1. Giới thiệu:

Nhà ở là một thành phần quan trọng quyết định đến sự phát triển đến nền kinh tế của một quốc gia. Khi một nền kinh tế phát triển, người dân di cư từ thành phố về nông thôn, điều này dẫn tới sự gia tăng dân số đô thị. Dân số đô thị tăng, nhu cầu nhà ở cũng tăng theo. Nhu cầu gia tăng thúc đẩy giá cả của nhà ở. Nhu cầu về nhà ở là một vấn đề đa chiều, đem đến cho con người một nơi để trú ẩn, chăm sóc sức khỏe, đảm bảo về an toàn của bản thân và nhiều những nhu cầu văn hóa xã hội khác. Mọi người muốn sở hữu riêng một căn nhà chỉ khi phúc lợi xã hội của họ đạt đến một mức độ nhất định. Tuy nhiên, điều mà mọi người đặc biệt quan tâm là lựa chọn một căn nhà phù hợp với ngân sách, đồng thời đáp ứng đủ nhu cầu cho các thành viên trong gia đình [2]. Các đặc điểm mà các vị khách khi đi mua nhà sẽ thường quan tâm như diện tích lô đất, số phòng ngủ, khoảng cách đến trung tâm, hay tuổi thọ của căn nhà… Khi đó, một căn nhà với mức giá hợp lý, nằm trong khả năng chi trả của khách mua nhà sẽ là lựa chọn tối ưu nhất.

Trong thời đại hiện đại, với sự bùng nổ của công nghệ thông tin, các phương pháp xử lý dữ liệu truyền thống không thể thích ứng với lượng dữ liệu lớn như vậy. So với các công cụ xử lý dữ liệu truyền thống, công nghệ Big Data có những đặc điểm đại chúng hóa, tính linh hoạt, giá trị, tính kịp thời và tính xác thực, đồng thời có thể xử lý dữ liệu lớn một cách hiệu quả và chính xác. Việc sử dụng hiệu quả học máy có thể nâng cao chức năng của công nghệ Dữ liệu lớn. Vì vậy, Machine learning đã được ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau của xã hội. Trong kỷ nguyên Dữ liệu lớn, Học máy cần có khả năng khái quát hóa và học nhanh, dễ hiểu và nhận thức được chi phí, sử dụng dữ liệu và chuyển giao kiến thức. Là một trong những nội dung quan trọng nhất của Big Data, Machine learning đã được sử dụng rộng rãi và phát triển sâu rộng trong việc dự đoán giá nhà đất [9].

Việc dự đoán giá nhà bằng các phương pháp học máy và trí tuệ nhân tạo đang trở thành một xu hướng không thể phủ nhận. Với sự phát triển mạnh mẽ của dữ liệu lớn và các thuật toán tiên tiến như Linear Regression, Neural Network, Random Forest, XGBoost, LightGBM, và các kỹ thuật như K-means Clustering, Word Embedding, cùng với việc sử dụng các đặc điểm văn bản và hình ảnh của nhà, các nghiên cứu trong lĩnh vực này đã đem lại những kết quả ấn tượng. Việc này không chỉ hỗ trợ người mua và người bán nhà trong việc đưa ra quyết định chính xác về giá cả mà còn giúp các nhà phát triển và chính phủ hiểu rõ hơn về thị trường bất động sản và điều chỉnh các chiến lược phát triển và quản lý thị trường. [13] Đây là một lĩnh vực nghiên cứu đầy tiềm năng với sự ứng dụng rộng rãi trong thực tiễn.

|  |  |  |
| --- | --- | --- |
| ID | Cite | Abstract |
| 1 | @article{Zou2023TheHP,  title={The House Price Prediction Using Machine Learning Algorithm: The Case of Jinan, China},  author={Chengke Zou},  journal={Highlights in Science, Engineering and Technology},  year={2023},  url={https://api.semanticscholar.org/CorpusID:258025150}  } | **The House Price Prediction Using Machine Learning Algorithm: The Case of Jinan, China**  Abstract: House prices increase substantially in China from 1998. Because of expensive house prices, most Chinese people have only one chance to select suitable houses. Therefore, building a house price prediction model based on housing conditions is significant for customers to make decisions. This paper collects the estate market data of Jinan city from the HomeLink website and performs several feature selection algorithms to get critical features for house price prediction. The paper compares the classical machine learning methods for the problem, including Multiple Linear Regression, Random Forest, and Catboost. After cross-validation tests, the CatBoost, algorithm with the lowest Mean Square Error (MSE) is regarded as the most accurate algorithm to predict house prices. *The analytic results show that the house price is dominated by the location features such as area and block.* |
| 2 | @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}  } | **House price prediction modeling using machine learning techniques: a comparative study**  Abstract: 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 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. |
| 3 | @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}  } | **Optimization house price prediction model using gradient boosted regression trees (GBRT) and xgboost algorithm**  Abstract: 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 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. |
| 4 | @article{Li2023HousePP,  title={House price prediction based on machine learning},  author={Hanwen Li},  journal={Applied and Computational Engineering},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259528857}  } | **House price prediction based on machine learning**  Abstract: Machine learning is commonly used in the real estate market. It is vital to apply the idea of machine learning in this field to predict house prices based on various features. The paper will focus on how to use the most appropriate machine learning models for house price prediction. It will use LightGBM(Light Gradient Boosting Machine), Gradient Boosting, and XGBoost(Extreme Gradient Boosting) to train models to predict house prices using the existing data from the Kaggle website. After three models make predictions, they will get an RMSE (root mean square error), whichis0.02975, 0.02537, and 0.01364. Based on the result, the XGBoost model is the best one among these three models used for house price prediction. |
| 5 | @article{Patel2023HOUSEPP,  title={HOUSE PRICE PREDICTION USING MACHINE LEARNING},  author={Deepanshu Patel and Parmeshwar Nayak and Shubham Gupta and Jayanth C},  journal={International Research Journal of Modernization in Engineering Technology and Science},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259865408}  } | **HOUSE PRICE PREDICTION USING MACHINE LEARNING**  Abstract: Forecasting the appropriate house pricing for real estate customers while taking into consideration their priorities and financial situation is the goal. By examining previous market patterns, price ranges, and approaching changes, future prices may be predicted. Research indicates that house price discrepancies are a common source of concern for both homeowners and the real estate industry. Several interrelated factors influence the price at which real estate sells in places like Bengaluru. The size, location as well as and amenities of the property are significant considerations that could have an impact on the price. The analysis' findings supported the use of boosting algorithms like Extreme Gradient Boost Regression (XG Boost), Support Vector Regression, Multiple Linear Regression (Least Squares), and Machine Learning Lasso and Ridge regression models among other regression techniques in modelling explorations |
| 6 | @article{Cui2023HousePP,  title={House Price Prediction with Big Data},  author={Ziyu Cui},  journal={Advances in Economics, Management and Political Sciences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259788957}  } | **House Price Prediction with Big Data**  Abstract: Many of the current models for house price prediction focus on the house price index itself or include too few factors. In this research, however, different aspects of housing properties (bedrooms, bathrooms, grade, view, etc.) will be incorporated to generate a more accurate prediction model. Doing so not only remedies for a lack of variables in current models but also benefits the consumers by providing an accurate estimation of housing values. This paper examines the estimation methodology of a multiple linear regression model. To get the optimal prediction power and keep the model as simple as possible, different variable combinations will be tested. Through comparing different regression models and analyzing the regression results, the predictive model introduced in this paper has a high house price prediction power. This paper provides a potential solution for the prediction of the house price in the King County, Washington. |
| 7 | @article{Palupi2023HousePP,  title={House Price Prediction Using Data Mining with Linear Regression and Neural Network Algorithms},  author={Endang Palupi},  journal={Jurnal Riset Informatika},  year={2023},  url={https://api.semanticscholar.org/CorpusID:266735854}  } | **House Price Prediction Using Data Mining with Linear Regression and Neural Network Algorithms**  Abstract: The need for housing in big cities is very high because most offices and economic centers are in big cities. Limited land and high demand cause house prices to rise. Many developers build housing on the outskirts of big cities with access to trains and toll roads to make transportation easier. Property developers compete by providing the best prices, various choices of house specifications, ease of the mortgage process, and attractive promotions such as no down payment. A house is a long-term investment whose price increases yearly, so proper analysis is needed to buy a place to live in. Several factors influence the price of a house, including location, land area, building area, building type, and so on. This research aims to create a house price prediction model using the Linear Regression Algorithm and Neural Network so that the results can be useful for property agents in predicting house sales or from the buyer's side in predicting house prices. The results of this research use the Linear Regression Algorithm RMSE 0.775, while the Neural Network Algorithm uses RMSE 0.645. From this research, modeling using the Linear Regression Algorithm has better results. Still, the Linear Regression Algorithm and Neural Network Algorithm have RMSE results that are close to accurate and have small errors. |
| 8 |  | **A House Price Prediction Model Based on K-means Clustering and Random Forest in Guangzhou**  Abstract: This paper addresses the key issues in house price forecasting from multiple perspectives by establishing a house price forecasting model for Guangzhou city, providing valuable information and decision support for home buyers, developers, and the government. First, this paper employs the Person coefficient, stepwise regression model and t-test to address the problem of quantifying data and exploring house price factors. By analyzing the correlation between the relevant variables and house prices, the key characteristics that have significant and strong correlation effects on house prices are obtained. Second, the K-means clustering model was used to classify houses into three categories: economic houses, comfortable houses and high-end houses. This classification result provides a more detailed data base for the subsequent construction of the house price prediction model. Finally, the random forest house price prediction model was established in this paper, and the model was validated by error analysis and stability analysis. The average absolute value error and goodness-of-fit obtained were 0.08 and 0.92, respectively, indicating that the model has high accuracy and reliability. The research in this paper has important implications for all parties, including home buyers, developers, and the government. For home buyers, the model can help them better understand the market situation; for developers, the model can guide their reasonable pricing and development strategies; for the government, the model can provide a scientific basis for real estate market regulation and policy control to promote market stability and sustainable development. |
| 9 | @article{Chen2023AnalysisAC,  title={Analysis and Comparison of House Price Prediction Based on XGboost and LightGBM},  author={Shengquan Chen and Huihui Jin and Ling Li},  journal={Advances in Economics, Management and Political Sciences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:265619181}  } | **Analysis and Comparison of House Price Prediction Based on XGboost and LightGBM**  Abstract: Real estate price prediction is one of the key research topics contemporarily. Based on the rapid development of Big Data, machine learning has gradually become the mainstream tool for housing price prediction. The XGboost and LightGBM models, as new advanced mod-els in recent years, have received widespread attention in the application in housing price prediction. Therefore, this study identifies the house price prediction based on XGboost model and LightGBM model and compares them with other models in order to obtain an analysis of the advantages and disadvantages of these two models in housing price predic-tion. According to the analysis, both models have ad-vantages such as high accuracy, high efficiency, and fast training speed. However, although XGboost has the smallest error pre-diction, it requires more computational time, thereby increasing computational costs. In ad-dition, LightGBM has disadvantages such as high overfitting risk in small sample sizes and increased sensitivity in noisy datasets. Therefore, besides the model studied in this article, feature selection methods such as Filter and Wrapper can also be introduced in subsequent studies to further improve the prediction accuracy. |
| 10 | @article{Parekh2023HousePP,  title={House Price Prediction Using Linear Regression Model},  author={Jaykumar Parekh},  journal={International Journal For Multidisciplinary Research},  year={2023},  url={https://api.semanticscholar.org/CorpusID:266805860}  } | **House Price Prediction Using Linear Regression Model**  Abstract: Machine learning is a subset of Artificial Intelligence. Artificial intelligence (AI) and machine learning (ML) are key technologies in solving problems and addressing a wide range of issues in many different fields. Due to its capacity to automate processes, analyze massive volumes of data, and make precise judgments. Generally, it is used in voice assistants, recommendation systems, autonomous vehicles as well as in fraud detection. It also plays a vital role in the real estate sector, accurate house price prediction helps buyers, sellers, and investors make accurate decisions. There is a need for technology to predict housing values because they rise annually. Predicting house prices can assist developers in setting a property's selling price as well as buyers in scheduling the ideal time to buy a home. Four factors influence the price of a house which are area, bedrooms, bathrooms, and location. This study uses a methodology to forecast the price of houses based on relevant features, specifically by applying a linear regression model. Through the use of machine learning methods such as Random Forest, K-Means, Decision Tree, and Linear regression. This strategy will make it easier for people to invest money in a legacy without going via a broker. The study's findings indicate that the Linear regression yields the best accuracy. |
| 11 | @inproceedings{Jachak2023HousePP,  title={House Price Prediction Using Texture and Visual Features},  author={Sweety G. Jachak and Dr. Sayantani Nath},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259769173}  } | **House Price Prediction Using Texture and Visual Features**  Abstract: Real estate sector has been growing at a rate never seen before. For this sector, a key role is played by the pricing of the property. Gone are the days when the price of the property was based on whims and facies of the real estate dealers. Machine learning has numerous applications in the domain of real estate, and one of the most popular ones is predicting house prices. The application of machine learning in house price prediction involves training a model on a dataset that includes a variety of visual and texture features related to the property. The model is then used to predict the price of a new property based on its features. This paper successfully explores machine learning based house price prediction. The methodology followed was to first use data sets to train the model. Later, using correlation-based hybrid GA-reinforcement strategy, a suitable set of features has been selected. In the end, these features are applied to a XG boost regressor to get results. The accuracies are compared with the cases of without feature selection of different regressors. This algorithm, if successfully deployed will be beneficial to both sellers and buyers, because it sets a data-based benchmarking for pricing the property. |
| 12 | @article{Zhang2023ARL,  title={A reinforcement learning-based weight fusion algorithm for house price prediction},  author={Yige Zhang and Zongwen Fan and Jin Gou},  journal={2023 IEEE 35th International Conference on Tools with Artificial Intelligence (ICTAI)},  year={2023},  pages={675-679},  url={https://api.semanticscholar.org/CorpusID:266438661}  } | **A reinforcement learning-based weight fusion algorithm for house price prediction**  Abstract: To address the limitations of existing house price prediction methods that struggle with handling mixed data types, we developed a novel model utilizing a reinforcement learning-based weight fusion algorithm. First, we reconstruct the feature fusion vectors from categorical feature variables. Then, we create individual house price prediction models using a variety of algorithms: Multiple Linear Regression, Regression Decision Tree, Extreme Random Regression Tree, and Multilayer Perceptron. These models' weights are dynamically adjusted using Q-Learning to minimize the Root Mean Square Error (RMSE). Our experiments demonstrate substantial gains in RMSE performance using our weight fusion approach. Specifically, on the small-scale Shenzhen dataset, RMSE improvements vary from 8.65% to 64.61%. On the large-scale Boston dataset, the improvements range between 53.06% and 76.40% in comparison to the standalone prediction algorithms. |
| 13 | @article{Zhang2023DescribeTH,  title={Describe the house and I will tell you the price: House price prediction with textual description data},  author={Han Zhang and Yansong Li and Paula Branco},  journal={Natural Language Engineering},  year={2023},  url={https://api.semanticscholar.org/CorpusID:260039713}  } | **Describe the house and I will tell you the price: House price prediction with textual description data**  Abstract: House price prediction is an important problem that could benefit home buyers and sellers. Traditional models for house price prediction use numerical attributes such as the number of rooms but disregard the house description text. The recent developments in text processing suggest these can be valuable attributes, which motivated us to use house descriptions. This paper focuses on the house asking/advertising price and studies the impact of using house description texts to predict the final house price. To achieve this, we collected a large and diverse set of attributes on house postings, including the house advertising price. Then, we compare the performance of three scenarios: using only the house description, only numeric attributes, or both. We processed the description text through three word embedding techniques: TF-IDF, Word2Vec, and BERT. Four regression algorithms are trained using only textual data, non-textual data, or both. Our results show that by using exclusively the description data with Word2Vec and a Deep Learning model, we can achieve good performance. However, the best overall performance is obtained when using both textual and non-textual features. An $R^2$ of 0.7904 is achieved by the deep learning model using only description data on the testing data. This clearly indicates that using the house description text alone is a strong predictor for the house price. However, when observing the RMSE on the test data, the best model was gradient boosting using both numeric and description data. Overall, we observe that combining the textual and non-textual features improves the learned model and provides performance benefits when compared against using only one of the feature types. We also provide a freely available application for house price prediction, which is solely based on a house text description and uses our final developed model with Word2Vec and Deep Learning to predict the house price. |
| 14 | @article{Fang2023MachineLM,  title={Machine learning models for house price prediction},  author={Lin Ming Fang},  journal={Applied and Computational Engineering},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259372016}  } | **Machine learning models for house price prediction**  Abstract: Housing prices have changed over the years, and there has been an increasing need to predict prices of future homes. This paper gives an overview of various machine learning models that can predict housing prices. There are numerous possible methods of pre-processing the data, so this paper explores ways to handle missing values and categorical data. In this study, the models of regression tree, random forest, XGBoost, gradient boosting, and LightGBM are described and used to predict housing prices. Machine learning models also have hyperparameters that can be adjusted, which can affect predictive accuracy. The methods are evaluated on several benchmark datasets. Based on the results, our approach is effective for the task of house price prediction. |
| 15 | @inproceedings{Kadakia2023ChatbotEH,  title={Chatbot Enhanced House Price Prediction Using Ensemble Technique},  author={Sarang Kadakia and Ishan Desai and Harshit Jain and Prof. Neha Katre},  year={2023},  url={https://api.semanticscholar.org/CorpusID:263215235}  } | **Chatbot Enhanced House Price Prediction Using Ensemble Technique**  Abstract: The rapid culmination of the real estate market has made accurate house price prediction an essential tool for homebuyers, sellers, and investors. The comprehensive house price prediction software uses cutting-edge machine learning algorithms to display property prices. To produce accurate price projections, the platform combines a various number of models to maximize the accuracy of the software. People get real-time price predictions which are calculated by formulating around 20 parameters. |
| 16 | @article{Wang2023JointGC,  title={Joint Gated Co-Attention Based Multi-Modal Networks for Subregion House Price Prediction},  author={Pengkun Wang and Chuancai Ge and Zhengyang Zhou and Xu Wang and Yuantao Li and Yang Wang},  journal={IEEE Transactions on Knowledge and Data Engineering},  year={2023},  volume={35},  pages={1667-1680},  url={https://api.semanticscholar.org/CorpusID:242356327}  } | **Joint Gated Co-Attention Based Multi-Modal Networks for Subregion House Price Prediction**  Abstract: Urban housing price is widely accepted as an economic indicator which is of both business and research interest in urban computing. However, due to the complex nature of influencing factors and the sparse property of transaction records, to implement such a model is still challenging. To address these challenges, in this work, we study an effective and fine-grained model for urban subregion housing price predictions. Compared to existing works, our proposal improves the forecasting granularity from city-level to mile-level, with only publicly released transaction data. We employ a feature selection mechanism to select more relevant features. Then, we propose an integrated model, JGC\_MMN (Joint Gated Co-attention Based Multi-modal Network), to learn all-level features and capture spatiotemporal correlations in all-time stages with a modified densely connected convolutional network as well as current ingredients and future expectations. Next, we devise a novel JGC based fusion method to better fuse the heterogeneous data of multi-stage models by considering their interactions in temporal dimension. Finally, extensive empirical studies on real datasets demonstrate the effectiveness of our proposal, and this fine-grained housing price forecasting has the potential to support a broad scope of applications, ranging from urban planning to housing market recommendations. |
| 17 | @article{ElMouna2023ACS,  title={A Comparative Study of Urban House Price Prediction using Machine Learning Algorithms},  author={Lale El Mouna and Hassan Silkan and Youssef Haynf and Mohamedade Farouk Nann and St{\'e}phane C. K. T{\'e}kouabou},  journal={E3S Web of Conferences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:261021595}  } | **A Comparative Study of Urban House Price Prediction using Machine Learning Algorithms**  Abstract: Accurate housing price forecasts are essential for several reasons. First, it allows individuals to make informed decisions about buying or selling real estate and to determine appropriate prices. Secondly, it helps real estate agents and investors make better investment decisions and negotiate contracts more effectively. In addition, housing prices are often an indication of the general state of the economy. A price decrease may indicate an economic recession, while an increase in prices may signal economic growth. In this study, we proposed to address this subject by predicting house prices using machine learning by choosing three types of machine learning: Linear Regression (LN), Random Forest (RF) and GradientBoosting (GB). We tested our models on the Melbourne real estate dataset, which includes 34,857 property sales and 21 features. |
| 18 | @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}  } | **House Price Prediction using Machine Learning Algorithms**  Abstract: House 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 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. |
| 19 | @article{Li2023BostonHP,  title={Boston House Price Prediction Based on Machine Leaning Methods},  author={Ze Li},  journal={BCP Business \& Management},  year={2023},  url={https://api.semanticscholar.org/CorpusID:257678430}  } | **Boston House Price Prediction Based on Machine Leaning Methods**  Abstract: A significant part of the U.S. economy is real estate and the housing industry. The level of housing prices influences the market's total price level to some extent. Data from Kaggle website is evaluated by using linear regression, random forest regressor and SVM regressor. There are 505 samples and the relationships between aim price with 13 different features. In terms of accuracy of the prediction, the method random forest is best method to predict house price. However, better model like XGboost could be chosen to improve prediction results. |
| 20 | @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}  } | **House Price Prediction Using Machine Learning**  Abstract: ION 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 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. |
| 21 | @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}  } | **Machine Learning Approach for House Price Prediction**  Abstract  In 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. |
| 22 | @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},  journal={Research Reports on Computer Science},  year={2023},  url={https://api.semanticscholar.org/CorpusID:258981174}  } | **An Analysis of House Price Prediction Using Ensemble Learning Algorithms**  Abstract: It 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. |
| 23 | @article{Zhang2023ApplicationOM,  title={Application of Machine Learning in Boston House Price Prediction},  author={Yuanheng Zhang},  journal={Advances in Economics, Management and Political Sciences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:265137316}  } | **Application of Machine Learning in Boston House Price Prediction**  Abstract: 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. |
| 24 | @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}  } | **House price prediction using hedonic pricing model and machine learning techniques**  Abstract: The 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 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%. |
| 25 | @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}  } | **House Price Prediction Based on Machine Learning: A Case of King County**  Abstract: This 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 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. |
| 26 | @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}  } | **Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism**  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 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-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. |
| 27 | @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}  } | **Imbalanced Multimodal Attention-Based System for Multiclass House Price Prediction**  Abstract: House 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. |
| 28 | @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}  } | **House Price Prediction Based On Deep Learning**  Abstract: Since 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 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. |
| 29 | @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},  url={https://api.semanticscholar.org/CorpusID:254301386}  } | **House Price Prediction Using Exploratory Data Analysis and Machine Learning with Feature Selection**  Abstract: In 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. |
| 30 | @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},  url={https://api.semanticscholar.org/CorpusID:253120714}  } | **House Price Prediction Model Using Bridge Memristors Recurrent Neural Network**  Abstract: In 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. |

1. Nghiên cứu liên quan

Trước thách thức khó khăn này, nhiều nghiên cứu đã được đưa ra, mỗi nghiên cứu đều tìm cách đưa ra những phương án để giải quyết những vấn đề trong việc dự đoán giá nhà. Từ việc mạng thần kinh nhân tạo (ANN) được quan sát đưa ra dự đoán với kết quả có ý nghĩa hơn so với máy vector hỗ trợ (SVR) và thuật toán hồi quy tuyến tính đa biến (MLR) [2]. Hay một nghiên cứu khác giúp tối ưu việc dự đoán giá nhà với mô hình Gradient Boost sử dụng mô hình cây quyết định (Decision Tree) và thuật toán XGBoost. Những tác phẩm này, cùng với hàng trăm 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.