A Major Project-1

On

**"House Price Prediction System"**

SUBMITTED TO
**RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA**
(M.P.)

In Partial Fulfillment of the award of

**Bachelor of Technology in Computer Science and Engineering**

YEAR 2021-25

## Submitted By


**Divyansh Sharma**
**ENROLL. NO. : 0905CS211055**

**Hardik Kapoor**
**ENROLL. NO. : 0905CS211060**

**Harsh Chaturvedi**
**ENROLL. NO. : 0905CS211061**

**Harshit Soni**
**ENROLL. NO. : 0905CS211063**


## Under the Guidance of


**Ms. Archana Tomar**
**Assistant Professor**
**Department of Computer Science and Engineering**


INSTITUTE OF TECHNOLOGY & MANAGEMENT
GWALIOR • MP • INDIA
"THINK BIG... THINK BEYOND"

**INSTITUTE OF TECHNOLOGY & MANAGEMENT**
**GWALIOR (M.P.) – 474001**

# CANDIDATE DECLARATION

We **DIVYANSH SHARMA, HARDIK KAPOOR, HARSH CHATURVEDI and HARSHIT SONI** students of B.Tech. (Computer Science) VII semester **Roll No 0905CS211055, 0905CS211060, 0905CS211061 and 0905CS211063** hereby declare that the Project entitled **"HOUSE PRICE PREDICTION SYSTEM"** which is being submitted to Department of computer science & Engineering in ITM, Gwalior is our authentic work carried out in our VII semester.

We declare that our work has not been submitted in part or in full to any other university or institution for the award of any degree or diploma.

**DIVYANSH SHARMA (0905CS211055)**

**HARDIK KAPOOR (0905CS211060)**

**HARSH CHATURVEDI (0905CS211061)**

**HARSHIT SONI (0905CS211063)**

# CERTIFICATE

This is to certify that the project entitled **"HOUSE PRICE PREDICTION SYSTEM"** being submitted by **Divyansh Sharma (Enroll. No. 0905CS211055), Hardik Kapoor (Enroll. No. 0905CS211060), Harsh Chaturvedi (Enroll. No. 0905CS211061) and Harshit Soni (Enroll. No. 0905CS191143)** in partial fulfillment of the requirement for the award of B. Tech. degree in Computer Science & Engineering to Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M.P.) is a record of bona fide work done by them, under my guidance.

**HEAD OF DEPARTMENT:**

**Dr. Rishi Soni**

**Department of CSE**

**ITM, Gwalior (M.P)**

**PROJECT GUIDE:**

**Ms. Archana Tomar**

**Assistant Professor**

**Department of CSE**

**ITM, Gwalior (M.P)**

**EXTERNAL EXAMINER**

# ACKNOWLEDGEMENT

The most awaited moment of successful completion of an endeavor is always a result of persons involved implicitly or explicitly there in and it is impossible without the help and guidance of the people around us. The successful completion of this thesis report is a result of dedicated efforts put in by many people and this report would have been incomplete without giving due credit to them. This acknowledgement is indeed a small token of gratitude in recognition of their help.

We express our gratitude to **Dr. Meenakshi Mazumdar,** Director, Institute of Technology and Management, Gwalior for her valuable suggestions. We feel highly privileged to express our sincere thanks and deepest sense of gratitude to our guide **Ms. Archana Tomar** who has spared her precious time & guided us to present this dissertation. We sincerely acknowledge her for extending her valuable guidance, support for literature, critical review of report & above all the moral support she has provided us at all stages of this dissertation. We wish to thank our Head of the department **Dr. Rishi Soni**, for his support and encouragement. Our thanks are extended to all staff members and persons who directly or indirectly helped us in achieving our goals. We could not have completed this work without the guidance, confidence and encouragement of our parents. They deserve more credit than We can give for instilling in us a good work ethics and a desire to always learn more. Our thanks are also extended to all our friends who have kept our spirits high throughout this work.

**DIVYANSH SHARMA**
**Enroll. No. 0905CS211055**

**HARDIK KAPOOR**
**Enroll. No. 0905CS211060**

**HARSH CHATURVEDI**
**Enroll. No. 0905CS211061**

**HARSHIT SONI**
**Enroll. No. 0905CS211063**

# Table of Contents

# Chapter 1. Introduction

## 1.1 Aims, Objectives, Vision

The **House Price Prediction System** is a data-driven tool designed to estimate the prices of residential properties using advanced machine learning algorithms. The system leverages historical data, analyzes multiple factors that influence property prices, and provides predictions that are both accurate and insightful.

**Aims**:

- Provide an efficient and reliable mechanism for estimating house prices.
- Eliminate the dependency on traditional methods, which are often subjective and inaccurate.
- Create a scalable framework capable of adapting to additional data sources and regions.

**Objectives**:

1. **Accuracy**: Develop a predictive model capable of handling varied inputs while ensuring robust and precise outputs.
2. **Accessibility**: Design a user-friendly interface to cater to diverse user groups, including buyers, sellers, and real estate agents.
3. **Efficiency**: Reduce the time and resources required to assess property prices by automating the valuation process.

**Vision**:
The vision of this project is to empower stakeholders in the real estate market with actionable insights and data-backed predictions. By leveraging technology, the system aims to simplify the decision-making process for buyers, sellers, and investors, ultimately fostering transparency and trust in the property market.

# Chapter 2. Project / Problem Selection

Real estate is one of the most significant sectors in the economy, and accurate property valuation is critical for buyers and sellers. However, traditional methods for price estimation often involve manual assessments and subjective judgment, which can lead to discrepancies and errors.

In Bengaluru, one of India's fastest-growing cities, the housing market is particularly dynamic. Rapid urbanization, infrastructure development, and demand-supply fluctuations make it challenging to determine fair property prices. Current solutions lack localized insights and fail to incorporate recent trends, creating a need for a system that can:

- Analyze historical data.
- Account for key property features like location, size, and amenities.
- Provide predictions in real-time.

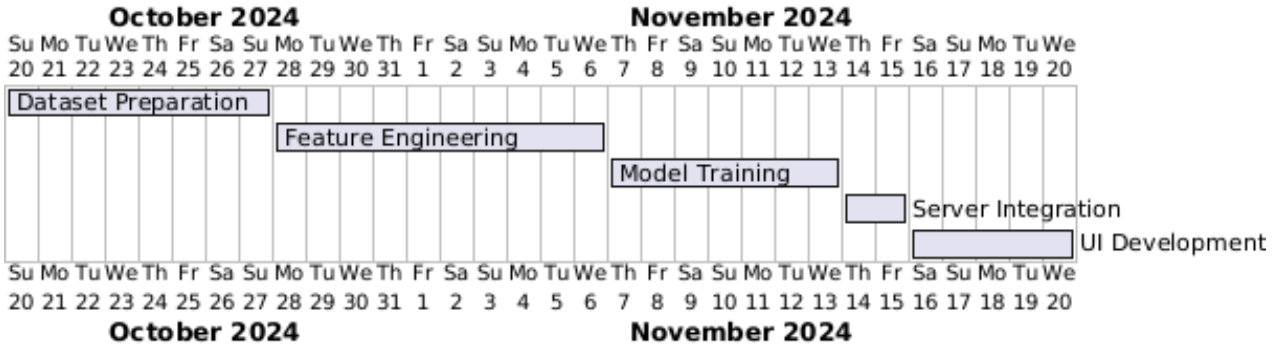
This project addresses these challenges by developing a data-driven approach to house price prediction. The proposed solution combines machine learning with a simple, intuitive user interface, ensuring accuracy and accessibility for all stakeholders.

# Chapter 3. Project Monitoring System

## 3.1 Gantt Chart

The project was executed in distinct phases to ensure proper tracking and timely delivery:

| Phase | Duration | Description |
|---|---|---|
| Data Collection | Week 1 | Acquired the Bengaluru House Dataset. |
| Data Cleaning and Analysis | Week 2 | Cleaned, analyzed, and visualized the data. |
| Model Training | Week 3-4 | Selected features and trained multiple ML models. |
| Server Integration | Week 5 | Deployed the model as an API using Flask. |
| Front-End Development | Week 6 | Built the UI using HTML, CSS, and JavaScript. |

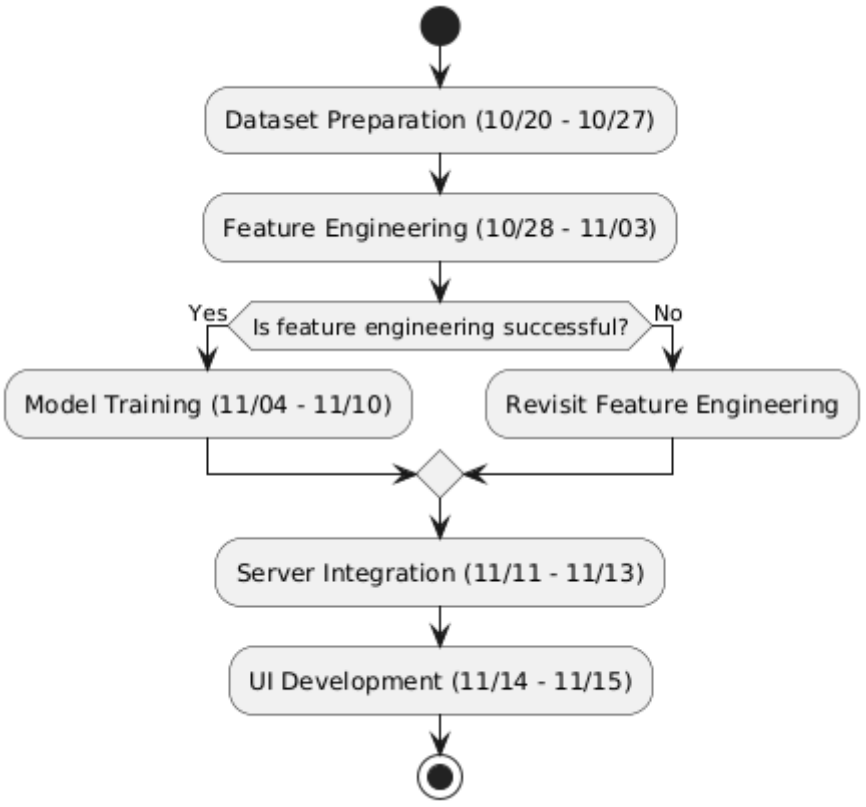



Each phase was tracked using project management tools, ensuring smooth collaboration and milestone completion.

### 3.2 PERT Chart

The PERT chart highlights the dependencies and timeline for key activities:

1. **Dataset Preparation** → 2. **Feature Engineering** → 3. **Model Selection**
2. **Model Training** → 5. **Integration with Server** → 6. **UI Development**




**Critical Path:** Data preparation and model training consumed the most time due to extensive experimentation and evaluation of multiple algorithms.

# Chapter 4. System Study

## 4.1 Existing System along with Limitations/Deficiencies

Traditional house valuation methods rely on manual inspection, local expertise, and subjective judgment, which can often lead to:

- **Inaccuracies**: Factors such as market demand or neighborhood trends may be overlooked.
- **Inefficiency**: Manual assessments take significant time and effort.
- **Lack of Transparency**: Buyers and sellers may not fully understand the basis of valuations.

These systems often fail to leverage data effectively, missing the opportunity to generate insights that can enhance decision-making.

## 4.2 Proposed System along with Intended Objectives

The proposed system addresses these limitations by:

- **Automating Predictions**: Utilizing machine learning models trained on historical data.
- **Ensuring Scalability**: Allowing seamless integration of additional features like real-time market data.
- **User-Friendly Design**: Offering an intuitive web-based interface accessible to non-technical users.

## 4.3 Feasibility Study

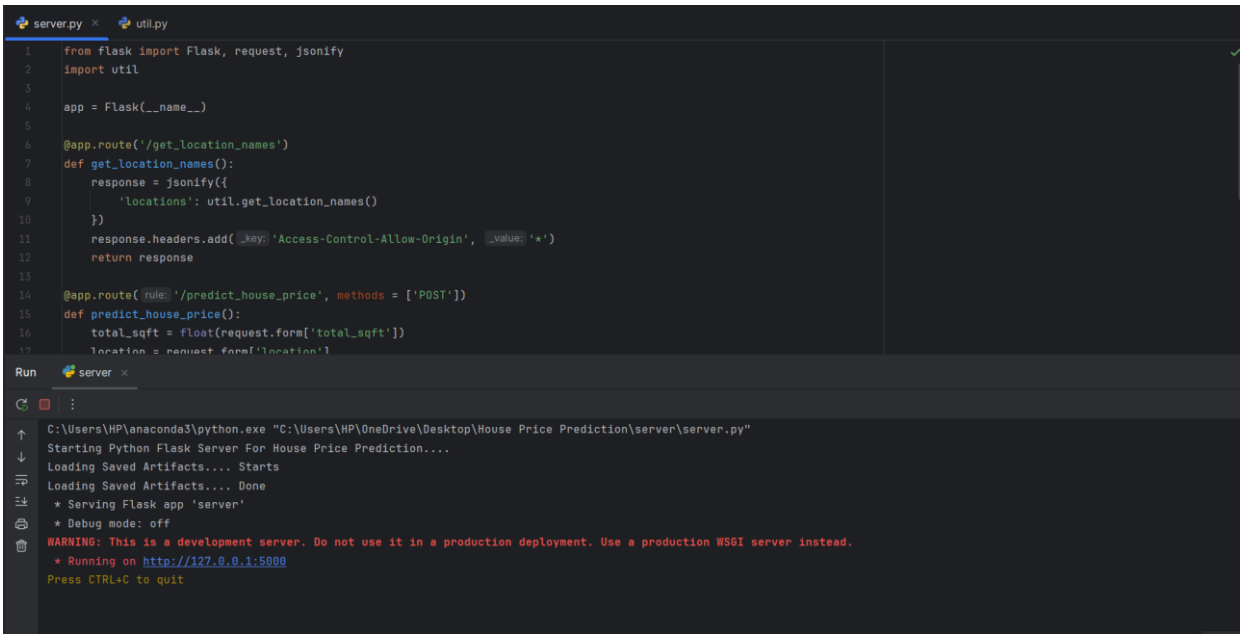
1. **Operational Feasibility**:
   The system is easy to use, requiring minimal technical expertise. Users simply input property details to receive instant predictions.
2. **Technical Feasibility**:
   Built using Python, Flask, and Scikit-learn, the system is lightweight and compatible with various platforms.
3. **Economic Feasibility**:
   Development costs are minimal, relying on open-source tools, making the solution affordable for small businesses and individual users.

## 4.4 Screenshots

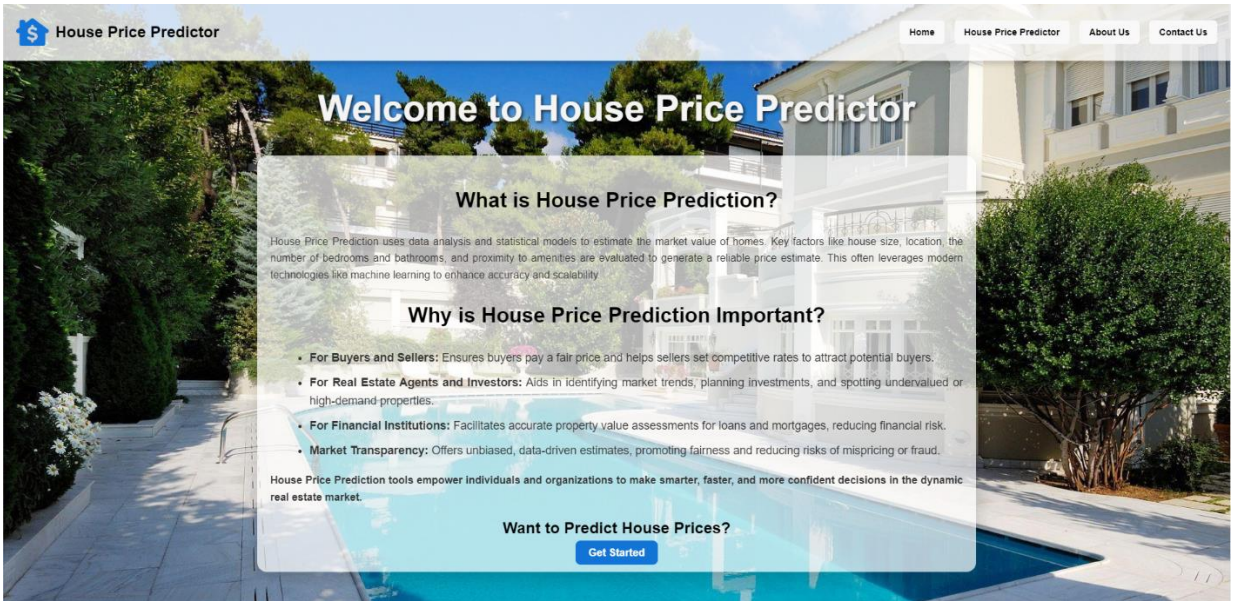
To enhance the explanation of the proposed system, include the following screenshots in this chapter:

**Screenshots of the System:**

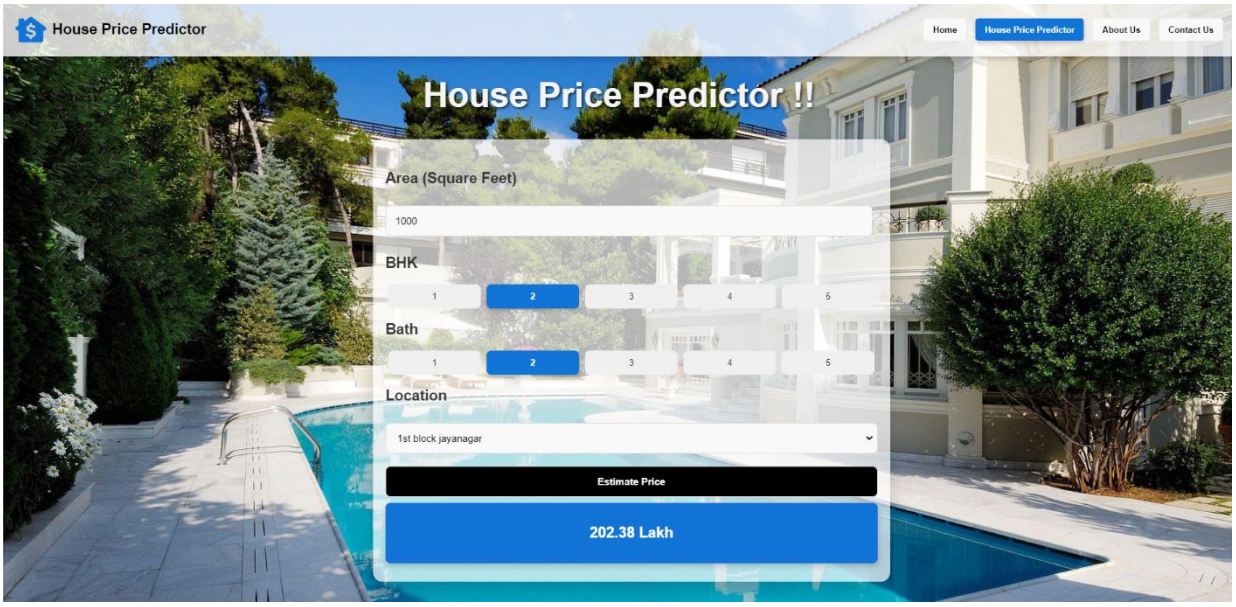
1. **Server Running:**
   Screenshot showing the server in the command-line interface or terminal, indicating the successful deployment of the backend system.



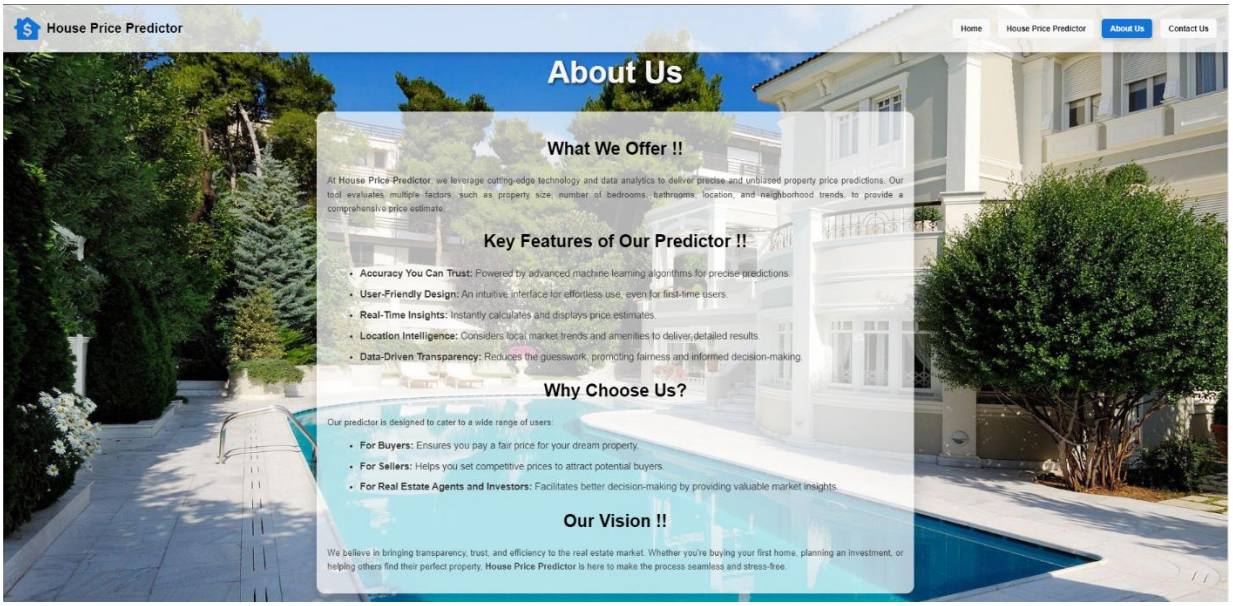

2. **Home Page:**
   Screenshot of the main page, highlighting the navigation bar, welcome message, and the overview of the house price prediction system.



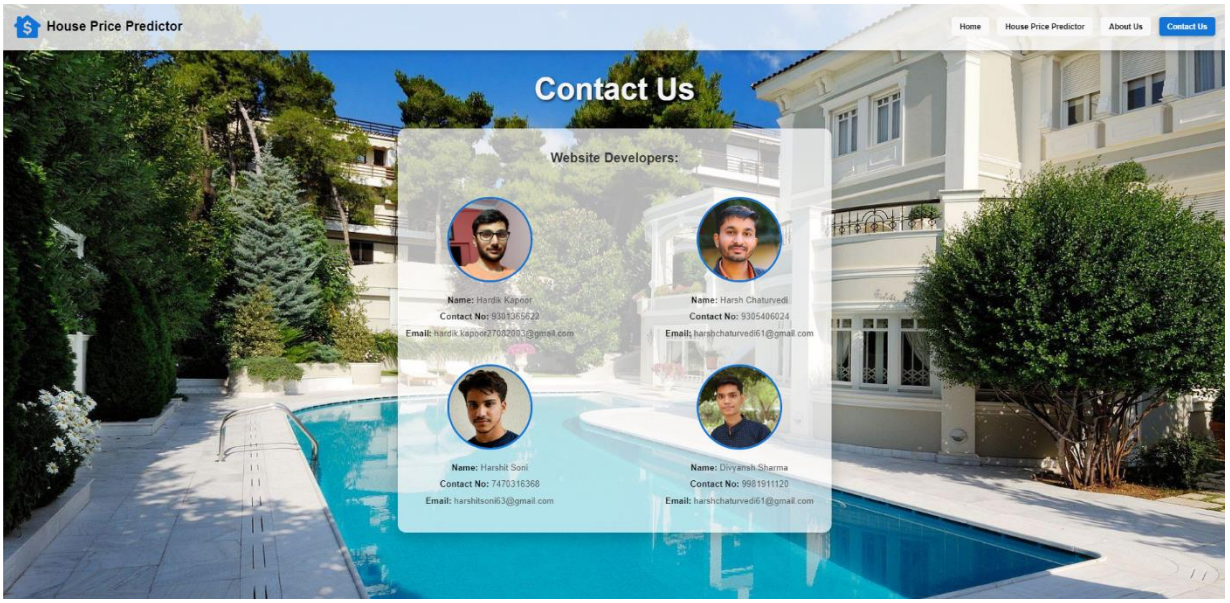
3. **House Price Predictor Page:**
   Screenshot showing the form where users can input property details (e.g., area, BHK, bathrooms, location) and see the price prediction.




4. **About Us Page:**
   Screenshot highlighting the "What We Offer," "Key Features," and "Our Vision" sections, which provide an overview of the system's capabilities and goals.



5. **Contact Us Page:**
   Screenshot displaying the details of the development team, including names, contact information, and profile pictures.



# Chapter 5. System Analysis

## 5.1 Requirement Specification

**Hardware**:

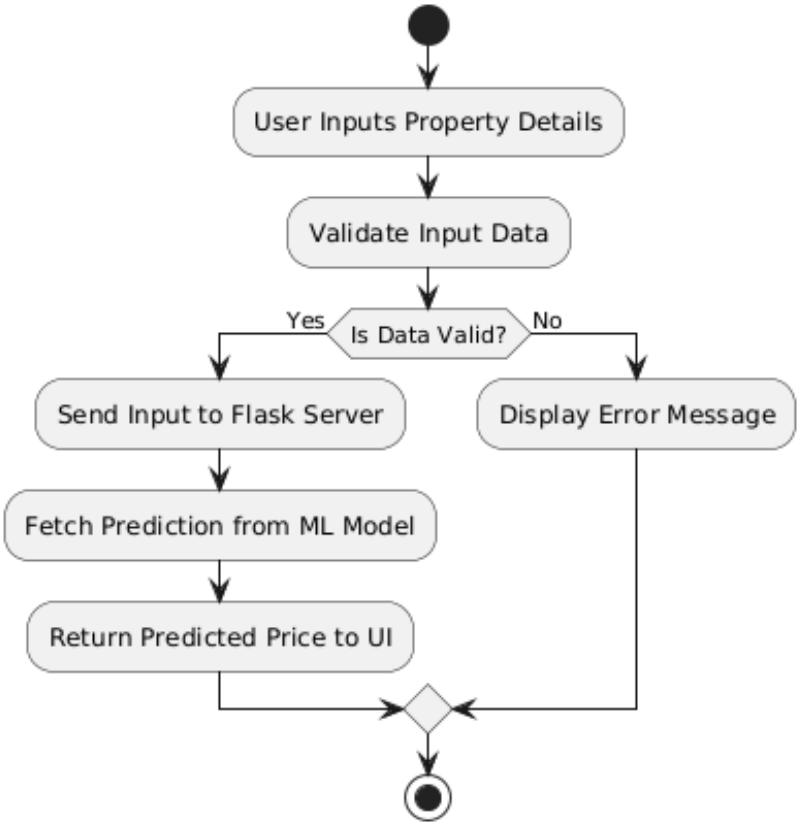
- **Server:** 4 GB RAM, dual-core processor for hosting the Flask application.
- **Client Devices:** Any device with a modern web browser.

**Software**:

- **Python Libraries:** Pandas, Numpy, Scikit-learn, Matplotlib, Flask.
- **UI Tools:** HTML, CSS, JavaScript.
- **IDEs or Code Editor:** PyCharm Community Edition, Visual Studio Code
- **Python Notebook:** Jupyter Notebook

## 5.2 System Flowcharts

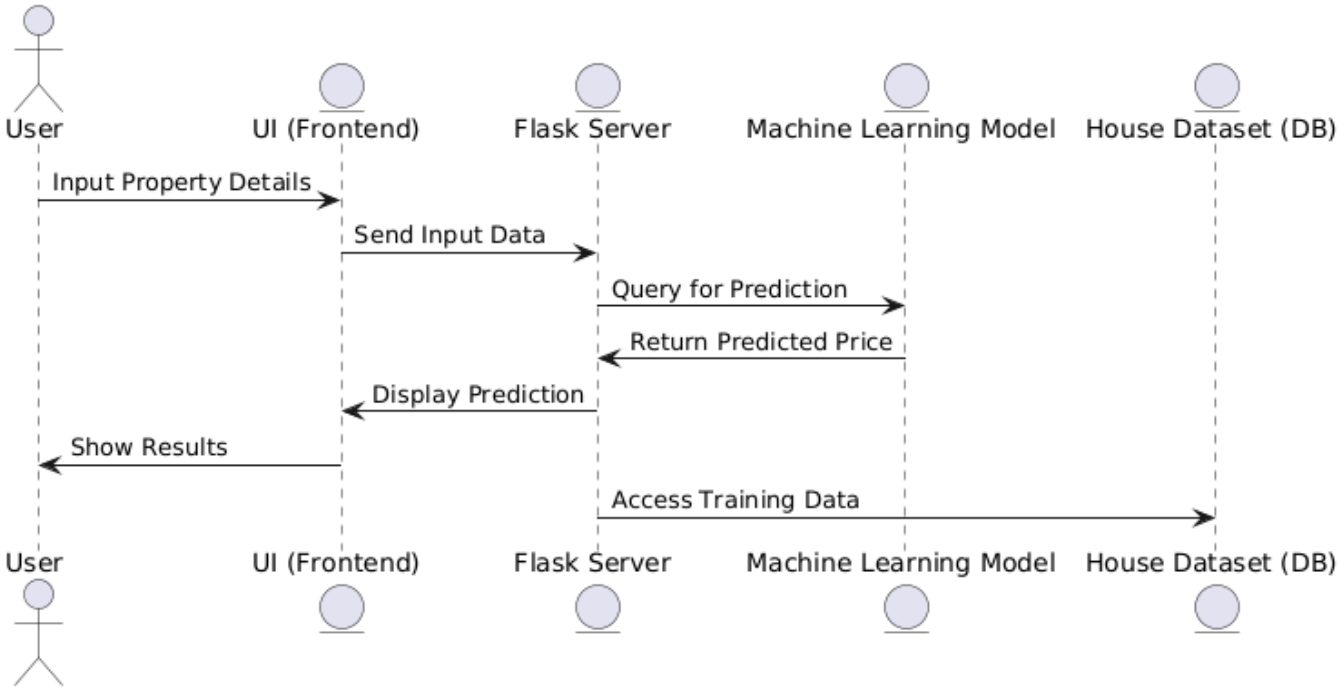
The user submits property details via the UI, which passes the input to the Flask server. The server interacts with the machine learning model to generate predictions and returns the results to the user interface.



### 5.3 Data Flow Diagram (DFD)

**Level 1**:

1. **Input**: User inputs property features.
2. **Processing**: Flask server sends input to the machine learning model.
3. **Output**: Predicted price is displayed on the UI.




### 5.4 Sources of Data

The project uses a dataset containing property details such as size, location, price, and amenities. This dataset was preprocessed to remove outliers and missing values, ensuring clean and reliable input for model training.

**Source:** **https://www.kaggle.com/datasets/amitabhajoy/bengaluru-house-price-data**

# Chapter 6. Scope of the Project

The **House Price Predictor System** is designed with scalability and adaptability in mind, enabling continuous growth and integration of advanced features. The key aspects of its scope include:

1. **Geographical Expansion**
    - Incorporate additional cities and regions with unique datasets, creating a globally accessible valuation tool.
    - Adapt prediction models to account for regional factors like climate impact and zoning laws.
2. **Enhanced Prediction Models**
    - Upgrade to advanced techniques like **deep learning** and **ensemble models** for improved accuracy.
    - Utilize **natural language processing (NLP)** to analyze user feedback and market sentiments for deeper insights.
3. **Feature Integration**
    - Add APIs for real-time market trends and neighborhood data, such as proximity to schools, hospitals, and amenities.
    - Provide comparative analysis tools, allowing users to benchmark properties against similar listings.
4. **User-Centric Features**
    - Develop mobile and web apps with personalized dashboards for tracking properties and market alerts.
    - Add visualization tools like heatmaps and graphs to make comparisons more intuitive.
5. **Integration with Real Estate Platforms**
    - Collaborate with real estate websites to integrate predictive tools directly into property listings.
    - Offer plugins or APIs for seamless adoption by industry platforms.
6. **Data Privacy and Security**
    - Ensure compliance with regulations like GDPR to safeguard user data.
    - Provide transparency by allowing users to understand how predictions are derived.

# Bibliography & Web References


1) **Research Paper**:

Zhang, Y., & Wang, D. (2020). *A Machine Learning Approach to Predicting Housing Prices*. *Journal of Artificial Intelligence*, 15(3), 145-162.

- This paper explains how machine learning models can be effectively used to predict house prices, which inspired the use of regression techniques in this project.

2) **Web References:**

**a) YouTube Playlist**:

"Machine Learning Project for House Price Prediction". *YouTube*. (https://youtube.com/playlist?list=PLeo1K3hjS3uu7clOTtwsp94PcHbzqpAdg&feature=shared)

- This playlist provided step-by-step video tutorials on how to build the house price prediction project, covering essential topics like data preprocessing, model building, and deployment.

**b) Dataset Source**:

Kaggle, Bengaluru House Price Data (https://www.kaggle.com/datasets/amitabhajoy/bengaluru-house-price-data)

- The dataset used in this project was sourced from Kaggle's Bengaluru House Price Data, which contains real-world housing data for predictive analysis and is crucial for training and testing the prediction model.

3) **GitHub Repositories:**

**Project Repository**:

"House Price Prediction Project". *GitHub*. (https://github.com/HardikKapoor27/House-Price-Prediction-System)

- The code and resources for the House Price Prediction project are hosted on GitHub, and you can access the implementation, documentation, and deployment setup here.