# ASSET PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

A Project Report Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

Submitted by

| | |
|---|---|
| **G. VIKRAM** | **18NR1A0532** |
| **P. VAMSI SAI** | **18NR1A0557** |
| **G. JYOTHSNA** | **18NR1A0534** |
| **D. YOCHANA RANI** | **18NR1A0522** |

Under the esteemed guidance of

**Mr. S Durga Prasad**

Associate Professor & Head of the Department

Department of Computer Science and Engineering

**BABA INSTITUTE OF TECHNOLOGY AND SCIENCES**

**(Approved by A.I.C.T.E and Affiliated to J.N.T.U, Kakinada)**

**P.M. PALEM, MADHURWADA**

**VISAKHAPATNAM-530041**

**(2021-2022)**

BITS VIZAG

BABA INSTITUTE OF TECHNOLOGY AND SCIENCES

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

## CERTIFICATE

This is to certify that the project work entitled **"ASSET PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES"** bonafide work done by **G. VIKRAM (18NR1A0532), P. VAMSI SAI (18NR1A0557), D. YOCHANA RANI (18NR1A0522), G. JYOTHSNA (18NR1A0534)** during the year 2021-2022 in partial fulfillment of the requirements for the award of the degree of **BACHELOR OF TECHNOLOGY** from **BABA INSTITUTE OF TECHNOLOGY AND SCIENCES**, (Affiliated to J.N.T.U. Kakinada), P.M.Palem, Madhurwada, Visakhapatnam.

| | |
|---|---|
| **Project Guide** | **Head of Department** |
| **Mr. S Durga Prasad** | **Mr. S Durga Prasad** |
| Associate Professor | Associate Professor |

**EXTERNAL EXAMINER**

# BITS-VIZAG

# DECLARATION

We hereby declare that this project entitled “**ASSET PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES**” submitted for the partial fulfillment of the requirements for the award of Bachelor of Technology under the guidance of **Mr**. **S DURGA PRASAD**, Associate Professor and Head of the Department, Department of Computer Science and Engineering. This work is not submitted to any university for the award of any degree/diploma.

| | |
|---|---|
| **G. VIKRAM** | **(18NR1A0532)** |
| **P. VAMSI SAI** | **(18NR1A0557)** |
| **G. JYOTHSNA** | **(18NR1A0534)** |
| **D. YOCHANA RANI** | **(18NR1A0522)** |

# ACKNOWLEDGMENT

We would like to express our special thanks and gratitude to the management of our esteemed institute “**BABA INSTITUTE OF TECHNOLOGY AND SCIENCES**”, who gave us the opportunity to study and do this project and also helped us in doing a lot of research and we came to know about so many things.

We also wish to express our deep gratitude to **Prof. N VICTOR BABU, Principal, BITS**. for creating a provision to undergo industry-oriented project training and development.

We would like to express our special thanks to **Mr. S DURGA PRASAD Associate Professor, Head of the Department, Department of Computer Science and Engineering,** and **project guide** who gave us moral support and guided us in different matters regarding the topic. He had been very kind and patient while suggesting to us the outlines of this project and correcting our doubts. We thank him for his overall support.

We would like to thank our parents for giving their support to undergo our graduation. We would also like to thank our friends, teaching, and non-teaching staff of the **Department of Computer Science and Engineering** for all their support at critical junctures throughout our study.

| | |
|---|---|
| **G. VIKRAM** | **(18NR1A0532)** |
| **P. VAMSI SAI** | **(18NR1A0557)** |
| **G. JYOTHSNA** | **(18NR1A0534)** |
| **D. YOCHANA RANI** | **(18NR1A0522)** |

# ABSTRACT

Every middle-class person has a dream to construct a new house within their specified budget. They have to know particular information about land or any house. Customers/Buyers are generally not aware of factors that influence land prices. So, there is a need for a system to predict land prices in the future and to know the projected price of an asset. This system helps in predicting the price of the asset based on the previous data of the locations that are located in Visakhapatnam city.

Asset price prediction determines the selling price of land or a house in a specific timeline and can help the customer who plans to buy an asset so that they can plan their finance well. Predicting land prices with real factors is the main aim of our project along with creating a web application. This system gives the functionality to buyers, allowing them to search for land by their preferences.

# TABLE CONTENT

# 1. INTRODUCTION

Nowadays Land price is drastically increasing as compared to earlier times. Most of the people have rough estimates regarding the land price etc. There is no predetermined method for calculating the cost of the asset. We have to know about the future land pricing and in buying the land if the cost is not well predicted it may incur a loss to the profit. There are many factors that have an impact on house prices, such as the number of bedrooms and bathrooms. Even the nearby location, a location with great accessibility to highways, schools, shopping malls, and local employment opportunities contribute to the rise in house prices. The sellers, do not want to sell their houses at a lower-than-average cost. So, they need to predict the accurate price of the asset. So that they can get or give the best price. Manual predication of land or price becomes more difficult, that's why we proposed a system which is used to forecast the price of land or a house. This project is used to predict prices by analyzing current land prices done using Machine learning Algorithms, thereby forecasting the future prices according to the user requirements based on location, the number of square feet (Sqft), Yards, and some other constraints like Wall Seepage, etc.

## Machine Learning

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people. Although machine learning is a field within computer science, it differs from traditional computational approaches. In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve. Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data to automate decision-making processes based on data inputs.

Any technology user today has benefitted from machine learning. Facial recognition technology allows social media platforms to help users tag and share photos of friends. Optical character recognition (OCR) technology converts images of text into movable types. Recommendation engines, powered by machine learning, suggest what movies or television shows to watch next based on user preferences. Self-driving cars that rely on machine learning to navigate may soon be available to consumers.

Machine learning is a continuously developing field. Because of this, there are some considerations to keep in mind as you work with machine learning methodologies or analyze the impact of machine learning processes. In machine learning, tasks are generally classified into broad categories. These categories are based on how learning is received or how feedback on the learning is given to the system developed. Two of the most widely adopted machine learning methods are supervised learning which trains algorithms based on example input and output data that is labeled by humans, and unsupervised learning which provides the algorithm with no labeled data to allow it to find structure within its input data.

### Supervised Learning

In supervised learning, the computer is provided with example inputs that are labeled with their desired outputs. The purpose of this method is for the algorithm to be able to "learn" by comparing its actual output with the "taught" outputs to find errors and modify the model accordingly. Supervised learning, therefore, uses patterns to predict label values on additional unlabeled data.

### Unsupervised Learning

In unsupervised learning, data is unlabeled, so the learning algorithm is left to find commonalities among its input data. As unlabeled data are more abundant than labeled data, machine learning methods that facilitate unsupervised learning are particularly valuable. The goal of unsupervised learning may be as straightforward as discovering hidden patterns within a dataset, but it may also have a goal of feature learning, which allows the computational machine to automatically discover the representations that are needed to classify raw data. Without being told a "correct" answer, unsupervised learning methods can look at complex data that is more expansive and seemingly unrelated to organize it in potentially meaningful ways. Unsupervised learning is often used for anomaly detection including for fraudulent credit card purchases, and recommender systems that recommend what products to buy next. In unsupervised learning, untagged photos of dogs can be used as input data for the algorithm to find likenesses and classify dog photos together.

### Semi-supervised learning

Semi-supervised machine learning is a combination of supervised and unsupervised machine learning methods. With more common supervised machine learning methods, you train a machine learning algorithm on a “labeled” dataset in which each record includes the outcome information. This allows the algorithm to deduce patterns and identify relationships between your target variable and the rest of the dataset based on its information. In contrast, unsupervised machine learning algorithms learn from a dataset without the outcome variable. In semi-supervised learning, an algorithm learns from a dataset that includes both labeled and unlabeled data, usually mostly unlabeled.

**Reinforcement Learning**

Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions. For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or a penalty. In Reinforcement Learning, the agent learns automatically using feedback without any labeled data, Since there is no labeled data, the agent is bound to learn by its experience only. Reinforcement learning is a type of machine learning method where an intelligent agent (computer program) interacts with the environment and learns to act within that.

## 1.1Feasibility Study

A preliminary investigation examines project feasibility; the likelihood that the system will be useful to the organization. The main objective of the feasibility study is to test the Technical, Operational, and Economical feasibility of adding new modules and debugging old running systems. All system is feasible if they are unlimited resources and infinite time. There are aspects in the feasibility study portion of the preliminary investigation.

A feasibility study assesses the economic, technical, and operational merits of the proposed project. A project is economically feasible if costs do not overshadow benefits. A project is technically feasible if the technology is available and capable of meeting users’ requests. A project is operationally feasible if the proposed system will be used once it is installed.

- Economical Feasibility
- Technical Feasibility
- Operational Feasibility

### 1.1.1 Economical Feasibility

As a system can be developed technically and will be used, if installed must be a good investment for the organization. In the economic feasibility, the development cost in creating the system is evaluated against the ultimate benefit derived from the new systems. Financial benefits must equal or exceed the costs. The system proposed is economically feasible. It does not require any additional hardware or software. Since the interface for this system is developed using the existing resources and technologies python which is open-source, there is nominal expenditure and economic feasibility for certain.

The system working is quite easy to use and learn due to its simple and attractive interface. In the economical feasibility, the development cost in creating the system is evaluated against the ultimate benefit derived from the new systems. Financial benefits must equal or exceed the costs. Hence the system is economically feasible.

### 1.1.2 Technical Feasibility

This assessment focuses on the technical resources available to the organization. It helps organizations determine whether the technical resources meet capacity and whether the technical team is capable of converting the ideas into working systems. Technical feasibility also involves the evaluation of the hardware, software, and other technical requirements of the proposed system. This assessment is based on an outline design of system requirements, to determine whether the company has the technical expertise to handle and complete the project. When writing a feasibility report, the following things should be taken into consideration:

- The part of the business being examined
- The human and economic factor
- The possible solutions to the problem

The technical feasibility assessment is focused on gaining an understanding of the present technical resources of the organization and their applicability to the expected needs of the proposed system. It is an evaluation of the hardware and software and how it meets the need of the proposed system. The currently proposed system is technically feasible, each of the tools and technologies that are used is open source, and the technical skills required are manageable.

Our project was planned to develop in HTML, with CSS Technology as the front-end, and Python as the back-end, we deploy the project in the local host. We thought of developing this project using the Windows platform. As, we team members, have knowledge of the above technologies, hence it is feasible to do this project.

### 1.1.3 Operational Feasibility

Proposed projects are beneficial only if they can be turned out into an information system. That will meet the organization’s operating requirements. Operational feasibility aspects of the project are to be taken as an important part of the project implementation. Some of the important issues raised are to test the operational feasibility of a project includes the following:

• Is there sufficient support for the management from the users?

• Will the system be used and work properly if it is being developed and implemented?

This system is targeted to be by the above-mentioned issues. Beforehand, the management issues and user requirements have been taken into consideration. The well-planned design would ensure the optimal utilization of the computer resources and would help in the improvement of performance status. The aspect of the study is to check the level of acceptance of the system by the user. This includes the process of training the user in the system efficiently. The level of acceptance by the users solely depends on the methods that 5 are employed to educate the user about the system and to make him familiar with it as he is the final user of the system.

This project is more reliable, maintainable, and affordable. This system would be used and works properly since it was implemented. User requirements have been taken into consideration all the specifications of the project could be realized and the specifications mentioned in the above points can be implemented. Hence the project is operationally feasible.

## 1.2 Existing System

Nowadays, Manual prediction of the price of a house or land is very hard. There are a lot of parameters that affect its price. Customers need to hire an agent to know the price of land or a house which again increases the cost of the process.

### Drawbacks:

- It doesn’t predict the future prices of a house or land at a specific location mentioned by the customer.

- Manual price prediction of an asset can lead to loss to the customers.

## 1.3 Proposed System

In our system, we provide an interface for customers to complete their needs with ease. The customers can easily know the price of land or a house at the preferred location with our system. Prediction is done based on the past three year’s data of land or a housing dataset. we use the Linear Regression algorithm to forecast the price of land or a house.

### Advantages:

- To eliminate the need of real estate agents to gain information regarding asset prices in the future.

# 2. SOFTWARE REQUIREMENTS SPECIFICATION

## 2.1 Introduction

The SRS is a specification for a specific software product, program, or set of applications that perform particular functions in a specific environment. It serves several goals depending on who is writing it. First, the SRS could be written by the client of a system. Second, the SRS could be written by a developer of the system. The two methods create entirely various situations and establish different purposes for the document altogether. The first case, SRS, is used to define the needs and expectations of the users. The second case, SRS, is written for various purposes and serves as a contract document between customer and developer. The software requirements specification document lists sufficient requirements for the project development. To derive the requirements, the developer needs to have a clear and thorough understanding of the products under development. This is achieved through detailed and continuous communications with the project team and customer throughout the software development process. The SRS may be one of a contract's deliverable data item descriptions or have other forms of organizationally-mandated content

### 2.1.1 Purpose

The purpose of this SRS is to specify the requirements to create a web application for the prediction of land or house price. This SRS document provides a detailed overview of our software product, its parameters, and goals.

### 2.1.2 Scope

The scope of this project is to enable the buyers to view the predicted price of land or a house. The motive for developing this web application is to design a feature that can predict the price of land or a house. We are comparing the scope of our study to the areas such as Pedawaltair, Chinna Waltair, Shivajipalem, Allipuram, Kancharapalem, Resapuvanipalem, Maddilapalem, Madhurawada.

### 2.1.3 Overview

For the rest of the document, a brief description of the application is given. Then, we have given the software requirements, followed by interface requirements and performance requirements. In the last section, we have provided the non-Functional attributes.

### 2.1.4 Document conventions

We used the “Times new roman” font for all the documents, in this bold letters are used for headings and normal letters are used for content.

### 2.1.5 Intended Audience and Reading Suggestions

The document is intended to be read by project managers, developers, testers, legal advisers, analysts, and any other related users. The document is organized into 5 parts.

- Introduction
- Overall Description
- External interface requirements
- System Features
- Other Non- functional requirements

### 2.1.6 Product Scope

The Scope of this product is to forecast the price of land or a house by machine learning techniques and to get the accurate price of land or a house, which reduce problems for the customers/buyers while buying a land or a house within their budget.

### 2.1.7 References

- Software Engineering Textbook by Pressman
- Wikipedia

### 2.1.8 Overall description

This section of the SRS describes the general requirements that drive the design of the software system. The goal is not to state-specific requirements, but rather to provide context to make those requirements easier to understand.

### 2.1.9 General Description

#### Product Perspective

The perspective of the product is to take the necessary inputs from the user and sends the data to the system. The system will start predicting the price and display it to the user.

#### Product Functions

The system must provide, at a minimum, the following functions by the other requirements described within the SRS document.

- Collects all details of land or house and stores them in the dataset.
- Providing a web application for users to view the predicted price.
- Provide a User Interface (UI) for a user to take all inputs about land or a house.
- List all the data and sends it to the system to start prediction.

#### User class and Characteristics

Users of the software system include researchers, practicing engineers, and software developers. All users should be familiar with the web application and familiar with real estate. Users are knowledgeable of both. However, users should have a good understanding of the tasks, activities, and artifacts of both processes.

#### Operating Environment

The system will operate in a Windows operating system with minimum specifications. The clients will operate the application within common web browser environments. The minimum set of browsers that must be supported are:

- Google Chrome 44+
- Microsoft Internet Explorer 10+

- Mozilla Firefox 40+
- Opera Mini 40+

### Design and Implementation Constraints

Sometimes the result will take more time to display the predicted price due to running more applications in the background. The graphical user interface of the web application won't display to the user if the network connection is slow.

### User Documentation

Any user software system target audience for user documentation generated about the software system.

### Assumptions and Dependencies

No specific assumptions or dependencies are considered at this time.

## 2.2 External Interface Requirements

This section of the SRS describes the interface requirements for the system. Requirements for the user, hardware, software, and communication interfaces are defined.

### 2.2.1 User Interfaces

This tells about user interfaces how they will work and how they will be displayed like that.

**Frontend:** HTML, CSS

**Backend:** JavaScript, Python,Django,Sqlite3

### 2.2.2 Hardware Interfaces

A Windows PC with mediocre specifications.

### 2.2.3 Software Interfaces

The software interface is web browsers in any operating system. The browser which supports all tools and APIs can use our web application. The following web browsers can be supported:

- Google Chrome
- Mozilla Firefox
- Internet Explorer
- Microsoft edge.

### 2.2.4 Networking Communication Interface:

- Client on Internet will be using HTTP/HTTPS Protocol.
- Client on the intranet will be using TCP/IP protocol.

## 2.3. System Features

### Modules:

- Admin:
    - Add/delete location: Admin can add locations.
    - Add Density: Admin can add density of land or a house per unit area.
    - View Location: Admin can view the added locations.
    - Admin can only maintain a dataset of a house or a land.
- User:
    - Search/Select Location: The user can select a specific location.
    - Select Type: The user can select house type.
    - Enter Density: The user can specify the density of the land.
    - View Predicted Price: The user can view the predicted price of land or a house.

## 2.3.1 Functional Requirements

### 2.3.1.1 The System shall be able to display the homepage.

**Description**

After being properly coded, the system shall be able to display a GUI of the Price Prediction System Website.

**Technical issues**

Pre-condition: website shall be properly coded.

Post-condition: the system shall properly display the Price Prediction System homepage.

**Risks**

The system may not properly display the homepage.

### 2.3.1.2 The System shall be able to display all the text fields.

**Description**

After displaying the homepage, the system shall be able to display all the text fields to take input from the user.

**Technical issues**

Pre-condition: website shall be properly coded Text fields must be displayed.

Post-condition: the system shall properly display all the fields.

**Risks**

The system may not properly display the homepage.

### 2.3.1.3 Admin can add/delete the location

**Description**

Here, the admin can update the location and change the parameters of the location.

**Response**

After accessing the dataset, the admin can add the location of land or a house.

**Functional Requirements**

After the admin changes the data, it will reflect in our system after the user refresh the Page.

### 2.3.1.4 User enters into the website

**Description**

Here, the user enters the website.

**Response**

After entering the website, a graphical user interface is provided to the user.

**Functional Requirements**

The User can choose the category type, after clicking the button. It will redirect to that page.

### 2.3.1.5 User Can View the Predicted Price

**Description**

Here, the user can view the predicted price.

**Response**

After entering all the required text fields, the User can view the predicted price based on his preferences.

**Functional Requirements**

The User can view the predicted price after clicking the prediction button. The predicted price will be displayed on the screen.

### 2.3.1.6 User Can Search/Select for A Location

**Description**

After entering the website, the user can select the required data.

**Response**

After entering the website, a flexible interface should be available to the user.

**Functional Requirements**

Users can select the data by entering the name of the location in the text field and selecting the preferred location.

## 2.4 Non-Functional Requirements

### Hardware Requirements

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware, a hardware requirements list is often accompanied by a hardware capability list (HCL), especially in the case of operating systems. An HCL lists tested compatible, and sometimes incompatible hardware devices for a particular operating system or application. The following subsections discuss the various aspects of hardware requirements.

### Hardware requirements for this project

| | | |
|---|---|---|
| RAM | : | 4GB and Higher |
| Processor | : | Intel i3 and above |
| Hard Disk | : | 500GB (Minimum) |
| Peripheral Device | : | Monitor, Mouse, and Keyboard |

### Software requirements:

Software requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the software installation package and need to be installed separately before the software works.

### Software requirements for this project

| | |
|---|---|
| Operating system | : Windows 7 or above |
| IDLE | : Visual studio/ Jupiter Notebook |
| Server-side Script | : Python 3.6+,Django,sqlite3 |
| Libraries Used | : Pandas, NumPy, Matplotlib, Scikit-Learn |
| | MS Excel/ Google Spreadsheet (for creating datasets) |

### 2.4.1 Performance Requirements

As it has already been mentioned, this module is not a resource hog and will run on almost every computer. Its functions and features are not personally intensive. It does not require a powerful processor or graphics card, much RAM, or disk space, for both the install and live version, and the program loads and runs fairly quickly. The only real requirement is the internet with associated browsers like Google Chrome, Mozilla Firefox, etc.

### 2.4.2 Security Requirements

The system's security shall be ensured by the admin only, the administrator shall be able to log in and modify/add the information of the various locations. General users shall not have access to modify content within the system however, under certain circumstances, users may give feedback to the system.

### 2.4.3 Maintainability Requirements

The system shall be easily maintainable by the administrator through the user interface. Also, other programmers shall be capable of easily modifying and updating code by using the documentation provided with the system.

### 2.4.4 Safety Requirements

As it stated with many warnings in the program, data loss is a serious probability, if you are not careful, backup your data and/or use the program with extreme caution. It is obvious that when you deal with partition changes, you could lose all your data by human or application error. This comes with absolutely no warranty and cannot be made responsible for any loss of data.

### 2.4.5 Software Quality Attributes

The application provides a quite friendly user interface with its operations accessible from the menu bar and the main toolbar. An average or casual user should not find any problem using the program to perform at least its main functions. Interoperability is guaranteed since this program runs on both Mac and PC (Linux, Windows, or other operating systems)

# 3. SYSTEM ANALYSIS

## Introduction:

Unified Modelling Language (UML):

The UML stands for Unified modelling language and is a standardized general-purpose visual modelling language in the field of Software Engineering. It is used for specifying, visualizing, constructing, and documenting the primary artifacts of the software system. It helps in designing and characterizing, especially those software systems that incorporate the concept of Object orientation. It describes the working of both the software and hardware systems.

The UML is a language for

- Visualizing
- Specifying
- Constructing
- Documenting

### Applications

The UML is intended primarily for the software-intensive system.

It has been used effectively for such domains as

- Enterprise information system
- Banking and financial services
- Telecommunications
- Defence / aerospace
- Retail
- Medical electronics
- Scientific
- Distributed Web-based services

To understand the UML, you need to form a conceptual model of the language, and this requires learning three major elements: the UML’s basic building blocks.

## Building Blocks of the UML

The vocabulary of the UML encompasses three kinds of building blocks:

- Things
- Relationships
- Diagrams

Things are the abstractions that are first-class citizens in a model; relationships tie these things together; diagrams group interesting collections of things.

## Things in the UML

There are four kinds of things in the UML

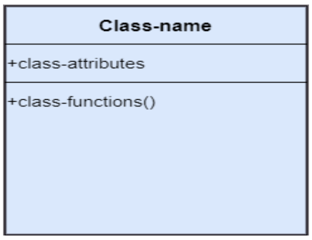
- Structural things
- Behavioural things
- Grouping things
- Annotational things

## Structural Things

Nouns that depict the static behaviour of a model are termed as structural things. They display the physical and conceptual components. They include class, object, interface, node, collaboration, component, and a use case.

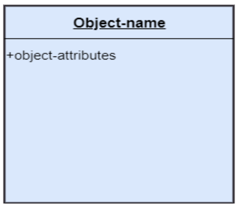
**Class:**

A Class is a set of identical things that outlines the functionality and properties of an object. It also represents the abstract class whose functionalities are not defined. Its notation is as follows;



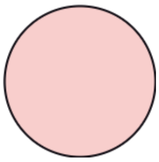
**Object:**

An individual that describes the behaviour and the functions of a system. The notation of the object is similar to that of the class; the only difference is that the object name is always underlined and its notation is given below;



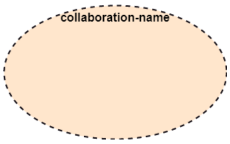
**Interface:**

A set of operations that describes the functionality of a class, which is implemented whenever an interface is implemented.



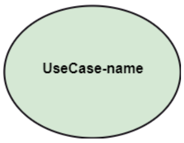
**Collaboration:**

It represents the interaction between things that are done to meet the goal. It is symbolized as a dotted ellipse with its name written inside it.



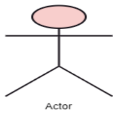
**Use case:**

Use case is the core concept of object-oriented modelling. It portrays a set of actions executed by a system to achieve the goal.



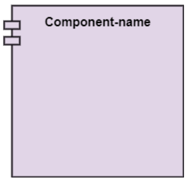
**Actor**:

It comes under the use case diagrams. It is an object that interacts with the system, for example, a user.



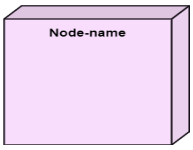
**Component:**

It represents the physical part of the system.



**Node:**

A physical element that exists at run time.

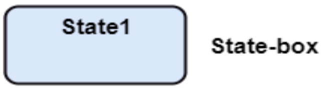


## Behavioural Things

They are the verbs that encompass the dynamic parts of a model. It depicts the behaviour of a system. They involve state machine, activity diagram, interaction diagram, grouping things, annotation things

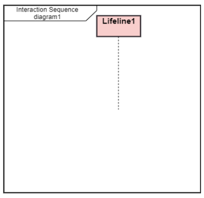
**State Machine:**

It defines a sequence of states that an entity goes through in the software development lifecycle. It keeps a record of several distinct states of a system component.



**Interaction:**

It is used to envision the flow of messages between several components in a system.

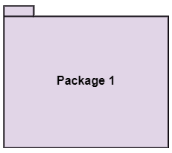


## Grouping Things

It is a method that together binds the elements of the UML model. In UML, the package is the only thing, which is used for grouping.

**Package:**

A package is the only thing that is available for grouping behavioural and structural things.

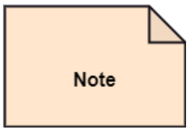


## Annotation Things

It is a mechanism that captures the remarks, descriptions, and comments of UML model elements. In UML, a note is the only Annotational thing.

**Note:**

It is used to attach the constraints, comments, and rules to the elements of the model. It is a kind of yellow sticky note.

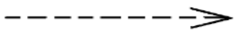


## Relationships

It illustrates the meaningful connections between things. It shows the association between the entities and defines the functionality of an application. There are four types of relationships given below:

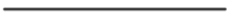
**Dependency:**

Dependency is a kind of relationship in which a change in the target element affects the source element, or simply we can say the source element is dependent on the target element. It is one of the most important notations in UML. It depicts the dependency from one entity to another. It is denoted by a dotted line followed by an arrow on one side as shown below,



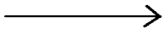
**Association:**

A set of links that associates the entities to the UML model. It tells how many elements are actually taking part in forming that relationship. It is denoted by a dotted line with arrowheads on both sides to describe the relationship with the element on both sides.



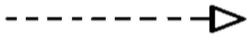
**Generalization:**

It portrays the relationship between a general thing (a parent class or superclass) and a specific kind of that thing (a child class or subclass). It is used to describe the concept of inheritance. It is denoted by a straight line followed by an empty arrowhead at one side.



**Realization:**

It is a semantic kind of relationship between two things, where one defines the behaviour to be carried out, and the other one implements the mentioned behaviour. It exists in interfaces. It is denoted by a dotted line with an empty arrowhead at one side.



**Diagrams**

The diagrams are the graphical implementation of the models that incorporate symbols and text. Each symbol has a different meaning in the context of the UML diagram. There are thirteen different types of UML diagrams that are available in UML 2.0, such that each diagram has its own set of a symbol. And each diagram manifests a different dimension, perspective, and view of the system.

UML diagrams are classified into three categories that are given below:

- Structural Diagram
- Behavioural Diagram
- Interaction Diagram

**Structural Diagram:**

It represents the static view of a system by portraying the structure of a system. It shows several objects residing in the system. Following are the structural diagrams given below:

- Class diagram
- Object diagram
- Component diagram
- Deployment diagram

**Behavioural Diagram:**

It depicts the behavioural features of a system. It deals with dynamic parts of the system. It encompasses the following diagrams:

- Activity diagram
- State machine diagram
- Use Case diagram

**Interaction diagram:**

It is a subset of behavioural diagrams. It depicts the interaction between two objects and the data flow between them. Following are the several interaction diagrams in UML:

- Sequence diagram
- Collaboration diagram

## 3.1 Use Case Diagram:

A use case diagram in the Unified Modelling Language (UML) is a type of behavioural diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. Use Case diagrams are formally included in two modelling languages defined by the Unified Modelling Language (UML).

This Use Case Diagram is a graphic depiction of the interactions among the elements of Asset Price Prediction System. It represents the methodology used in system analysis to identify, clarify, and organize system requirements.

The main actors of this use case diagram are:- User and Admin, who perform the different types of use cases are select the asset features, view predicted price, add/delete locations, update features of the asset.

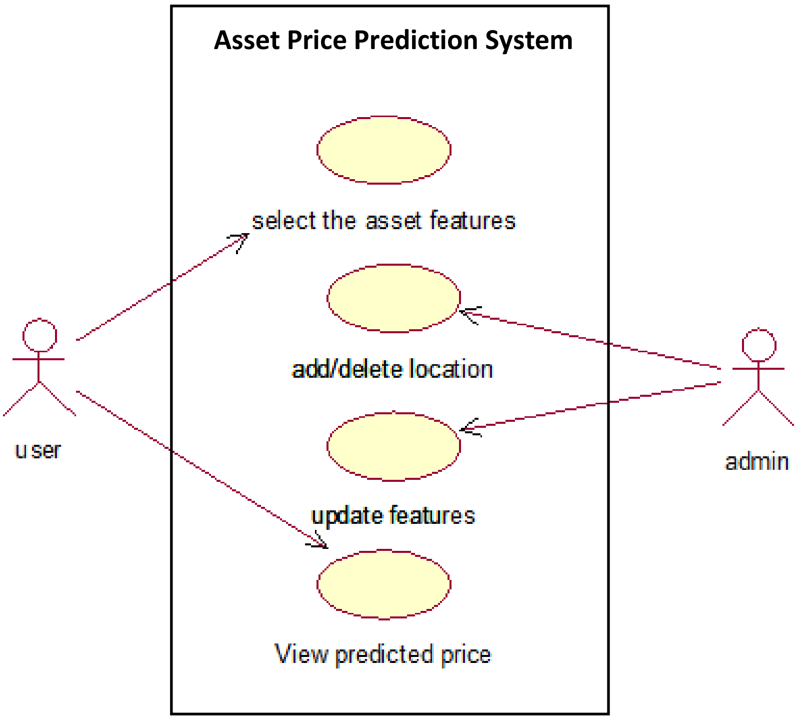


**Fig. 3.1 Use Case Diagram for Asset Price Prediction System**

## 3.2 Sequence Diagram:

A sequence diagram in Unified Modelling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

Initially, the user will enter into the website, a GUI is provided to the user to choose the asset type. After selecting the asset type, the web application redirects to the selected type page. After redirecting to the page, the user needs to enter all the Asset features fields on the page. After entering the required fields, the user needs to click the prediction button to start the prediction. The web application sends data to the model and starts predicting the value based on the user request. After prediction, the predicted value is displayed to the user. Similarly, the admin can add/delete the location and update the values or features.

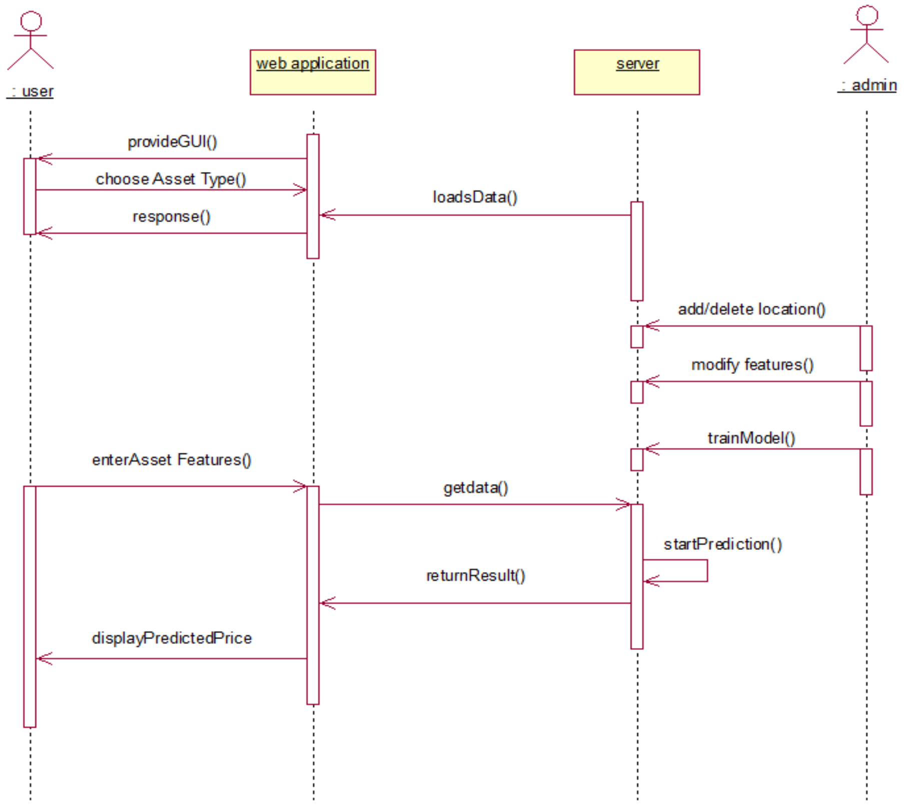


**Fig. 3.2 Sequence Diagram for Asset Price Prediction System**

## 3.3 Collaboration Diagram:

A Collaboration diagram is very similar to a Sequence diagram in the purpose it achieves; in other words, it shows the dynamic interaction of the objects in a system. A distinguishing feature of a Collaboration diagram is that it shows the objects and their association with other objects in the system apart from how they interact with each other. The association between objects is not represented in a Sequence diagram. A Collaboration diagram is easily represented by modelling objects in a system and representing the associations between the objects as links. The interaction between the objects is denoted by arrows. To identify the sequence of invocation of these objects, a number is placed next to each of these.

Initially, the user will enter into the website, a GUI is provided to the user to choose the asset type. After selecting the asset type, the web application redirects to the selected page. After redirecting to the page, the user needs to enter all the displayed fields on the page such as select location, choose house type, choose a floor, enter dimensions, and choose dimensions type. After entering the required fields, the user needs to click the prediction button to start the prediction. The web application sends data to the model and starts predicting the value based on the user request. After prediction, the predicted value is displayed to the user. Similarly, the admin can add/delete the location in the dataset. Whenever the admin adds a location, he needs to train the model. The admin can also modify the plot values and its price. The admin can only maintain the dataset.

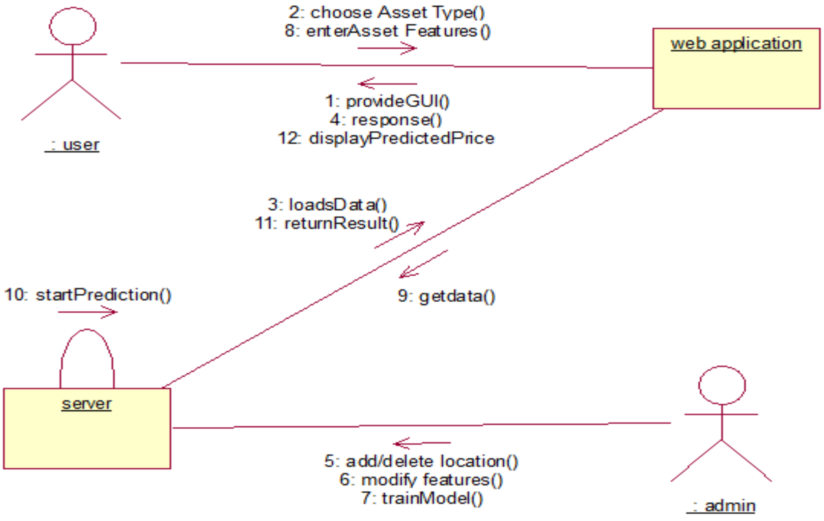


**Fig. 3.3 Collaboration Diagram for Asset Price Prediction System**

# 4. SYSTEM DESIGN

## 4.1 Architecture Diagram

An architectural diagram is a visual representation that maps out the physical implementation of components of a software system. It shows the general structure of the software system and the associations, limitations, and boundaries between each element. An architecture diagram is a visual representation of all the elements that make up part, or all, of a system. Above all, it helps the engineers, designers, stakeholders and anyone else involved in the project understand a system or app's layout. There are many kinds of architecture diagrams, like a software architecture diagram, system architecture diagram, application architecture diagram, security architecture diagram, etc.

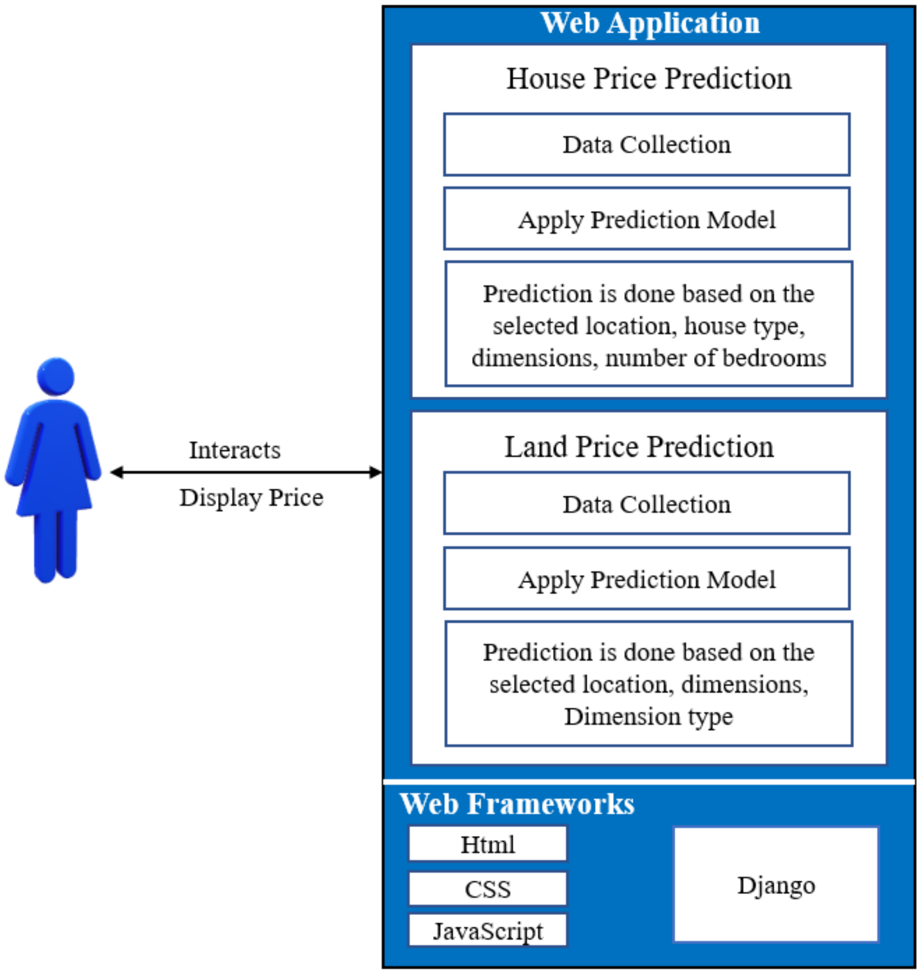


**Fig. 4.1 Architecture Diagram for Asset Price Prediction System**

## 4.2 Project Work Flow Diagram

### Project Flow

A Process Flow Diagram (PFD) is a type of flowchart that illustrates the relationships between major components at an industrial plant. It's most often used in chemical engineering and process engineering, though its concepts are sometimes applied to other processes as well. It’s used to document a process, improve a process or model a new one. Depending on its use and content, it may also be called a Process Flow Chart, Flowsheet, Block Flow Diagram, Schematic Flow Diagram, Macro Flowchart, Top-down Flowchart, Piping and Instrument Diagram, System Flow Diagram or System Diagram. They use a series of symbols and notations to depict a process. The symbols vary in different places, and the diagrams may range from simple, hand-drawn scrawls or sticky notes to professional-looking diagrams with expandable detail, produced with software.

A Process Flow Diagram has multiple purposes:

- To document a process for better understanding, quality control and training of employees.
- To standardize a process for optimal efficiency and repeatability.
- To study a process for efficiency and improvement. It helps to show unnecessary steps, bottlenecks and other inefficiencies.
- To model a better process or create a brand-new process.
- To communicate and collaborate with diagrams that speak to various roles in the organization or outside of it.

Our project workflow consists of the below steps:

**Step-1:** Collect the Asset raw data file in the form of a CSV file or Excel File.

**Step-2:** Understand the data for pre-processing and remove the noise data.

**Step-3:** 100% of the data set is divided into 2 half 80% and 20%.80% are used for the train set and 20% is used for the test set.

**Step-4:** Apply the Linear Regression algorithm to the training dataset and train the model.

**Step-5**: Predict the result and the test set is applied to the trained model.

**Step-6**: deploy the model into the web application and displays the predicted value to the user

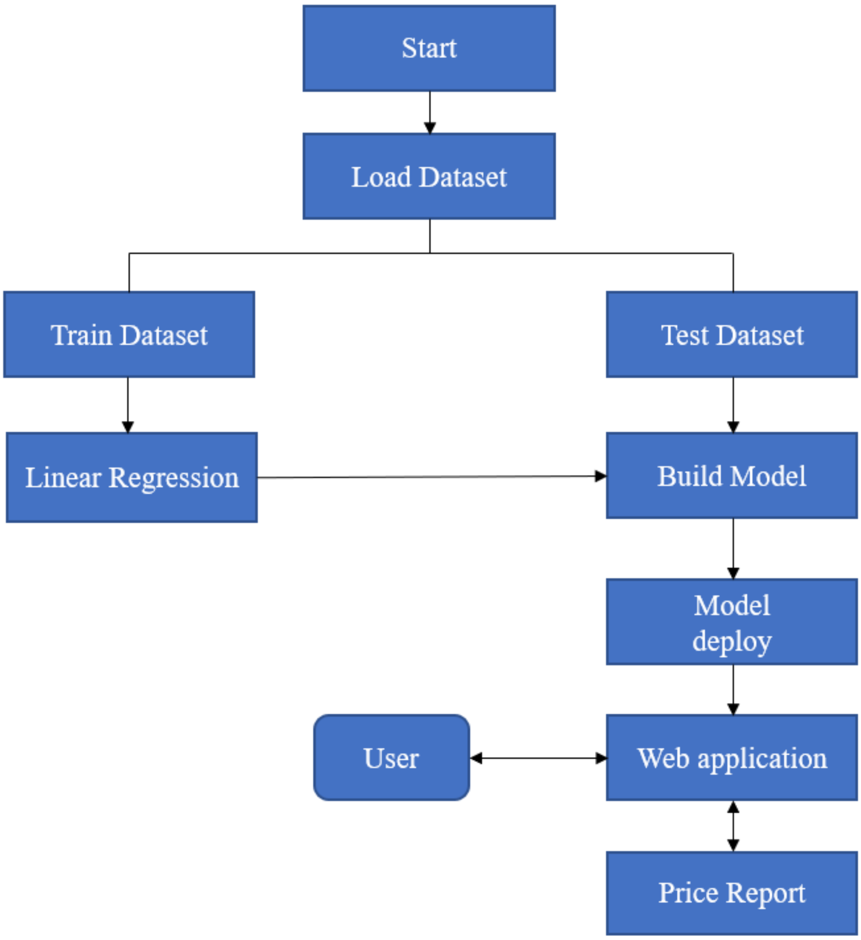


**Fig. 4.2 Project Flow Diagram for Asset Price Prediction System**

## 4.3 Data Flow Diagram

The DFD is also known as a bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data generated by the system. It maps out the flow of information for any process or system, and how data is processed in terms of inputs and outputs. It uses defined symbols like rectangles, circles and arrows to show data inputs, outputs, storage points and the routes between each destination. They can be used to analyse an existing system or model a new one. A DFD can often visually "say" things that would be hard to explain in words and they work for both technical and non-technical.

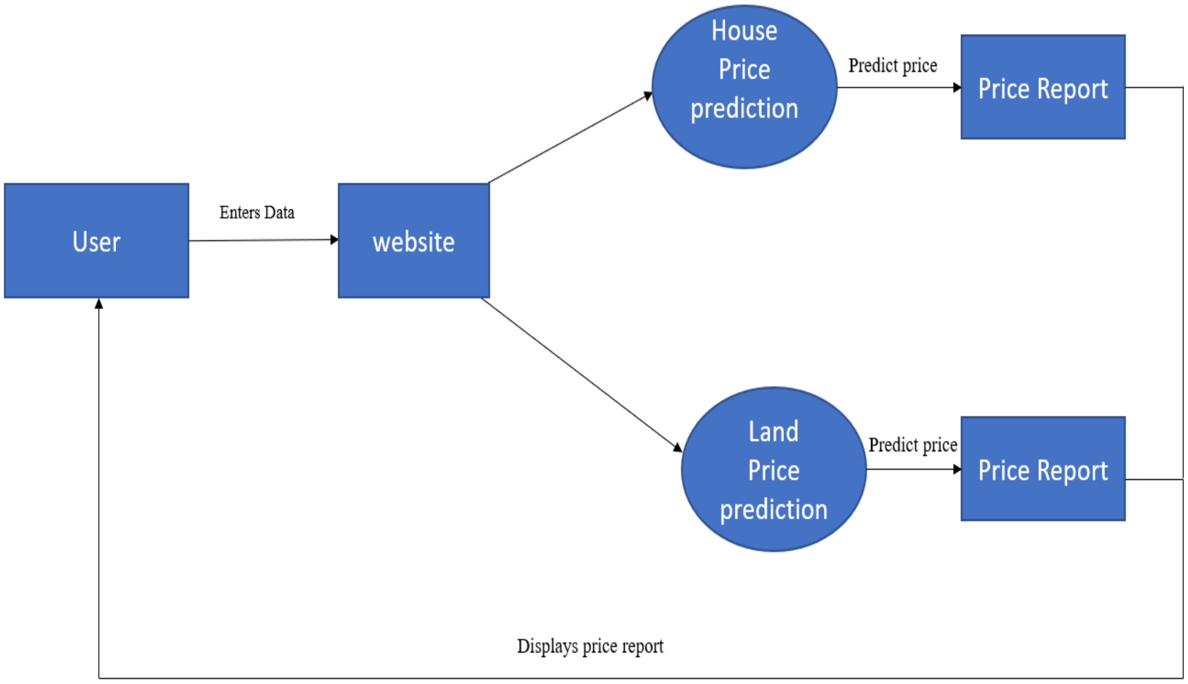


**Fig. 4.3 Data Flow Diagram for Asset Price Prediction System**

The Above diagram gives a quick overview of the data flow diagram. Initially, the user enters the website and enters all the detail regarding the house or land. The web application sends all the information to the prediction model. The prediction model predicts the value and generates a price report for the future 3 years and displays the price report to the user.

## 4.4 Class Diagram:

In software engineering, a class diagram in the Unified Modelling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, and the relationships between the classes. The class diagram is the main building block in object-oriented modelling. They are being used both for general conceptual modelling of the systematics of the application and for detailed modelling translating the models into programming code. The classes in a class diagram represent both the main objects or interactions in the application and the objects to be programmed.

In this class diagram, there are different classes which are user, web application, admin, house prediction, and land prediction. Each class contain different operations and attributes.

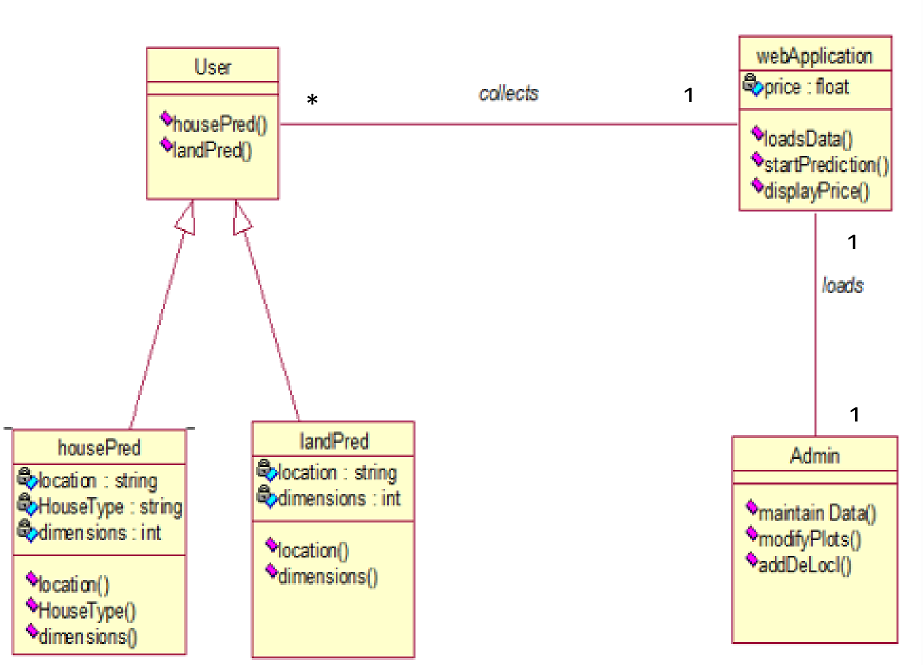


Fig. 4.4 Class Diagram for Asset Price Prediction System

## 4.5 Deployment Diagram

A deployment diagram is a UML Diagram type that shows the execution architecture of a system, including nodes such as hardware or software execution environments, and the middleware connecting them. Deployment diagrams are typically used to visualize the physical hardware and software of a system. Using it you can understand how the system will be physically deployed on the hardware. Deployment diagrams help model the hardware topology of a system compared to other UML Diagram types which mostly outline the logical components of a system.

This web application is used for the prediction of land or a house price. The artifact contains the main.py file, which contains the prediction model. Whenever the client wants to predict the price, the web application automatically runs the main.py and displays the result from the prediction model to the user. Any client can access our web application in their system through a webserver.

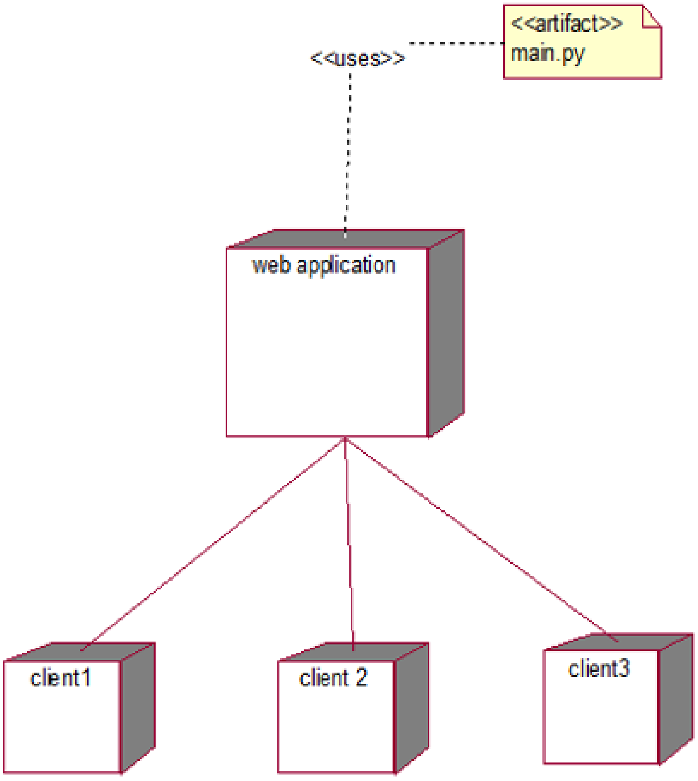


**Fig. 4.5 Deployment Diagram for Asset Price Prediction System**

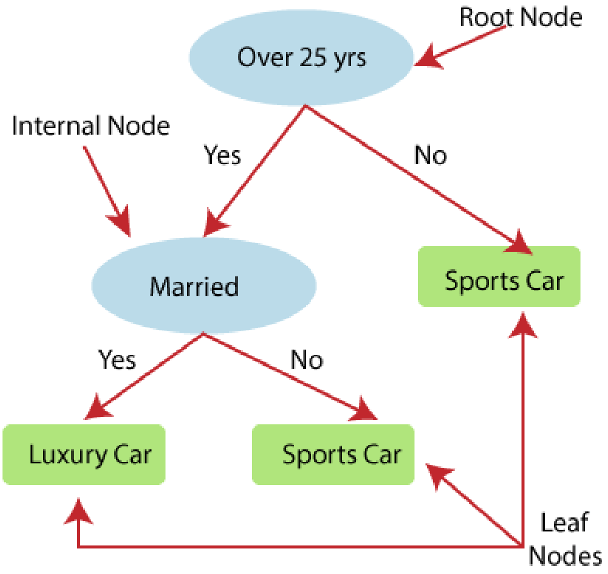
# 5. IMPLEMENTATION

## Decision Tree:

Decision Tree is a supervised learning algorithm that can be used for solving both classification and regression problems. It can solve problems for both categorical and numerical data.

Decision Tree regression builds a tree-like structure in which each internal node represents the "test" for an attribute, each branch represents the result of the test, and each leaf node represents the final decision or result.

A decision tree is constructed starting from the root node/parent node (dataset), which splits into left and right child nodes (subsets of the dataset). These child nodes are further divided into their child's nodes, and themselves become the parent node of those nodes. Consider the below image:



The above image shows the example of Decision Tee regression, here, the model is trying to predict the choice of a person between Sports cars or Luxury cars.

While making a decision tree, calculate the following values.

### Information Gain

Information gain is used to decide which feature to split on at each step in building the tree. Simplicity is best, so we want to keep our tree small. To do so, at each step we should choose the split that results in the purest daughter nodes. A commonly used measure of purity is called information gain. For each node of the tree, the information value measures how much information a feature gives us about the class. The split with the highest information gain will be taken as the first split and the process will continue until all children nodes are pure, or until the information gain is 0.

### Gini Impurity

Gini Impurity is a measurement of the likelihood of incorrect classification of a new instance of a random variable if that new instance were randomly classified according to the distribution of class labels from the data set.

### Pure

Pure means, that in a selected sample of dataset all data belongs to the same class (PURE).

### Impure

Impure means, data is a mixture of different classes.

If our dataset is Pure then the likelihood of incorrect classification is 0. If our sample is a mixture of different classes then the likelihood of incorrect classification will be high.

### Steps for making a decision tree

- Get the list of rows (dataset) that are taken into consideration for making a decision tree (recursively at each node).
- Calculate uncertainty of our dataset or Gini impurity or how much our data is mixed up etc.
- Generate a list of all question that needs to be asked at that node.
- Partition rows into True rows and false rows based on each question asked.
- Calculate information gain based on Gini impurity and partition of data from the previous step.
- Update the highest information gain based on each question asked.
- Update best question based on information gain (higher information gain).
- Divide the node by the best question. Repeat from step 1 until we get pure nodes (leaf nodes).

### Advantages of Decision Tree

- Easy to use and understand.
- Can handle both categorical and numerical data.
- Resistant to outliers, hence requiring little data preprocessing.

### Disadvantages of Decision Tree

- Prone to overfitting.
- Require some kind of measurement as to how well they are doing.
- Need to be careful with parameter tuning.
- Can create biased learned trees if some classes dominate.

## Linear Regression:

Linear regression is a statistical regression method that is used for predictive analysis. It is one of the very simple and easy algorithms which works on regression and shows the relationship between the continuous variables. It is used for solving the regression problem in machine learning.

Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), hence called linear regression. If there is only one input variable (x), then such linear regression is called simple linear regression. And if there is more than one input variable, then such linear regression is called multiple linear regression.

Below is the mathematical equation for Linear regression:

$$Y= aX+b$$

Here, Y = dependent variables (target variables),

X= Independent variables (predictor variables),

a and b are the linear coefficients

**Applications of linear regression are:**

- Analyzing trends and sales estimates
- Salary forecasting
- Real estate prediction
- Arriving at ETAs in traffic.

**Advantages of Linear Regression:**

- Linear Regression is simple to implement and easier to interpret the output coefficients.
- When you know the relationship between the independent and dependent variable has a linear relationship, this algorithm is the best to use because it’s less complexity to compared to other algorithms.

**Disadvantages of Linear Regression:**

- outliers can have huge effects on the regression

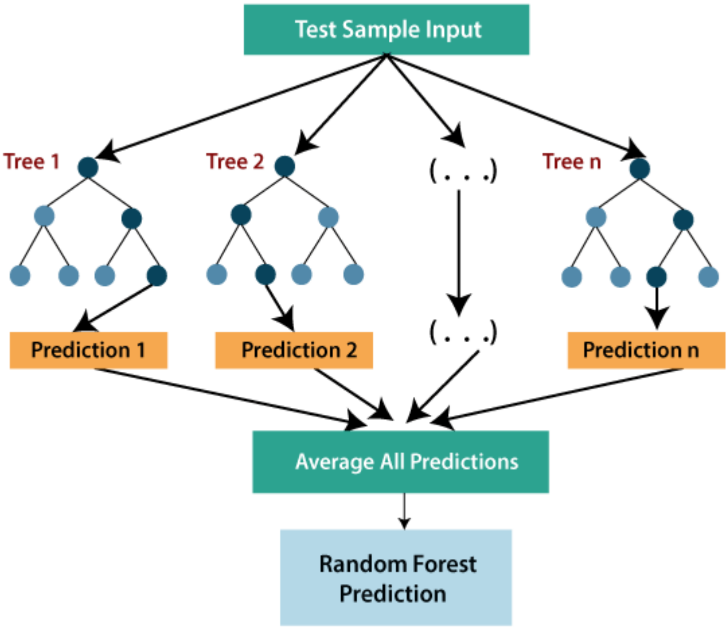
## Random Forest

Random Forest is one of the most powerful supervised learning algorithms which is capable of performing regression as well as classification tasks. The Random Forest regression is an ensemble learning method that combines multiple decision trees and predicts the final output based on the average of each tree output. The combined decision trees are called base models, and they can be represented more formally as:

$$g(x)= f0(x)+ f1(x)+ f2(x)+....$$

Random forest uses the Bagging or Bootstrap Aggregation technique of ensemble learning in which aggregated decision tree runs in parallel and do not interact with each other. With the help of Random Forest regression, we can prevent Overfitting in the model by creating random subsets of the dataset. We can prevent Overfitting in the model by creating random subsets of the dataset.

Random Forest works in two-phase first is to create the random forest by combining N decision trees, and the second is to make predictions for each tree created in the first phase.
The Working process can be explained in the below steps and diagram:
**Step-1:** Select random K data points from the training set.
**Step-2:** Build the decision trees associated with the selected data points (Subsets).
**Step-3:** Choose the number N for the decision trees that you want to build.
**Step-4:** Repeat Step 1 & 2.
**Step-5:** For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority of votes.



### Applications of Random Forest

There are mainly four sectors where Random forest is mostly used:

- **Banking:** The banking sector mostly uses this algorithm for the identification of loan risk.
- **Medicine:** With the help of this algorithm, disease trends and risks of the disease can be identified.
- **Land Use:** We can identify the areas of similar land use by this algorithm.
- **Marketing:** Marketing trends can be identified using this algorithm.

### Advantages of Random Forest

- Random Forest is capable of performing both Classification and Regression tasks.
- It is capable of handling large datasets with high dimensionality.
- It enhances the accuracy of the model and prevents the overfitting issue.

### Disadvantages of Random Forest

- Although random forest can be used for both classification and regression tasks, it is not more suitable for Regression tasks.

# 5.1 Introduction to Technology

## Python Programming:

Python is often used as a support language for software developers, for building control and management, for testing, and in many other ways. The role of machine learning is to recognize patterns in data. A machine learning engineer is responsible for extracting, processing, refining, cleaning, arranging, and making sense of data to develop intelligent algorithms. Python is easy to understand. While linear algebra or calculus concepts can be so complex, they take the maximum amount of effort. Python can be implemented quickly, which helps machine learning engineers to validate an idea promptly.

One of the main reasons why Python is the preferred language for machine learning is its access to many libraries. A library is a collection of functions and routines that a programming language can use. Having access to various libraries allows developers to perform complex tasks without the need to rewrite many code lines. Since machine learning heavily relies on mathematical optimization, probability and statistics, Python libraries help data scientists perform various studies easily.

## Python Web Frameworks

A web framework is a code library that makes a developer's life easier when building reliable, scalable, and maintainable web applications.

Why are web frameworks useful?

Web frameworks encapsulate what developers have learned over the past twenty years while programming sites and applications for the web. Frameworks make it easier to reuse code for common HTTP operations and to structure projects so other developers with knowledge of the framework can quickly build and maintain the application.

Common web framework functionality

Frameworks provide functionality in their code or through extensions to perform common operations required to run web applications. These common operations include:

- URL routing
- HTML, XML, JSON, and other output format templating
- Database manipulation
- Session storage and retrieval

## Django Web Framework

Django is a Python-based web framework that allows you to quickly create efficient web applications. It is also called batteries included framework because Django provides built-in features for everything including Django Admin Interface, default database – SQLlite3, etc. When you're building a website, you always need a similar set of components: a way to handle user authentication (signing up, signing in, signing out), a management panel for your website, forms, a way to upload files, etc. Django gives you ready-made components to use and that too for rapid development.

## Why Django Framework?

- Django is a rapid web development framework that can be used to develop fully fleshed web applications in a short period.
- It's very easy to switch databases in the Django framework.
- It has a built-in admin interface which makes it easy to work with it.
- Django is a fully functional framework that requires nothing else.
- It has thousands of additional packages available.
- It is very scalable.

## Installation of Django

- Install python3 if not installed in your system (according to the configuration of your system and OS). Try to download the latest version of python it's python3.10. this time.

  **Note-** Installation of Django in Linux and Mac is similar, here I am showing it in windows for Linux and mac just open the terminal in place of the command prompt and go through the following commands.

- **Install pip-** Open the command prompt and enter the following command-

  ```
  python -m pip install -U pip
  ```

- **Install virtual environment**- Enter the following command in cmd-

  ```
  pip install virtualenv
  ```

- **Set Virtual environment-** Setting up the virtual environment will allow you to edit the dependency which generally your system wouldn't allow.

  Follow these steps to set up a virtual environment-

- Create a virtual environment by giving this command in cmd-

```
virtualenv env_site
```

- Change directory to env_site by this command-

```
cd env_site
```

- Go to Script directory inside env_site and activate virtual environment-

```
cd Scripts
activate
```

- **Install Django**- Install Django by giving the following command-

```
pip install Django
```

## CSS

CSS works by creating rules. These rules are simultaneously applied to multiple elements within the site. Eliminating the repetitive coding style of HTML makes development work faster and less monotonous. Errors are also reduced considerably. Since the content is completely separated from the design, changes across the website can be implemented all at once. This reduces delivery times and costs of future edits. The only major limitation of CSS is that its performance depends largely on browser support. However, in case your CSS styling isn’t fully supported by a browser, people will still be able to experience the HTML functionalities. Therefore, you should always have a well-structured HTML along with good CSS.

## HTML

HTML is the code used to structure a web page and its content. For example, content could be structured within a set of paragraphs, a list of bulleted points, or using images and data tables. HTML is a markup language that defines the structure of your content. HTML consists of a series of elements, which you use to enclose, or wrap, different parts of the contents to make it appear a certain way. The enclosing tags can make a word or image hyperlink to somewhere else can italicize words, can make the font bigger or smaller, and so on. "Hypertext" refers to links that connect web pages, either within a single website or between websites. Links are a fundamental aspect of the Web. By uploading content to the Internet and linking it to pages created by other people, you become an active participant in the World Wide Web.

## USED MODULES

These are the modules used in this project:

- **NumPy**

  NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. NumPy can be used to perform a wide variety of mathematical operations on arrays. It adds powerful data structures to Python that guarantee efficient calculations with arrays and matrices and it supplies an enormous library of high-level mathematical functions that operate on these arrays and matrices.

- **Pandas**

  Pandas is an open-source Python library providing high-performance data manipulation and analysis tools using its powerful data structures. Python was majorly used for data munging and preparation. It had very little contribution to data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, statistics, analytics, etc.

- **Sckit-Learn**

  Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering, and dimensionality reduction via a consistence interface in Python. This library, which is largely written in Python, is built upon NumPy, SciPy, and Matplotlib.

- **Mathplotlib**

  Matplotlib is a Python 2D plotting library that produces publication-quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter notebook, web application servers, and four graphical user interface toolkits. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatter plots, etc., with just a few lines of code.

## 5.2 Sample Code

### home.html

```
{% load static %}
<!DOCTYPE html>
<html>
<head>
<title>Asset Price</title>
<style>
 body{
 background-image:url("{%static '14.jpg'%}");
 background-repeat:no-repeat;
 background-attachment:fixed;
 background-size:100% 100%;
 }
input[type=submit]{
background-color:#0dcaf0;
color : white;
cursor : pointer;
padding: 16px 32px;
border-radius:40px;
}
#c1{
padding-top:18pc;
}
#heading{
font-size:5.5rem;
}
</style>
<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css"
rel="stylesheet"integrity="sha3841BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrC
EXSU1oBoqyl2QvZ6jIW3" crossorigin="anonymous">
```

```
<scriptsrc="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js"integrity="sha384ka7Sk0Gln4gmtz2MlQnikT1wXgYsOg+OMhuP+IlRH9sENBO0LRn5q+8nbTov4+1p" crossorigin="anonymous"></script>
<scriptsrc="https://kit.fontawesome.com/552998d52c.js"crossorigin="anonymous"></script>
</head>
<body style="text-align:center">
<div class="container" id="c1">
<h1>Welcome to </h1>
<h1 id="heading">Asset Price Prediction</h1><br><br>
<div class="row">
<div class="col">
<form action="/house_predict">
<h3>For House Price Prediction,<br>click the below button</h3>
<input type="submit" id="but" value="START">
</form>
</div>
<div class="col">
<form action="/land_predict">
<h3>For Land Price Prediction,<br>click the below button</h3>
<input type="submit" id="but" value="START">
</form>
</div>
</div>
</div>
</body>
</html>
```

## Land_predict.html

```
{% load static %}
<!DOCTYPE html>
<html>
<head>
<title>Asset Price</title>
```

```
<style>
html{
overflow-x: hidden;
}
body{
background-image: url("{%static '14.jpg'%}");
background-repeat: no-repeat;
background-attachment: fixed;
background-size:100% 100%;
}
h1{
font-size:60px;
}
input[type=submit]{
background-color:#000000;
color: white;
cursor: pointer;
}
.card{
width:54pc;
height:100%;
margin-top:12pc;
margin-left:12pc;
}
.card-header{
text-align: center;
border-radius:40px;
width:54pc;
}
#r2{
text-align: center;
margin-top:20px;
}
```

```
#btn1{
font-size: 20px;
}
#location{
width:50pc;
}
#bhk{
border-radius:200pc;
width:50pc;
}
#bedrooms{
border-radius:200pc;
width:50pc;
}
#total_sqft{
width:22pc;
}
#dim{
width:23.4pc;
}
#card1{
border-radius:40px;
}
#type{
width:50pc;
}
.icons{
padding-top:2pc;
padding-left:2pc;
color: black;
}
i {
text-shadow: 2px 4px 6px black;
}
```

```
</style>
 <linkhref="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css"rel="styl
esheet"integrity="sha3841BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrCEXSU1o
Boqyl2QvZ6jIW3" crossorigin="anonymous">
<scriptsrc="https://kit.fontawesome.com/552998d52c.js"crossorigin="anonymous"></script>
<scriptsrc="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js"inte
grity="sha384ka7Sk0Gln4gmtz2MlQnikT1wXgYsOg+OMhuP+IlRH9sENBO0LRn5q+8nb
Tov4+1p" crossorigin="anonymous"></script>
</head>
<body class="bg-dark">
<div class="icons">
<a href="/home"> <i class="fa-solid fa-house fa-2x"></i></a>
</div>
<div class="container">
<div class="card" id="card1" >
<div class="card-header">
<h1>Welcome to Land Price Predictor</h1>
</div>
<br>
<div class="card-body">
<form action="result">
{% csrf_token %}
<div class="row">
<div class="col form-group">
<label><b>Select the Location:</b></label>
<select class="select-picker form-control" id="location" name="location"
required="0">
{% for location in locations %}
<option value="{{location}}">{{location}}</option>
{% endfor %}
</select>
</div>
</div>
<br>
```

```
<div class="row">
<div class="col form-group" >
<label><b>Enter Dimensions:</b></label><br>
<input type="number"  class="form-control" id="total_sqft"
name="total_sqft"  required="1"/>
</div>
<div class="col form-group">
<label><b>Choose Dimensions:</b></label><br>
<select class="form-control" id="dim" name="dim">
<option value="1">Sq_Yard</option>
<option value="2">Acre</option>
<option value="3">Cents</option>
</select>
</div>
</div>
<br>
<center>
<div class="col">
<input type="submit" class="btn btn-primary"  id="btn1" value="PredictPrice"/>
</div>
</center>
</div>
</form>
</div>
</div>
<div class="row" id="r2">
<h2>{{result2}}</h2>
</div>
</div>
</body>
</html>
```

## HousePrediction.html

```
{% load static %}
<!DOCTYPE html>
<html>
<head>
<title>Asset Price</title>
<style>
body{
background-image: url("{%static '14.jpg'%}");
background-repeat: no-repeat;
background-attachment: fixed;
background-size:100% 100%;
}
h1{
font-size:60px;
}
input[type=submit]{
background-color:#000000;
color: white;
cursor: pointer;
padding: 16px 32px;
}
.card{
width:54pc;
height:100%;
margin-top:75px;
margin-left:15pc;
}
.card-header{
text-align: center;
border-radius:40px;
width:54pc;
}
```

```
#r2{
text-align: center;
margin-top:20px;
}
#btn1{
border-radius:200pc;
font-size: 25px;
width:50pc;
}
#location{
width:50pc;
}
#bhk{
border-radius:200pc;
width:50pc;
}
#bedrooms{
border-radius:200pc;
width:50pc;
}
#total_sqft{
width:22pc;
}
#dim{
width:23.4pc;
}
#card1{
border-radius:40px;
}
#type{
width:50pc;
}
.icons{
padding-top:2pc;
padding-left:2pc;
color: black;
}
i {
text-shadow: 2px 4px 6px black;
}
</style>
<script type="text/javascript">
function t(x){
if(x==0) document.getElementById('ftt').style.display="";
else document.getElementById('ftt').style.display="none";
return ;
}
</script>
<linkhref=”https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css”rel="style
sheet"integrity="sha3841BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrCEXSU1o
Boqyl2QvZ6jIW3" crossorigin="anonymous">
<scriptsrc="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js"intgr
ity="sha384ka7Sk0Gln4gmtz2MlQnikT1wXgYsOg+OMhuP+IlRH9sENBO0LRn5q+8nbTo
v4+1p" crossorigin="anonymous"></script>
<scriptsrc="https://kit.fontawesome.com/552998d52c.js"crossorigin="anonymous"></script>
</head>
<body class="bg-dark">
<div class="icons">
<a href="/home"> <i class="fa-solid fa-house fa-2x"></i></a>
</div>
<div class="container">
<div class="card" id="card1" >
<div class="card-header">
<h1>Welcome to House Price Predictor</h1>
</div>
<br>
<div class="card-body">
<form  action="result1">
```

```
{% csrf_token %}
<div class="row">
<div class="col form-group">
<label><b>Select the Location:</b></label>
<select class="select-picker form-control" id="location" name="location"   required="1">
{% for location in locations %}
<option value="{{location}}">{{location}}</option>
{% endfor %}
</select>
</div>
</div>
<div class="row">
<div class="col form-group">
<br>
<label><b>Choose Type:</b></label><br>
<select class="select-picker form-control" id="ct" name="ct" required="1">
<option value="Individual House" onclick(1)>Individual House</option>
<option value="Appartment" onclick=t(0)>Appartment</option>
</select>
</div>
</div>
<div class="row" >
<div class="col form-group" id="ftt" style="display:none;">
<br>
<label><b>Choose floor:</b></label><br>
<select class="select-picker form-control" id="ft" name="ft" required="1">
<option value="1">Ground Floor</option>
<option value="2">First Floor</option>
<option value="3">Second Floor</option>
<option value="4">Above </option>
</select>
</div>
</div>
<br>
```

```html
<div class="row">
<div class="col form-group">
<label><b>Select Number of Bedrooms:</b></label><br>
<select class="select-picker form-control" id="bhk" name="bhk" required="1">
<option value="1">1</option>
<option value="2">2</option>
<option value="3">3</option>
</select>
</div>
</div>
<br>
<div class="row">
<div class="col form-group">
<label><b>Select Number of Bathrooms:</b></label><br>
<select class="select-picker form-control" id="bathrooms" name="bath" required="1">
<option value="1">1</option>
<option value="2">2</option>
<option value="3">3</option>
</select>
</div>
</div>
<br>
<div class="row">
<div class="col form-group" >
<label><b>Enter Dimensions:</b></label><br>
<inputtype="number  class="form-control"  id="total_sqft" name=”total_sqft”  required="1"/>
</div>
<div class="col form-group">
<label><b>Choose Dimensions:</b></label><br>
<select class="form-control" id="dim" name="dim">
<option value="1">Sq_Feet</option>
<option value="2">Sq_Yard</option>
</select>
</div>
```

```
</div>
<br>
<div class="col form-group">
<input type="submit" class="btn btn-primary form-group" id="btn1" value="Predict Price"/>
</div>
</form>
<div class="row" id="r2">
<h2>{{result2}}</h2>
</div>
</div>
</div>
</body>
</html>
```

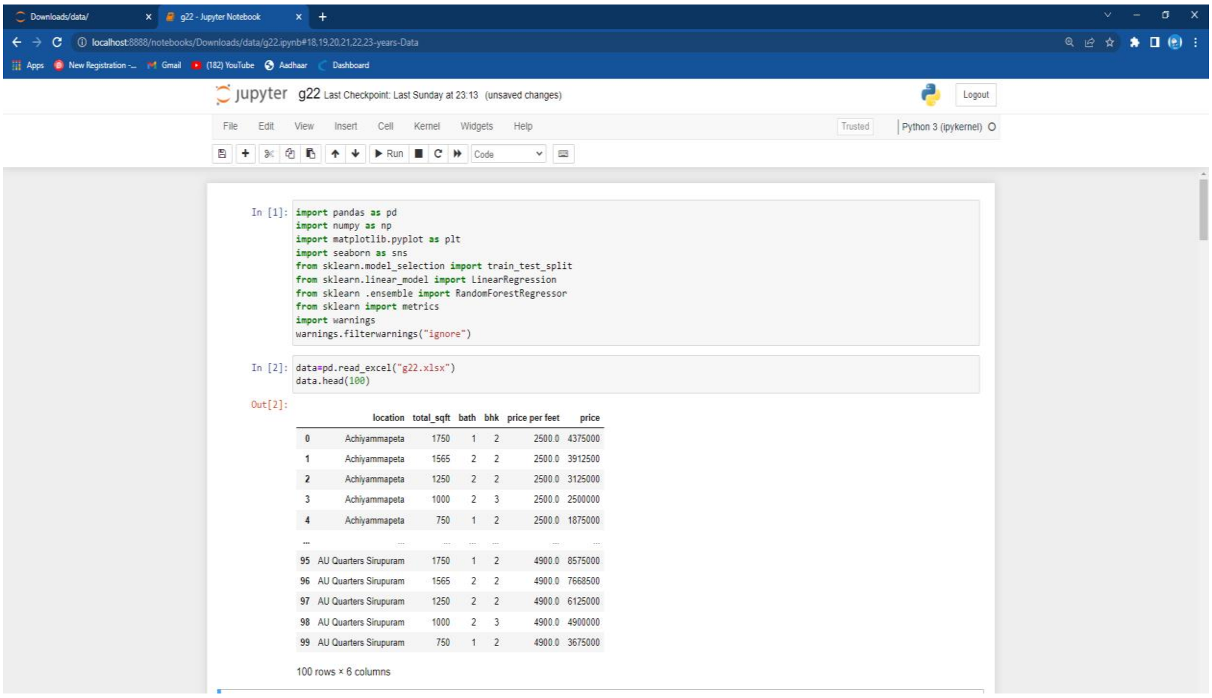
## View.py

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from django.shortcuts import render
import warnings
warnings.filterwarnings("ignore")

def get24(location,total_sqft,bath,bhk):
    data=pd.read_excel(r'C:/Users/VIKRAM/Downloads/data/g24.xlsx')
    dummies=pd.get_dummies(data.location)
    df=pd.concat([data,dummies],axis='columns')
    #df1=df.drop('price per feet',axis='columns')
    df2=df.drop('location',axis='columns')
    X=df2.drop('price',axis="columns")
    y=df2['price']
    X_train,X_test,y_train,y_test =train_test_split(X,y,test_size=0.2,random_state=10)
    lr=LinearRegression()
    lr.fit(X_train,y_train)
    loc=np.where(X.columns==location)[0][0]
    x=np.zeros(len(X.columns))
    x[0]=total_sqft
    x[1]=bath
    x[2]=bhk
    if loc>=0:
        x[loc]=1
    return lr.predict([x])[0]

def get23(location,total_sqft,bath,bhk):
    data=pd.read_excel(r'C:/Users/VIKRAM/Downloads/data/g23.xlsx')
    dummies=pd.get_dummies(data.location)
    df=pd.concat([data,dummies],axis='columns')
    df1=df.drop('price per feet',axis='columns')
    df2=df1.drop('location',axis='columns')
    X=df2.drop('price',axis="columns")
    y=df2['price']
    X_train,X_test,y_train,y_test =train_test_split(X,y,test_size=0.2,random_state=10)
    lr=LinearRegression()
    lr.fit(X_train,y_train)
    loc=np.where(X.columns==location)[0][0]
    x=np.zeros(len(X.columns))
    x[0]=total_sqft
    x[1]=bath
    x[2]=bhk
    if loc>=0:
        x[loc]=1
    return lr.predict([x])[0]

def get22(location,total_sqft,bath,bhk):
    data=pd.read_excel(r'C:/Users/VIKRAM/Downloads/data/g22.xlsx')
    dummies=pd.get_dummies(data.location)
    df=pd.concat([data,dummies],axis='columns')
    df1=df.drop('price per feet',axis='columns')
    df2=df1.drop('location',axis='columns')
    X=df2.drop('price',axis="columns")
    y=df2['price']
    X_train,X_test,y_train,y_test =train_test_split(X,y,test_size=0.2,random_state=10)
    lr=LinearRegression()
    lr.fit(X_train,y_train)
    loc=np.where(X.columns==location)[0][0]
    x=np.zeros(len(X.columns))
    x[0]=total_sqft
    x[1]=bath
    x[2]=bhk
    if loc>=0:
        x[loc]=1
    return lr.predict([x])[0]

def convert(n,var5):
    if n==1:
      var5*=9
    elif n==2:
      var5*=43560
    elif n==3:
        var5*=435.6
    return int(var5)
def c(n):
     if n==1:
       a="Sq_Yards"
     elif n==2:
      a="Acre"
     elif n==3:
        a="Cents"
     return a
```

```
def c1(n):
     if n==1:
       a="Sq Feets"
     elif n==2:
      a="Sq Yards"
     return a
def home(request):
    return render(request,'home.html')
def house_predict(request):
    data=pd.read_excel(r'C:/Users/VIKRAM/Downloads/data/upto 4 years.xlsx')
    locations=sorted(data['location'].unique())
    return render(request,'house_predict.html',{'locations':locations})
def land_predict(request):
    data=pd.read_excel(r'C:/Users/VIKRAM/Downloads/data/upto 4 years.xlsx')
    locations=sorted(data['location'].unique())
    return render(request,'land_predict.html',{'locations':locations})
def result1(request):
    var1=str(request.GET['location'])
    var2=str(request.GET['ct'])
    #var8=str(request.GET['HT'])
    #var7=int(request.GET['ft'])
    var3=int(request.GET['bath'])
    var4=int(request.GET['bhk'])
    var5=int(request.GET['total_sqft'])
    var6=int(request.GET['dim'])
    n5=var5
    print(var2)
    if var6==2:
      var5=var5*9
    print("sq_feets",+var5)
    g22=get22(var1,var5,var3,var4).round(0)
    g23=get23(var1,var5,var3,var4).round(0)
    g24=get24(var1,var5,var3,var4).round(0)
    n=c1(var6)
    d1=" "+str(n5)+" "+n
    g2=" Present price  is"+" "+str(g22)+" "+"INR"
    g3="At 2023,the Predicted price  is"+" "+str(g23)+" "+"INR"
    g4="At2024,thePredictedpriceis"+""+str(g24)+""+"INR"
returnrender(request,'HouReport.html',{"result22":g2,"result23":g3,"result24":g4,"location":v
ar1,"dm":d1,"ht":var2,"cf":c,"bat":var3,"bed":var4})
#Land Function()
def result(request):
    loc=str(request.GET['location'])
    var5=int(request.GET['total_sqft'])
    dim=int(request.GET['dim'])
    dm=convert(dim,var5)
    print("sq_feets",+dm)
    g22=get22(loc,var5,1,1).round(0)
    g23=get23(loc,var5,1,1).round(0)
    g24=get24(loc,var5,1,1).round(0)
    n=c(dim)
    d=" "+str(var5)+" "+n
    g2=" Present price  is"+" "+str(g22)+" "+"INR"
    g3="The Predicted price at 2023 is"+" "+str(g23)+" "+"INR"
    g4="The Predicted price at 2024 is"+" "+str(g24)+" "+"INR"
  return render(request,'landReport.html',{"result22":g2,"result23":g3,"result24":g4,
 "location":loc,"dm":d})
```

## 5.3 Screenshots

### Screenshot for uploading Asset dataset



```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        from sklearn.model_selection import train_test_split
        from sklearn.linear_model import LinearRegression
        from sklearn .ensemble import RandomForestRegressor
        from sklearn import metrics
        import warnings
        warnings.filterwarnings("ignore")
```

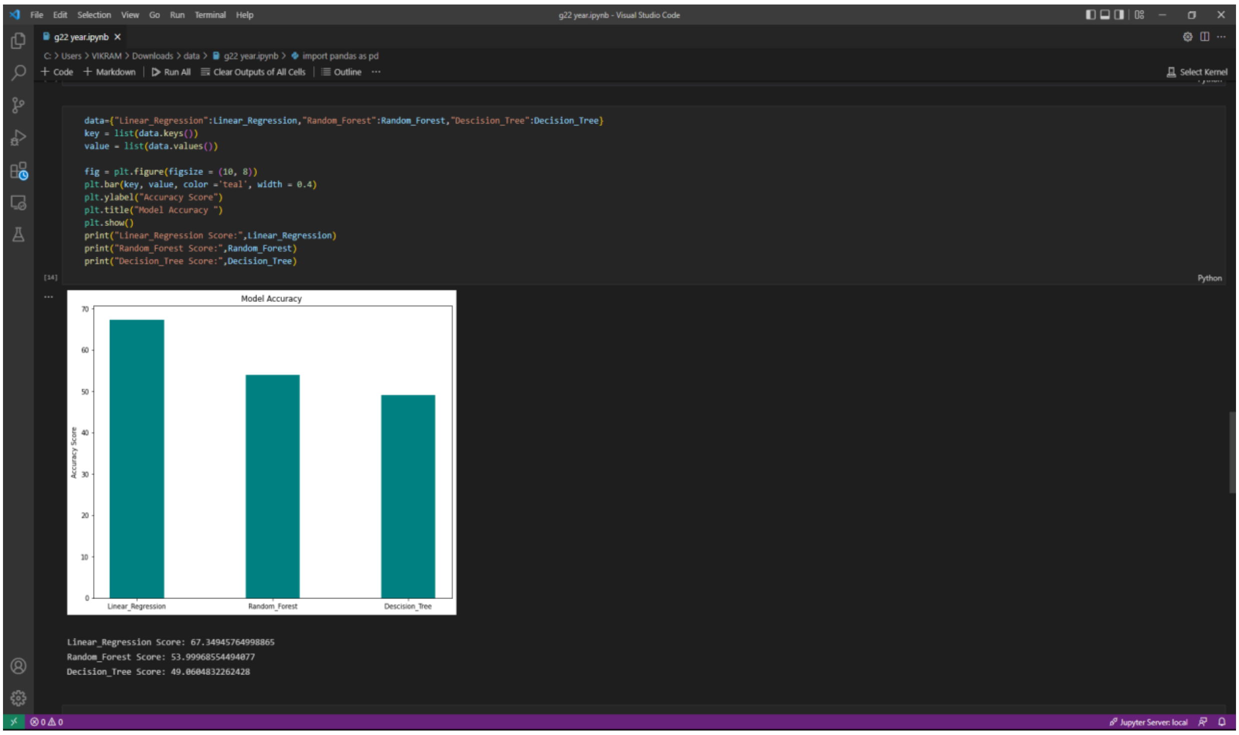
```
In [2]: data=pd.read_excel("g22.xlsx")
        data.head(100)
```

```
Out[2]:
                  location  total_sqft  bath  bhk  price per feet    price
0            Achiyammapeta        1750     1    2          2500.0  4375000
1            Achiyammapeta        1565     2    2          2500.0  3912500
2            Achiyammapeta        1250     2    2          2500.0  3125000
3            Achiyammapeta        1000     2    3          2500.0  2500000
4            Achiyammapeta         750     1    2          2500.0  1875000
...                    ...         ...   ...  ...             ...      ...
95  AU Quarters Sirupuram        1750     1    2          4900.0  8575000
96  AU Quarters Sirupuram        1565     2    2          4900.0  7668500
97  AU Quarters Sirupuram        1250     2    2          4900.0  6125000
98  AU Quarters Sirupuram        1000     2    3          4900.0  4900000
99  AU Quarters Sirupuram         750     1    2          4900.0  3675000

100 rows × 6 columns
```

The above screenshot describes how the asset dataset is loaded into the jupyter notebook. we use the pandas module to read the dataset. Pandas are used mainly for removing noisy data or cleaning the dataset. The dataset can be in the form of an excel sheet or CSV format file. Pandas consist of the read_excel( ) function to read the excel sheet file and read_csv( ) function to read the CSV format file.

The Asset dataset consists of six columns i.e., location, total_sqft, bath, bhk, price per feet, and price. In this Asset dataset, independent variables are location, total_sqft, bath, and bhk, and the dependent variable is the price column.

# Accuracy score for model-1



File Edit Selection View Go Run Terminal Help
g22 year.ipynb - Visual Studio Code
g22 year.ipynb
C: > Users > VIKRAM > Downloads > data > g22 year.ipynb > import pandas as pd
+ Code + Markdown | Run All | Clear Outputs of All Cells | Outline ···
Select Kernel
[14]
Python
Model Accuracy
Accuracy Score
0
10
20
30
40
50
60
70
Linear_Regression
Random_Forest
Descision_Tree
⊗ 0 ⚠ 0
Jupyter Server: local

```python
data={"Linear_Regression":Linear_Regression,"Random_Forest":Random_Forest,"Descision_Tree":Decision_Tree}
key = list(data.keys())
value = list(data.values())

fig = plt.figure(figsize = (10, 8))
plt.bar(key, value, color ='teal', width = 0.4)
plt.ylabel("Accuracy Score")
plt.title("Model Accuracy ")
plt.show()
print("Linear_Regression Score:",Linear_Regression)
print("Random_Forest Score:",Random_Forest)
print("Decision_Tree Score:",Decision_Tree)
```

```
Linear_Regression Score: 67.34945764998865
Random_Forest Score: 53.99968554494077
Decision_Tree Score: 49.0604832262428
```

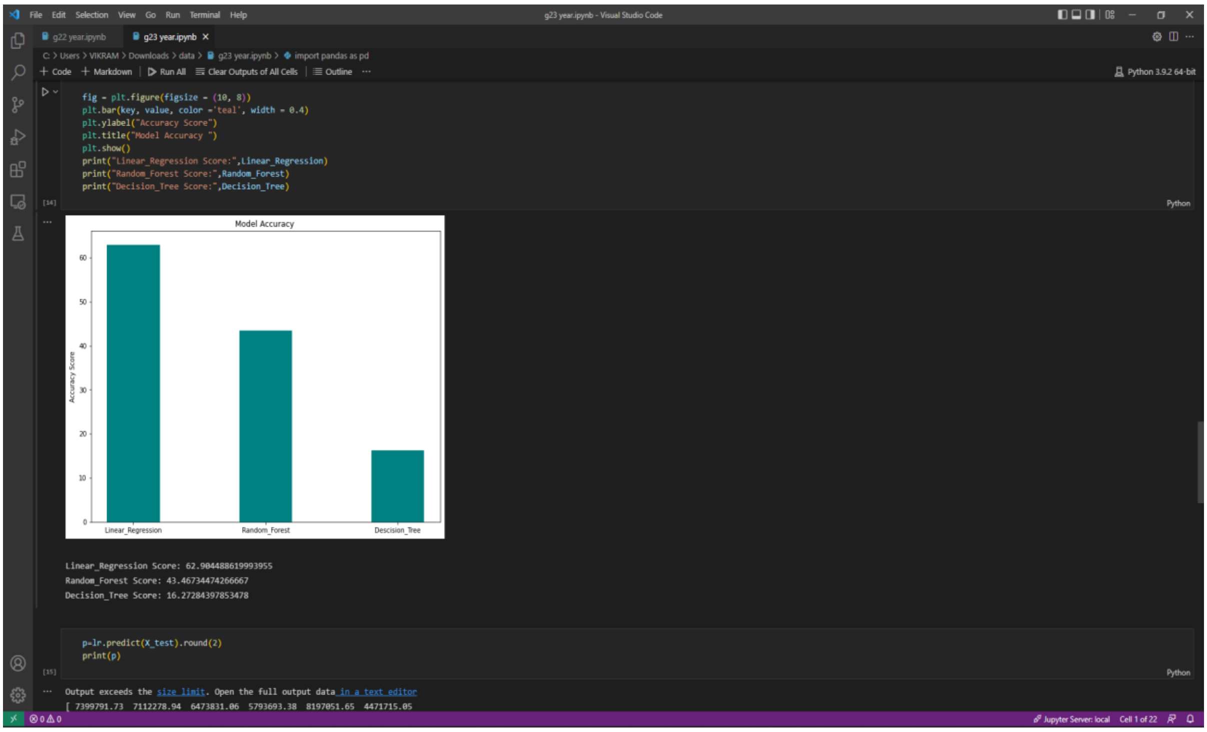
The above screenshot describes the accuracy of model-1, we used mathplotlib library to visualize the accuracy score in the form of bar plots.

This model is used to predict the price of the asset. In this model, we use three machine learning algorithms i.e., Linear Regression, Random Forest, and Decision Tree algorithm. The Accuracy score for these algorithms :

- Linear_Regression Algorithm : 67.349576499
- Random_Forest Algorithm : 53.999685449
- Decision_Tree Algorithm : 49.060483226

From these accuracy scores of model-1, we can say that linear Regression has the highest accuracy score among the other two algorithms.

## Accuracy score for model-2



```
fig = plt.figure(figsize = (10, 8))
plt.bar(key, value, color ='teal', width = 0.4)
plt.ylabel("Accuracy Score")
plt.title("Model Accuracy ")
plt.show()
print("Linear_Regression Score:",Linear_Regression)
print("Random_Forest Score:",Random_Forest)
print("Decision_Tree Score:",Decision_Tree)
```

[14]

```
Linear_Regression Score: 62.904488619993955
Random_Forest Score: 43.46734474266667
Decision_Tree Score: 16.27284397853478
```

```
p=lr.predict(X_test).round(2)
print(p)
```

[15]

```
Output exceeds the size limit. Open the full output data in a text editor
[ 7399791.73  7112278.94  6473831.06  5793693.38  8197051.65  4471715.05
```

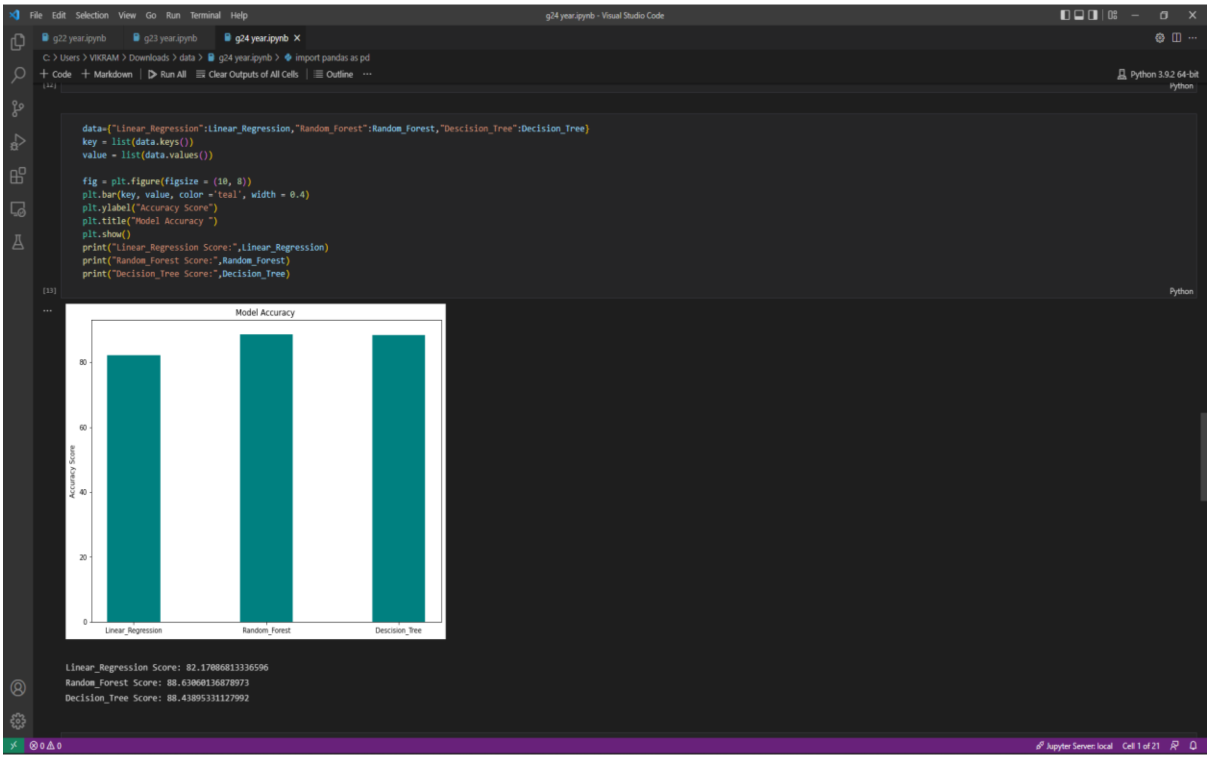
The above screenshot describes the accuracy of model-2, we used mathplotlib library to visualize the accuracy score in the form of bar plots.

This model is used to predict the price of the asset. In this model, we use three machine learning algorithms i.e., Linear Regression, Random Forest, and Decision Tree algorithm. The Accuracy score for these algorithms :

- Linear_Regression Algorithm : 62.984488676
- Random_Forest Algorithm : 43.467344742
- Decision_Tree Algorithm : 16.272842397

From these accuracy scores of model-2, we can say that linear Regression has the highest accuracy score among the other two algorithms.

## Accuracy score for model-3



```
data={"Linear_Regression":Linear_Regression,"Random_Forest":Random_Forest,"Descision_Tree":Decision_Tree}
key = list(data.keys())
value = list(data.values())

fig = plt.figure(figsize = (10, 8))
plt.bar(key, value, color ='teal', width = 0.4)
plt.ylabel("Accuracy Score")
plt.title("Model Accuracy ")
plt.show()
print("Linear_Regression Score:",Linear_Regression)
print("Random_Forest Score:",Random_Forest)
print("Decision_Tree Score:",Decision_Tree)
```

```
Linear_Regression Score: 82.17086813336596
Random_Forest Score: 88.63060136878973
Decision_Tree Score: 88.43895331127992
```

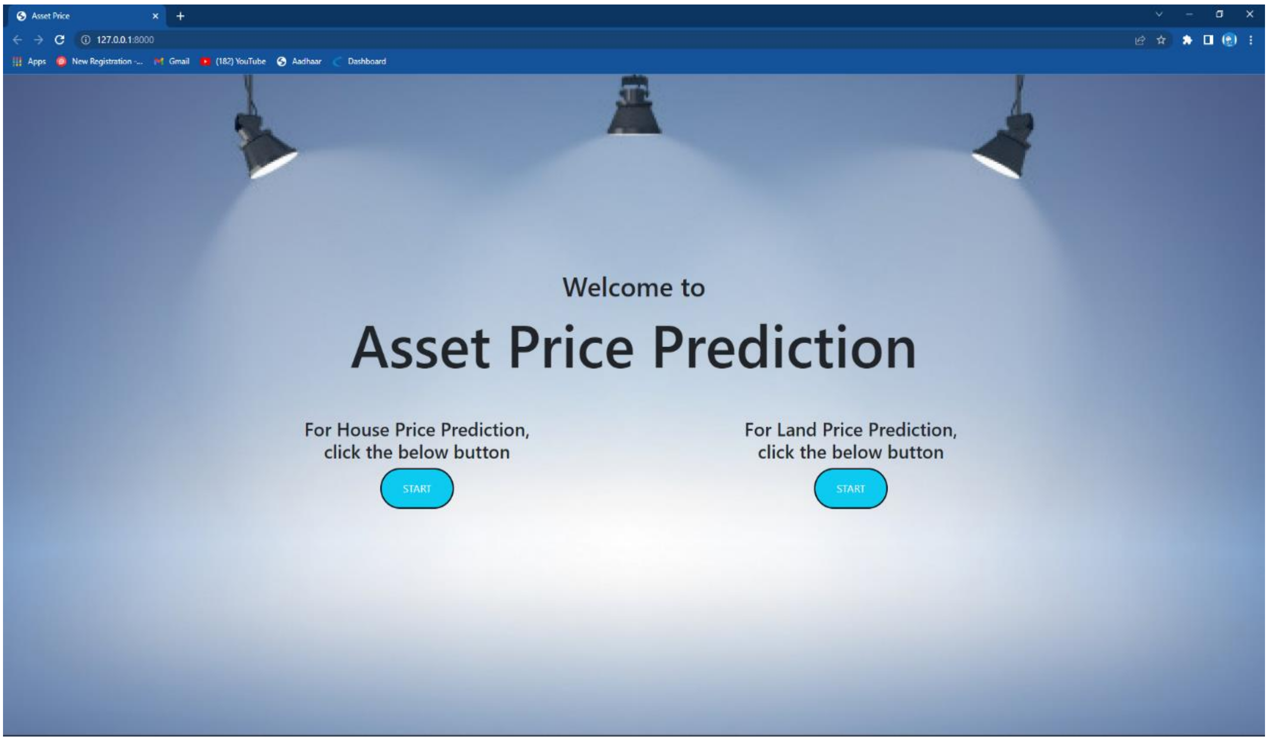
The above screenshot describes the accuracy of model-3, we used mathplotlib library to visualize the accuracy score in the form of bar plots.

This model is used to predict the price of the asset. In this model, we use three machine learning algorithms i.e., Linear Regression, Random Forest, and Decision Tree algorithm. The Accuracy score for these algorithms :

- Linear_Regression Algorithm : 82.1708681333
- Random_Forest Algorithm : 88.6306013687
- Decision_Tree Algorithm : 88.4389533112

From these accuracy scores of model-3, we can say that Random Forest has the highest accuracy score among the other two algorithms.

## Screenshot for  Homepage



Welcome to

Asset Price Prediction

For House Price Prediction,
click the below button

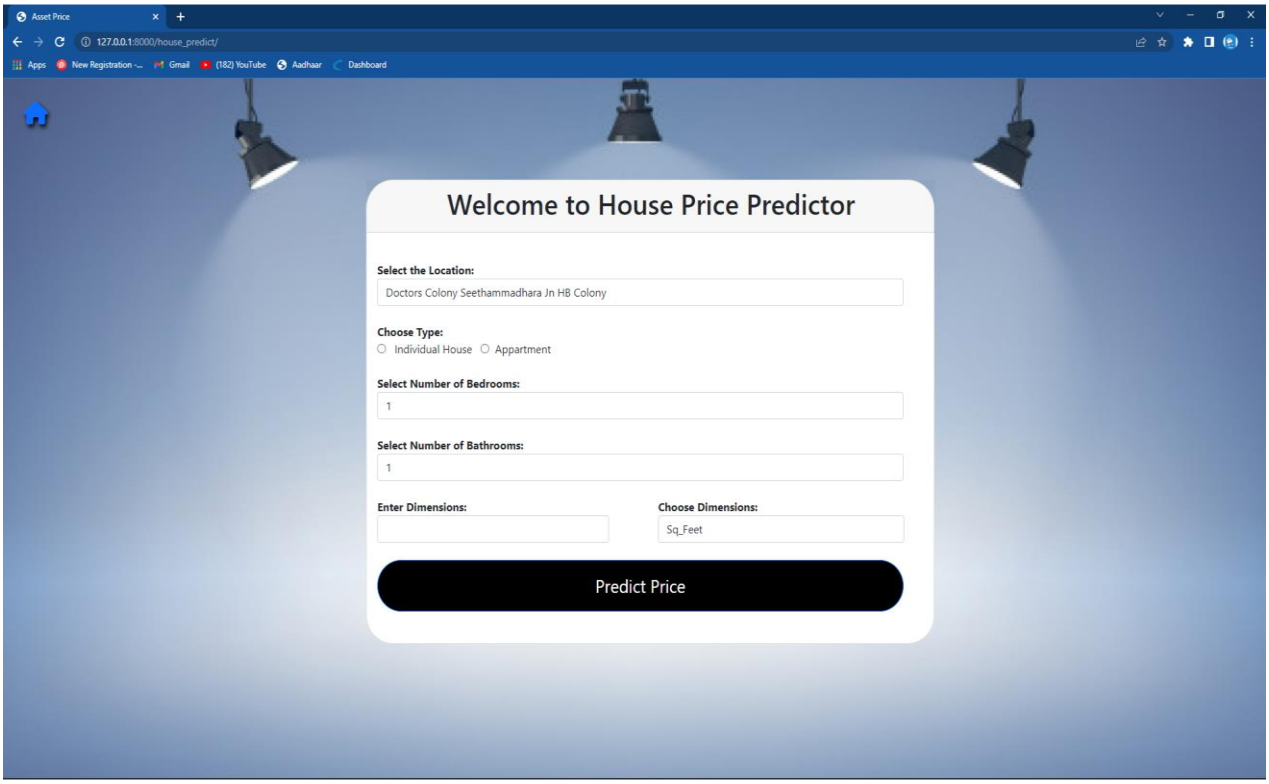
For Land Price Prediction,
click the below button

The above screenshot describes the home page of the website. The home page consists of asset features like House and Land. The user needs to select a feature as per their requirement.

If the user selects house price prediction, it will redirect to the house price prediction page or if the user selects land price prediction it will redirect to the land price prediction page.

Based on the user selection, the website redirects either to the house price prediction page or the land price prediction page.

## Screenshot for House Prediction Page



Welcome to House Price Predictor

Select the Location: Doctors Colony Seethammadhara Jn HB Colony

Choose Type: ○ Individual House ○ Appartment

Select Number of Bedrooms: 1

Select Number of Bathrooms: 1

Enter Dimensions:

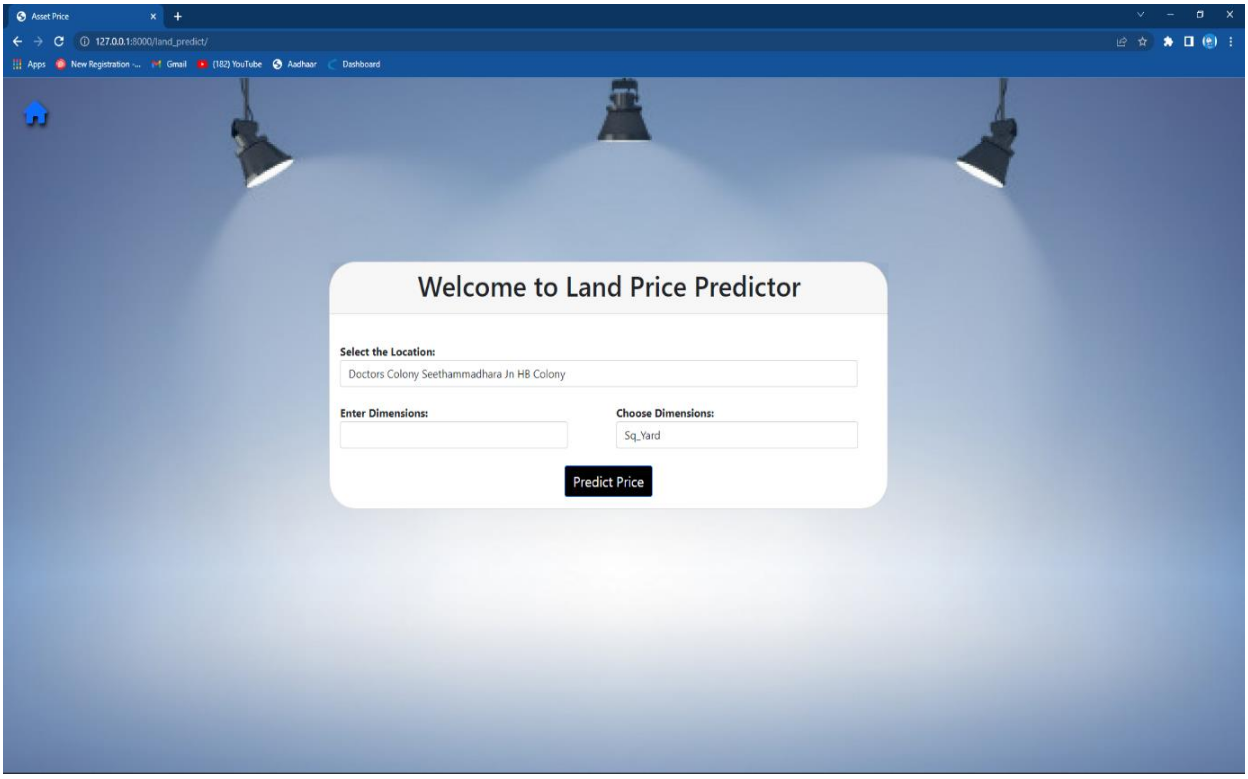
Choose Dimensions: Sq_Feet

The above screenshot describes the House price prediction page, After the selection of asset feature by the user. The home page redirects to this page.

On this page, the user needs to enter all the required fields for the prediction of the house price. The user needs to choose a specific location, and choose the house type. After choosing the house type, the user needs to select the number of bhk, number of bath, and enter the dimensions of the house for the specific location.

After entering all the fields, the user needs to click on the predict price button. The button verifies whether all the fields are selected or not, if not it gives an error message like the field is required. After clicking the button, it will redirect to the house price report page and display all the fields entered by the user and display the predicted price for the future three years.

## Screenshot for Land Prediction Page



Welcome to Land Price Predictor

Select the Location: Doctors Colony Seethammadhara Jn HB Colony

Enter Dimensions:

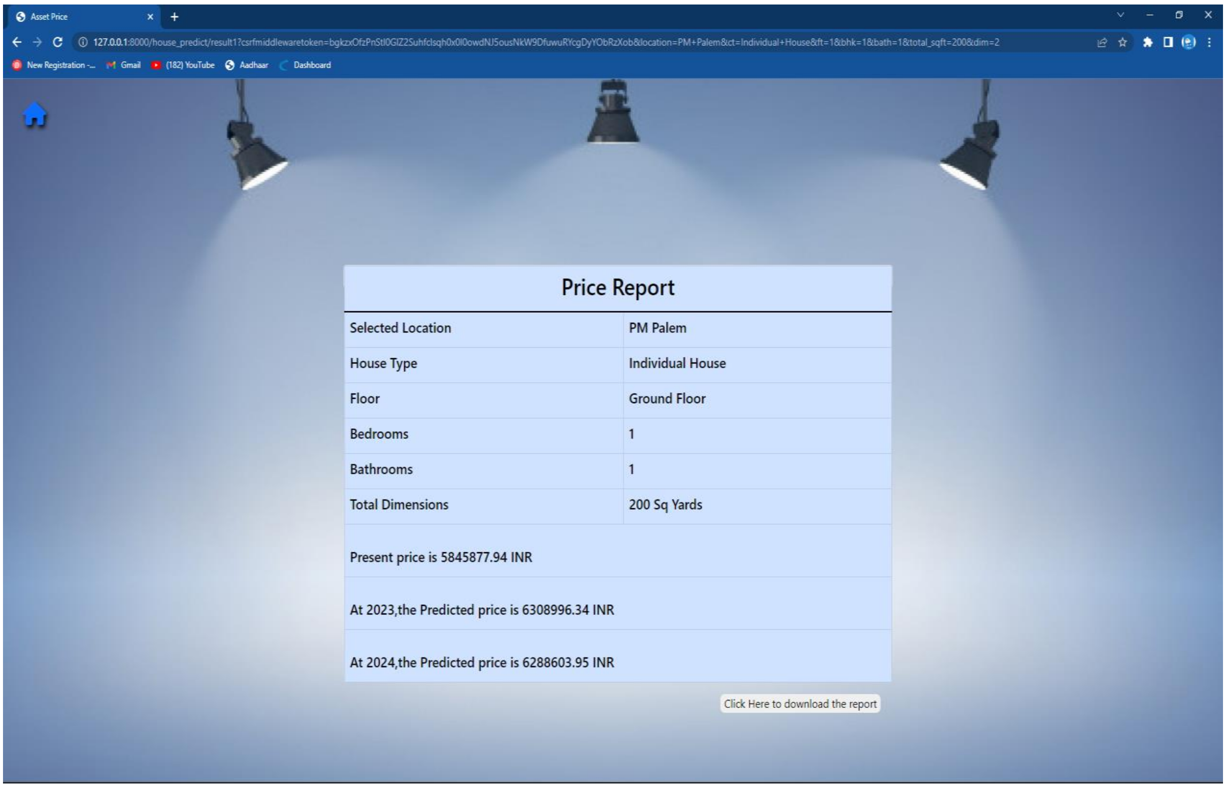
Choose Dimensions: Sq_Yard

The above screenshot describes the Land price prediction page, After the selection of asset feature by the user. The home page redirects to this page.

On this page, the user needs to enter all the required fields for the prediction of the land price. The user needs to choose a specific location, enter the dimensions of the land and choose the dimension type for the selected location.

After entering all the fields, the user needs to click on the predict price button. The button verifies whether all the fields are selected or not, if not it gives an error message like the field is required. After clicking the button, it will redirect to the land price report page and display all the fields entered by the user and display the predicted price for the future three years.

## Screenshot for House Price Report



127.0.0.1:8000/house_predict/result1?csrfmiddlewaretoken=bgkzxOfzPnStl0GlZ2Suhfclsqh0x0l0owdNJ5ousNkW9DfuwuRYcgDyYObRzXob&location=PM+Palem&ct=Individual+House&ft=1&bhk=1&bath=1&total_sqft=200&dim=2

Price Report

| | |
|---|---|
| Selected Location | PM Palem |
| House Type | Individual House |
| Floor | Ground Floor |
| Bedrooms | 1 |
| Bathrooms | 1 |
| Total Dimensions | 200 Sq Yards |

Present price is 5845877.94 INR

At 2023,the Predicted price is 6308996.34 INR

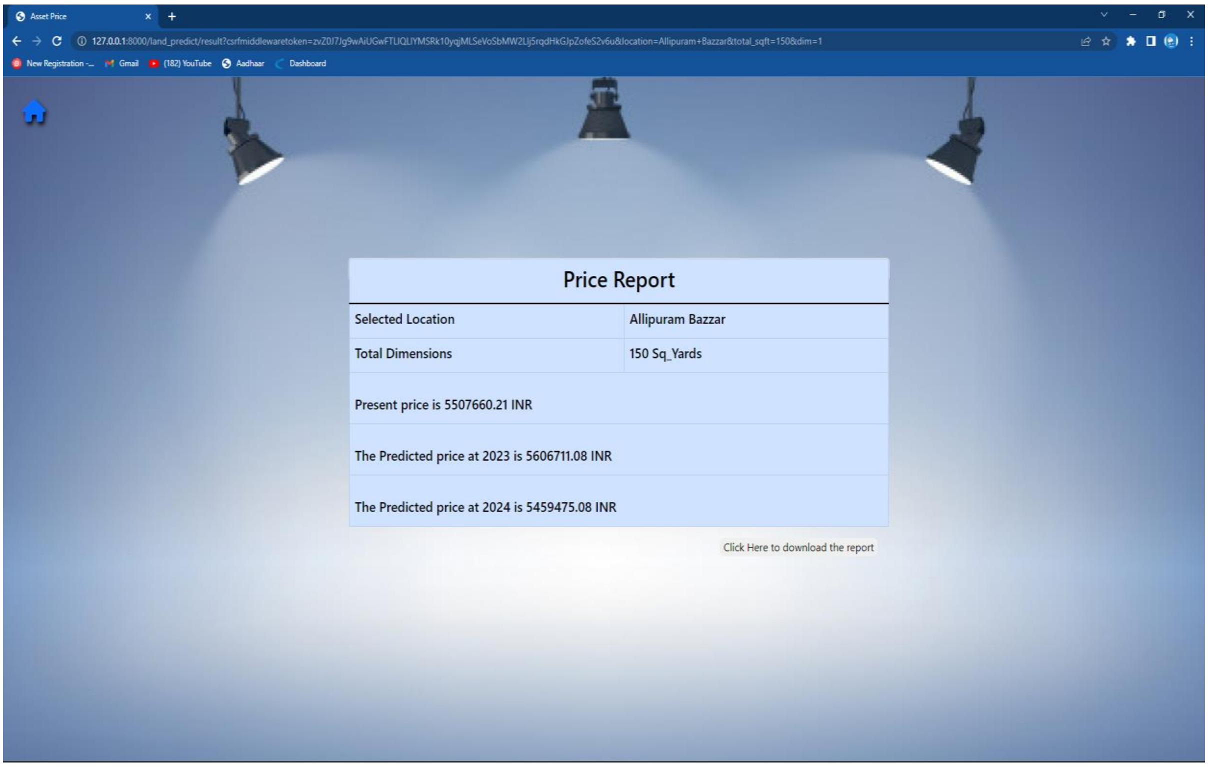
At 2024,the Predicted price is 6288603.95 INR

Click Here to download the report

The above screenshot describes the house price report page. After clicking predict price button on the house price prediction page, it will automatically redirect to this page.

It displays the values that are selected by the user on the house price prediction page and also displays the predicted price of the house for the future three years. If the user wants to download the price report, they need to click on the download button which is displayed at the bottom of the page. After clicking on the button, it downloads the house price report in the form of a pdf file.

## Screenshot for Land Price Report


Asset Price
127.0.0.1:8000/land_predict/result?csrfmiddlewaretoken=zvZ0J7Jg9wAiUGwFTLIQLIYMSRk10yqjMLSeVoSbMW2Llj5rqdHkGJpZofeS2v6u&location=Allipuram+Bazzar&total_sqft=150&dim=1
New Registration -...
Gmail
(182) YouTube
Aadhaar
Dashboard

Price Report

| Selected Location | Allipuram Bazzar |
|---|---|
| Total Dimensions | 150 Sq_Yards |

Present price is 5507660.21 INR

The Predicted price at 2023 is 5606711.08 INR

The Predicted price at 2024 is 5459475.08 INR

Click Here to download the report

The above screenshot describes the land price report page. After clicking predict price button on the land price prediction page, it will automatically redirect to this page.

It displays the values that are selected by the user on the land price prediction page and also displays the predicted price of land for the future three years. If the user wants to download the price report, they need to click on the download button which is displayed at the bottom of the page. After clicking on the button, it downloads the land price report in the form of a pdf file.

# 6. TESTING

## 6.1 Software Testing

This is a critical element of software quality assurance and represents the ultimate service of specification design and coding. The Increasing Visibility of software as a system element and the attended costs associated with the software failure and motivating forces for well-planned, thorough testing.

### Testing Objectives

The following are the testing objectives

- Testing is the process of executing a program with the intent of finding an error.
- A good test case has a high probability of finding an as-yet-undiscovered error.
- A successful test uncovers a yet undiscovered error.

The above objectives imply a dramatic change in viewpoint. They move counter to the commonly held view that a successful test is one in which no errors are found. Our Objective is to design tests that systematically different clauses of errors and do so with a minimum amount of time and effort. If testing is conducted successfully, it will uncover errors in the software. As a secondary benefit, testing demonstrates that software functions appear to be working according to specification and that performance requirements appear to have been met. In addition, data collected as testing is conducted provides a good indication of software. Testing can't show the absence of defects, it can only show that software errors are present. It is important to keep this state in mind as testing is being conducted.

### Testing principles

Before applying methods to design effective test cases, a software engineer must understand the basic principles that guide software testing.

- All tests should be traceable to customer requirements.
- Tests should be planned long before testing begins.
- Testing should begin "in the small" and progress towards testing "in the large
- Exhaustive testing is not possible.

## Testing Strategies

A strategy for software testing indicates software test case design methods into a well-planned series of steps that results in the successful construction of software. The strategy provides a road map that describes the steps to be conducted as part of testing when these steps are planned and then undertaken, and how much effort, time, and p resources will be required. It must be rigid enough to promote reasonable planning and management tracking as the project progresses.

## Types of Testing

The primary objective of test case design is to derive a set of tests that have the highest likelihood of uncovering errors in the software. To accomplish this objective two different categories of test case design techniques are used.

## White-Box Testing

White box testing sometimes called glass box testing is a test case design that focuses on the program control structure. Test cases are derived to ensure that

- Guarantee that all independent paths within a module have been exercised at least once.
- Exercise all logical designs on their true and false sides.
- Executes all loops at their boundaries and within their operational boundaries.
- Exercise internal data structure to ensure their validity. Several methods are used in the white box testing.

## Black Box Testing

Tests can be conducted at software interface by knowing the specified function that a product has been designed to perform, tests can be conducted that demonstrate each function is fully operational, at the same time searching for errors in called black-box testing, sometimes called behavioral testing. Black box testing is not an alternative to white-box techniques. Rather it is a complementary approach that is likely to uncover a different class of errors than the white-box method. Black box tests are designed to uncover errors in functional requirements without regard to the internal workings of a program. Black box testing techniques focus on the information domain of the software, Black Box testing attempts to find errors in the following categories.

- Incorrect or missing functions.
- Interface errors.
- Errors in the data structures or external database access.
- Initialization and termination errors.

## 6.2 Sample Test Cases

### Test case-1

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | The webpage should display all the fields (Textbox, Radio Button, dropdown, etc) and should be properly aligned. | Display all the fields properly. | same | pass |

**Risk:** The web page may not be displayed properly.

### Test case-2

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | Users need to enter all the mandatory fields displayed on the page. | All the data can be processed. | same | pass |

**Risk:** It displays an error message like” enter all the required fields”.

### Test case-3

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | Users need to enter the numerical values into the numeric field. | Takes numerical values from the keyboard. | same | pass |

**Risk:** The field should not accept the alphabet entered by the user from the keyboard.

## Test case-4

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | The page must be redirected to the other page, When the user clicks on the hyperlink button | The page redirects to the new page. | same | pass |

**Risk:** It displays the error message like "Page not found or not defined".

## Test case-5

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | The connection must be stable. | Webpage runs and loads all the data efficiently. | same | pass |

**Risk:** Otherwise, it displays the message like "Unable to connect to internet or no internet".

## Test case-6

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | Loading the dataset in the form of CSV. | Successfully loaded the dataset | same | pass |

**Risk:** It displays a message like "Invalid Format".

## Test case-7

| S.no | Description | Expected Output | Actual Output | Result |
|---|---|---|---|---|
| **1** | The system can display the price. | Displays the price on the page. | same | pass |

**Risk:** Otherwise, it takes more and redirects to the same page.

# 7. CONCLUSION

The project “Asset Price Prediction Using Machine Learning Techniques” has been developed to predict the price of a house and land that are located in the cities of Visakhapatnam. In this project, the website allows the user to know the price of the asset located in Visakhapatnam.

In this project, we used linear regression algorithm to predict the asset price according to the user requirement. The linear regression algorithm gives an accuracy score of an average of 75% compared with two other machine learning algorithms i.e., Random Forest and Decision Tree Algorithm. The main objective is to achieve a system that will reduce the human effort to predict the price of a house or land. The Asset price forecasting system predicts the price of an asset for the future three years.

# 8. FUTURE ENHANCEMENT

There are some improvements and additions which can be done. The first is the ability to increase and update the dataset on a regular basis. The user interface will be improved with animated CSS. Provides a download option for the user to download the price report of the asset and to add images of the asset for every location.  So, it becomes different from other prediction systems. It helps people to know the price of the asset in the future.

# 9. REFERENCES

1. Furia, Palak, and Anand Khandare."Real Estate Price Prediction Using Machine Learning Algorithm." E Conference on Data Science and Intelligent Computing. 2020.
2. Grinberg, Miguel. Developing web applications with python. " O'Reilly Media, Inc.", 2018.
3. Debanjan Banerjee, Suchibrota Dutta Predicting the housing price direction using machine learning techniques2017 IEEE International Conference on Power, Control, Signals, and Instrumentation Engineering,2017.
4. N N Ghosalkar, S N Dhage Real Estate Value Prediction Using Linear Regression2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), p.
5. Yu, H., and J. Wu. "Real estate price prediction with regression and classification CS 229 Autumn 2016 Project Final Report 1–5." (2016).
6. G. Naga Satish, Ch. V. Raghavendran, M.D.Sugnana Rao, Ch. Srinivasulu House Price Prediction Using Machine Learning International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-9, July 2019.
7. Introduction to Machine Learning with Python by Andreas Muller and Sarah Guido.
8. Pandas: http://pandas.pydata.org/
9. NumPy: http://www.numpy.org/
10. Scikit-learn: http://scikitlearn.org/
11. Python: https://www.python.org/
12. Matplotlib: https://matplotlib.org/
13. Dataset: https://www.kaggle.com/datasets/vikramvicky00/visakhapatnam-data-set