaction to discover useful models for house buyers and sellers. Revealed is the high discrepancy between house prices in the most expensive and most affordable suburbs. Moreover, experiments demonstrate that the combination of stepwise and support vector machine that is based on mean squared error measurement is a competitive</p>
27	<pre>@article{Sibindi2022ABE, title={A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices}, author={Racheal Sibindi and Ronald Waweru Mwangi and Anthony Waititu},</pre>	<p>A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices</p>

	<pre> journal={Engineering Reports}, year={2022}, volume={5}, url={https://api.semanticscholar.org/CorpusID:253864306} } </pre>	<p>This study focuses on developing a hybrid LGBM and XGBoost model in order to prevent overfitting through minimizing variance whilst improving accuracy.</p> <p>AbstractThe implementation of tree-ensemble models has become increasingly essential in solving classification and prediction problems. Boosting ensemble techniques have been widely used as individual machine learning algorithms in predicting house prices. One of the techniques is LGBM algorithm that employs leaf wise growth strategy, reduces loss and improves accuracy during training which results in overfitting. However, XGBoost algorithm uses level wise growth strategy which takes time to compute resulting in higher computation time. Nevertheless, XGBoost has a regularization parameter, implements column sampling and weight reduction on new trees which combats overfitting. This study focuses on developing a hybrid LGBM and XGBoost model in order to prevent overfitting through minimizing variance whilst improving accuracy. Bayesian hyperparameter optimization technique is implemented on the base learners in order to find the best combination of hyperparameters. This resulted in reduced variance (overfitting) in the hybrid model since the regularization parameter values were optimized. The hybrid model is compared to LGBM, XGBoost, Adaboost and GBM algorithms to evaluate its performance in giving accurate house price predictions using MSE, MAE and MAPE evaluation metrics. The hybrid LGBM and XGBoost model outperformed the other models with MSE, MAE and MAPE of 0.193, 0.285, and 0.156 respectively.</p>
28	<pre> @article{Zhang2022AnalysisAP, title={Analysis and prediction of second-hand house price based on random forest}, </pre>	<p>Analysis and prediction of second-hand house price based on random forest</p>

	<pre> author={Yan Zhang and Jin-Xin Huang and Jiahui Zhang and Shuying Liu and Samer Shorman}, journal={Applied Mathematics and Nonlinear Sciences}, year={2022}, volume={7}, pages={27 - 42}, url={https://api.semanticscholar.org/CorpusID:250118370} } </pre>	<p>Using Python language and combined with data analysis and mining technology, the authors capture and clean the housing source data of second-hand houses in Chengdu from Beike Network, and visually analyse the cleaned data to establish a Random Forest model.</p> <p>Abstract Abstract Using Python language and combined with data analysis and mining technology, the authors capture and clean the housing source data of second-hand houses in Chengdu from Beike Network, and visually analyse the cleaned data. Then, a Random Forest (RF) model is established for 38,363 data elements. According to the visual analysis results, the model variables are revalued, the key factors affecting house prices are studied and the optimised model is used to predict house prices. The experiment shows that the deviation between the house price predicted by the RF model and that predicted by the real house price is small; it also indicates the accuracy of the RF model and demonstrates its good application value.</p>
29	<pre> @article{Wang2021PredictionOH, title={Prediction of House Price Index Based on Bagging Integrated WOA-SVR Model}, author={Xiang Wang and Shen Gao and Shiyu Zhou and Yibin Guo and Yonghui Duan and Daqing Wu}, journal={Mathematical Problems in Engineering}, year={2021}, url={https://api.semanticscholar.org/CorpusID:240259377} } </pre>	<p>Prediction of House Price Index Based on Bagging Integrated WOA-SVR Model</p> <p>A whale algorithm optimized support vector regression model based on bagging ensemble learning method is proposed, and the results show that the prediction accuracy of the proposed integrated model is better than the comparison model in all cases.</p> <p>Abstract Aiming at the shortcomings of a single machine learning model with low model prediction accuracy and insufficient generalization ability in house price index prediction, a whale algorithm optimized support vector</p>

		<p>regression model based on bagging ensemble learning method is proposed. Firstly, gray correlation analysis is used to obtain the main influencing factors of house prices, and the segmentation forecasting method is used to divide the data set and forecast the house prices in the coming year using the data of the past ten years. Secondly, the whale optimization algorithm is used to find the optimal parameters of the penalty factor and kernel function in the SVR model, and then, the WOA-SVR model is established. Finally, in order to further improve the model generalization capability, a bagging integration strategy is used to further integrate and optimize the WOA-SVR model. The experiments are conducted to forecast the house price indices of four regions, Beijing, Shanghai, Tianjin, and Chongqing, respectively, and the results show that the prediction accuracy of the proposed integrated model is better than the comparison model in all cases.</p>
30	<pre>@article{JureguiVelarde2023WebAW, title={Web Application with Machine Learning for House Price Prediction}, author={Ra{\u}l J{\a}uregui-Velarde and Laberiano Andrade-Arenas and Domingo Hern{\a}ndez Celis and Roberto Carlos D{\a}vila-Mor{\a}n and Michael Cabanillas-Carbonell}, journal={International Journal of Interactive Mobile Technologies (iJIM)}, year={2023}, url={https://api.semanticscholar.org/CorpusID:266308973} }</pre>	<p>Web Application with Machine Learning for House Price Prediction</p> <p>The developed prediction system facilitates real estate agencies and their customers the accurate prediction of the price of urban and rural housing, minimizing accuracy errors in price prediction, Benefiting all people interested in the real estate world.</p> <p>AbstractEvery year, the price of a house changes due to different aspects, so accurately estimating the buying and selling price is a problem for real estate agencies. Therefore, the research work aims to build a Machine Learning (ML) model in Azure ML Studio and a web application to predict the buying and selling price of two types of houses: urban and rural houses, according to their characteristics, to minimize the forecast error in prediction. Following the basic stages of machine learning construction, we build the prediction model and</p>

		<p>the Rational Unified Process (RUP) methodology to build the web application. As a result, we obtained a model trained with a linear regression algorithm and a predictive ML model with a coefficient of determination of 95% and a web application that consumes the prediction model through an Application Programming Interface (API) that facilitates price prediction to customers. The quality of the prediction system was evaluated by expert judgment; they evaluated efficiency, usability, and functionality. After the calculation, they obtained an average quality of 4.88, which indicates that the quality is very high. In conclusion, the developed prediction system facilitates real estate agencies and their customers the accurate prediction of the price of urban and rural housing, minimizing accuracy errors in price prediction. Benefiting all people interested in the real estate world.</p>
31	<pre>@article{M2023HousePP, title={House Price Prediction using Machine Learning Algorithms}, author={Angulakshmi M and Deepa M and Mala Serene I and Thilagavathi M and Aarthi P}, journal={International Journal on Recent and Innovation Trends in Computing and Communication}, year={2023}, url={https://api.semanticscholar.org/CorpusID:266148818} }</pre>	<p>House Price Prediction using Machine Learning Algorithms</p> <p>This study will be made using four machine learning algorithms such as linear regression, polynomial regression, random forest, decision tree to understand and estimate the pricing of their houses.</p> <p>AbstractHouse prices are a major financial decision for everyone involved in the housing market, including potential home buyers. A major part of the real estate industry is housing. An accurate housing price prediction is a valuable tool for buyer and seller as well as real estate agents. The study is done for the purpose of knowledge among the people to understand and estimate the pricing of their houses. The prediction will be made using four machine learning algorithms such as linear regression, polynomial regression, random</p>

		<p>forest, decision tree. Linear Regression has good interpretability. Decision tree is a graphical representation of all possible solutions. Polynomial regression can be easily fitted to a wide variety of curves. Regression and classification issues are resolved with random forests .Among the given algorithm, Random forest provides better accuracy of about 89% for given dataset.</p>
32	<pre>@article{Jyothsna2023HousePP, title={House Price Prediction Using Machine Learning}, author={Robbi Jyothsna}, journal={REST Journal on Banking, Accounting and Business}, year={2023}, url={https://api.semanticscholar.org/CorpusID:266148817} }</pre>	<p>House Price Prediction Using Machine Learning</p> <p>The way the house rating model works with the assistance of machine learning and that dataset is employed in the projected model, and the house worth trend is expected supported the ARIMA model.</p> <p>AbstractION There is a rise in demand for renting a house and buying house therefore , determining a more efficient to calculate the house rents is crucial. House rent increases once a year, So there's a desire to predict house rents within the future .House rent prediction has gained lots of focus nowadays. House rent prediction system studies behaviour of your time series data and reflects the long run rents. Forecasting foreign countries is vital to understand the house trends in an exceedingly particular country. Software implementations for the experiment were selected from python libraries .Data preprocessing and preparation techinques so as to get clean data. To make machine learning models ready to predict house price supported house features.to research and compare models performance so as to decide on the simplest model. We applied three different Machine Learning algorithms: Decision tree, Random forest and XG Bootsting on the training data 3. 2020] This paper provides an summary concerning the way to predict house prices utilizing totally different regression</p>

		<p>ways with the help of python libraries. The projected technique thought-about the additional refined aspects used for the calculation of house value and provided a additional correct prediction. It conjointly provides a quick concerning varied graphical and numerical techniques which is able to be needed to predict the worth of a house. This paper contains what and the way the house rating model works with the assistance of machine learning and that dataset is employed in our projected model. different like of predict individual housing worth. There has been a significantly sizable amount of papers adopting ancient machine learning approaches to predict housing costs however they the performance of individual models and sixteen neglect the less fashionable nonetheless complicated models. As a to explore numerous impacts of options on prediction strategies, this paper can apply each ancient and advanced machine learning approaches to research the distinction among many advanced models. The nonlinear relationship between powerful factors and house worth and inadequate variety of sample sizes may be the for the poor performance of the normal models. Meanwhile, the daily information of the \$64000 estate market is extremely vast and it's increasing chop-chop. the normal house worth prediction approaches lack capability for large information analysis, inflicting low utilization of information. to deal with these considerations, a house worth prediction model supported deep learning is planned during this paper, enforced on the TensorFlow framework. Adam optimizer is employed to coach the model, wherever the Relu operate is adopted to be the activation operate. Then the house worth trend is expected supported the ARIMA model.</p>
--	--	--

33	<pre>@article{Chowhaan2023MachineLA, title={Machine Learning Approach for House Price Prediction}, author={M. Jagan Chowhaan and Domale Nitish and Gopaluni Sai Akash and Nelli Sreevidya and Subhani Shaik}, journal={Asian Journal of Research in Computer Science}, year={2023}, url={https://api.semanticscholar.org/CorpusID:259486519} }</pre>	<p>Machine Learning Approach for House Price Prediction</p> <p>The analyses' findings supported the usage of artificial neural networks, support vector regression, and linear regression as the most effective modeling techniques and imply that real estate agents and geography play important roles in determining property prices.</p> <p>AbstractIn our ecosystem, real estate is clearly a distinct industry. Predicting house prices, significant housing characteristics, and many other things is made a lot easier by the capacity to extract data from raw data and extract essential information. Daily fluctuations in housing costs are still present, and they occasionally rise without regard to calculations. According to research, changes in property prices frequently have an impact on both homeowners and the real estate market. To analyze the key elements and the best predictive models for home prices, literature research is conducted. The analyses' findings supported the usage of artificial neural networks, support vector regression, and linear regression as the most effective modeling techniques. Our results also imply that real estate agents and geography play important roles in determining property prices. Finding the most crucial factors affecting housing prices and identifying the best machine learning model to utilize for this research would both be greatly aided by this study, especially for housing developers and researchers.</p>
34	<pre>@article{Boyapati2023AnAO, title={An Analysis of House Price Prediction Using Ensemble Learning Algorithms}, author={Sai Venkat Boyapati and Maddirala Sai Karthik and Konakanchi Subrahmanyam and B. Ramachandra Reddy},</pre>	<p>An Analysis of House Price Prediction Using Ensemble Learning Algorithms</p>

	<pre>journal={Research Reports on Computer Science}, year={2023}, url={https://api.semanticscholar.org/CorpusID:258981174} }</pre>	<p>The principal purpose of the study is the prediction of house prices based on current conditions and the newly developed model resolves such issues and provides a rich user interface with a better model.</p> <p>AbstractIt is very important to understand the market drifts in the wake of booming civilization and ever-changing market requirements. The principal purpose of the study is the prediction of house prices based on current conditions. From historical data on property markets, literature attempts to draw useful insights. Business trends must be understood so that individuals may prepare their budgetary needs accordingly. A society that is ever-expanding is driven by the growing real estate industry. A lot of clients have been duped by agents setting up a fake market rate. As a result, the real estate industry has become less transparent in recent years. Due to decreased accuracy and overfitting of data, the previous model reduced efficiency, whereas the newly developed model resolves such issues and provides a rich user interface with a better model. An important part of this study is to develop an extensive model that is beneficial to both business societies and individuals. This is the main objective of this study. In order to simplify the client's fieldwork and free up his time and money, this software is intended to assist him. Machine learning algorithms enable models to be enlightened such as root mean square error, random forest, support vector machine, k-nearest neighbors, mean squared error, extreme gradient boost, mean absolute error, R-squared score, linear regression, AdaBoost, CatBoost.</p>
	<pre>@article{Zhang2023ApplicationOM, title={Application of Machine Learning in Boston House Price Prediction}, author={Yuanheng Zhang},</pre>	<p>Application of Machine Learning in Boston House Price Prediction</p>

	<pre> journal={Advances in Economics, Management and Political Sciences}, year={2023}, url={https://api.semanticscholar.org/CorpusID:265137316} } </pre>	<p>House prices are an important economic indicator for a country or region, rising house prices are often associated with economic growth and increased employment opportunities, while a decline in house prices may indicate an economic slowdown or other unfavorable factors. Governments and relevant agencies need to understand the dynamics of housing prices to formulate appropriate housing policies and plan urban development. As one of the major economic centers in the United States, the fluctuations in Boston's housing prices can reflect the local economic conditions and development trends. This paper selects data from Boston in 1970. Each record in this database describes a Boston suburb or town. The proposed method was evaluated using 5 metrics. By comparison, the paper ends up with the result that XGBoost works best out of the four regression models.</p>
	<pre> @article{Gawali2023HousePP, title={House Price Prediction using R language}, author={Prof. Palomi Gawali}, journal={INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT}, year={2023}, url={https://api.semanticscholar.org/CorpusID:264315491} } </pre>	<p>House Price Prediction using R language</p> <p>This project highlights the effectiveness of using R language and machine learning techniques to predict house prices while providing valuable insights into the key factors that influence house prices.</p> <p>AbstractThe purpose of this project is to use the R language to predict house prices based on various features, such as location, size, number of rooms, and other characteristics, using a dataset from Kaggle. In order to gain a better understanding of the data, we conducted exploratory data analysis, which included assessing the distribution of variables, identifying missing values, and detecting potential outliers. Additionally, we examined the correlation between the features and the target variable After performing data cleaning and pre- processing, we employed multiple</p>

		<p>machine learning algorithms, such as linear regression, random forests, and XGBoost, to construct predictive models. We evaluated the performance of these models using various metrics, such as mean absolute error, mean squared error, and root mean squared error. Our findings revealed that XGBoost had the best performance, achieving a root mean squared error of 0.13, indicating high predictive accuracy. We also performed a feature importance analysis to identify the most critical variables for predicting house prices. In summary, our project highlights the effectiveness of using R language and machine learning techniques to predict house prices while providing valuable insights into the key factors that influence house prices. Keywords— House price Prediction, Machine Learning , Regression</p>
	<pre>@article{Zaki2022HousePP, title={House price prediction using hedonic pricing model and machine learning techniques}, author={John F. W. Zaki and Anand Nayyar and Surjeet Dalal and Zainab Hasan Ali}, journal={Concurrency and Computation: Practice and Experience}, year={2022}, volume={34}, url={https://api.semanticscholar.org/CorpusID:252499186} }</pre>	<p>House price prediction using hedonic pricing model and machine learning techniques</p> <p>This work explores the ability of machine learning techniques (MLTs) in enhancing economic activities by increasing the accuracy of house price prediction by integrating XGBoost algorithm with outlier sum-statistic (OS) approach.</p> <p>AbstractThe problem with property valuation is that it is extremely complex. It is difficult to objectively model the pricing process or fairly estimate a property value. Many factors can contribute to this complexity such as spatial and time factors. Evaluators and researchers have been trying to model the process for centuries. Up until recently, when computer-aided valuation systems provided better solutions in the data evaluation and real estate valuation. Nevertheless, they may suffer from low transparency, inaccuracy, and inefficiency. This work</p>

		<p>explores the ability of machine learning techniques (MLTs) in enhancing economic activities by increasing the accuracy of house price prediction. In this article, XGBoost algorithm has been integrated with outlier sum-statistic (OS) approach. In the real estate industry, the price of property plays a crucial role in economic growth. The research attempts to predict the price of a house using MLTs. Here, the price of the property is predicted using Extreme Gradient (XG) boosting algorithm and hedonic regression pricing. Both XGBoost and hedonic pricing models use 13 variables as inputs to predict house prices. The contribution of this research lies in the practicality of using XGboost technique to predict house prices. Finally, the accuracy of the prediction algorithms is reported with XGBoosting showing the highest accuracy of 84.1% while the accuracy of the hedonic regression algorithm is 42%.</p>
	<pre>@article{Wang2022HousePP, title={House Price Prediction Based on Machine Learning: A Case of King County}, author={Yijia Wang and Q Zhao}, journal={Proceedings of the 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022)}, year={2022}, url={https://api.semanticscholar.org/CorpusID:252154311} }</pre>	<p>House Price Prediction Based on Machine Learning: A Case of King County</p> <p>It is found that Catboost performs the best among all models and can be used for house price prediction and makes necessary supplement to the exploration of features that influence house price from a microscope. AbstractThis paper focuses on formulating a feasible method for house price prediction. A dataset containing features and house price of King County in the US is used. During the data preprocessing, extreme values are winsorized and highly correlated features are removed. Eight models including Catboost, lightGBM and XGBoost serve as candidate models. They are evaluated by several evaluation indicators, including rooted mean square error, R-squared score, adjusted R-squared score and K-fold cross validation score. The model that has</p>

		<p>low RMSE, achieves a high R-squared score and adjusted R-squared score, especially in the test set, and acquires a high score in cross validation is considered a better model. This paper finds out that Catboost performs the best among all models and can be used for house price prediction. Location, living space and condition of the house are the most important features influencing house price. After comparison and contrast with other papers, it is attested that findings in this paper conform to real life. This paper formulates a model that fits better than preceding studies for house price prediction and makes necessary supplement to the exploration of features that influence house price from a microscope.</p>
	<pre>@article{Wu2022HousePP, title={House Price Prediction Based On Deep Learning}, author={Yuying Wu and Youshan Zhang}, journal={ArXiv}, year={2022}, volume={abs/2204.09050}, url={https://api.semanticscholar.org/CorpusID:248266837} }</pre>	<p>House Price Prediction Based On Deep Learning</p> <p>The existing methods of house price prediction are summarized and a novel MVTs model based on mixed depth vision and text features is proposed, which has higher prediction accuracy and is more suitable for the house price Prediction.</p> <p>AbstractSince ancient times, what Chinese people have been pursuing is very simple, which is nothing more than "to live and work happily, to eat and dress comfortable". Today, more than 40 years after the reform and opening, people have basically solved the problem of food and clothing, and the urgent problem is housing. Nowadays, due to the storm of long-term rental apartment intermediary platforms such as eggshell, increasing the sense of insecurity of renters, as well as the urbanization in recent years and the scramble for people in major cities, this will make the future real estate market</p>

		<p>competition more intense. In order to better grasp the real estate price, let consumers buy a house reasonably, and provide a reference for the government to formulate policies, this paper summarizes the existing methods of house price prediction and proposes a house price prediction method based on mixed depth vision and text features. real estate related analyzes the development status the real estate at home M2, GRP, commonly existing resnet50 are used to preprocess the index attributes. The flow chart of the MVTs model is constructed, and the implementation steps of the proposed model are described in detail. Finally, we select 75% of housing samples as training, and the remaining 25% of the samples as test. We compare the prediction results of our model with other five models: autoregressive integrated moving average mode (ARIMA), grey prediction model (GM(1,1)), support vector regression (SVR), BP neural network and artificial neural network (ANN), the results show that the proposed novel MVTs model has higher prediction accuracy. Therefore, MVTs is more suitable for the house price prediction.</p>
	<pre>@inproceedings{Singh2022HousePP, title={House Price Prediction Analysis using Machine Learning}, author={Aniket Singh and Adarsh Singh and Aditya Raj and Harshit Jain and Mrs. Asha}, year={2022}, url={https://api.semanticscholar.org/CorpusID:252178860} }</pre>	<p>House Price Prediction Analysis using Machine Learning</p> <p>This work proposes an end to end joint self-attention model for house prediction that considers the interaction between two different features to learn the complicated relationship between features in order to increase prediction precision.</p>

		<p>Abstract— House price prediction is a popular topic, and research teams are increasingly performing related studies by using deep learning or machine learning models. However, because some studies have not considered comprehensive information that affects house prices, prediction sufficiently precise. Therefore, we propose an end to end joint self-attention model for house prediction. In this model, we import data on public facilities such as parks, schools, and mass rapid transit stations to represent the availability of amenities, and we use satellite maps to analyze the environment surrounding houses. We adopt attention mechanisms, which are widely used in image, speech, and translation tasks, to identify crucial features that are considered by prospective house buyers. The model can automatically assign weights when given transaction data. Our proposed model differs from self-attention models because it considers the interaction between two different features to learn the complicated relationship between features in order to increase prediction precision. We conduct experiments to demonstrate the performance of the model. Experimental data include actual selling prices in real estate transaction data for the period from 2017 to 2018, public facility data acquired from the Taipei and New Taipei governments, and satellite maps crawled using the Google Maps application programming interface.</p>
	<pre>@article{Li2022ImbalancedMA, title={Imbalanced Multimodal Attention-Based System for Multiclass House Price Prediction}, author={Yansong Li and Paula Branco and Han Zhang}, journal={Mathematics}, year={2022}, url={https://api.semanticscholar.org/CorpusID:255216442} }</pre>	<p>Imbalanced Multimodal Attention-Based System for Multiclass House Price Prediction</p> <p>The IMAS makes use of an oversampling strategy that operates on multimodal data, namely using text, numeric, categorical, and boolean data types, which allows for the determination of the features that are the</p>

		<p>most relevant and adapts the weights used according to that information when performing inference.</p> <p>AbstractHouse price prediction is an important problem for individuals, companies, organizations, and governments. With a vast amount of diversified and multimodal data available about houses, the predictive models built should seek to make the best use of these data. This leads to the complex problem of how to effectively use multimodal data for house price prediction. Moreover, this is also a context suffering from class imbalance, an issue that cannot be disregarded. In this paper, we propose a new algorithm for addressing these problems: the imbalanced multimodal attention-based system (IMAS). The IMAS makes use of an oversampling strategy that operates on multimodal data, namely using text, numeric, categorical, and boolean data types. A self-attention mechanism is embedded to leverage the usage of neighboring information that can benefit the model's performance. Moreover, the self-attention mechanism allows for the determination of the features that are the most relevant and adapts the weights used according to that information when performing inference. Our experimental results show the clear advantage of the IMAS, which outperforms all the competitors tested. The analysis of the weights obtained through the self-attention mechanism provides insights into the features' relevance and also supports the importance of using this mechanism in the predictive model.</p>
	@article{Basysyar2022HousePP, title={House Price Prediction Using Exploratory Data Analysis and Machine Learning with Feature Selection}, author={Fadhil Muhammad Basysyar and Gifthera Dwilestari}, journal={Acadlore Transactions on AI and Machine Learning}, year={2022},	House Price Prediction Using Exploratory Data Analysis and Machine Learning with Feature Selection

	<pre>url={https://api.semanticscholar.org/CorpusID:254301386} }</pre>	<p>This study investigates the exploratory data analysis based on linear regression, ridge regression, Lasso regression, and Elastic Net regression, with the aid of machine learning with feature selection, and learns that the proposed prediction model for house prices is satisfactory in predicting house prices.</p> <p>AbstractIn many real-world applications, it is more realistic to predict a price range than to forecast a single value. When the goal is to identify a range of prices, price prediction becomes a classification problem. The House Price Index is a typical instrument for estimating house price discrepancies. This repeat sale index analyzes the mean price variation in repeat sales or refinancing of the same assets. Since it depends on all transactions, the House Price Index is poor at projecting the price of a single house. To forecast house prices effectively, this study investigates the exploratory data analysis based on linear regression, ridge regression, Lasso regression, and Elastic Net regression, with the aid of machine learning with feature selection. The proposed prediction model for house prices was evaluated on a machine learning housing dataset, which covers 1,460 records and 81 features. By comparing the predicted and actual prices, it was learned that our model outputted an acceptable, expected values compared to the actual values. The error margin to actual values was very small. The comparison shows that our model is satisfactory in predicting house prices.</p>
	<pre>@article{Shi2022HousePP, title={House Price Prediction Model Using Bridge Memristors Recurrent Neural Network}, author={Wenzhao Shi}, journal={Proceedings of the 7th International Conference on Cyber Security and Information Engineering}, year={2022},</pre>	<p>House Price Prediction Model Using Bridge Memristors Recurrent Neural Network</p> <p>A novel Bridge Memristors Recurrent Neural Network is proposed to forecast the house price prediction and carried out outstanding performance than existing</p>

	<pre>url={https://api.semanticscholar.org/CorpusID:253120714} }</pre>	<p>models to attain the high prediction rate by analyzing the correlation coefficient.</p> <p>AbstractIn recent decay, the house price prediction plays important role because of it's the volatile of house price which makes significant impact on property valuation and economic growth. It characterizes are attracted the numerous researchers, businessman and people who buy or sell house towards it. The volatile of house price is occurred based on various factors like location, facility, neighborhood, etc. In this way, researchers are evaluating the factors using machine and deep learning process to analysis the information. Although, regression-based analysis has problem due to its nonlinear and linear information in neural network. Thus, we have proposed a novel Bridge Memristors Recurrent Neural Network to forecast the house price prediction in this paper. In addition, RBP algorithm is used on Bridge Memristors RNN for train the neural network in efficient manner. Besides, our proposed model carried out outstanding performance than existing models to attain the high prediction rate by analyzing the correlation coefficient.</p>
	<pre>@article{Hjort2022HousePP, title={House price prediction with gradient boosted trees under different loss functions}, author={Anders Hjort and Johan Pensar and Ida Scheel and Dag einar Sommervoll}, journal={Journal of Property Research}, year={2022}, volume={39}, pages={338 - 364}, url={https://api.semanticscholar.org/CorpusID:249056193} }</pre>	<p>House price prediction with gradient boosted trees under different loss functions</p> <p>This work investigates whether introducing a loss function closer to the AVM actual loss measure improves performance in machine learning approaches, specifically for a gradient boosted tree approach, and finds an improvement from to of predictions within of the true value on a data set of transactions from the Norwegian housing market.</p>

		<p>AbstractABSTRACT Many banks and credit institutions are required to assess the value of dwellings in their mortgage portfolio. This valuation often relies on an Automated Valuation Model (AVM). Moreover, these institutions often report the models accuracy by two numbers: The fraction of predictions within and range from the true values. Until recently, AVMs tended to be hedonic regression models, but lately machine learning approaches like random forest and gradient boosted trees have been increasingly applied. Both the traditional approaches and the machine learning approaches rely on minimising mean squared prediction error, and not the number of predictions in the and range. We investigate whether introducing a loss function closer to the AVMs actual loss measure improves performance in machine learning approaches, specifically for a gradient boosted tree approach. This loss function yields an improvement from to of predictions within of the true value on a data set of transactions from the Norwegian housing market between 2013 and 2015, with the biggest improvements in performance coming from the lower price segments. We also find that a weighted average of models with different loss functions improves performance further, yielding of the observations within of the true value.</p>
	@article{Wiradinata2022POSTPANDEMICA, title={POST-PANDEMIC ANALYSIS OF HOUSE PRICE PREDICTION IN SURABAYA: A MACHINE LEARNING APPROACH},	POST-PANDEMIC ANALYSIS OF HOUSE PRICE PREDICTION IN SURABAYA: A MACHINE LEARNING APPROACH

<pre>author={Trianggoro Wiradinata and Felicia Graciella and Rinabi Tanamal and Yosua Setyawan Soekamto and Theresia Ratih Dewi Saputri}, journal={Journal of Southwest Jiaotong University}, year={2022}, url={https://api.semanticscholar.org/CorpusID:253516826} }</pre>	<p>This work aims to develop a machine learning-based prediction model for real estate agents to use in determining property prices, with the expectation that the resulting predictions will be more accurate and supported by the data, increasing seller and buyer confidence.</p> <p>AbstractThe COVID-19 outbreak caused a slowdown in the Indonesian economy, as it did in many other impacted nations. Consequently, the housing market in Indonesia, along with other industries, deteriorated. Other post-pandemic issues displace the property industry's priorities in Indonesia. Determining a fair property price is a problem occurring because of the economic slowdown. Property sellers expected their property selling prices to be the same before the pandemic or even increase, but property agents hoped the properties would be selling fast, creating a sense of distrust between the seller and the property agents. This work aims to develop a machine learning-based prediction model for real estate agents to use in determining property prices, with the expectation that the resulting predictions will be more accurate and supported by the data, increasing seller and buyer confidence. Following the suggestion from previous studies, several supervised algorithms such as Linear Regression, Decision Tree, and Random Forest were used to develop the model. Training data were collected from five property agents in Surabaya and as well as web scraping from the online home sales portals. Findings from the study show that Random Forest performs best in predicting with the highest coefficient of determination and lowest error. Using evaluation measures such as Mean Absolute Percent Error (MAPE), the error was calculated to be 23%, which is acceptable for prediction.</p>
---	---

	<pre>@inproceedings{Dange2022MachineLB, title={Machine Learning based House Price Prediction using Regression Techniques}, author={Trupti K. Dange and Amritesh Ranjan Mishra and Abhishek Jagtap and Shubham Chavhan and Niraj R. Chavan}, year={2022}, url={https://api.semanticscholar.org/CorpusID:250086242} }</pre>	<p>Machine Learning based House Price Prediction using Regression Techniques</p> <p>This book aims to create a comparable system for establishing housing values to that utilised by huge corporations like 99acres and others, and to address the problems with current market-based house price forecasts.</p> <p>- Buying a house is without a doubt one of the most important decisions one can make. A house's price is decided by a variety of factors, including the house's location, features, and the demand and supply in the real estate market. The housing industry is equally vital to the entire economy. To create a comparable system for establishing housing values to that utilised by huge corporations like 99acres and others. To address the problems with current market-based house price forecasts. To give services that are superior to earlier ones while taking all variables into consideration. To create an accurate pricing prediction model. Customization and the abolition of manual labour.</p>
	<pre>@article{Lakshmi2022AnER, title={An Enhanced Regression Technique for House Price Prediction}, author={K. Lakshmi. and P. V. Narayana and P. Durga Bhavani and V. V. S. Madhavacharyulu and N. Lavanya and S. Pousia}, journal={2022 5th International Conference on Contemporary Computing and Informatics (IC3I)}, year={2022}, pages={457-462}, url={https://api.semanticscholar.org/CorpusID:257667687} }</pre>	<p>An Enhanced Regression Technique for House Price Prediction</p> <p>The results in this disclosure show that a model combining linear regression, random forest, and K Nearest Neighbors (KNN) yields the lowest inaccuracies, which eliminates the need to rely on realtors to determine a fair price for a home based on key features.</p> <p><u>Abstract</u>The real estate market is one of the least transparent sectors of society as real estate prices change daily and are often overvalued rather than valued. Homebuyers use budget and market methods to find new homes. However, a</p>

		<p>fundamental problem with the current approach is the inability to predict future market trends that will lead to price spikes. It is very important for researchers to base their house price proposals on empirical studies. In order to accurately predict the price of a home, customers need to carefully evaluate factors related to the home, which is very difficult. Using machine learning (ML) to solve this problem seems like a viable solution. To address this problem, ML models such as Linear Regression (LR), K Nearest Neighbors (KNN), Random Forests (RF); Ensembles (LR, KNN, RF) are used. A number of error metrics are used to select the best model, including mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE). The results in this disclosure show that a model combining linear regression (LR), random forest (RF) and K Nearest Neighbors (KNN) yields the lowest inaccuracies. A successful regression model should have a minimal error value. This eliminates the need to rely on realtors to determine a fair price for a home based on key features.</p>
	<pre>@article{Mubarak2022AMR, title={A Map-Based Recommendation System and House Price Prediction Model for Real Estate}, author={Maryam Mubarak and Ali Tahir and Fizza Waqar and Ibraheem Haneef and Gavin Mcardle and Michela Bertolotto and Muhammad Tariq Saeed}, journal={ISPRS Int. J. Geo Inf.}, year={2022}, volume={11}, pages={178}, url={https://api.semanticscholar.org/CorpusID:247304050} }</pre>	<p>A Map-Based Recommendation System and House Price Prediction Model for Real Estate</p> <p>A novel framework for recommending real estate to users is presented, which can make personalised recommendations of real estate based on content, collaboration and location by monitoring user interactions through an online real estate portal.</p> <p>AbstractIn 2015, global real estate was worth \$217 trillion, which is approximately 2.7 times the global GDP; it also accounts for roughly 60% of all conventional global resources, making it one of the key factors behind</p>

		<p>any country's economic growth and stability. The accessibility of spatial big data will help real estate investors make better judgment calls and earn additional profit. Since location is deemed necessary for real estate and consequent decision-making, digital maps have become a prime resource for real estate purchases, planning and development. Personalisation can assist in making judgments by identifying user desires and inclinations, which can then be recorded or captured as a user performs some interactions with a digital map. A personalised real estate portal can use this information to suggest properties, assist homeowners and provide valuable real estate analytics. This article presents a novel framework for recommending real estate to users. By monitoring user interactions through an online real estate portal, the framework can make personalised recommendations of real estate based on content, collaboration and location. The effectiveness of the recommendations was tested by the user feedback mechanism through a method of mean absolute precision, and the results show that 79% precise suggestions were generated, i.e., out of 5 recommendations produced, users were interested in at least 3. Along with that, a separate house price prediction model based on neural networks and classical regression techniques was also implemented to assist users in making an informed decision regarding prospects of real estate purchase.</p>
	<pre>@article{Sharma2024AnOH, title={An Optimal House Price Prediction Algorithm: XGBoost}, author={Hemlata Sharma and Hitesh Harsora and Bayode Ogunleye}, journal={Analytics}, year={2024}, url={https://api.semanticscholar.org/CorpusID:266853666} }</pre>	<p>An Optimal House Price Prediction Algorithm: XGBoost</p> <p>This work addressed the house price prediction problem as a regression task and thus employed various machine learning techniques capable of expressing the significance of independent variables to show that</p>

		<p>XGBoost is the best performing model for house price prediction.</p> <p>AbstractAn accurate prediction of house prices is a fundamental requirement for various sectors, including real estate and mortgage lending. It is widely recognized that a property's value is not solely determined by its physical attributes but is significantly influenced by its surrounding neighborhood. Meeting the diverse housing needs of individuals while balancing budget constraints is a primary concern for real estate developers. To this end, we addressed the house price prediction problem as a regression task and thus employed various machine learning (ML) techniques capable of expressing the significance of independent variables. We made use of the housing dataset of Ames City in Iowa, USA to compare XGBoost, support vector regressor, random forest regressor, multilayer perceptron, and multiple linear regression algorithms for house price prediction. Afterwards, we identified the key factors that influence housing costs. Our results show that XGBoost is the best performing model for house price prediction. Our findings present valuable insights and tools for stakeholders, facilitating more accurate property price estimates and, in turn, enabling more informed decision making to meet the housing needs of diverse populations while considering budget constraints.</p>
	@article{Liu2024DatasetAA, title={Dataset Analysis and House Price Prediction}, author={Junjie Liu}, journal={Highlights in Science, Engineering and Technology}, year={2024}, url={https://api.semanticscholar.org/CorpusID:267286835} }	<p>Dataset Analysis and House Price Prediction</p> <p><u>This study has made preliminary progress in data cleaning, dataset analysis, and predictive modeling, and the use of charts provides a more intuitive representation of dataset characteristics, which has implications for the fields of data cleaning and machine learning.</u></p>

		<p><u>Abstract</u>The prediction of house prices through the analysis of data using machine learning and charts is a crucial and significant topic. Many scholars have conducted research in this area, providing valuable insights for both academic learning and real-world applications. The goal of this study is to predict house prices and thoroughly analyze the dataset. The methodology includes data cleaning techniques to ensure data quality. Additionally, three types of charts are employed to analyze the dataset effectively. Finally, two popular models are utilized to predict house prices, and their accuracy is evaluated. The results indicate that the Random Forest Regressor model is more suitable for the specific dataset, and the impact of each factor on house price prediction varies. Looking ahead, future research will involve the utilization of more advanced models to further enhance prediction accuracy. This will enable realistic simulations and contribute to the ongoing development of the society. This study has made preliminary progress in data cleaning, dataset analysis, and predictive modeling. The use of charts provides a more intuitive representation of dataset characteristics. The findings have implications for the fields of data cleaning and machine learning.</p>
	<pre>@article{Gokalani2022HousePP, title={House Price Prediction of Real Time Data (DHA Defence) Karachi Using Machine Learning}, author={Lata Bai Gokalani and Bhagwan Das and Dilip Kumar Ramnani and Mahender Kumar and Mazhar Ali Shah}, journal={Sir Syed University Research Journal of Engineering \& Technology},</pre>	<p>House Price Prediction of Real Time Data (DHA Defence) Karachi Using Machine Learning</p> <p>This research paper mainly focuses on real time Defense Housing Authority (DHA) Karachi data, applying different regression algorithms like Decision tree, Random forest and linear regression to find the sales</p>

	<pre> year={2022}, url={https://api.semanticscholar.org/CorpusID:255170981} } </pre>	<p>price prediction of the house and compare the performance of these models.</p> <p>AbstractPakistan's real estate market has a large impact in GDP growth. Investment in real estate sector in Pakistan is encumbered with lucrative opportunities. The market demand for housing is ever increasing year by year. House sales prices keep on changing and increasing frequently, so there is a need for a system to forecast house sales prices in the future. Several factors that influence house sales price includes; location, physical attributes, number of bedrooms as well as several other economic factors. One of the main motivation of choosing Karachi for the house prediction is that Karachi is capital of Sindh and it has significant importance in country's economic as it is the major commercial and industrial center of Sindh. It is one of the main contribution of the work is that through this the house prediction model based on DHA Karachi data is developed and as per best of our knowledge till today there is no prediction of housing for the country's important has been developed. has This research paper mainly focuses on real time Defense Housing Authority (DHA) Karachi data, applying different regression algorithms like Decision tree, Random forest and linear regression to find the sales price prediction of the house and compare the performance of these models. Random Forest algorithm gives 98% of accuracy. The proposed work will be very much helpful for the common people, real-estate people, investors and builders to inform them about making decision of selling or buying at Defense Housing Authority (DHA) Karachi.</p>
	<pre> @inproceedings{Hjort2022HousePP, title={House Price Prediction with Confidence: Empirical Results from the Norwegian Market}, author={Anders Hjort}, </pre>	<p>House Price Prediction with Confidence: Empirical Results from the Norwegian Market</p> <p>The results indicate that the methods using conformalized quantile regression create narrower</p>

	<pre>booktitle={International Symposium on Conformal and Probabilistic Prediction with Applications}, year={2022}, url={https://api.semanticscholar.org/CorpusID:254534112} }</pre>	<p>confidence regions than split conformal inference, which is a major challenge in this empirical contribution. AbstractAutomated Valuation Models are statistical models used by banks and other financial institutions to estimate the price of a dwelling, typically motivated by financial risk management purposes. The preferred choice of model for this task is often tree based machine learning models such as gradient boosted trees or random forest, where uncertainty quantification is a major challenge. In this empirical contribution, we compare split conformal inference, conformalized quantile regression and Mondrian conformalized quantile regression on data from the Norwegian housing market, and use random forest as a point prediction. The data consists of $N = 29\,993$ transactions from Oslo (Norway) from the time period 2018-2019. The results indicate that the methods using conformalized quantile regression create narrower confidence regions than split conformal inference.</p>
	<pre>@article{Semnani2021HousePP, title={House Price Prediction using Satellite Imagery}, author={Sina J. Semnani and Hoormazd Rezaei}, journal={ArXiv}, year={2021}, volume={abs/2105.06060}, url={https://api.semanticscholar.org/CorpusID:234482655} }</pre>	<p>House Price Prediction using Satellite Imagery</p> <p>Using Los Angeles County's property assessment dataset, by transferring learning from an Inception-v3 model pretrained on ImageNet, an improvement of ~10% in R-squared score is achieved compared to two baseline models that only use non-image features of the house.</p> <p>AbstractIn this paper we show how using satellite images can improve the accuracy of housing price estimation models. Using Los Angeles County's property assessment dataset, by transferring learning from an Inception-v3 model pretrained on ImageNet, we could achieve an improvement of ~10% in R-squared score compared to two baseline models that only use non-image features of the house.</p>

HOUSE PRICES PREDICTION

(DỰ ĐOÁN GIÁ NHÀ)

I. Giới thiệu

Những năm gần đây, cùng với sự phát triển của ngành bất động sản, giá nhà đất liên tục tăng cao. Cả nước, xã hội và các cá nhân đều quan tâm đến mức giá này. Đối với giá nhà ở hàng hóa, có rất nhiều yếu tố ảnh hưởng đến giá nhà ở. Ngoài các quy định quốc gia, bản thân các yếu tố như ánh sáng, cách bố trí, môi trường của ngôi nhà cũng có tác động nhất định đến giá cả, dẫn đến thị trường bất động sản có những biến động đáng kể. Vì vậy, việc nghiên cứu một mô hình dự báo giá nhà đất chính xác có ý nghĩa thực tiễn. Nó có thể hướng dẫn người dân trong việc tiêu dùng nhà ở và đưa ra các khuyến nghị chính sách về điều tiết giá của chính phủ [1] [2]. Dự đoán giá nhà là một bài toán cơ bản trong lĩnh vực bất động sản và kinh tế. Việc ước tính chính xác giá nhà là rất quan trọng đối với các bên liên quan khác nhau bao gồm chủ sở hữu nhà, người mua, người bán và các nhà hoạch định chính sách. Trong bài viết này, chúng tôi khám phá ứng dụng hồi quy tuyến tính, một kỹ thuật học máy đơn giản nhưng hiệu quả để dự đoán giá nhà. Chúng tôi thảo luận về phương pháp, tiền xử lý dữ liệu, lựa chọn tính năng, đào tạo mô hình và số liệu đánh giá. Chúng tôi cũng nêu bật những điểm mạnh, hạn chế và những cải tiến tiềm năng của phương pháp hồi quy tuyến tính trong bối cảnh dự đoán giá nhà [2]. Nhà nghiên cứu sử dụng các kỹ thuật học tập tổng hợp, kết hợp Cây hồi quy tăng cường độ dốc (GBRT) và thuật toán XGBoost, để phát triển mô hình dự đoán có thể ước tính chính xác giá nhà dựa trên các tính năng hoặc biến số có liên quan [4]. Trong tài liệu, có hai cách tiếp cận cơ bản liên quan đến việc xác định giá nhà. Một trong số đó là dự đoán giá nhà bằng cách sử dụng các biến số kinh tế vĩ mô ở quốc gia sản xuất ra ngôi nhà, và một trong số đó là mô hình dự đoán giá, mà chúng ta có thể biểu thị dưới dạng các biến vi mô, bằng cách xem xét các đặc điểm của ngôi nhà. Trong nghiên cứu này, giá của ngôi nhà đã được cố gắng dự đoán bằng phương pháp học máy bằng cách thiết lập một mô hình với các biến vi mô tiết lộ các đặc điểm của ngôi nhà. Nghiên cứu được thực hiện tại tỉnh Antalya của Thổ Nhĩ Kỳ, nơi nhu cầu nhà ở hộ gia đình của người nước ngoài cũng rất cao. Quảng cáo nhà ở những địa điểm thuộc nhóm thu nhập thấp, trung bình và cao được chọn làm mẫu [6]. Trong học tập trí tuệ nhân tạo, bài toán phân tích dự báo giá nhà ở Boston là một bài toán hồi quy cổ điển. Dựa trên thông tin về giá nhà đất do Cục Điều tra Dân số Hoa Kỳ tại Boston, Massachusetts thu thập. Bài viết này chia tập dữ liệu giá nhà ở Boston và xây dựng mô hình hồi quy hồi quy tuyến tính, hồi quy cây quyết định và hồi quy máy vectơ hỗ trợ SVR và huấn luyện tập dữ liệu để thu được mối quan hệ giữa các dữ liệu khác nhau liên quan đến giá nhà ở Boston và sử dụng Mối quan hệ này để kết nối tất cả dữ liệu, cuối cùng nó có thể dự đoán xu hướng giá nhà trong tương lai ở Boston và hiển thị nó thông qua hoạt động trực quan. Thông qua ba giá trị dự đoán của mô hình hồi quy, lần lượt được so sánh với giá trị thực tế, xu hướng của giá trị tổng thể, giá trị thực tế và giá trị ước tính là như nhau, nhưng có một sai số nhất

định, đặc biệt là khi giá giao ngay cao hơn hoặc thấp hơn, thường không thể dự báo chính xác, do đó, Dữ liệu cho việc lựa chọn giá trị đặc trưng vẫn còn chỗ cần cải thiện, nghiên cứu trong tương lai cần lấy thêm dữ liệu và đặc điểm có dữ liệu phong phú [8]. Hồi quy trọng số địa lý (GWR) là một công cụ phổ biến để mô hình hóa tính không đồng nhất về không gian trong mô hình hồi quy. Tuy nhiên, hàm trọng số hiện tại được sử dụng trong GWR chỉ xem xét khoảng cách địa lý, trong khi tính tương tự thuộc tính hoàn toàn bị bỏ qua. Trong nghiên cứu này, chúng tôi đã đề xuất hàm trọng số đồng biến kết hợp khoảng cách địa lý và khoảng cách thuộc tính. Hồi quy có trọng số khoảng cách đồng biến (CWR) là phần mở rộng của GWR bao gồm khoảng cách địa lý và khoảng cách thuộc tính. Giá nhà bị ảnh hưởng bởi nhiều yếu tố, chẳng hạn như tuổi nhà, diện tích sàn và mục đích sử dụng đất. Mô hình dự đoán được sử dụng để giúp hiểu được đặc điểm của giá nhà khu vực. CWR được sử dụng để hiểu mối quan hệ giữa giá nhà và các yếu tố kiểm soát. CWR có thể xem xét khoảng cách địa chất và thuộc tính, đồng thời đưa ra ước tính chính xác về giá nhà để duy trì ma trận trọng số cho các hàm khoảng cách địa chất và thuộc tính. Kết quả cho thấy các thuộc tính/điều kiện của ngôi nhà và đặc điểm của ngôi nhà, chẳng hạn như diện tích sàn và tuổi nhà, có thể ảnh hưởng đến giá nhà. Sau khi lựa chọn yếu tố, trong đó chỉ xem xét tuổi nhà và diện tích sàn của tòa nhà, RMSE của mô hình CWR có thể được cải thiện từ 2,9% -26,3% cho các tòa nhà chọc trời khi so sánh với GWR. CWR có thể giảm các lỗi ước tính một cách hiệu quả từ các mô hình hồi quy không gian truyền thống và cung cấp các mô hình mới và khả thi để ước tính không gian [9].

II. Nghiên cứu liên quan

Trước thách thức khó khăn này, một loạt các nghiên cứu đã xuất hiện, mỗi nghiên cứu đều tìm cách giải mã những bí mật nằm trong lĩnh vực dự đoán giá nhà. Từ các kỹ thuật hồi quy nâng cao được khám phá trong "Kỹ thuật hồi quy nâng cao dự đoán giá nhà [1]" đến phương pháp hồi quy theo trọng số hiệp phương sai được đề cao trong "Hồi quy trọng số hiệp phương sai (CWR.): Một nghiên cứu điển hình về ước tính giá nhà [9]", các nhà nghiên cứu đã không ngừng theo đuổi những con đường hiểu biết mới. Những tiếng nói đáng chú ý khác trong bản giao hưởng này bao gồm "Dự báo giá bán lại nhà bằng cách sử dụng phương pháp học tập hợp [10]" và "Dự đoán giá nhà bằng phương pháp học tập sâu và rừng ngẫu nhiên [15, 20]". Những tác phẩm này, cùng với hàng triệu tác phẩm khác, làm sáng tỏ các cách tiếp cận đa dạng và các phương pháp đang phát triển được triển khai nhằm mục đích dự đoán giá nhà chính xác. . Sau đó, mô hình dự đoán được đánh giá bằng phương pháp Lỗi bình phương gốc (RMSE). Kết quả RMSE từ một nghiên cứu trước đây sử dụng kết hợp Lasso và XGBoost là 0,11260, trong khi kết quả RMSE từ nghiên cứu này là 0,00480. Điều này cho thấy giá trị RMSE giảm, cho thấy mức độ lỗi trong mô hình thấp hơn. Điều đó cũng có nghĩa là sự kết hợp giữa thuật toán GBRT và XGBoost đã cải thiện thành công độ chính xác dự đoán của mô hình nghiên cứu trước đó [4]. Bài viết này đề xuất một mô hình kết hợp dựa trên Python và các gói xgboost, DF21 và Geatpy của nó để dự đoán giá bán lại ở Singapore có thể đạt được sai số phần trăm tuyệt đối trung bình (MAPE) là 7,36% trong xu hướng tổng thể cố định của xu hướng giá nhà, bài viết này đề xuất một mô hình kết hợp dựa trên Python và

các gói xgboost, DF21 và Geatpy của nó để dự đoán giá bán lại ở Singapore. Đầu tiên, các thuộc tính phân loại có lượng số cao được xử lý trước bằng phương pháp mã hóa trung bình. Sau đó, chúng tôi đề xuất phương pháp trộn tuyến tính bao gồm GA-HL-Xg-Boost, GARandom Forest (GA-RF), deep-Random Forest (DRF) và lightGBM, với tạp chất Gini để xác định tầm quan trọng của các tính năng. Cuối cùng, kết quả cho thấy rằng nó có thể đạt được sai số phần trăm tuyệt đối trung bình (MAPE) là 7,36% trong xu hướng chung ổn định của xu hướng giá nhà. Nghiên cứu của chúng tôi có thể cung cấp một công cụ dự báo mạnh mẽ về giá bán lại nhà ở các môi trường kinh tế khác nhau [11]. Hồi quy vector hỗ trợ (SVR) với nhân Gaussian được phát triển để dự đoán giá nhà đất và chứng minh độ chính xác và hiệu quả vượt trội so với SVR với nhân tuyến tính, KNN và cây quyết định [12].