

B. TECH CSE IIIrd year
SEMESTER: VIth
Real Estate Price Prediction

MINI PROJECT II

SYNOPSIS



Department of Computer Science &
Application

Institute of Engineering & Technology

SUBMITTED TO: -

Mr. Ankit Arora

SUBMITTED BY:

Ayush Kumar (201500179)

Acknowledgement

It gives us a great sense of pleasure to present the synopsis of the B. Tech mini project undertaken during B. Tech III Year. This project is going to be an acknowledgement to the inspiration, drive and technical assistance will be contributed to it by many individuals. We owe special debt of gratitude to Md. Amir Khan, Technical Trainer, for providing us with an encouraging platform to develop this project, which thus helped us in shaping our abilities towards a constructive goal and for his constant support and guidance to our work.

His sincerity, thoroughness and perseverance has been a constant source of inspiration for us. We believe that he will shower us with all his extensively experienced ideas and insightful comments at different stages of the project & also taught us about the latest industry-oriented technologies. We also do not like miss the opportunity to acknowledge the contribution of all faculty members of the department for their kind guidance and co-operation.

Ayush Kumar

Declaration

I hereby declare that the project entitled “**Real Estate Price Prediction**”, which is being submitted as Mini Project of 6th semester in Computer Science & Engineering to **GLA University**, Mathura is authentic record of our genuine work done under the guidance of Mr. **Ankit Arora**, Dept. of Computer Science and Engineering, GLA

University, Mathura.

GitHub link:

<https://github.com/Ayushkumar1240/miniproject2>

Date: 25/04/2023

Place: Mathura

Ayush Kumar

ABSTRACT

Real estate price prediction is a critical task for buyers, sellers, and investors to make informed decisions in the dynamic real estate market. In this project, we utilize data analytics and machine learning techniques to develop accurate price prediction models for residential properties. We collect extensive data on various factors that affect real estate prices, including location, square footage, number of bedrooms and bathrooms, and nearby amenities. The collected data is then analyzed using advanced data analytics techniques to identify patterns and trends. We apply various machine learning algorithms, such as linear regression, decision trees, random forests, and neural networks, to create predictive models that can accurately forecast real estate prices.

My project aims to provide valuable insights into the real estate market by leveraging the power of big data and machine learning. Our models are trained on historical data, allowing them to capture market dynamics and trends. We evaluate the performance of our models using rigorous statistical measures and cross-validation techniques to ensure their accuracy and reliability. We also explore feature engineering techniques to enhance the predictive power of our models.

The results of my project are expected to benefit a wide range of stakeholders, including homebuyers looking for accurate price estimates, sellers seeking optimal pricing strategies, and investors interested in making informed investment decisions. My project contributes to the growing field of real estate analytics and demonstrates the potential of data-driven approaches in predicting real estate prices.

Contents

Abstract

Declaration

Acknowledgement

1. Introduction

2. Objective

3. Tools and Technologies Used

4. Working and Implementation

5. Future Scope

6. References

INTRODUCTION

Real estate price prediction is a fascinating and challenging field that involves using statistical and machine learning techniques to forecast the prices of properties. With the rise of big data and the availability of real estate market data, researchers and practitioners are increasingly turning to data-driven methods to make accurate predictions of housing prices. In this mini project, we will explore the topic of real estate price prediction and build a machine learning model to forecast the prices of houses based on various features such as location, square foot, number of bedrooms, and more. By the end of this project, you will have a better understanding of the techniques used in real estate price prediction and the ability to apply them to your own data.

Objective: -

The objective of this mini project is to develop a machine learning model that can accurately predict the prices of real estate properties based on relevant features such as location, size, age, and amenities. The model will be trained on a dataset of historical real estate prices, and its performance will be evaluated based on its ability to accurately predict the prices of new, unseen properties.

Specifically, the project will involve the following steps:

1.Data collection: Collecting a dataset of real estate properties with relevant features such as location, size, age, and amenities, as well as their corresponding sale prices.

2.Data preprocessing: Cleaning and preparing the data for analysis, including handling missing data, encoding categorical variables, and scaling numerical features.

3.Feature selection: Identifying the most important features that are strongly correlated with the sale prices of real estate properties.

4. Model training: Developing a machine learning model that can predict the sale prices of real estate properties based on the selected features.

5. Model evaluation: Evaluating the performance of the model using appropriate metrics such as mean absolute error (MAE) and mean squared error (MSE).

6. Model deployment: Deploying the trained model into a userfriendly application that can take in input data about a new real estate property and output a predicted sale price.

The ultimate goal of this mini project is to create a reliable and accurate real estate price prediction model that can help buyers and sellers make informed decisions about the value of properties in the market.

Tools and Technologies:

- Python programming language
- Scikit-learn library for machine learning
- Pandas library for data manipulation
- Matplotlib library for data visualization
- Flask to create and scale web apps quickly and simply
- Tableau for dashboard designing.

Working and Implementation

Data Collection:

- Collect data on real estate properties in a particular area from online sources or through scraping websites like Zillow, Redfin, or Realtor.com.
- Data should include attributes such as square footage, number of bedrooms/bathrooms, location, year built, lot size, etc.
- It is important to ensure the data is clean and free of outliers.

Data Preprocessing:

- Split the data into training and testing sets.
- Normalize or standardize the numerical features.
- Encode the categorical features using techniques such as one-hot encoding or label encoding.

Model Selection:

- Experiment with different regression models, such as linear regression, decision trees, random forests, and gradient boosting.
- Evaluate each model's performance using metrics such as mean squared error, R-squared, and cross-validation.

Model Tuning:

- Fine-tune the selected model by adjusting hyperparameters such as learning rate, regularization strength, and number of estimators.
- Evaluate the tuned model's performance on the test set.

Results and Visualization:

- Visualize the performance of the model using various charts such as scatter plots, residual plots, and feature importance plots.
- Compare the model's predicted values to the actual sale prices using metrics such as root mean squared error and mean absolute error.

Implementation:

Step 1. Cleaning and processing the data on Google Collab using python libraries.

Step 2. Creating a pickle file and a json file for further implementation.

Step 3. Making a python flask server

Step 4. Making a UI using html, CSS and JavaScript.

Future Scope

Real estate price prediction is a challenging task that requires careful analysis of multiple factors that affect the market. As such, there are several areas where future research can be focused to improve the accuracy and usefulness of real estate price prediction models. Here are some potential future scopes for a real estate price prediction project:

Integration of Alternative Data Sources: In addition to traditional sources of real estate data such as transaction records and property characteristics, future projects could incorporate alternative data

sources such as social media data, satellite imagery, and sensor data to improve the accuracy of real estate price prediction models.

Use of Machine Learning Techniques: While traditional statistical models have been used extensively in real estate price prediction, machine learning techniques such as deep learning and reinforcement learning have shown promise in recent years. Future projects could explore the use of these techniques to improve prediction accuracy.

Incorporation of Real-time Data: Real estate markets are dynamic and constantly changing, making real-time data critical for accurate price prediction. Future projects could explore the use of real-time data such as web scraping, IoT sensors, and news articles to improve the timeliness and accuracy of predictions.

Integration of Spatial Analysis: The location of a property is a critical factor in determining its price. Future projects could incorporate spatial analysis techniques such as geospatial data mining, spatial econometrics, and spatial machine learning to account for spatial dependencies in real estate data.

RESOURCES:

Google

Flask Documentation

Pandas Documentation

Kaggle

