iJETRM

International Journal of Engineering Technology Research & Management

HOUSE PRICE PREDICATION USING MACHINCE LEARNING AND FLASK FRAMEWORK

P.Rekha
Assistant Professor Sri Muthukumaran Institute of Technology, Anna University, Chennai.
K.Purusothaman
Sri Muthukumaran Institute of Technology, Anna University, Chennai.

ABSTRACT

The real estate market has become increasingly complex, necessitating accurate and reliable house price predictions for effective decision-making by buyers, sellers, and investors. This project presents a machine learning-based solution that employs the Random Forest algorithm to predict house prices, utilizing a comprehensive dataset that includes various property attributes such as location, size, number of bedrooms, and local amenities. By leveraging advanced data analytics techniques, the model aims to provide precise estimations that reflect current market conditions, thereby assisting stakeholders in navigating the intricacies of property transactions.

To enhance accessibility and user engagement, a web application has been developed using the Flask framework, allowing users to input property details and receive instant price predictions. The application features an intuitive interface that caters to users with varying levels of technical expertise, ensuring that vital insights are readily available to all stakeholders. In addition to price predictions, the system incorporates data visualization tools that present trends and correlations within the data, facilitating a deeper understanding of the factors influencing property values.

INTRODUCTION

The real estate market plays a crucial role in the economy, and accurate house price predictions are essential for buyers, sellers, and investors aiming to make informed decisions. Traditional valuation methods often rely on subjective assessments, comparative market analyses, and historical trends, which can lead to inaccuracies and inefficiencies. As the market becomes increasingly dynamic, the need for more sophisticated and objective tools to predict property values has emerged. Machine learning algorithms, particularly ensemble methods like Random Forest, have shown great promise in effectively analyzing large datasets and capturing complex relationships between various factors influencing house prices.

This project aims to address the challenges associated with traditional house pricing methods by developing a machine learning-based solution that employs the Random Forest algorithm. By utilizing a diverse dataset that includes key property attributes—such as location, size, number of bedrooms, and nearby amenities—our model seeks to provide accurate and reliable price estimates. The Random Forest algorithm is particularly well-suited for this task due to its ability to handle high-dimensional data, mitigate overfitting, and provide insights into feature importance. Through this approach, we aim to empower stakeholders in the real estate market with data-driven insights that enhance their decision-making processes.

### 2.1.   ARTIFICIAL INTELLINGENCE:

Artificial intelligence (AI) is the ability of a computer program a machine to think and learn. It is also field of study which tries to make computers "smart". As machines become increasingly capable, mental facilities once thought require intelligence are removed from the definition. AI is an area of computer sciences emphasizes the creation of intelligent machines that work and reacts humans. Some of the activities computers with artificial intelligence are designed for include: Face recognition, Learning, Planning, Decision making etc.,

Artificial intelligence is the use computer science programming to imitate human thought and action by analyzing data and surroundings, solving anticipating problems and learning or self-teaching to adapt a variety of tasks.

iJETRM

International Journal of Engineering Technology Research & Management

2.2.   MACHINE LEARNING

Machine learning growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms building mathematical models and making predictions using historical data information. Currently, it is being used for various tasks such as image recognition, speech recognition, Facebook auto-tagging, recommender system, and many more.

Machine Learning is said as a subset of artificial intelligence that is mainly concerned with the development of algorithms which allow a computer learn from the data and past experiences on their own. The term machine learning first introduced by Arthur Samuel in 1959. We can define it in a summarized way as: "Machine learning enables machine to automatically learn from data, improve performance from experiences, and predict things without being explicitly programmed".
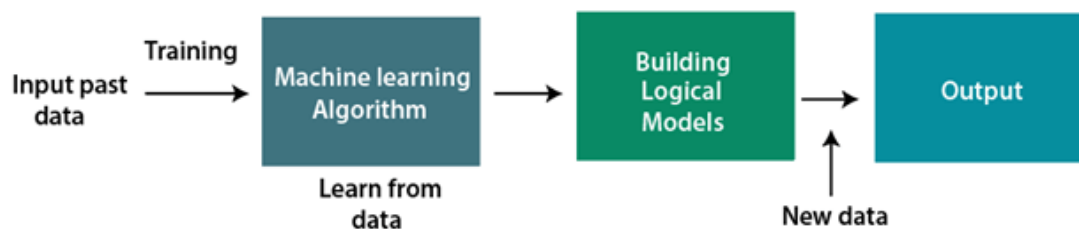
Figure 2.2 Machine Learning

2.2.1.   Features of Machine Learning:
- Machine learning uses data to detect various patterns in a given dataset.
- It can learn from past data and improve automatically.
- It is a data-driven technology.
- Machine learning is much similar to data mining as it also deals with the huge amount of the data.

2.2.2.   Classification of Machine Learning
At a broad level, machine learning classified into three types:
1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

1) Supervised Learning

Supervised learning is a type of machine learning method in which we provide sample labeled data to the machine learning system in order train it, and on that basis, it predicts the output.

The system creates a model using label data to understand the datasets and learn about each data, once the training and processing are done we test the model providing a sample data to check whether it predicting the exact output or not.

The goal of supervised learning is map input data with the output data. The supervised learning is based on supervision, and it the same as when a student learns things in the supervision of the teacher. The example supervised learning is spam filtering.

Supervised learning be grouped further in two categories of algorithms:

- Classification
- Regression

2)Unsupervised Learning

Unsupervised learning is learning method in which a machine learns without any supervision. The training is provided to the machine with the set of data that has not been labeled, classified, or categorized, and the algorithm needs act on that data without any supervision. The goal of unsupervised learning to restructure the input data into new features or a group of objects similar patterns.

iJETRM

International Journal of Engineering Technology Research & Management

In unsupervised learning, we don't have predetermined result. The machine tries to find useful insights from the huge amount data.
It can be further classifieds into two categories algorithms:

- Clustering
- Association

2.3.   DEEP LEARNING

Deep learning is branch of machine learning which is completely based artificial neural networks, as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. In deep learning, we need to explicitly program everything. The concept of deep learning not new. It has been around couple of years now. It's on hype nowadays because earlier we did not have that much processing power and lot of data. As in the last 20 years, processing power increases exponentially, deep learning and machine learning came in the picture.
In human brain approximately 100 billion neurons all together this a picture of an individual neuron and each neuron   is   connected   through   thousands   their   neighbours.
The question here is how do we recreate these neurons in computer. So, we create an artificial structure called an artificial neural net where have nodes or neurons. We have some neurons for input value and some for-output value and between, there may be lots of neurons interconnected the hidden layer.

Architectures:

1. Deep Neural Network – It is a neural network with a certain level of complexity (having multiple hidden layers in between input and output layers). They are capable modeling and processing non-linear relationships.
2. Deep Belief Network (DBN) – It is a class of Deep Neural Network. It is multi-layerbelief networks.

Steps for performing DBN:
a. Learn   a   layer   features   from   visible   units   using   Contrastive   Divergence   algorithm.
b. Treat   activations   of   previously   trained   features   visible   units   and   then   learn   features   features.
c. Finally, the whole DBN trained when learning for final hidden layer is achieved.
Recurrent (perform same task for every element of sequence) Neural Network – Allows for parallel sequential computation. Similar the human brain (large feedback network of connected neurons). They are able to remember important things about input they received and hence enables them to be more precise.

LITERATURE REVIEW:

[1] Satish, G. N., Raghavendran, C. V., Rao, M. S., & Srinivasulu, C. (2019). House price prediction using machine learning. Journal of Innovative Technology and Exploring Engineering, 8(9), 717-722.Machine learning plays a major role from past years in image detection, spam recognition, normal speech command, product recommendation, and medical diagnosis. Present machine learning algorithms help us enhance security alerts, ensure public safety, and improve medical enhancements. Machine learning systems also provide better customer service and safer automobile systems. In the present paper, we discuss the prediction of future housing prices generated by machine learning algorithms. For the selection of prediction methods, we compare and explore various techniques. We utilize lasso regression our model because of adaptable and probabilistic methodology on model selection. Our results exhibit that our approach to the issue is successful and has the ability to process predictions that are comparable to other house price prediction models. Moreover, housing value indices and the advancement of housing cost prediction tend to improve real estate policy schemes. This study utilizes machine learning algorithms as a research method to develop housing price prediction models. We create a housing cost prediction model based on machine learning algorithms, such as XGBoost, lasso regression, and neural networks, to examine their classification accuracy performance. We then recommend a housing cost prediction model to support home sellers or real estate agents with better information based on the valuation of homes. These analyses demonstrate that the lasso regression algorithm, in terms of accuracy, consistently outperforms alternative models in housing cost prediction.
[2] Thamarai, M., & Malarvizhi, S. P. (2020). House Price Prediction Modeling Using Machine Learning. International Journal of Information Engineering & Electronic Business, 12 Machine Learning is seeing its growth more rapidly in this decade. Many applications and algorithms evolve in Machine Learningday by day. One such application found journals house price prediction. House prices increasing every year, which has

iJETRM
International Journal of Engineering Technology Research & Management

necessitated the modeling house price prediction. These models constructed help customers purchase a house suitable for their needs. The proposed work makes use of the attributes or features of the houses, such as the number of bedrooms available in the house, the age of the house, traveling facilities from the location, school facilities available nearby, and shopping malls accessible near the house location. House availability based desired features and house price prediction are modeled in the proposed work, and the model is constructed for small town in the West Godavari district Andhra Pradesh. The work involves decision tree classification, decision tree regression, and multiple linear regression and is implemented using the Scikit-Learn Machine Learning tool.

[3] Zulkifley, N. H., Rahman, S. A., Ubaidullah, N. H., & Ibrahim, I. (2020). House price prediction using a machine learning model: a survey of literature. International Journal of Modern Education and Computer Science, 12(6), 46-54Data mining is now commonly applied in the real estate market. Data mining's ability extract relevant knowledge from raw data makes very useful for predicting house prices, key housing attributes, and more. Research has stated that fluctuations in house prices are often a concern for homeowners and the real estate market. A survey of literature is carried out to analyze the relevant attributes and the most efficient models to forecast house prices. The findings this analysis verified the use of Artificial Neural Networks, Support Vector Regression, and XGBoost as the most efficient models compared others. Moreover, findings also suggest locational attributes and structural attributes prominent factors predicting house prices.

[4] Park, B., & Bae, J. K. (2015). Using machine learning algorithms housing price prediction: The Fairfax County, Virginia housing data. Expert systems with applications, 42(6), 2928-2934. House sales are determined based on the Standard & Poor's Case-Shiller home price indices and the housing price index of the Office of Federal Housing Enterprise Oversight (OFHEO). These reflect the trends of US housing market. In addition to these housing price indices, the development of a housing price prediction model can greatly assist in the prediction of future housing prices the establishment of real estate policies. This study uses machine learning algorithms a research methodology to develop housing price prediction model. To improve the accuracy housing price prediction, this paper analyzes the housing data of 5359 townhouses Fairfax County, Virginia, gathered by Multiple Listing Service (MLS) of the Metropolitan Regional Information Systems (MRIS). We develop housing price prediction model based on machine learning algorithms such C4.5, RIPPER, Naïve Bayesian, and AdaBoost compare their classification accuracy performance. We then propose an improved housing price prediction model to assist a house seller a real estate agent make better informed decisions based on house price valuation. The experiments demonstrate the RIPPER algorithm, based accuracy, consistently outperforms the other models in the performance housing price prediction.

[5] Ravikumar, A. S. (2017). Real estate price prediction using machine learning (Doctoral dissertation, National College of Ireland). The below document presents the implementation of a price prediction project for the real estate markets and housing. Many algorithms are used here to effectively increase accuracy percentage. Various researchers have worked on this project and implemented algorithms like hedonic regression, artificial neural networks, AdaBoost, and J48 tree, which are considered some of the best models in price prediction. These are regarded as the base models, and with the help of advanced data mining tools, algorithms such as random forest, gradient boosted trees, multilayer perceptron, and ensemble learning models are employed, achieving a higher rate of prediction accuracy. The results and evaluation these models using machine learning advanced data mining tools like Weka and RapidMiner will have significant influence price prediction.

SYSTEM STUDY
FEASIBILITY STUDY
The feasibility the project is analyzed this phase and business proposal put forth N with a very general plan for project and some cost estimates. During system analysis the feasibility study of the proposed system is be carried out. This is to ensure that the proposed system is not burden to the company. For feasibility analysis, some understanding major requirements for the system essential. Three key considerations involved the feasibility analysis are

i. Economic Feasibility
ii. Technical Feasibility

IJETRM

International Journal of Engineering Technology Research & Management

iii. Social Feasibility

Economic Feasibility

This study is carried out check the economic impact that the system will have on the organization. The amount of fund that the company pour into the research and development of the system is limited. The expenditures must be justified. Thus, developed system as well within budget and this was achieved because most of the technologies used are freely available. Only customized products had purchased.

Technical Feasibility

This study carried out to check the technical feasibility, the technical requirements the system. Any system developed must not have a high demand the available technical resources. This will lead high demands on the available technical resources. This will lead high demands being placed the client. The developed system must have a modest requirement, as only minimal null changes are required for implementing this system.

Social Feasibility

The aspect of study is check the level of acceptance the system by the user. This includes the process training the user use the system efficiently. The user must not feel

threatened the system, instead must accept it a necessity. The level of acceptance the users solely depends on the methods that are employed educate the user about the system and to make him familiar with it. His level confidence must be raised that he is also able to make some constructive criticism, which is welcomed, as he the final user of the system.

EXISTING SYSTEM

The real estate market is characterized by its complexity and volatility, making it challenging for buyers, sellers, and investors to accurately assess property values. Traditional methods of valuation, such as comparative market analysis and appraisals, often rely on subjective judgments and may not adequately account for the multitude of factors influencing house prices. This can lead to significant discrepancies in pricing, resulting in financial losses for stakeholders and diminished trust in the market. Furthermore, as housing markets evolve due to economic shifts, demographic changes, and urban development, existing valuation techniques struggle to adapt, further complicating the decision-making process.

In addition, the reliance on outdated data and simplified models in many existing online price estimation tools contributes to inaccuracies in property valuation. Many of these systems lack the ability to process large volumes of data or to analyze the intricate relationships between various property features, leading to misleading price estimates. As a result, potential buyers may overpay for properties, while sellers may undervalue their homes, negatively impacting their financial outcomes. The absence of an effective and accurate predictive model presents a significant barrier to informed decision-making within the real estate sector.

To address these challenges, there is a pressing need for a more sophisticated approach to house price prediction that leverages advanced machine learning techniques. The Random Forest algorithm offers a powerful solution by effectively handling high-dimensional datasets, capturing complex relationships among various factors, and providing reliable predictions based on current market conditions. By developing a user-friendly web application that utilizes this algorithm, this project aims to equip stakeholders in the real estate market with accurate, data-driven insights, ultimately improving the efficiency and transparency of property transactions.

3.2. DISADVANTAGES

- Traditional Valuation Methods
- Limited Predictive Models
- Time-Consuming Processes
- Inaccurate Estimation

PROPOSED SYSTEM

The proposed system aims to develop an advanced house price prediction tool using the Random Forest algorithm, implemented through a user-friendly web application built with the Flask framework. This system will address the limitations of existing methods by utilizing a comprehensive dataset that encompasses a wide range of property features, such as location, size, number of bedrooms, and local amenities. Users will benefit from real-time price predictions through an intuitive interface that allows easy input of property details, enhancing the decision-making process for buyers, sellers, and investors. Additionally, the application will incorporate data visualization tools to present insights and trends effectively, ensuring that users can easily

iJETRM

International Journal of Engineering Technology Research & Management

understand the factors influencing property values. By continuously learning from new data, the model will adapt to market changes, improving its accuracy over time. Ultimately, this proposed system aims to empower stakeholders in the real estate market with accurate, efficient, and accessible tools for house price prediction.

ADVANTAGES
- ➢ Real-Time Analysis
- ➢ Data-Driven Decision Making
- ➢ Visual Representation
- ➢ Accuracy

CONCLUSION

In conclusion, the development of a house price prediction system utilizing the Random Forest algorithm represents a significant advancement in the application of machine learning to the real estate sector. By addressing the limitations of traditional valuation methods, this project offers a robust solution that leverages comprehensive datasets and advanced analytical techniques to provide accurate and reliable property value estimations. The integration of a user-friendly web application enhances accessibility, allowing stakeholders—including buyers, sellers, and investors—to easily obtain real-time predictions and insights based on current market conditions.

Furthermore, the project emphasizes the importance of continuous improvement through data collection and model retraining, ensuring that the predictions remain relevant and reflective of the ever-changing dynamics of the housing market. The incorporation of data visualization tools further enriches the user experience, enabling a clearer understanding of the factors influencing property values. Ultimately, this project not only contributes to better decision-making in real estate transactions but also promotes transparency and efficiency within the market.

By harnessing the power of machine learning and making sophisticated predictive analytics accessible to a wider audience, this initiative has the potential to transform the way property valuations are conducted. As the system evolves and adapts to new data and market trends, it stands to provide invaluable support to stakeholders in making informed, strategic decisions that align with their financial goals, thereby fostering a more reliable and efficient real estate market.

REFERENCES

[1] Satish, G. N., Raghavendran, C. V., Rao, M. S., & Srinivasulu, C. (2019). House price prediction using machine learning. Journal of Innovative Technology and Exploring Engineering, 8(9), 717-722.

[2] Rawool, A. G., Rogye, D. V., Rane, S. G., & Bharadi, V. A. (2021). House price prediction using machine learning. Int. J. Res. Appl. Sci. Eng. Technol, 9, 686-692.

[3] Truong, Q., Nguyen, M., Dang, H., & Mei, B. (2020). Housing price prediction via improved machine learning techniques. Procedia Computer Science, 174, 433-442.

[4] Thamarai, M., & Malarvizhi, S. P. (2020). House Price Prediction Modeling Using Machine Learning. International Journal of Information Engineering & Electronic Business, 12(2).

[5] Zulkifley, N. H., Rahman, S. A., Ubaidullah, N. H., & Ibrahim, I. (2020). House price prediction using a machine learning model: survey of literature. International Journal Modern Education and Computer Science, 12(6), 46-54.

[6] Park, B., & Bae, J. K. (2015). Using machine learning algorithms housing price prediction: The case Fairfax County, Virginia housing data. Expert systems with applications, 42(6), 2928-2934.

[7] Vineeth, N., Ayyappa, M., & Bharathi, B. (2018). House price prediction using machine learning algorithms. In Soft Computing Systems: Second International Conference, ICSCS 2018, Kollam, India, April 19–20, 2018, Revised Selected Papers 2 (pp. 425-433). Springer Singapore.

[8] Singh, A. P., Rastogi, K., & Rajpoot, S. (2021, December). House price prediction using machine learning. In 2021 3rd International Conference on Advances Computing, Communication Control and Networking (ICAC3N) (pp. 203-206). IEEE.

[9] Ahtesham, M., Bawany, N. Z., & Fatima, K. (2020, November). House price prediction using machine learning algorithm-the case Karachi city, Pakistan. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-5). IEEE.

[10] Modi, M., Sharma, A., & Madhavan, P. (2020). Applied research on house price prediction using diverse machine learning techniques. International Journal of Scientific & Technology Research, 9(04).

iJETRM

International Journal of Engineering Technology Research & Management

[11] Manasa, J., Gupta, R., & Narahari, N. S. (2020, March). Machine learning predicting house prices using regression techniques. In 2020 2nd International conference innovative mechanisms for industry applications (ICIMIA) (pp. 624-630). IEEE.

[12] Phan, T. D. (2018, December). Housing price prediction using machine learning algorithms: The case Melbourne city, Australia. In 2018 International conference machine learning and data engineering (iCMLDE) (pp. 35-42). IEEE.

[13] Wang, P. Y., Chen, C. T., Su, J. W., Wang, T. Y., & Huang, S. H. (2021). Deep learning model house price prediction using heterogeneous data analysis along joint self-attention mechanism. IEEE access, 9, 55244-55259.

[14] Chen, Y., Xue, R., & Zhang, Y. (2021, September). House price prediction based machine learning and deep learning methods. In 2021 International conference on electronic information engineering and computer science (EIECS) (pp. 699-702). IEEE.

[15] Kang, Y., Zhang, F., Peng, W., Gao, S., Rao, J., Duarte, F., & Ratti, C. (2021). Understanding house price appreciation using multi-source big data machine learning. Land use policy, 111, 104919.