

Prediction of House Price Using Machine Learning

Aryan Rana, Dev Raval

Artificial Intelligence and Data Science, Shah and Anchor Kutchhi Engineering College, Chembur, Mumbai, Maharashtra, India

Email: aryan.17465@sakec.ac.in

Artificial Intelligence and Data Science, Shah and Anchor Kutchhi Engineering College, Chembur, Mumbai, Maharashtra, India

Email: dev.17010@sakec.ac.in

Abstract:

The trend of the sudden drop or constant rising of housing prices has attracted interest from the researcher as well as many other interested people. There have been various research works that use different methods and techniques to address the question of the changing of house prices. This work considers the issue of changing house price as a classification problem and discuss machine learning techniques to predict whether house prices will rise or fall using available data. This work applies various feature selection techniques such as variance influence factor, Information value, principle component analysis, and data transformation techniques such as outlier and missing value treatment as well as different transformation techniques. The performance of the machine learning techniques is measured by the four parameters of accuracy, precision, specificity, and sensitivity. The work considers two discrete values 0 and 1 as respective classes. If the value of the class is 0 then we consider that the price of the house has decreased and if the value of the class is 1 then we consider that the price of the house has increased.

Keywords — House Price Prediction, Real estate market, Traditional machine learning.

INTRODUCTION :-

House is most important aspect of human life. Development of civilization is the beginning of the increase in demand for houses day by day. Predicting and analyzing the real estate market can evaluate the stability of the real estate market and can facilitate the government to carry out control on house prices and maintain the healthy and stable development. Many people invest in real estate, because sometimes, it can bring a lot of capital income. Investment in real estate is also a reflection of a local real estate development situation to a certain extent. The purpose of house price prediction and analyses is to provide basis for pricing between buyers and sellers. Different groups of buyers may focus on different house attributes so if a house's features and attributes identified through an attention mechanism match the needs of a buyer, then the house can be

recommended to the buyer. Many researchers have already worked to understand and break the mysteries of the prediction of house prices.

LITERATURE REVIEW :-

1. Paper Name: Virtual Reality for Real Estate
Author: Bogdan, Alexandru Deaky, Luminita Parv.

This paper presents the results of the VR4RE (Virtual Reality for Real Estate) project, which aims at saving time and money for both real estate sellers and buyers by employing modern technologies. VR4RE is one of the innovative projects developed by Bluemind Software and it is in an advanced state. This paper also illustrates the history of in-house technological attempts at creating appropriate presentation tools for real estate properties with 3D and VR (Virtual Reality).

2. Paper Name: Developing Smart Commercial Real Estate

Author: Peter Ekman

In this paper, CNN-based detection and evaluation of infected patients. Tailored CNN models: A set of tailored models based on CNN have been designed to take three sets of image categories (e.g.; normal case, viral pneumonia case, and bacterial case). DenseNet169 architecture, RNN-based architectures are used.

To evaluate the potential of smart commercial real-estate (CRE) we studied a Swedish commercial real estate firm that has developed and deployed a technology-based self-service (TBSS) to help tenants reduce energy consumption.

3. Paper Name; An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting

Author :Nor Azizah Hitama, Amelia Ritahani Ismailb, Faisal Saeed

In this, we have studied various Support Vector Machine(SVM) and Particle Swarm Optimization (PSO) for forecasting cryptocurrency. Forecasting accurate future value is very important in the financial sector. An optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) is introduced in forecasting the cryptocurrency's future price. It is part of Artificial Intelligence (AI) that uses previous experience to forecast the future price.

4. Paper Name: Comparison of Ensemble Methods for Real Estate Appraisal

Author::- Prathamesh Kumar, Ishan Madan, Ashutosh Kale

In this paper, four ensemble methods, namely Bagging, Random Forest, Gradient Boosting, and Extreme Gradient Boosting were analyzed and compared in terms of their efficiency in the appraisal of real estate in Mumbai. The property listings available on the real estate website 99acres were used as the data source for this study. The analysis showed that Extreme Gradient Boosting (XGBoost) model performed the best as compared to the rest of the ensemble

models. The results confirm that ensemble models can be useful for estimating real estate prices.

5. Paper Name: Prediction of House Pricing Using Machine Learning with Python

Author: Mansi Jain, Himani Rajput, Neha Garg

This paper provides an overview of how to predict house costs utilizing different regression methods with the assistance of python libraries. The proposed technique considered the more refined aspects used for the calculation of house prices and provide a more accurate prediction. It also provides a brief about various graphical and numerical techniques which will be required to predict the price of a house. This paper contains what and how the house pricing model works with the help of machine learning and which dataset is used in our proposed model.

LIMITATIONS OF EXISTING SYSTEMS :-

In India, there are multiple real estate classified websites where properties are listed for sell/buy/rent purposes such as 99acres, no broker, housing, magic bricks, and many more. However, in each of these websites, we can see a lot of inconsistencies in terms of pricing of a house and there are some cases when similar properties are priced differently and thus there is a lack of transparency and accuracy. Sometimes the customers may feel the value is not justified for a particular listed house but there is no way to confirm and check the data is accurate or not.

Proper evaluations and justified prices of properties can bring in a lot of transparency and trust back to the real estate industry, which is very important as for most consumers, especially in India the transaction prices are quite high, and addressing this issue will help both the customers and the real estate industry in the long run.

We propose to use machine learning and artificial intelligence techniques to develop an algorithm that can predict housing prices based on certain input attributes. The business application of this algorithm is that classified websites can directly use this algorithm to predict prices of new properties that are going to be listed by taking some input variables and predicting the

correct and justified price i.e. avoiding taking the wrong valuation for the house. This study is a proof-of-concept (POC) and can be treated as a valuation Report.

The purpose is to raise awareness about the correct valuation of property by accurate valuation. price inputs from customers and thus not letting any error creep into the system. This study on the proactive pricing of houses in the Indian context has never been reported earlier in the literature to the best of our knowledge. However, the problem of house price prediction is quite old and there have been many studies and competitions addressing the same including the Boston housing price challenge on Kaggle.

As far as housing price prediction in India is concerned, using machine learning techniques such as XGBoost for the prediction of housing prices in Bengaluru.

MachineHack conducted a hackathon on predicting housing prices in Bengaluru in 2018. The problem statement was to predict the price of houses in Bengaluru given 9 features such as area type, availability, location, price, size, society, total square foot, number of bathrooms, and bedrooms. Moreover, there have been other studies for house price prediction in other cities of India such as Mumbai as well.

PROPOSED SYSTEM :-

Machine Learning is a field of Artificial Intelligence that enables PC frameworks to learn and improve in execution with the assistance of information. It is used to study the construction of algorithms that make predictions on data. Machine learning is used to perform a lot of computing tasks. It is also used to make predictions with the use of computers. Machine learning is sometimes also used to devise complex models. The principle point of machine learning is to permit the PCs to learn things naturally without the assistance of people. Machine learning is very useful and is widely used around the whole world. The process of machine learning involves providing data and then training the computers by building machine learning models with the help of various

algorithms. Machine learning can be used to make various applications such as face detection applications, etc. Machine Learning is a field in software engineering that has changed the way of examining information colossally.

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. SVM Classifiers offer good accuracy and perform faster prediction compared to the Naïve Bayes algorithm. They also use less memory because they use a subset of training points in the decision phase. SVM works well with a clear margin of separation and with high dimensional space. SVMs don't output probabilities natively, but probability calibration methods can be used to convert the output to class probabilities. ... In the binary case, the probabilities are calibrated using Platt scaling: logistic regression on the SVM's scores, fit by additional cross-validation on the training data. So, here we will be using the machine learning technique of SVM to predict house prices by using various attributes to get the optimal and accurate house prices for the consumer.

SYSTEM ARCHITECTURE :-

1. In this system we take data as text input from the user and then we pre-process data of the user
2. Next we extract the required information from the data and then it is sent for classification.
3. In classification data is classified using train data set available in the system and using various algorithm price is predicted.

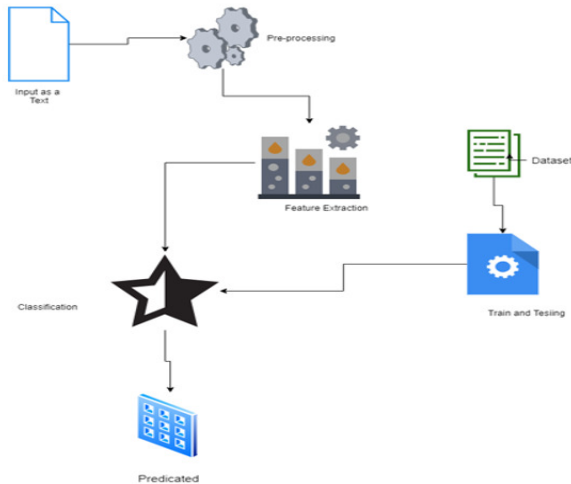


Fig 1. System Architecture

showed competitive accuracy during this time frame. Looking ahead, the paper proposes enhancements to the model with a focus on improving the accuracy of future price forecasts. One of the key areas for future work involves improving the data preprocessing phase. The authors plan to incorporate sentiment analysis into the input data, which would provide an additional layer of contextual understanding before the model undergoes training and testing. By combining structured numerical data with unstructured sentiment data—such as user reviews, market news, or public opinions—the model is expected to deliver even more precise and reliable predictions in subsequent iterations.

CONCLUSION :-

The paper focuses on the application of the Support Vector Machine (SVM) algorithm in the domain of machine learning for the purpose of house price prediction. It emphasizes how the model utilizes user-provided data, which is processed and analyzed in conjunction with an existing, pre-available dataset. By leveraging various classification algorithms, the system is able to categorize the input data and predict the likely market value of a property. The methodology highlights the importance of accurate classification, and how predictive performance is influenced not just by the algorithm itself, but significantly by the quality and comprehensiveness of the training dataset used.

The study also underscores that both the population size and the diversity of the training data play a critical role in enhancing prediction accuracy. A comparative analysis was conducted between standard SVM and an optimized version of SVM, which revealed meaningful insights into their performance differences. The optimized SVM showed improved results, demonstrating the potential of fine-tuning machine learning algorithms for better predictive outcomes. Interestingly, SVM's performance was also compared against models typically used for forecasting trends in cryptocurrencies, and it

REFERENCES :-

1. <https://towardsdatascience.com/a-data-science-web-app-to-predict-real-estate-price-d2366df2a4fd>
2. An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting Nor Azizah Hitam, Amelia Ritahani Ismail, Faisal Saeed [6] PropTech for Proactive Pricing of Houses in Classified Advertisements in the Indian Real Estate Market Sayan Putatunda Member, IEEE
3. <https://www.turcomat.org/index.php/turkbilmat/article/download/6435/5333>