

Predicting Home Prices Using Machine Learning Algorithms

Siddhant Sambodhi Rahul¹, Shubham Kumgonda Patil², Prasad Shamu Kamble³,

Siddharth Mukul Goenka⁴, F. A. Patel⁵

Students^{1,2,3,4}, Assistant Professor⁵

Department of Computer Science and Engineering

Sharad Institute of Technology College of Engineering

Email id: -prasadkamblepsk@gmail.com³

Abstract

Real estate properties are a very profitable and healthy future investment, as it can also be risky, cause it's hard to predict the value of the land as it can go higher or lower, as our main goal is to achieve the accuracy to predict the value of that property. As our main idea is to give users the most accurate value of the property they desire. We are using technology like NumPy and scikit learn to predict the datasets. We can say research in these topics can open a wide variety of fields to explore and get an advantage in the future.

Keywords: - *Real estate properties, Machine Learning Algorithms, Home Prices, Artificial intelligence*

INTRODUCTION

Buying a property is a big decision in a person's life and needs a practical amount of thought and research. One would like to buy a house at the best rate and no risk and would like it to be the perfect investment for the future. There are online websites, real estate agents, and realtors try to guide home buyers by letting them compare different houses available for purchase. As our project is completely based on the prediction of raw data of real estate that

we are collecting from previous data. There are a lot of people who want to buy properties or at least want to think of buying them, it is not ideal any buyer would buy any property without checking others or comparing to others, As before buying any property private investigation is preferred, this is where our projects come in, We are trying to predict the house prices using Machine learning algorithms Linear Regression. Our project aims to provide the best areas to invest in for a real

estate developer, individual buyers looking for a place to develop a new apartment building or to purchase. Another goal is to predict the house prices in a county in the next few years.

LITERATUREREVIEW

Machine learning is a form of artificial intelligence that prepares skilfully to train available computers without being programmed correctly. Machine learning interest in the expansion of computer programs that are capable enough to be modified when new data is insecure. Machine learning algorithms are broadly classified into three divisions, namely; supervised learning, unusable learning and reinforcement learning. Supervised learning is an education in which we teach or train a machine using data that is properly labeled, which means that some data is already tagged with the correct answer. After that, the machine is provided with a new set of examples so that the supervised learning algorithm analyzes the training data and produces the correct results from the labeled data.

Unhelpful learning is the training of a machine using information that is neither classified nor labeled, and the algorithm is allowed to act on that information without guidance. Here the function of the machine

is to group unresolved information according to similarity, pattern and difference without any prior training of data. Unlike supervised teaching, no teacher is provided, i.e., no training will be given to the machine. In this way, the machine is limited by us to track down the secret construction in unlabeled information.

Reinforcement learning is an area of machine learning. It is about taking appropriate action for the maximum reward in a particular situation. It is employed by various software and machines to find the best possible behavior or path in a specific situation. Reinforcement learning differs from supervised learning in that the training data in supervised learning is the answer key, so the model is trained with the correct answer, whereas reinforcement learning has no answer, but the reinforcement agent decides it does what to do to perform the given task. Without a preparation dataset, it will undoubtedly gain from its experience.

Machine learning has many applications, one of which is the prediction of application real estate. The real estate market is the most competitive in terms of pricing and varies greatly depending on a

lot of factors, forecasting the price of the property to support budget allocation, making decisions for both buyers and investors in finding a property. Therefore, locating strategies and determining appropriate policies have become one of the major areas of applying machine learning concepts to optimize and predict prices with high accuracy. Studies on land price trend have been considered important to support decisions in urban planning.

The land framework is a shaky stochastic interaction. Investors' decisions are based on market trends to achieve maximum returns. Developers are interested in knowing the future trends to make their decisions. In order to accurately estimate property prices and future trends, analysis, modeling, and forecasting require a large amount of data affecting the price of land.

Eric Sloan (2014) developed relationships between different household characteristics, and the asking price of a residential property was analyzed using both simple linear regression and multiple linear regression using the method of simple least squares. Home square footage was used as an explanatory variable in simple linear regression, and multiple linear regression included land parcel size, number of bedrooms, year of construction,

and other explanatory variables.

The results of multi-linear regression proved to be biased due to omissions of important factors in simple linear regression. Home square footage was found to be the most important factor in determining the price of a residential property, while garage capacity proved to be the weakest factor. [2]

OBJECTIVES

- To predict the price of a given home according to the market price.
- Easily known the future price home which is suitable for our budget
- Minimizing the difference between the actual price and predicted price.
- Easily accessible each and every information on given location within just a few minutes

METHODOLOGY

Idea: Our basic idea is to create an application which will help user to find their desired property. To achieve this goal, we are using machine learning and data science. These are the steps followed:

Step1: To get the raw data for Kaggle. Kaggle provides the largest datasets and

resources for data science and machine learning projects.

Step 2: From Kaggle, we downloaded the dataset and started doing an analysis of that dataset. For analysis, we used different tools like jupyter Notebook and libraries like Numpy, Pandas scikit-learn etc. with the use of tools and libraries, we separated useful and unusual data.

Step 3: Later, we separated the dataset into training and testing's datasets. Afterwards, we used different algorithms like Linear regression, Decision Tree, Lasso (least check the accuracy & from the all algorithms that we checked the highest accuracy is shown absolute shrinkage and selection operator) algorithm. These algorithms we used to check the accuracy, and from all algorithms that we checked, the highest accuracy is shown by linear regression.

Step 4: Further on, we used a linear regression algorithm on the actual data to predict the price of that real estate, for the price prediction of the real estate using Machine learning. Our team is currently using the website that we have created. For Front-end, we used technologies like bootstrap, CSS & HTML. For Back-end, we used Flask technology.

CONCLUSION

In this paper, the linear regression model is used to predict the selling price of any real estate property. Decision tree, linear regression out of two this linear regression model has best accurate accuracy than a decision tree. We use 80% data for training and 20% data for testing. An accurate calculated value this value is closer to the actual value, which means easily predict the accurate value closer to the actual value. Using Flask framework developed user interface. And it has best advantages is easily known what's future price.

REFERENCES

1. Akshay Babu¹, Dr. Anjana S Chandran², "Literature Review on Real Estate value prediction using Machine learning", (3 March 2019)./www.ijcsma.com
2. Rohan Bafna¹,Anirudh Dhole²,Ankit Jagtap³,Asif Kazi⁴,Arbaz Kazi⁵ "Prediction of Residential Property Prices – A State of the Art" (3 march 2018).DOI.10.17148/IARJSET
3. Nihar Bhagat¹, Ankit Mohokar²,Shreyash Mane³, "House price Forecasting using

Data mining”,
October2016/International journal
of computer application(0975-
8887)

4. Atharva Chogle¹,Priyanka
Khaire², Akshata Gaud³,Jinal
Jain⁴, International journal of
advanced research in computer and
communication engineering,
“House price forecasting using
data mining techniques” Vol. 6.
Issue 12,
December2017/IJARCCE