**House price prediction**

***Submitted in partial fulfillment of the requirements for the award of the Degree***

***of***

**Bachelor of Technology (B. Tech)**

**in**

**INFORMATION TECHNOLOGY**

**By**

**Ramavath Swamy(22AG1A1252)**

**Rayala Charanya(22AG1A1253)**

**Suryawanshi Steevan(22AG1A1259)**

**Under the esteemed Guidance of**

Mr. G. Prasad

Assistant Professor



**Department of Information Technology**

**ACE ENGINEERING COLLEGE**

**An Autonomous Institution**

All the courses are Accredited by NBA and NAAC with A Grade

**(Affiliated to Jawaharlal Nehru Technological University, Hyderabad, Telangana)**

Ankushapur, Ghatkesar, Medchal, Hyderabad – 501 301

ACE

**Engineering College**

**An Autonomous Institution**

**All the courses are Accredited by NBA and NAAC with A Grade**

**(Affiliated to Jawaharlal Nehru Technological University, Hyderabad, Telangana)**

**Website : www.aceec.ac.in E-mail: info@aceec.ac.in**This e-mail address is being protected from spambots, you need JavaScript enabled to view it

### CERTIFICATE

This is to certify that the Project work entitled “**HOUSE PRICE PREDICTION”** is being

Submitted by **Ramavath Swamy(22AG1A1252), Rayala Charanya(22AG1A1253), Suryawanshi Steevan(22AG1A1259**) in partial fulfilment for the award of Degree of **BACHELOR OF TECHNOLOGY in INFORMATION TECHNOLOGY** to the Jawaharlal Nehru Technological University, Hyderabad during the academic year 2023-2024 is a record of bona-fide work carried out by them under our guidance and supervision.

The results embodied in this report have not been submitted by the students to any other university or institution for the award of any degree or diploma.

**Internal Guide Head of the Department**

**Mr. G. PRASAD Prof. K. JAYA BHARATHI**

**Assistant professor Professor and Head Dept. of IT**

# ACKNOWLEDGEMENT

We would like to express our gratitude to all the people behind the screen who have helped us to transform an idea into a real time application.

We would like to express our heart-felt gratitude to our parents without whom we would not have been privileged to achieve and fulfil our dreams.

A special thanks to our Secretary, **Prof. Y. V. GOPALA KRISHNA MURTHY** and Joint Secretary, **Mrs. M. PADMAVATHI** for having founded such an esteemed institution. We are also grateful to our beloved principal, **Dr. B. L. RAJU** for permitting us to carry out this project.

We profoundly thank **Prof. K. JAYA BHARATHI**, Head of the Department of Information Technology, who has been an excellent guide and a great source of inspiration to our work.

We are very thankful to our internal guide Mr. **G. Prasad**, who has been an excellent and given continuous support for the completion of our project work.

The satisfaction and euphoria that accompany the successful completion of the task would be great, but incomplete without the mention of the people who made it possible, whose guidance and encouragement crown all the efforts with success. In this context, we would like to thank all the other staff members, teaching and non- teaching, who have extended their timely help and eased our task.

**Ramavath Swamy(22AG1A1252)**

**Rayala Charanya(22AG1A1253)**

**Suryawanshi Steevan(22AG1A1259)**

# DECLARATION

We hereby declare that the project entitled “**HOUSE PRICE PREDICTTION**” submitted in partial fulfilment of the requirements for the award of degree of Bachelor of Technology in Information Technology. This dissertation is our original work and the project has not formed the basis for the award of any degree, associate ship, fellowship or any other similar titles and no part of it has been published or sent for the publication at the time of submission.

**Ramavath Swamy(22AG1A1252)**

**Rayala Charanya(22AG1A1253)**

**Suryawanshi Steevan(22AG1A1259)**

# ABSTRACT

**HOUSE PRICE PREDICTION**

The project titled house price prediction is a machine learning project mainly using python. Predicting house prices is a critical task in the realms of real estate and finance, significantly impacting decisions made by homeowners, investors, and policymakers. Traditional prediction methods rely heavily on historical data and expert judgment.

However, recent advancements in machine learning offer the potential for more accurate and efficient forecasting. This project presents a novel approach to predicting house prices using state-of-the-art machine learning algorithms. The evaluation of model performance incorporates metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R2) to gauge predictive accuracy and generalization capability.

Robustness is ensured through cross-validation techniques, validating the models on different subsets of data to guarantee reliability and performance consistency.

New algorithms in machine learning like random forest and decision tree. Most of previous one used linear regression it gives less accuracy and high prices. Using random forest our prediction has done.

# INDEX

**CONTENTS Page No.**

**List Of Figures** viii

**List Of Abbreviations** ix

**CHAPTER 1: INTRODUCTION 1-3**

1.1 Overview 1

* 1. Purpose 2

1.3 Scope 3

**CHAPTER 2: LITERATURE SURVEY 4-6**

2.1 Existing System 5

2.2 Proposed System 6

**CHAPTER 3: SOFTWARE REQUIREMENT ANALYSIS 7-9**

3.1 Software Requirement Specification 7

3.1.1 Functional Requirements 8

3.1.2 Non-Functional Requirements 9

**CHAPTER 4: SOFTWARE DESIGN 10-13**

4.1 Introduction 10

4.2 Architecture 11

4.3 Dataflow Diagrams  **12-13**

**CHAPTER 5**: **Software &Hardware Requirements**

5.1 Hardware Requirements 14

5.2 Software Requirements

**CHAPTER 6: MODULES**

# 6.1 Data Collection and Preprocessing

# 6.2 Feature Selection and Engineering 15

# 6.3 model testing

# 6.4 model evaluation

**CHAPTER 7: CODING TEMPLATES 16-18**

**CHAPTER 8: TESTING AND VALIDATION 19-20**

8.1Software testing 19

8.2 Accuracy evaluation

8.3 Cross validation 20

8.4 Sensitivity analysis

**CHAPTER 9: OUTPUT SCREENS 21-22**

**CHAPTER 10: CONCLUSIONS 23**

**CHAPTER 11: FUTURE ENHANCEMENTS 24-25**

**CHAPTER 12: REFERENCES 26**

# 

# LIST OF FIGURES

**List of Figures Page No.**

1. Architecture………………………………………………………………....11
2. DFD Level – 0……………………………………………………………...12
3. DFD Level – 1………………………………………………………………13

**LIST OF ABBREVIATIONS**

**Abbreviation**

SRS SOFTWARE REQUIREMENT SPECIFICATIONS

DFD Data Flow Diagram

# CHAPTER 1

# INTRODUCTION

## 

### The housing market is a cornerstone of the global economy, affecting a wide range of stakeholders from individual homeowners and real estate investors. Traditional methods of price prediction often depend on historical data analysis and expert opinions, which, while valuable, can be limited in their ability to adapt to rapidly changing market conditions and diverse influencing factors.

### Machine learning models can process vast amounts of data and identify complex patterns and relationships that are not immediately apparent through traditional methods.

### This project aims to leverage advanced machine learning algorithms to predict house prices based on a variety of factors including area, number of bedrooms, amenities, and location.

### By incorporating extensive datasets and employing rigorous evaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R2), the project ensures a comprehensive analysis of model performance.

### 1.1 Overview:

This project on house price prediction leverages cutting-edge machine learning techniques to forecast real estate values.

The primary objective is to develop a robust predictive model that accurately estimates house prices based on various influential factors such as area, number of bedrooms, amenities, and location.

By combining comprehensive data collection, meticulous preprocessing, advanced machine learning techniques, and interactive deployment, this project aims to provide a reliable and insightful tool for predicting house prices.

This tool can significantly aid homeowners, investors, and policymakers in making informed decisions, ultimately contributing to a more transparent and efficient real estate market.

### 1.2 Purpose:

The purpose of this project is to develop a sophisticated machine learning-based system for accurately predicting house prices. This system is designed to serve multiple stakeholders, including homeowners, real estate investors, policymakers, and urban planners, by providing them with reliable and actionable insights into the housing market.

key purpose of this project is Enhancing Decision-Making, Supporting Policymaking and Urban Planning, Advancing Technological Integration.

By fulfilling these purposes, this project aims to not only improve the accuracy of house price predictions but also to empower various stakeholders with the tools and insights needed to navigate the complex landscape of the real estate market effectively.

* 1. **Scope**

The scope of this project encompasses the comprehensive development and deployment of a machine learning-based house price prediction system.

This project covers various aspects, from data collection to model deployment, ensuring a robust and user-friendly tool for stakeholders.

The detailed scope includes data collecting and management, data preprocessing, deployment, Model Evaluation and Validation, Scalability and Maintenance.

By encompassing these aspects, the scope of this project ensures a holistic approach to house price prediction, delivering a reliable and efficient tool for various stakeholders in the real estate market.

**CHAPTER 2**

## LITERATURE SURVEY

* **www.streamlit.com**

Author: Anubhav Kulshrestha, Akas deep Gupta

Publisher: Matheus de Mattos July 2022

Technology: Machine learning

Summary: This project Showing high amount for houses and only focus on some areas only.

But it covering metropolitan areas and doing it best based on linear regression.

* **www.housing.com**

Author: Vidhya, Hema

Publisher: REA India Year: 2012

Technology: Machine learning

Summary: This project Showing high amount for houses and only focus on some areas only.

But it covering metropolitan areas and doing it best based on linear regression. It is having drawback like old data present not uploading new things.

* **House price prediction**

Author: Rushab Sawant, Yashwant Jangid, Tushar Tiwari, Saurabh Jain, Ankita Gupta

Publisher: 2018

Technology: Machine learning

Summary: This project showing better accuracy for house price. In this project random forest and decision tree used Random forest provides best R2 score 0.9996. It not holding high amount of data.

#### **EXISTING SYSTEM**

### The existing systems for house price prediction primarily focus on utilizing historical data, property features, and economic indicators. These systems depend heavily on past data to forecast future prices, analyzing historical trends and patterns to make predictions. They consider various property characteristics such as size, number of bedrooms, location, and amenities to create predictive models. Additionally, broader economic factors like interest rates, employment rates, and inflation are incorporated to understand the market conditions influencing house prices.

### The techniques commonly used in these systems include linear regression, time series analysis, and data mining. Linear regression is employed to identify relationships between property features and prices, while time series analysis helps in understanding trends over time.

### However, these existing systems have several limitations. They often struggle with capturing complex nonlinear relationships within the data and rely on predefined features, which may not fully capture the intricacies of the housing market. Moreover, handling large and diverse datasets poses significant challenges, limiting the system's ability to integrate and analyze new and varied data effectively. As a result, there is a need for more advanced approaches that can overcome these limitations and provide more accurate and comprehensive house price predictions.

### LIMITATIONS:

* Non-linear data prediction is difficult.
* Difficult to hold large data for prediction.
* Not covering entire locations.
  1. **PROPOSED SYSTEM**

The proposed system introduces data enhance new techniques and advanced data preprocessing to improve accurate prediction. Utilizing new models like random forest and neural networks. Collection of data not only based on surveys we can use social media and satellite images. Extracting data from platforms like Twitter, Facebook, or Instagram to capture sentiment analysis, public opinions, or trends related to real estate markets.

Using satellite imagery to gather spatial data such as land use, infrastructure development, environmental factors, or proximity to amenities. This data can provide valuable insights into property values and location-specifictrends**.** Ensemble learning method that builds multiple decision trees and aggregates their predictions, providing robustness and handling non-linear relationships well.

Deep learning models capable of learning complex patterns and relationships in data through multiple layers of interconnected neurons. They excel in tasks requiring high-dimensional data and nonlinear transformations.

**ADVANTAGES:**

# Improved Prediction Accuracy.

# Robustness and Generalization

# Reduce house cost.

# CHAPTER 3

**SOFTWARE REQUIREMENT ANALYSIS**

## SOFTWARE REQUIREMENT SPECIFICATIONS (SRS)

## Functional requirements

**Data Collection and Integration**

The system should be able to collect and integrate data from multiple sources, including structured datasets (e.g., surveys, real estate listings) and unstructured sources (e.g., social media, satellite images).

Implement data ingestion pipelines that can handle different formats and sources, ensuring data quality and consistency.

**Data Preprocessing and Cleaning**

Perform preprocessing tasks to clean and transform raw data into a usable format for modelling. Functionality: Include functionalities for handling missing values, outlier detection, normalization, and feature engineering. Ensure robustness in handling diverse data types and quality issues.

**Model Development and Training**

Develop predictive models using advanced techniques such as Random Forests and Neural Networks. Implement algorithms for model training, hyperparameter tuning, and model evaluation using techniques like cross-validation. Ensure scalability for handling large datasets and complex models.

**Visualization and Reporting**

Visualize model predictions, insights, and data trends to facilitate decision-making.

Develop dashboards or reporting tools that present model outputs, feature importance, and data analytics in an intuitive manner. Support interactive exploration of results and drill-down capabilities.

**3.1.2 Non-functional requirements**

**Performance**

* **Requirement**: The system should handle large volumes of data efficiently.
* **Criteria**: Ensure that data processing, model training, and prediction responses meet defined performance benchmarks (e.g., response time, throughput) under expected load conditions.

**Scalability**

* **Requirement**: The system should scale to accommodate increasing data sizes and user demands.
* **Criteria**: Design components to horizontally scale (e.g., distributed processing for data ingestion, parallel model training) and handle concurrent user requests without degradation in performance.

**Reliability**

* **Requirement**: The system should operate reliably without frequent failures or downtime.
* **Criteria**: Implement fault-tolerant mechanisms (e.g., redundancy in data storage, automatic recovery from failures) and perform rigorous testing (e.g., stress testing, failure simulation) to ensure robustness.

**Security**

* **Requirement**: Ensure data confidentiality, integrity, and availability throughout the system.
* **Criteria**: Implement encryption for data in transit and at rest, enforce access control mechanisms (e.g., authentication, authorization), and adhere to industry standards and regulations (e.g., GDPR, HIPAA) for data protection.

**Usability**:

* **Requirement**: The system should be intuitive and easy to use for stakeholders, including data scientists, analysts, and decision-makers.
* **Criteria**: Design user interfaces that are user-friendly, provide clear feedback, and support customization (e.g., dashboards, visualization tools) to facilitate effective data exploration and interpretation.

**Maintainability**:

* **Requirement**: The system should be easy to maintain and support for ongoing enhancements and updates.
* **Criteria**: Use modular and well-documented code, adhere to coding best practices, and provide version control for models and software components. Implement logging and monitoring tools for diagnostics and troubleshooting.

**Compatibility**:

* **Requirement**: Ensure compatibility with existing IT infrastructure, software systems, and data sources.
* **Criteria**: Validate interoperability with commonly used platforms, databases, and data formats. Provide APIs or integration points to facilitate data exchange and interoperability.

# CHAPTER 4

# SOFTWARE DESIGN

### 4.1 INTRODUCTION

The design phase of the **house price prediction** project involves creating a detailed plan for the predicting and implementation of the system. This includes identifying the functional and technical requirements for the system, as well as determining the appropriate hardware and software components needed to support it.

The design phase also involves creating mock- ups and prototypes of the system to facilitate user testing and feedback, as well as developing a project timeline and budget. The end result of the design phase will be a comprehensive plan for the creation of house price prediction.

### 4.2 ARCHITECTURE

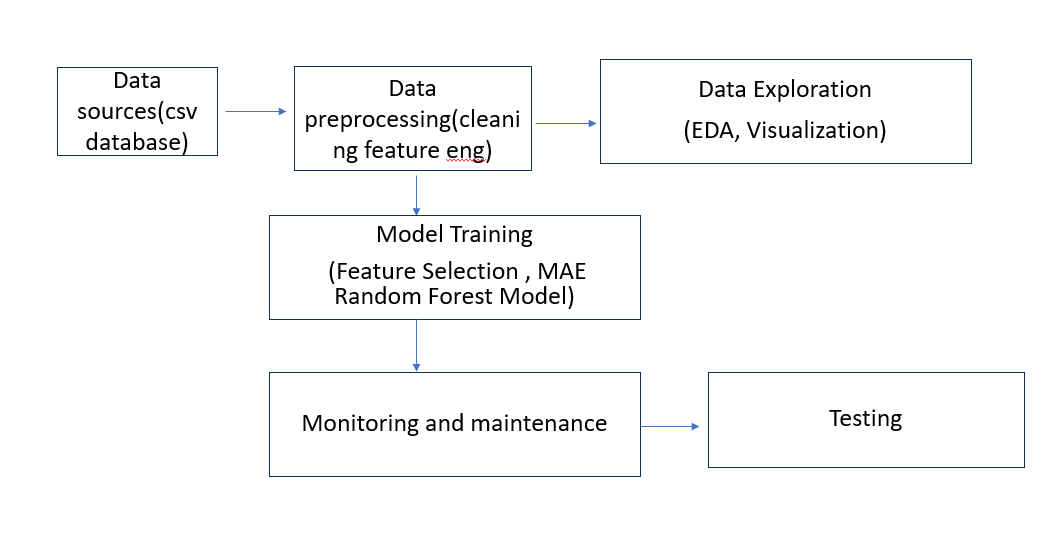
****

Fig: 4.1 Architecture

### 4.3 DATA FLOW DIAGRAM

DFD is a visual representation of the flow of data within a system, showing how data is transformed and moved from one process or entity to another. A DFD consists of a set of symbols and connectors that represent the various processes, entities, and data flows in the system. The symbols are used to represent different types of elements in the system, such as processes, data stores, and external entities and the connectors are used to show the flow of data between these elements. DFDs are used to model and design systems, as well as to document and communicate the structure and behavior of existing systems. They can be used to identify and analyze the data flows within a system, as well as to identify potential bottlenecks or other issues that may need to be addressed.

**4.3.1 DFD diagram**

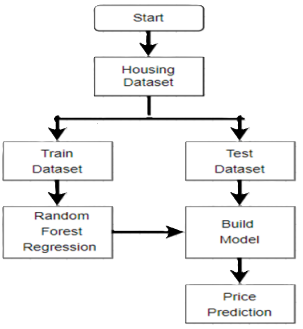


Fig: 4.2 data flow diagram

**4.3.2 DFD diagram**

# 

Fig: 4.3 data flow diagram

# CHAPTER 5

# Software and Hardware requirements

**5.1 Software requirements**

* Programming language: Python 3.9
* Machine learning libraries: scikit-learn, seaborn, widgets.
* Data analysis libraries: pandas, numPy

**5.2 Hardware requirements:**

* Processor: Intel core i5
* RAM:16 GB DDR4
* Operating system: windows 11
* Storage:1TB SSD

# CHAPTER 6

**MODULES**

# 6.1 Data Collection and Preprocessing

# pandas: For data manipulation and preprocessing, pandas offer functionalities such as reading data from various file formats, cleaning data, and handling missing values. Common methods include pd. read\_ csv () to read CSV files, df. Fillna () to fill missing values, and pd.get\_ dummies () to convert categorical variables to dummy/indicator variables.

# numpy: This library is used for efficient array computations and handling numerical data. Key methods include np. Array () to create an array, np. Mean () to compute the mean, and np. Nan

# to represent missing values.

# 6.2 Feature Selection and Engineering

# This module is used for selecting the most important features based on statistical tests. Common classes/functions include Select KBest, which selects features according to the k highest scores and f\_ regression, which performs an F-test for regression.

# Sklearn. preprocessing: In addition to data preprocessing, this module also offers tools for feature scaling and normalization. Key classes/functions include Min Max Scaler, which transforms features by scaling each feature to a given range.

# 6.3 Model Training

# This module contains ensemble methods, including the RandomForest Regressor for training models. The RandomForest Regressor is a commonly used regressor within this module.

# 6.4 Model Evaluation

# This module provides various metrics to evaluate the performance of a model. Common classes/functions include mean\_ squared\_ error for mean squared error regression loss and r2\_score for the R^2 (coefficient of determination) regression score.

# In summary, pandas is used for data manipulation and preprocessing with functionalities for reading data, cleaning, and handling missing values. Numpy assists in

# providing efficient array computations and handling numerical data.

# CHAPTER 7

**CODING TEMPLATES**

**7.1**

import pandas as pd

import NumPy as np

from sklearn. impute import Simple Imputer

from sklearn. Pipeline import Pipeline

from sklearn. Preprocessing import Standard Scaler, One Hot Encoder

from sklearn. Compose import Column Transformer

from sklearn. Ensemble import RandomForest Regressor

# Load datasets

datasets = {

    "Hyderabad": pd. read\_ csv('C:/Users/chara/Downloads/Hyderabad.csv'),

    "Delhi": pd. read\_ csv('C:/Users/chara/Downloads/Delhi.csv'),

    "Chennai": pd. read\_ csv('C:/Users/chara/Downloads/chennai.csv'),

}

# Combine datasets

combined\_ df = pd. Concat (datasets. Values ())

# Extract unique locations

all\_ locations = sorted (set (combined\_ df['Location']. Unique ()))

# Define features and target

features = ["Area", "No. of Bedrooms", "Swimming Pool", "Shopping Mall", "ATM",

            "Power Backup", "Car Parking", "AC", "Wifi", "Lift Available", "TV", "Location"]

target = "Price"

# Prepare training data

X\_ train = combined\_ df[features]

y\_ train = combined\_ df[target]

# Define preprocessing pipelines

numeric\_ features = X\_ train. select\_ d types (include= [np. Number]).columns. tolist ()

categorical\_ features = ["Location"]

numeric\_ transformer = Pipeline (steps= [

    ('imputer', Simple Imputer (strategy='most\_ frequent')),

    ('scaler', Standard Scaler ())

])

categorical\_ transformer = Pipeline (steps= [

    ('imputer', Simple Imputer (strategy='most\_ frequent')),

    ('onehot ', One Hot Encoder (handle\_ unknown='ignore'))

])

preprocessor = Column Transformer (

    transformers= [

        ('num', numeric\_ transformer, numeric\_ features),

        ('cat', categorical\_ transformer, categorical\_ features)

])

# Define model pipeline

model\_ pipeline = Pipeline (steps= [

    ('preprocessor', preprocessor),

    ('regressor', RandomForest Regressor ())

])

# Train the model

model\_ pipeline. Fit (X\_ train, y\_ train)

def process\_ data (area, bedrooms, Swimming Pool, shopping\_ mall, atm, power\_ backup, car\_ parking,

                 ac, wifi, lift\_ available, tv, location):

data = {

        "Area": area,

        "No. of Bedrooms": bedrooms,

        "Swimming Pool": Swimming Pool,

        "Shopping Mall": shopping\_ mall,

        "ATM": atm,

        "Power Backup": power\_ backup,

        "Car Parking": car\_ parking,

        "AC": ac,

        "Wifi": wifi,

        "Lift Available": lift\_ available,

        "TV": tv,

        "Location": location

    }

    # Create Data Frame from input data

    input\_ df = pd. Data Frame (data, index= [0])

    # Predict using the model pipeline

    prediction = model\_ pipeline. Predict (input\_ df) [0]

    return f" Predicted apartment price: ${prediction:.2f}"

from ipy widgets import interact, widgets

@interact (

model\_ pipeline=widgets. Dropdown (options=models. Keys (), description='Model'),

area=widgets. Float Text (description='Area'),

bedrooms=widgets. Intext(description='Bedrooms'),

Swimming Pool=widgets. Checkbox(description='Swimming Pool'),

shopping\_ mall=widgets. Checkbox (description='Shopping Mall'),

atm=widgets. Checkbox (description='ATM'),

power\_ backup=widgets. Checkbox (description='Power Backup'),

car\_ parking=widgets. Checkbox (description='Car Parking'),

ac=widgets. Checkbox (description='AC'),

wifi =widgets. Checkbox (description='Wifi'),

lift\_ available=widgets. Checkbox (description='Lift Available'),

tv=widgets. Checkbox (description='TV'),

location=location\_ dropdown

)

def interactive\_ prediction (model\_ pipeline, area, bedrooms, Swimming Pool, shopping\_ mall, atm, power\_ backup,

car\_ parking, ac, wifi, lift\_ available, tv, location):

# Get the selected model pipeline from the model dictionary

model = models [model\_ pipeline]

# Call make\_ prediction with all parameters

prediction = make\_ prediction (model, area, bedrooms, Swimming Pool, shopping\_ mall, atm, power\_ backup,

car\_ parking, ac, wifi,

lift\_ available, tv, location)

return prediction

## CHAPTER 8

## TESTING AND VALIDATION

## 8.1 SOFTWARE TESTING

Software testing is a critical element of software quality assurance and represents the ultimate review of specification, design and coding. In fact, testing is the one step in the software engineering process that could be viewed as destructive rather than constructive.

A strategy for software testing integrates software test case design methods into a well-planned series of steps that result in the successful construction of software. Testing is the set of activities that can be planned in advance and conducted systematically.

The underlying motivation of program testing is to affirm software quality with methods that can economically and effectively apply to both strategic to both large and small-scale systems.

### 8.2 Accuracy Evaluation

The model's performance is assessed using metrics like mean squared error (MSE), root mean squared error (RMSE), and R-squared to determine its predictive accuracy.

**8.3 Cross-Validation**

The model is tested on multiple subsets of the data to ensure its performance is consistent and not overfitted to the training data.

Cross-validation is employed to ensure that the model's performance metrics are reliable and not overly optimistic due to overfitting.

It involves partitioning the data into multiple subsets, training the model on a subset (training set), and evaluating it on the remaining subsets (validation sets).

This process is repeated multiple times, typically using different partitions, to obtain a more accurate estimate of the model's performance on unseen data.

* 1. **Sensitivity Analysis:**

The model is tested under different scenarios and with various input values to understand its sensitivity to changes in features. Helps identify potential weaknesses or uncertainties in the model's predictions.

Provides insights into improving the model's robustness and reliability. Assists in making informed decisions based on the model's sensitivity to different factors.

### CHAPTER 9

**OUTPUT SCREENS OF THE PROJECT**

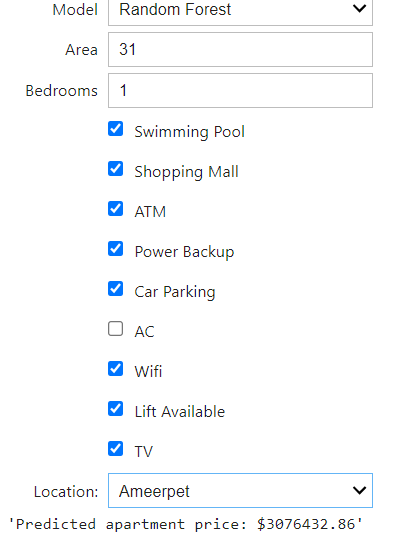
****

Fig 9.1

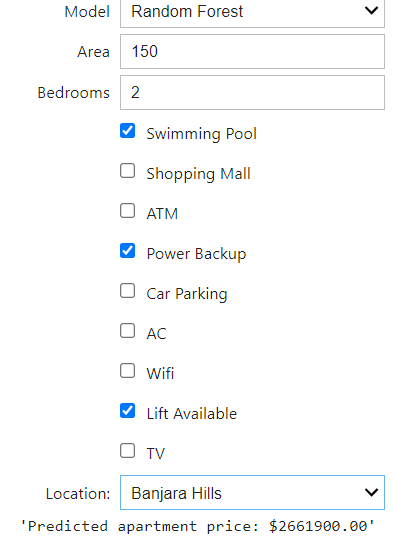


Fig 9.2

# 

# CHAPTER 10

## 

## CONCLUSION

* Every year, house prices rise, necessitating the creation of a mechanism to forecast future house prices. Land owners, estate valuers, and policymakers may use house price prediction to calculate the valuation of a home and the acceptable sale price.

* This will assist potential buyers in determining the right time to purchase a home. While physical conditions, styles, and location are the three main factors that influence a house's price, the individual variables that influence a house's price vary.
* A comparison of the predicted and actual prices sown in Table 1 revealed that the model achieved a prediction difference of ±5.
* This showed that the model can be used to predict house prices. Several other machine learning models especially deep learning models can also be explored for house price prediction
* Given the positive results, further exploration of advanced machine learning techniques, particularly deep learning models, may enhance prediction accuracy even more.
* Implementing such models can provide deeper insights and more precise forecasts, ultimately benefiting the entire real estate market.
* Accurate price prediction not only supports fair and competitive pricing but also aids in strategic planning and investment decisions, reinforcing the value of data-driven approaches in the real estate industry.

# 

# 

# CHAPTER 11

## FUTURE ENHANCEMENT

* Incorporate additional features like school district ratings, proximity to amenities, and neighbourhood crime statistics.
* Explore more complex machine learning algorithms, such as support vector machines (SVMs) or ensemble methods, for improved accuracy.
* Develop a system that provides real-time house price predictions based on up-to-date market data and user-defined preferences
* Incorporating up-to-date market data, including recent sales, listing prices, economic indicators, and demographic trends.

 Allowing users to input preferences such as desired location, property size, amenities, and

budget constraints.

 Utilizing models trained on current data to generate instant predictions based on user inputs and

real-time market conditions.

 Creating a user-friendly interface where users can input criteria and receive predicted house prices,

potentially integrating with maps or interactive dashboards for visualization.

 Factors like distance to shopping centers, parks, public transportation, and healthcare facilities influence convenience and appeal, thereby affecting property values.

 Crime rates and safety perceptions are crucial considerations for buyers, making this data valuable for predicting house prices accurately.

# CHAPTER 12

## REFERENCES

* Malang, C. S., Java, E., & Febrita, R. E. (2017). Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization. International Journal of Advanced Computer Science and Applications, 8(10), 323–326.
* Kang, Y., Zhang, F., Peng, W., Gao, S., Rao, J., Duarte, F., & Ratti, C. (2020). Land Use Policy Understanding house price appreciation using multi-source big geo-data and machine learning. Land Use Policy, July, 104919. <https://doi.org/10.1016/j.landusepol.2020.104919>
* Greenaway- mcgrevy, R., & Sorensen, K. (2021). A Time-Varying Hedonic Approach to quantifying the effects of loss aversion on house prices. Economic Modelling, 99(March), 105491. <https://doi.org/10.1016/j.econmod.2021.03.010>
* Filip F.G., Zamfirescu CB., Ciurea C. (2017) Collaboration and Decision-Making in Context. In: ComputerSupported Collaborative Decision-Making. Automation, Collaboration, & E-Services, vol 4. Springer, Cham. <https://doi.org/10.1007/978-3-319-47221-8_1>
* Kayode Anthonia Aderonke, Akande Noah Oluwatobi, Saheed O Jabaru, Oladele O Tinuke (2020), “An Empirical Investigation of the Prevalence of Osteoarthritis in South West Nigeria: A Population-Based Study”, International Journal of Online and Biomedical Engineering (iJOE), 16(1), 100-114.
* Oluwatobi Noah Akande, Oluwakemi Christiana Abikoye, Aderonke Anthonia Kayode, and Yema Lamari (2020), “Implementation of a Framework for Healthy and Diabetic Retinopathy Retinal Image Recognition”, Scientifica, Volume 2020, Article ID 4972527.