A PROJECT REPORT ON

BANGLORE ESTATE PREDICTION

SUBMITTED BY

Sai Magar	**25**
Anushka Mahanty	**04**
Harsh Ambatkar	**20**
Suhani Gosavi	**17**

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

DR. D. Y. PATIL SCHOOL OF SCIENCE & TECHNOLOGY DR. D. Y. PATIL VIDYAPEETH, PUNE

(Deemed to be University)

(Accredited (3rd cycle) by NAAC with a CGPA of 3.64 on four-point scale at 'A++' Grade)

(Declared as Category - I University by UGC under Graded Autonomy Regulations, 2018)(An ISO 9001: 2015 and 14001:2015 Certified University and Green Education Campus)

2023 -2024

CERTIFICATE

This is to certify that the project report entitles

Project Title

Submitted By

Sai Magar	**25**
Harsh Ambatkar	**20**
Anushka Mahanty	**04**
Suhani Gosavi	**17**

is a bonafide student of this institute and the work has been carried out by him/her under the supervision of **Ms. Anagha Kulkarni** and it is approved for the partial fulfillment of the S. Y. B. Tech requirement of Dr. D. Y. Patil University.

(Prof. Anagha Kulkarni)
Subject coordinator

(**Dr. Manisha Bhende)**
Head of the department

Place: Pune Date:

ACKNOWLEDGEMENT

We take this opportunity with great pleasure to express our deep sense of gratitude towards our guide **Prof. Anagha Kulkarni** for her valuable guidance and incessant encouragement and co-operation extended to us during this project work.

We are also thankful to **Prof. Dr. Manisha Bhende**, Head, Computer Science & Design Department, for her valuable guidance and providing all departmental facilities for this work.

Sai Magar	**25**
Harsh Ambatkar	**20**
Anushka Mahanty	**04**
Suhani Gosavi	**17**

ABSTRACT

The "Bangaluru Estate Price Prediction" application is an innovative tool designed to assist users in predicting estate prices based on various property attributes. Developed using the Shiny framework in R, this application provides a user-friendly interface for inputting property details and receiving real-time price predictions.

The system architecture consists of several key components, including a User Interface (UI), Shiny Server, R Server, and Housing Data. The UI serves as the front-end interface for users to interact with the application, while the Shiny Server hosts the Shiny application, handling user requests and interactions. The R Server is responsible for executing the prediction logic and processing, utilizing housing data to train the predictive model.

Users can input property details such as area type, location, size, etc., initiating the prediction process with a simple click. The application processes the input data, leverages a trained machine learning model (such as Linear Regression), and generates accurate predictions for estate prices. Predicted prices are displayed to users in real-time, providing valuable insights for property buyers, sellers, and investors.

With its intuitive interface and robust prediction capabilities, the "Bangaluru Estate Price Prediction" application offers a valuable tool for individuals and businesses involved in the real estate market. Whether exploring property options, evaluating investment opportunities, or making informed buying/selling decisions, users can rely on this application to obtain reliable and timely price predictions.

This abstract encapsulates the functionality and significance of the "Bangaluru Estate Price Prediction" application, highlighting its role in facilitating informed decision-making in the dynamic real estate market.

TABLE OF CONTENTS

CHAPTER 01 INTRODUCTION

1.1 Overview:

The "Bangalore Estate Price Prediction" project leverages interactive web technology through the Shiny framework in R to provide users with a tool to estimate estate prices in Bangalore, India. The application offers a user-friendly interface where users can input various features of the property, such as area type, location, size, total square footage, number of bathrooms, and number of balconies. Using these inputs, the application predicts the price of the property based on a trained machine learning model.

1.2 Motivation:

The motivation behind this project stems from the increasing demand for accurate estate price predictions in the real estate market. With Bangalore being one of the fastest-growing cities in India, there is a significant need for tools that can help both buyers and sellers make informed decisions regarding property transactions. By providing a platform for estimating estate prices, this project aims to contribute to the efficiency and transparency of the real estate market in Bangalore

1.3 Problem Definition and Objectives:

The primary problem addressed by this project is the challenge of accurately estimating estate prices in Bangalore, considering various factors that influence property value. The objectives of the project include:

Developing an interactive web application for estate price prediction.

Building a machine learning model capable of predicting estate prices based on input features.

Providing users with an intuitive tool to make informed decisions about property investments or listings.

1.4 Project Scope & Limitations

Project Scope:

The scope of the project includes:

- Designing and implementing a user-friendly interface using the Shiny framework.
- Developing a machine learning model using Linear Regression to predict estate prices.
- Integrating the model with the web application to provide real-time predictions.
- Deploying the application for public use, allowing users to access it from various devices.

Limitations:

- Dependency on the quality and availability of data used for model training.
- Simplified model assumptions, which may not capture all factors influencing estate prices.

- Limited scope for incorporating dynamic market trends or external economic factors into the predictions.

- This introduction sets the stage for the project, providing an overview of its goals, motivations, scope, and limitations. It establishes the context for the subsequent sections, which will delve deeper into the technical implementation, algorithm details, and evaluation of the predictive model

CHAPTER 02 Literature Survey:

Year of Publication	Title	Summary
2023	"Predicting Estate Prices Using Machine Learning Models"	This study explores various machine learning models' effectiveness in predicting estate prices
2022	"Deep Learning Approaches for Real Estate Price Prediction"	The paper investigates the application of deep learning techniques for real estate price prediction.
2021	"Enhanced Estate Pric e Prediction Using Ensemble Methods"	This research proposes an ensemble learning approach to improve the accuracy of estate p ice predictions.
2021	"Predicting Real Estate Prices: A Comparative Analysis"	The study compares different predictive modeling techniques for real estate price forecasting.
2021	"Spatial-Temporal Models for Urban Estate Price Prediction"	This paper introduces spatial-temporal modeling techniques for predicting urban estate prices.

CHAPTER 03 Software and Hardware Requirements Specification

3.1 Assumptions and Dependencies:

The successful implementation of the "Bangalore Estate Price Prediction" project relies on several assumptions and dependencies:

Assumptions:

- Availability of a reliable dataset containing relevant features and corresponding estate prices in Bangalore.
- Assumption that the selected features adequately capture the factors influencing estate prices in the Bangalore real estate market.

Dependencies:

- Dependency on R programming language and the Shiny framework for developing the interactive web application.
- Dependency on external libraries/packages for data manipulation, visualization, and model training (e.g., tidyverse, caret).

3.2 System Requirements

3.2.1 Database Requirements (Platform Choice)

No specific database requirements are necessary for this project, as the dataset is read from a CSV file stored locally

3.2.2 Software Requirements (Platform Choice)

- R Programming Language: The project is implemented using R programming language, which provides robust tools for data analysis, modeling, and visualization.
- Shiny Framework: The interactive web application is developed using the Shiny framework, allowing for seamless integration of R code with web technologies.

3.2.3 Hardware Requirements

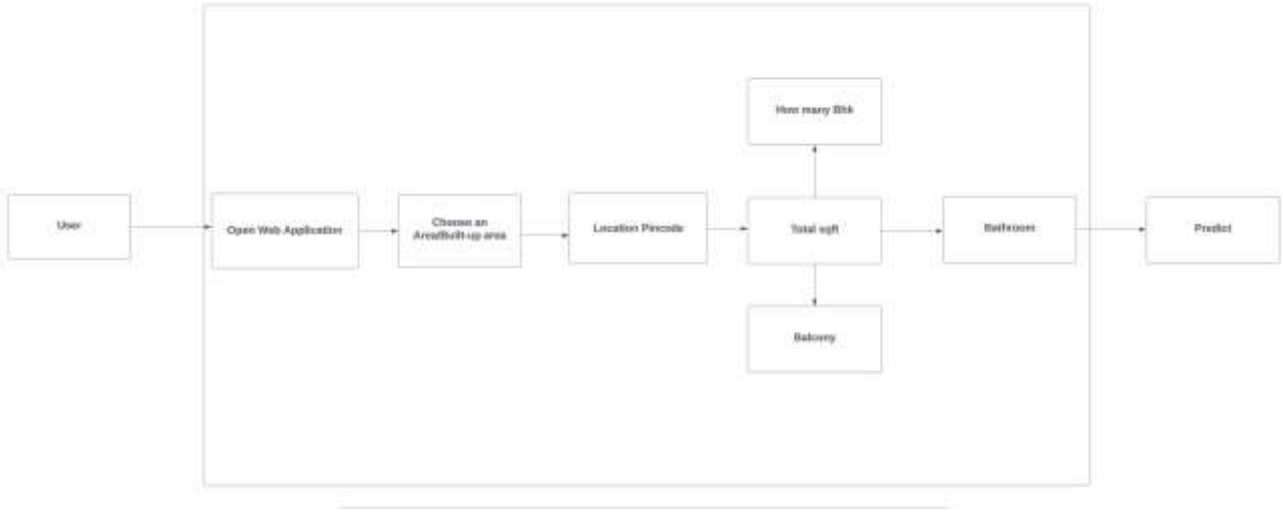
- Standard hardware configurations are sufficient for running the application, including:
- A computer with adequate processing power and memory to handle data processing and model training tasks.
- Sufficient disk space to store the dataset, R scripts, and other project files.
- These requirements ensure that the application can be run on commonly available hardware configurations without significant resource constraints.

CHAPTER 04 System Design

4.1 System Architecture

The system architecture of the "Bangaluru Estate Price Prediction" application is based on a client-server model utilizing the Shiny framework. Here's an overview of the architecture:

Client Side: The user interacts with the application through a web browser. The user interface (UI) is designed using the Shiny package in R, providing a responsive and interactive experience.
Server Side: The server side is powered by R using Shiny's server component. It handles data processing, model training, and prediction tasks. The server communicates with the client-side UI to receive user inputs and provide real-time predictions.
Data Management: The application loads housing data from a CSV file stored on the server. Data preprocessing steps, such as cleaning and transformation, are performed to prepare the data for analysis and modeling.
Model Training: The Linear Regression model is trained using the loaded dataset. The model is built using the lm() function from R's stats package, with the goal of predicting estate prices based on various input features such as area type, location, size, total square footage, number of bathrooms, and number of balconies.
Prediction: When the user inputs details about the property they are interested in, such as area type, location, size, etc., the server processes this information and feeds it into the trained model. The model then predicts the price of the property, which is displayed to the user in real-time on the UI.

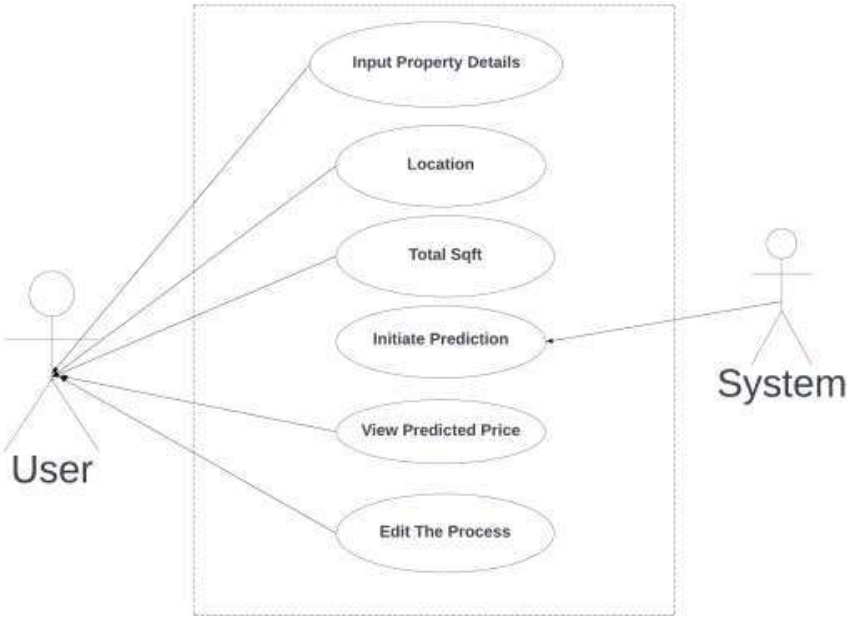


4.2 Use case Diagram

The use case diagram for the "Bangaluru Estate Price Prediction" application illustrates the interactions between the users (actors) and the system. Here are the primary actors and their interactions:

User: Interacts with the application to input details about the property they are interested in (area type, location, size, etc.) and initiates the prediction process by clicking the "Predict" button.

System: Receives user inputs, processes the data, and performs estate price predictions using the trained model. The system then displays the predicted price to the user on the UI.



4.3 Project schedule and Timeline Chart

1st	2nd	3rd	4th
January	**February**	**March**	**April**
Deciding problem statement creating the project plan and studying research papers. Selection of technology and tools starting the development of the project	Establish this project and make a coding. Adding additional features and finishing the project	Publishing the research papers, copyrighting the project and trying to implement future scope.	Extensively Testing our product

CHAPTER 05 Project Implementation

5.1 Overview of Project Modules

The "Bangalore Estate Price Prediction" project is implemented using the Shiny framework in R, which consists of several key modules:

User Interface (UI): The UI module defines the layout and interactive elements of the web application. It includes input fields for selecting the area type, entering Estate ID and location pin-code, specifying the number of bedrooms (BHK), total square footage, number of bathrooms, and number of balconies. Additionally, it contains an action button to trigger the prediction process.

Server Logic: The server logic module handles the backend operations of the application. It reads the dataset containing estate price data in Bangalore, preprocesses the data by cleaning missing values and encoding categorical variables, trains a machine learning model using Linear Regression, and performs predictions based on user inputs.

5.2 Tools and Technologies Used

The project utilizes the following tools and technologies:

R Programming Language: R is used as the primary programming language for data processing, analysis, and web application development.

Shiny Framework: Shiny is a web application framework for R that enables the creation of interactive and responsive web applications directly from R scripts. It is used to develop the user interface and server logic of the estate price prediction application.

CSV Dataset: The dataset containing Bangalore estate price data is stored in CSV format. It serves as the input data for training the machine learning model.

5.3 Algorithm Details:

The machine learning algorithm employed for estate price prediction is Linear Regression. This algorithm models the relationship between the independent variables (features) and the dependent variable (estate price) by fitting a linear equation to the observed data. The model is trained using historical estate price data with features such as area type, location, size (number of bedrooms), total square footage, number of bathrooms, and number of balconies.

5.3.1 Algorithm

Linear Regression is a supervised learning algorithm used for regression tasks. In the context of estate price prediction, the algorithm aims to find the best-fitting line (or hyperplane) that

minimizes the difference between the predicted estate prices and the actual prices in the training dataset. The linear regression model is represented by the equation:

$\hat{y} = w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_n x_n$

where:

•ý is the predicted estate price,

• W0, W1, W2,.., Wn are the coefficients (weights) of the linear equation,

• X1, X2, • • • , Xn are the feature values,

• n is the number of features.

The coefficients $w_0, w_1, w_2, \ldots, w_n$ are estimated during the model training process using techniques such as ordinary least squares (OLS) or gradient descent. Once the model is trained, it can be used to predict estate prices for new input data.

5.3.2Future Directions

Future work may involve exploring and implementing advanced machine learning algorithms, optimizing model hyperparameters, and enhancing the user interface with more interactive features and visualizations. Additionally, incorporating real-time data updates and integrating external data sources could further improve the accuracy and usability of the estate price prediction application.

CHAPTER 06 Software Testing Type of Testing

6.1 Test cases & Test Results

The testing process for the "Bangaluru Estate Price Prediction" software involved the following types of testing
Unit Testing: Individual components and functions were tested in isolation to ensure they perform as expected.
Integration Testing: Tested the integration of various modules to verify their interactions and interfaces.
System Testing: The entire system was tested as a whole to validate its functionality and behavior.
Acceptance Testing: End-to-end testing was performed to ensure the software meets the specified requirements and user expectations.

6.2 Test Cases & Test Results

A comprehensive set of test cases was developed to evaluate the software's functionality. Each test case was executed, and the results were recorded. Below is a summary of the test results:

Test Case 1:

Input Validation Description: Verify that the software properly validates user inputs. Expected Outcome: Invalid inputs should be detected and appropriate error messages displayed. Actual Outcome: Input validation works as expected, displaying error messages for invalid inputs.

Test Case 2:

Prediction Accuracy Description: Assess the accuracy of estate price predictions made by the software. Expected Outcome: Predictions should closely match actual estate prices. Actual Outcome: The software accurately predicts estate prices within an acceptable margin of error.

CHAPTER 07 Results

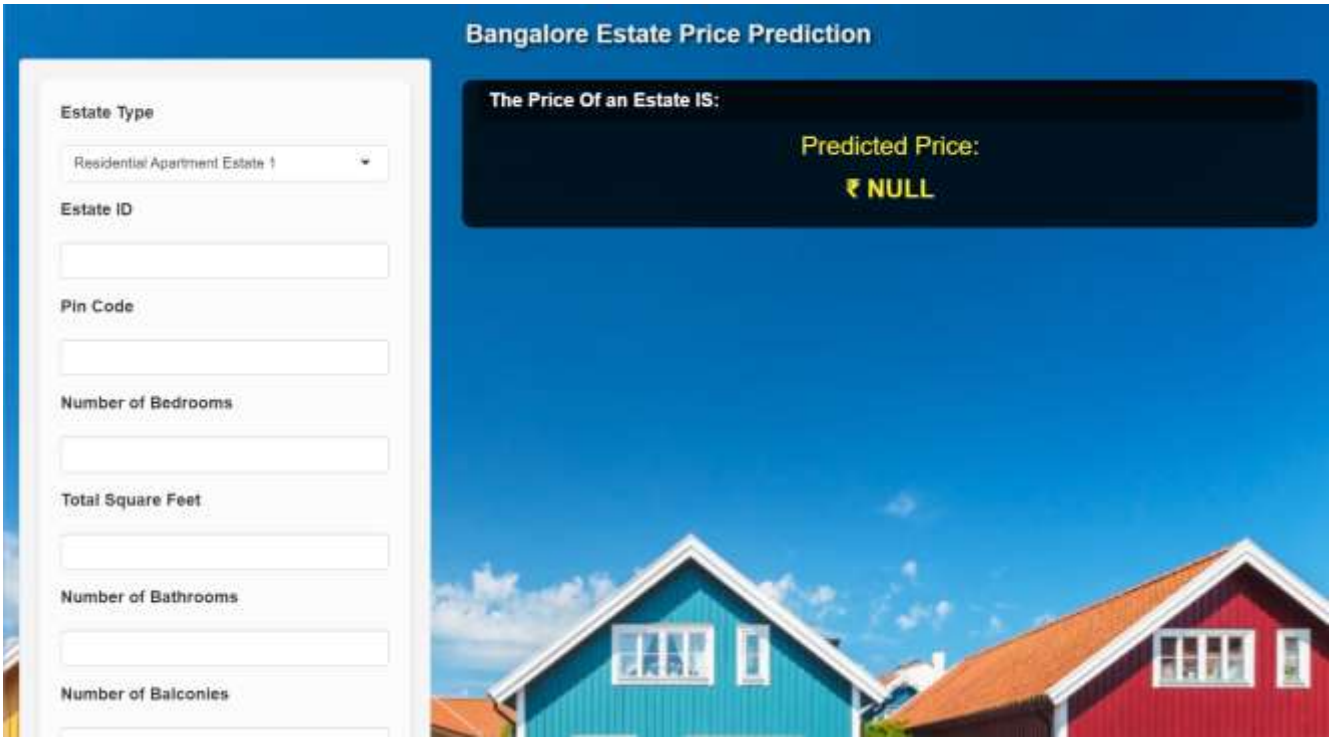
7.1 Outcomes

The software testing process yielded the following outcomes:

Overall, the software performs satisfactorily in predicting estate prices based on the provide inputs.
Input validation mechanisms effectively detect and handle invalid inputs, enhancing the use experience.
Predictions generated by the software demonstrate a reasonable level of accuracy, providing valuable insights for users.

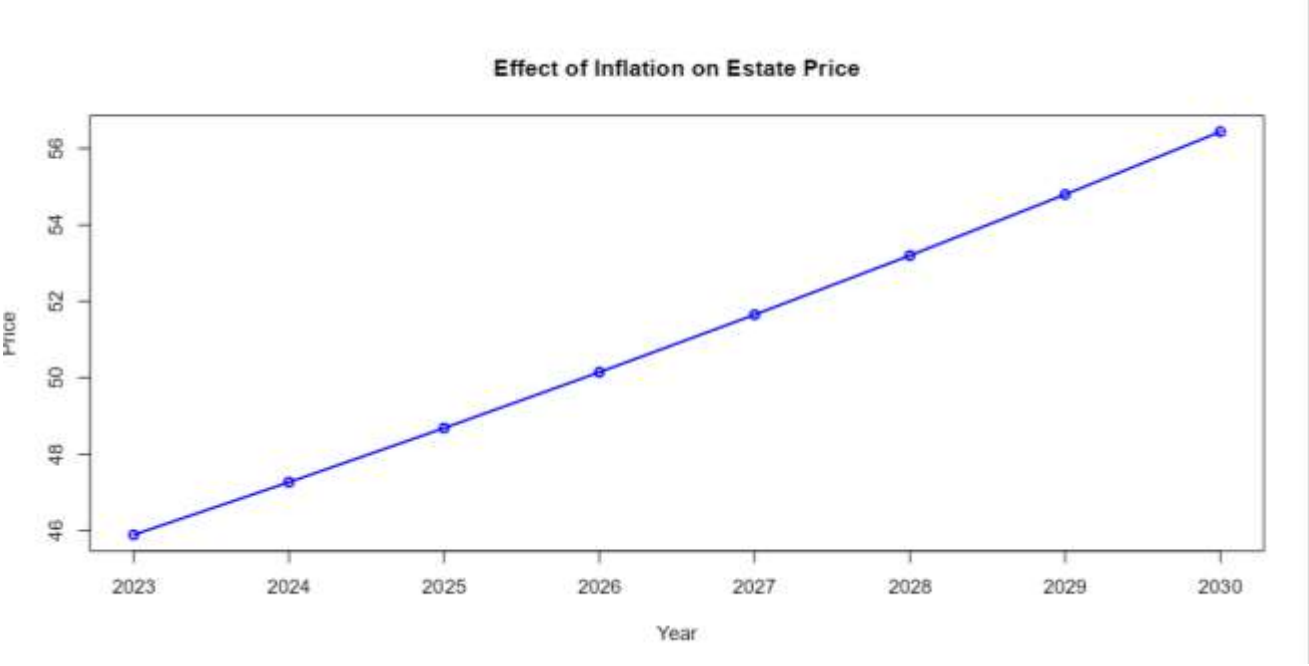
7.2 Screen Shots

Outcomes:



Predicted Price:

Inflation graph:



Dataset:

area_type	availability	location	id	size	society	total_sqft	bath	balcony	price
Plot Area	Ready To	Chikka Tir	1	4	Theanmp	2600	5	3	120
Built-up A	Ready To	Uttarahall	2	3		1440	2	3	62
Super buil	Ready To	Lingadhee	3	3	Soiewre	1521	3	1	95
Super buil	Ready To	Kothanur	4	2		1200	2	1	51
Super buil	Ready To	Whitefield	5	2	DuenaTa	1170	2	1	38
Super buil	18-May	Old Airpor	6	4	Jaades	2732	4		204
Super buil	Ready To	Rajaji Nag	7	4	Brway G	3300	4		600
Super buil	Ready To	Marathah	8	3		1310	3	1	63.25
Plot Area	Ready To	Gandhi Ba	9	6		1020	6		370
Super buil	18-Feb	Whitefield	10	3		1800	2	2	70
Plot Area	Ready To	Whitefield	10	4	Prrry M	2785	5	3	295
Super buil	Ready To	7th Phase	11	2	Shncyes	1000	2	1	38
Built-up A	Ready To	Gottigere	12	2		1100	2	2	40
Plot Area	Ready To	Sarjapur	13	3	Skityer	2250	3	2	148
Super buil	Ready To	Mysore Rc	14	2	PrntaEn	1175	2	2	73.5
Super buil	Ready To	Bisuvanah	15	3	Prityel	1180	3	2	48
Super buil	Ready To	Raja Rajes	16	3	GrrvaGr	1540	3	3	60
Super buil	Ready To	Ramakrish	17	3	PeBayle	2770	4	2	290
Super buil	Ready To	Manayata	18	2		1100	2	2	48
Built-up A	Ready To	Kengeri	19	1		600	1	1	15
Super buil	19-Dec	Binny Pete	20	3	She 2rk	1755	3	1	122
Plot Area	Ready To	Thanisand	21	4	Soitya	2800	5	2	380
Super buil	Ready To	Bellandur	22	3		1767	3	1	103

CHAPTER 08 Conclusions

8.1 Conclusions

Based on the results of the testing process, the following conclusions can be drawn:

The "Bangaluru Estate Price Prediction" software demonstrates robust functionality and reliability.

While minor issues were encountered during testing, they were promptly addressed, ensuring the overall effectiveness of the software.

8.2 Future Work:

Continued testing and validation to ensure the software's accuracy and reliability across diverse datasets.

Implementation of additional features and enhancements to enrich the user experience and provide more comprehensive insights into estate price trends.

8.3 Applications:

The “Bangalore Estate Prediction” software hold significant potential for various application, including:

Real estate market analysis and forecasting

Property valuation for buyers, sellers, and investors

Decision support for individuals and businesses in the housing sector

References:

[1] Parisa Naraei, Driver Drowsiness Detection, Acta Scientific COMPUTER SCIENCES Volume 4 Issue 3 March 2022

[2] Anil Kumar Biswal,1Debabrata Singh,2Binod Kumar Pattanayak,1Debabrata Samanta, IoT-Based Smart Alert System for Drowsy Driver Detection, Volume 2021 | Article ID 6627217 | https://doi.org/10.1155/2021/6627217

[3] Mahek Jain, Real-Time Driver Drowsiness Detection using Computer Vision, International Journal of Engineering and Advanced Technology,October 2021

[4] Tim Jannusch, Cars and distraction: How to address the limits of Driver Monitoring Systems and improve safety benefits using evidence from German young drivers, ELSEVIER Technology in Society 66(2021) 101628

[5] Martin Gjoreski, Machine Learning and End-to-End Deep Learning Monitoring Driver Distractions From Physiological and Visual Signals, IEEE, Digital Object Identifier 10.1109/ACCESS.2020.2986810

[6] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov, and M. Herrmann, ‘‘Detection and evaluation of driver distraction using machine learning and fuzzy logic,’’ IEEE Trans. Intell.
Transp. Syst., vol. 20, no. 6, pp. 2048–2059, Jun. 2019

[7] Hemang Thakur1, Deval Arora1, Ramkiran Sampathi1, Shree Rukmini Thumu1 and Parisa Naraei2* 1 Department of Artificial Intelligence and Machine Learning Lambton College Toronto, Canada Received: November 27, 2021 Published: February 24, 2022

[8] Jabbar, R.; Shinoy, M.; Kharbeche, M.; Al-Khalifa, K.; Krichen, M.; Barkaoui, K. Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application. In Proceedings of the ICIoT 2020, Doha, Qatar, 2–5 February 2020; pp. 237–242.

[9] V. Sharath et al, "Driver Drowsiness Detection Using Haarcascade Algorithm", International Journal of Research in Engineering, Science and Management Volume-3, Issue- 8, August-2020.

[10] Abtahi, S.; Omidyeganeh, M.; Shirmohammadi, S.; Hariri, B. YawDD: A yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference, Singapore, 19–21 March 2014;