



BSIT-S20-001

03-135171-019 ABDUL MOUEED

Airbnb Price Prediction and Revenue Forecasting using Machine Learning

In partial fulfilment of the requirements for the degree of
Bachelor of Science in Information Technology

Supervisor: Dr. Ghulam Mustafa

Department of Computer Sciences
Bahria University, Lahore Campus

January 2021

Certificate



We accept the work contained in the report titled
“Airbnb Price Prediction and Revenue Forecasting using Machine Learning
written by
ABDUL MOUEED

as a confirmation to the required standard for the partial fulfilment of the degree of
Bachelor of Science in Information Technology

Approved by:

Supervisor: Dr. Ghulam Mustafa

(Signature)

January 18, 2021

DECLARATION

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Enrolment	Name	Signature
03-135171-019	ABDUL MOUEED	

Date : January 18, 2021

Specially dedicated to
My beloved Father, Mother, Elder Brother and Sisters.

ABDUL MOUEED

ACKNOWLEDGEMENTS

In the name of ALLAH Almighty, the Most Gracious and the Most Merciful. Alhamdulillah, all praises to ALLAH for the strengths and His blessing in completing this project.

We would like to express our deepest gratitude to our Supervisor Dr. Ghulam Mustafa for his deep patience, providing project information, helping for getting the necessary data, scholarly inspiration guidance during the project.

We would like to express many special thanks to our parents and family, who were always there to give all sorts of understanding and support.

ABDUL MOUEED

Airbnb Price Prediction and Revenue Forecasting using Machine Learning

ABSTRACT

Airbnb is an online marketplace for short-term home and apartment rentals. It offers a service for people to rent out their homes or living space for a short period while they are away or spare space to travellers. While Airbnb is an exciting service to earn some extra cash while renting out their home, it is very challenging for property owners to determine the price of the property as Airbnb determines the price based on the number of guests.

This work aims to develop a model with a Machine Learning (ML) framework, which will predict the price of the property and forecast revenue based on other rental information in the locality. Airbnb historical data will be scrapped and used to train two types of Machine Learning models 1) Hedonic Model Regression, and 2) XGBoost and compare their results for best accuracy. These Machine Learning models will later be hosted on a web server. A mobile compatible web client will be able to send queries to the webserver with the address of the property for 1) Property rent price prediction, 2) Periodic revenue forecast, and 3) Features of super-host in the neighbourhood. The final product of this project will be a web application that Airbnb property renters will able to use for the above-mentioned features.

TABLE OF CONTENTS

DECLARATION	ii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF SYMBOLS / ABBREVIATIONS	xiv

CHAPTERS

1	INTRODUCTION	1
	1.1 Background	1
	1.2 Problem Statements	2
	1.3 Aims and Objectives	2
	1.4 Scope of Project	2
2	LITERATURE REVIEW	4
	2.1 Introduction	4
	2.2 Background	4
	2.3 Airbnb PRF Functions	5
	2.3.1 User Profile	5
	2.3.2 Admin Profile	5
	2.3.3 Property Rent Recommendation	5
	2.3.4 Revenue Forecast	6
	2.3.5 Super-host Features	6

2.4	Operating Environment	6
2.4.1	Hardware Environment	6
2.4.2	Software Environment	6
3	DESIGN AND METHODOLOGY	7
3.1	Methodology	7
3.1.1	Model Development	7
3.2	Application	8
3.2.1	Server Application	9
3.2.2	Client Application	9
3.3	Feasibility Plan	9
3.3.1	Resource Requirement	9
3.4	Use-case Description	10
3.4.1	Sign Up	10
3.4.2	Sign In	11
3.4.3	Logout	12
3.4.4	View Profile	13
3.4.5	Update Profile	14
3.4.6	Reset Password	15
3.4.7	Property Rent Recommendation	16
3.4.8	Revenue Forecast	17
3.4.9	Superhost Features	18
3.4.10	Manage Users	19
3.5	System Requirements Chart	20
3.6	Use-case Diagram	23
3.6.1	Main Use Case Diagram	23
3.6.2	User Use Case Diagram	24
3.6.3	Admin Use Case Diagram	25
3.7	Domain Model	26
3.8	Sequence Diagram	27
3.8.1	Sign In	27
3.8.2	Sign Up	28
3.8.3	Log out	29

3.8.4	Reset Password	30
3.8.5	Update	31
3.8.6	View Profile	32
3.8.7	Property Rent Recommendation	33
3.8.8	Revenue Forecast	34
3.8.9	Superhost Feature	35
3.8.10	User Management	36
3.9	Collaboration Diagram	37
3.10	Design Class Diagram	37
3.11	Data Model	39
4	DATA AND EXPERIMENTS	40
4.1	Languages Used	40
4.1.1	HTML	40
4.1.2	JavaScript	40
4.1.3	Nodejs	41
4.1.4	Pug	41
4.1.5	Python	41
4.2	Framework	41
4.2.1	Express.js	41
4.2.2	Bootstrap	42
4.3	Tools	42
4.4	Algorithm	42
4.5	Project Workflow	43
5	RESULTS AND DISCUSSIONS	44
5.1	USER MANUAL	44
5.1.1	Main Page	44
5.1.2	Register	45
5.1.3	User Registration	45
5.1.4	Login	46
5.1.5	Dashboard	46
5.1.6	Profile	48

5.1.7	Update User Information	48
5.1.8	Change Password	49
5.1.9	Forecasting and Revenue	50
5.1.10	History	50
5.1.11	Admin Panel	51
5.2	Model Results	51
5.2.1	Numerical Features	51
5.2.2	Categorical features	53
5.2.3	Boolean Features	54
5.2.4	Important Features	62
5.2.5	Results	63
6	CONCLUSION AND RECOMMENDATIONS	64
	REFERENCES	65

LIST OF TABLES

TABLE	TITLE	PAGE
	Table 3.1:Requirement Traceability	20
	Table 5.1:Results	63

LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure 3.1:	Working Process of the Proposed Model	8
Figure 3.2:	Main Use case	23
Figure 3.3:	User Use case	24
Figure 3.4:	Admin Use case	25
Figure 3.5:	Domain Model	26
Figure 3.6:	Sign In	27
Figure 3.7:	Sign Up	28
Figure 3.8:	Log out	29
Figure 3.9:	Reset Password	30
Figure 3.10:	Update	31
Figure 3.11:	View Profile	32
Figure 3.12:	Property Rent Recommendation	33
Figure 3.13:	Revenue Forecast	34
Figure 3.14:	Superhost Feature	35
Figure 3.15:	User Management	36
Figure 3.16:	Collaboration Diagram	37
Figure 3.17:	Class Diagram	38
Figure 3.18:	Data Model	39
Figure 5.1:	Main Page	44

Figure 5.2: Register	45
Figure 5.3: User Register	45
Figure 5.4: Login	46
Figure 5.5: Dashboard 1	46
Figure 5.6: Dashboard 2	47
Figure 5.7: Dashboard 3	47
Figure 5.8: Dashboard 4	47
Figure 5.9: Profile	48
Figure 5.10: Update User Information	48
Figure 5.11: Updated User Information	49
Figure 5.12: Change Password	49
Figure 5.13: Revenue and Forecasting	50
Figure 5.14: History	50
Figure 5.15: Admin Panel	51
Figure 5.16: Numerical Features	51
Figure 5.17: Numerical Features 1	52
Figure 5.18: Numerical Features 2	52
Figure 5.19: Numerical Features 3	53
Figure 5.20: Property Type Figure 5.21: Room Type	53
Figure 5.22: Host is Superhost	54
Figure 5.23: Host Identity Verified	54
Figure 5.24: Instant Bookable	55
Figure 5.25: Balcony	55
Figure 5.26: Bed Linen	55
Figure 5.27: Breakfast	56

Figure 5.28: Child Friendly	56
Figure 5.29: Coffee Machine	56
Figure 5.30: Cooking Basics	57
Figure 5.31: Elevator	57
Figure 5.32: Host Greeting	57
Figure 5.33: Internet	58
Figure 5.34: Outdoor Space	58
Figure 5.35: Parking	58
Figure 5.36: TV	59
Figure 5.37: white Goods	59
Figure 5.38: Long-term Stays	59
Figure 5.39: Pets Allowed	60
Figure 5.40: Private Entrance	60
Figure 5.41: Secure	60
Figure 5.42: Self Check in	61
Figure 5.43: Important Feature	62

LIST OF SYMBOLS / ABBREVIATIONS

Term	Definition
AirbnbPRF	Airbnb Price Prediction and Revenue Forecast
User	Someone who interacts with the web-based application
Admin/Administrator	A system administrator who is given specific permission for managing and controlling the system
VS Code	It is an IDE that is used to develop applications and programs.
ML	Machine Learning
Jupyter notebook	Used to develop Machine Learning Model.
U	Use Case
Dataset	A collection of Data.
XGBoost	eXtreme Gradient Boosting
Airbnb Rent Prediction	Website of the Project.
RMSE	Root Mean Square Error
R^2/r^2	The coefficient of determination denoted R^2 or r^2 and pronounced "R squared"

CHAPTER 1

INTRODUCTION

1.1 Background

Airbnb is an online short stay property rental service, which offers some advantages over the hotels. For a traveller, travelling to a place, they can find a space which best suits their circumstance. They can find a place to stay according to their need, for example, a place to stay near a ski spot or finding a place in a specific neighbourhood. Often travellers like to book a shared space with Airbnb as it helps them interact with local people to get useful insights for that city or neighbourhood. Additionally, Airbnb places are many times much cheaper than hotels. From a renter perspective, they can rent out their place for a small period, for example, if they are travelling somewhere and no one is utilizing their place, they can make some cash by renting it out, or if they have extra rooms in their place, they can make money for their utility bills, or save it for other purposes.

Airbnb has over 180 million users worldwide, with 2.9 million hosts and 6 million property listings. It had an economic impact of \$100 billion in 2018. It is operating in 81,000 cities and 191 countries. Users do an average of 2 million bookings daily. The number of users and renters is increasing at a steady speed. For a renter, it is still a challenge for them to determine the price for their property, as Airbnb only offers an initial price based on the number of guests that a property can host.

1.2 Problem Statements

Currently, while listing property Airbnb provides hosts with a general guideline for determining the best price for a place to rent, it presents with a listing of neighbourhood property list to help host determine rent but it becomes a tedious job at times. Also, since the market is dynamic, a host would want to update the price frequently and this method can help predict rent price depending on the trends.

1.3 Aims and Objectives

1. Scrap data for training Machine Learning model
2. Train Machine Learning model for price prediction
3. Train Machine Learning model for revenue forecasting
4. Train Machine Learning model to find features of superhost
5. Develop a web server for hosting the Machine Learning model
6. Develop Rest APIs that will be used from client application to send and retrieve data from an application server
7. Develop a front-end web application for following result visualization
8. Price Prediction for the user input address
9. Revenue forecast for the user input address
10. Features of superhost in the user's neighbourhood

1.4 Scope of Project

Airbnb is an online marketplace for renting and property for a small stay. This service has been quite popular with travellers; they can easily find a space which suits best for their circumstances. A Renter can rent their place for a small-time if they are travelling somewhere, or they want to share their space to earn some cash for managing their living expenses. Since Airbnb has been quite popular recently, it has become a challenge for a renter to determine the price for their property, as Airbnb only offers an initial price based on the number of guests a property can host. It is a very difficult

job for a renter to go Airbnb and try to find nearby Airbnb places and analyze price trends in the neighbourhood. This work aims to develop a service that renters will be able to use to find price predictions for their property. Machine Learning model will make a price prediction after training from data for a certain locality. Renters will be able to use this service to find forecast revenue for property throughout times (average daily, weekly, monthly, yearly). This service will also provide features of super-host in their neighbourhood; a renter can learn them to provide a better experience for their guests.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This Software Requirement Specification (SRS) provides a complete description of all the functionalities and the specifications of the Airbnb PPRF (Airbnb Price Prediction and Revenue Forecast) project. The developers and the testers can use this document as a guideline for developing the design and test plan documents. For this substance, we are going to work on Airbnb PPRF (Airbnb Price Prediction and Revenue Forecast). However, it will explain different functions as well as non-functional requirements of the system, the interfaces of the system, what the system will do or how the system will interact with the external users, what are the constraints under which it will operate. This Software Requirement Specification document will provide a clear understanding of what is expected by the user in the proposed system. This will give a clear idea and guideline to the development team that how to develop software for end-users. This SRS will provide a basic flow and solid base for the project as well as help us to design low-level design documents and also helpful to create test plans for testing purposes

2.2 Background

This work aims to develop a website portal that will predict the price for a property using a machine learning model based on its locality and features. It will also forecast revenue for property throughout times. i.e., average daily, monthly, and yearly.

It will also provide useful insights into other properties in the neighbourhood and the kind of features super-hosts are offering for them.

2.3 Airbnb PRF Functions

2.3.1 User Profile

- User can sign up
- User can sign In
- User can update profile
- User can view profile

2.3.2 Admin Profile

- Admin can sign up
- Admin can sign in
- Admin can view users
- Admin can enable new users
- Admin can disable users

2.3.3 Property Rent Recommendation

While a user is logged-in with valid credentials, they can input the house address and its other related information in the rent price prediction section and can get a prediction for the rent of the house/apartment.

2.3.4 Revenue Forecast

While a user is logged-in with valid credentials, they can input the house address and its other housing information. Select time (weekly, monthly, yearly) and after generation of results will give to the user in the form of Revenue forecast.

2.3.5 Super-host Features

While a user is logged-in with valid credentials, they can input the house address and its other housing information. Select an area to find super-host features and after that model will predict the super-host features.

2.4 Operating Environment

2.4.1 Hardware Environment

While a user is logged-in with valid credentials, they can input the house address and its other housing information. Select an area to find super-host features and after that model will predict the super-host features.

2.4.2 Software Environment

Python frameworks are used to develop ML Model and for server-side website deployment, for building website JavaScript and its frameworks are being used. The supported web browsers are commonly Chrome, Firefox, and Microsoft Edge.

CHAPTER 3

DESIGN AND METHODOLOGY

3.1 Methodology

3.1.1 Model Development

In this work, I will develop a machine learning model to predict the price of the property. For training the machine learning model.

3.1.1.1 Data Collection

I will use Airbnb data, which could be accessed at this portal: (Inside Airbnb, 2020). This portal provides scrapped data for many popular cities around the world. Additional data can be scrapped using (Airbnb Data Collection, 2019) an open-source tool available on GitHub.

3.1.1.2 Designing Model

I will train two different types of Machine learning models and compare their accuracy and performance:

- Hedonic Model Regression: It is a common technique used in real estate appraisal and real estate economics. This model will produce a hedonic price estimation that includes spatial and locational features.
- XGBoost model: This is an implementation of gradient boosted decision trees.

3.1.1.3 Implementation

I will develop these models using Python programming language on Jupiter Notebook. For evaluation metrics, I will use mean squared error (for loss) and r-squared (for accuracy).

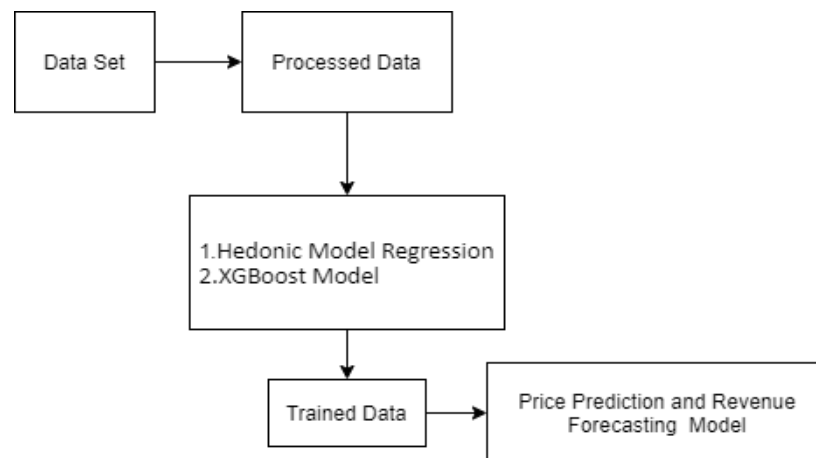


Figure 3.1: Working Process of the Proposed Model

3.2 Application

A client-server web application will develop to get user input, test machine learning models, and show results

3.2.1 Server Application

I will develop a server application where the Machine Learning model will host as well. This server application will provide REST APIs for getting information from the client application and result for visualization.

3.2.2 Client Application

A web client application will develop, which can take input from a user, send queries to web servers, get results, and visualize them. This web client application will be mobile-friendly and will be made with a focus to be useable on mobile phones.

3.3 Feasibility Plan

3.3.1 Resource Requirement

3.3.1.1 The expertise of the Team

I have already taken the following courses:

- Web Systems and Technologies
- Web Engineering

I have already taken a Python Programming course online and have developed beginner level expertise. I am taking the Data Mining course this semester, which will be helpful for this project. I have already started the Machine Learning course online learning techniques required to complete this project.

3.3.1.2 Tools / Technology

- An ubuntu-based computer for hosting the web application and machine learning model
- Nvidia GPU with at least 4GB VRAM and CUDA support
- Jupyter Notebook
- Visual Studio Code

3.4 Use-case Description

A use-case description is a text that captures the detailed functionality of a use-case. Description of all use-cases is written down in this section.

3.4.1 Sign Up

1. **U1**
2. **Objective** – To create an account on Airbnb PRF to access information
3. **Priority** – High
4. **Source** – Client
5. **Actors** – User
6. **Flow of Event**

6.1. Basic Flow

- 6.1.1. User opens the Airbnb PRF website
- 6.1.2. User Click on the Sign-up button
- 6.1.3. A user enters the required information
- 6.1.4. The system will confirm the username and password to the user again
- 6.1.5. User click on the create account button
- 6.1.6. User Redirected to the home page of Airbnb PRF

- 6.2. **Alternative Flow 1** – Step 6.1.3 user enter wrong information

or missing field

621. Show Redline on the border of each missing field

622. Show message, please fill the required field

6.3. **Alternative Flow 2** – Step 6.1.4 entered username and password not matched with the first attempt

631. Sorry, your username and password did not match with the first attempt

632. Please re-enter the password

633. On the third time enter an incorrect password return to step 6.1.4

6.4. **Alternative Flow 3** - Step 6.1.5 failed submission.

641. Please check the connection

642. Please go to help

643. Please re-enter the information

6.5. **Exception Flow 1** -

651. Server down.

7. **Includes** – None

8. **Preconditions** – The user must have an internet connection and a computer or Smartphone to access this Airbnb PRF.

9. **Postconditions** – User Successfully Created Account / Sign Up.

10. **Notes/Issues** – None

3.4.2 Sign In

1. **U2**

2. **Objective** – Give authorization to a user for the access of Airbnb PRF information

3. **Priority** – High
4. **Source** – Client
5. **Actors** – User, Admin
6. **Flow of Events**
 - 6.1. **Basic Flow**
 - 6.1.1. User opens the Airbnb PRF website
 - 6.1.2. Please enter username and password
 - 6.1.3. Press Sign in Button
 - 6.2. **Alternative Flow 1** – Step 6.1.3 user enter wrong username and password
 - 6.2.1. Show Warning for entering either Wrong username or password
 - 6.2.2. Show message, please re-enter username or password
 - 6.3. **Alternative Flow 2** –Step 6.1.2 User forget Password
 - 6.3.1. The user clicks on the password reset button
 - 6.3.2. Go to email to reset the account password
 - 6.4. **Exception Flow 1-**
 - 6.4.1. Server down
7. **Includes-** U1
8. **Preconditions** – User must already have the Airbnb PRF account
9. **Postconditions** – User Successfully logged in to Airbnb PRF and access the information
10. **Notes/Issues** – None

3.4.3 Logout

1. **U3**
2. **Objective** – User able to Logout from Airbnb PRF
3. **Priority** – Medium
4. **Source** – Client
5. **Actors** – User, Admin
6. **Flow of Events**

6.1. Basic Flow

- 6.1.1. The logout button is fixed in the menu bar
- 6.1.2. User can click on the Logout button to leave the Airbnb PRF
- 6.1.3. The system will destroy the session
- 6.1.4. The user will be redirected to the main page

6.2. Alternative Flow 1 – N/A

6.3. Exception Flow 1-

- 6.3.1. Server down

7. Includes – U2

8. **Preconditions** – User must be logged in to Airbnb PRF.

9. **Postconditions** – The user will be logged out from Airbnb PRF.

10. **Notes/Issues** – None

3.4.4 View Profile

1. **U4**

2. **Objective** – User able to View his/her Airbnb PRF profile.

3. **Priority** – Medium

4. **Source** – Client

5. **Actors** – User, Admin

6. **Flow of Events**

6.1. Basic Flow

- 6.1.1. User opens the Airbnb PRF website
- 6.1.2. User Login to Airbnb PRF
- 6.1.3. View Profile option is fixed on the menu bar
- 6.1.4. User will be redirected to the main page
- 6.1.5. A system extracts the information from the database and shows it on the new page

6.2. **Alternative Flow 1** – The Step 6.1.4 system failed to extract the information.

- 621. System show error
- 622. The system shows a message to try again

6.3. Exception Flow 1-

- 631. Server down.
- 7. **Includes** – U2
- 8. **Preconditions** – User must be logged in to Airbnb PRF
- 9. **Postconditions** – The system will show the information of the user
- 10. **Notes/Issues** – None

3.4.5 Update Profile

- 1. **U5**
- 2. **Objective** – User can Update his/her profile.
- 3. **Priority** – High
- 4. **Source** – Client
- 5. **Actors** – User, Admin
- 6. **Flow of Events**
 - 6.1. **Basic Flow**
 - 61.1. User opens the Airbnb PRF website
 - 61.2. User click on the update profile button
 - 61.3. User Profile information is shown on the page
 - 61.4. User refill the option that wants to update and fill the blank options
 - 61.5. User click on the update button
 - 61.6. User information will be immediately updated
 - 6.2. **Alternative Flow 1** – The Step 6.1.5 system failed to update the information due to false entry or due to an empty data field
 - 621. System show error message
 - 622. The system takes the user to step 6.1.4

6.3. Exception Flow 1

631. Server down

7. **Includes** – U2
8. **Preconditions** – User must be logged in to Airbnb PRF
9. **Postconditions** – The system will update the user information
10. **Notes/Issues** – None

3.4.6 Reset Password

1. **U6**
2. **Objective** – User can Update his/her profile.
3. **Priority** – Medium
4. **Source** – Client
5. **Actors** – User, Admin
6. **Flow of Events**

6.1. Basic Flow

- 61.1. The user opens the Airbnb PRF website
- 61.2. Enter the username or password and click on sign in
- 61.3. System show error
- 61.4. The user forgot the password and click on the Reset password button
- 61.5. A new page opens where the user enters their email for reset password and click on the button
- 61.6. The user receives the code in the mail
- 61.7. A user enters the code in the code field
- 61.8. User will be directed to a page where they enter a new password
- 61.9. Click on confirm will take the user to the main page

- 6.2. **Alternative Flow 1** – The Step 6.1.8 system failed to reset the password due to a mismatch password in fields

- 621. System show error message
- 622. The system takes the user to step 6.1.8

6.3. Exception Flow 1 –

- 631. Server down
- 7. **Includes** – U1
- 8. **Preconditions** – User must open the Airbnb PRF.
- 9. **Postconditions** – User Successfully logged in to Airbnb PRF and access the information.
- 10. **Notes/Issues** – None

3.4.7 Property Rent Recommendation

- 1. **U7**
- 2. **Objective** – Property Prices are recommended to the user.
- 3. **Priority** – High
- 4. **Source** – Client
- 5. **Actors** – User, Admin
- 6. **Flow of Events**
 - 6.1. **Basic Flow**
 - 61.1. User opens the Airbnb PRF website
 - 61.2. User Login to Airbnb PRF
 - 61.3. User Click on the Property Rent Recommendation button
 - 61.4. A new Form Appears
 - 61.5. User input the house address and its other housing information
 - 61.6. Click the generated result
 - 61.7. Model Predict the Rent Price and show it to the user on a new page
 - 6.2. **Alternative Flow 1** – The Step 6.1.6 system failed to generate the results due to false information

- 621. System show error message
- 622. The system takes the user to step 6.1.5

6.3. Exception Flow 1 –

- 631. Server down.
7. **Includes** – U1
 8. **Preconditions** – User must be logged in to Airbnb PRF
 9. **Postconditions** – User successfully gets recommended rent price of their property
 10. **Notes/Issues** – None

3.4.8 Revenue Forecast

1. **U8**
2. **Objective** – Here revenue is forecast for the user's property.
3. **Priority** – High
4. **Source** – Client
5. **Actors** – User, Admin
6. **Flow of Events**

6.1. Basic Flow

- 611. User opens the Airbnb PRF website
- 612. User Login to Airbnb PRF
- 613. User Click on the Revenue Forecast button
- 614. A new Form Appears
- 615. User input the house address and its other housing information
- 616. Select time (weekly, monthly, yearly)
- 617. Click generate result
- 618. Model Forecast the Revenue and shown it to a user on the new page

6.2. Alternative Flow 1 – Step 6.1.6 system failed to generate the results

due to false information

- 621. The system shows an error message
- 622. The system takes the user to step 6.1.5

6.3. Exception Flow 1 –

- 631. Server down.

- 7. **Includes** – U1
- 8. **Preconditions** – User must be logged in to Airbnb PRF
- 9. **Postconditions** – User successfully gets revenue forecast of their property
- 10. **Notes/Issues** – None

3.4.9 Superhost Features

- 1. **U9**
- 2. **Objective** – Here a user gets the information of superhost features
- 3. **Priority** – High
- 4. **Source** – Client
- 5. **Actors** – User, Admin
- 6. **Flow of Events**

6.1. Basic Flow

- 61.1. User opens the Airbnb PRF website
- 61.2. User login to Airbnb PRF
- 61.3. User clicks on the Super-host feature button.
- 61.4. A new form appears
- 61.5. User input the house address and its other housing information.
- 61.6. Select an area to find superhost features
- 61.7. Click the generate result
- 61.8. Model Predict the Super-host features and show it to the user on the new page

6.2. Alternative Flow 1 – The Step 6.1.6 system failed to generate the

results due to false information.

621. System show error message.

622. System takes the user to step 6.1.5.

6.3. Exception Flow 1-

631. Server down.

7. **Includes** – U1

8. **Preconditions** – User must be logged in to Airbnb PRF

9. **Postconditions** – User Successfully got the super-host features

10. **Notes/Issues** – None

3.4.10 Manage Users

1. **U10**

2. **Objective** – Admin can manage user

3. **Priority** – High

4. **Source** – Client

5. **Actors** – Admin

6. **Flow of Events**

6.1. Basic Flow

61.1. Admin opens the Airbnb PRF website

61.2. Admin Login to Airbnb PRF

61.3. Click on the view profile button

61.4. A new form appears

61.5. Admin click on manage user button

61.6. A new Form Appears where the system extracts the data of users and show it to the admin

61.7. Admin can enable and disable a user from this page

6.2. **Alternative Flow 1** – The Step 6.1.3 system failed to generate the results due to false information

621. System show error message of server down

622. System show error message of loss of internet connection problem

6.3. Exception Flow 1-

631. Server down

7. **Includes** – U1, U4

8. **Preconditions** – User must be logged in to Airbnb PRF

9. **Post conditions** – Admin Successfully perform the actions

10. **Notes/Issues** – None

3.5 System Requirements Chart

Table 3.1:Requirement Traceability

ID	Priority	Type NF=Non Functional F=Functiona l	Source	Used in Use case	Description
1	High	F	Client	U1	The system shall allow a user to create a profile and set his credential
2	High	F	Client	U2	Login into Airbnb PRF.
3	Medium	F		U3	User and Admin both can log out of the website and system session will be destroyed.

4	Medium	F	Client	U4	The system shall authenticate user credentials to view the profile.
5	High	F	Client	U5	The system shall allow a user to update the profile information
6	Medium	F	Client	U6	The system will send an email with a code to reset the profile password.
7	High	F	Client	U7	Here user Input the details and get recommended prices.
8	High	F	Client	U8	Here User Input details of area w.r.t to Property info and get Revenue forecast on a weekly, monthly, and daily basis.
9	High	F	Client	U9	Here User Input details of Property alongside super-host info and get the datils of super-host for given details.

10	High	F	Client	U10	Here Admin Can manage user. Admin can view, add, update, and delete users.
11	Low	NF	Client	N/A	The performance shall depend upon hardware components of the client/customer
12	High	NF	Client	N/A	The system's back-end servers shall only be accessible to authenticated administrators
13	High	NF	Client	N/A	The system's back-end databases shall be encrypted
14	Medium	NF	Client	N/A	The system provides use of icons according to the conditions and their usage
15	Low	NF	Client	N/A	The system provides a smooth look and flow between the different pages

3.6 Use-case Diagram

3.6.1 Main Use Case Diagram

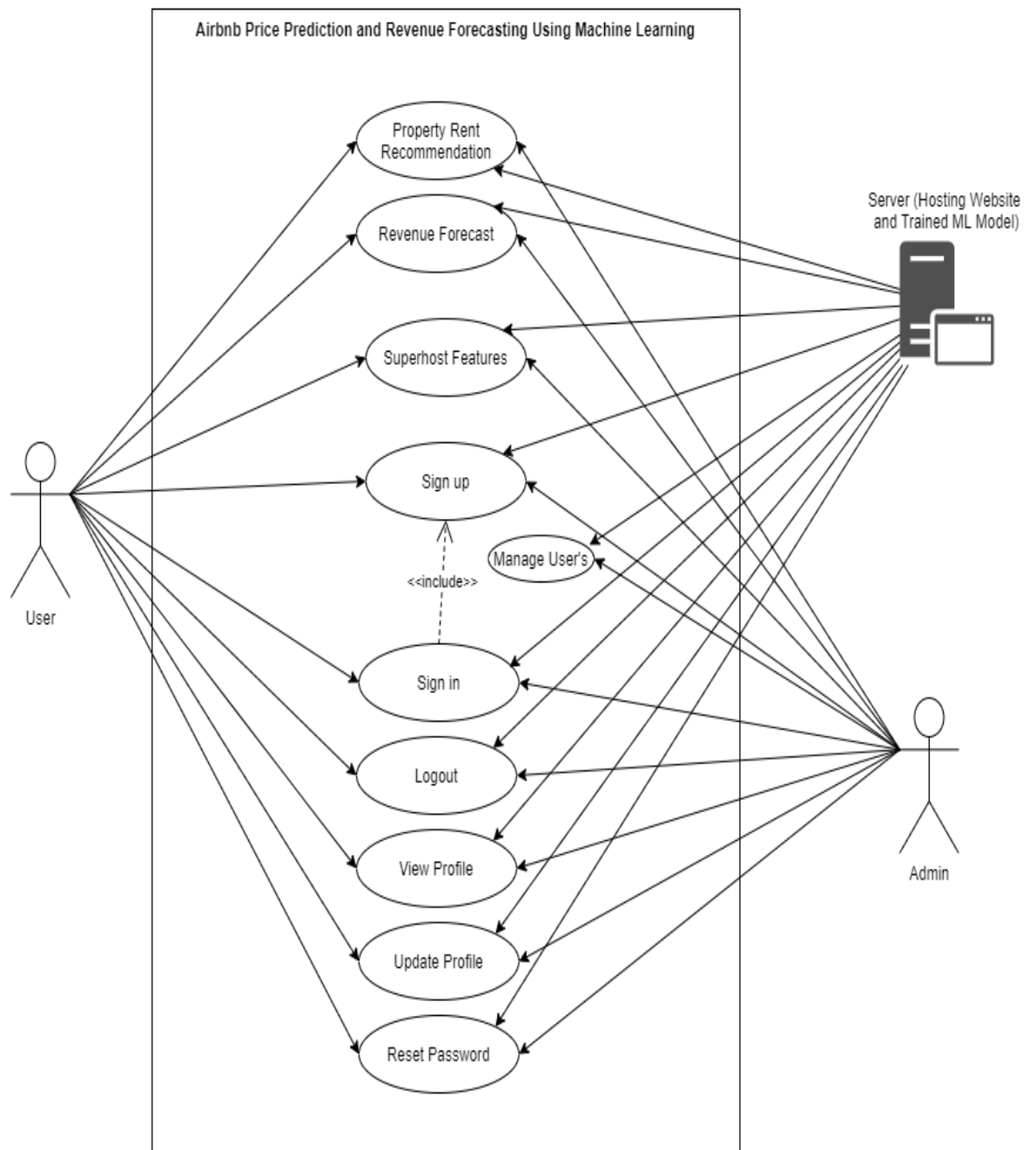


Figure 3.2: Main Use case

3.6.2 User Use Case Diagram

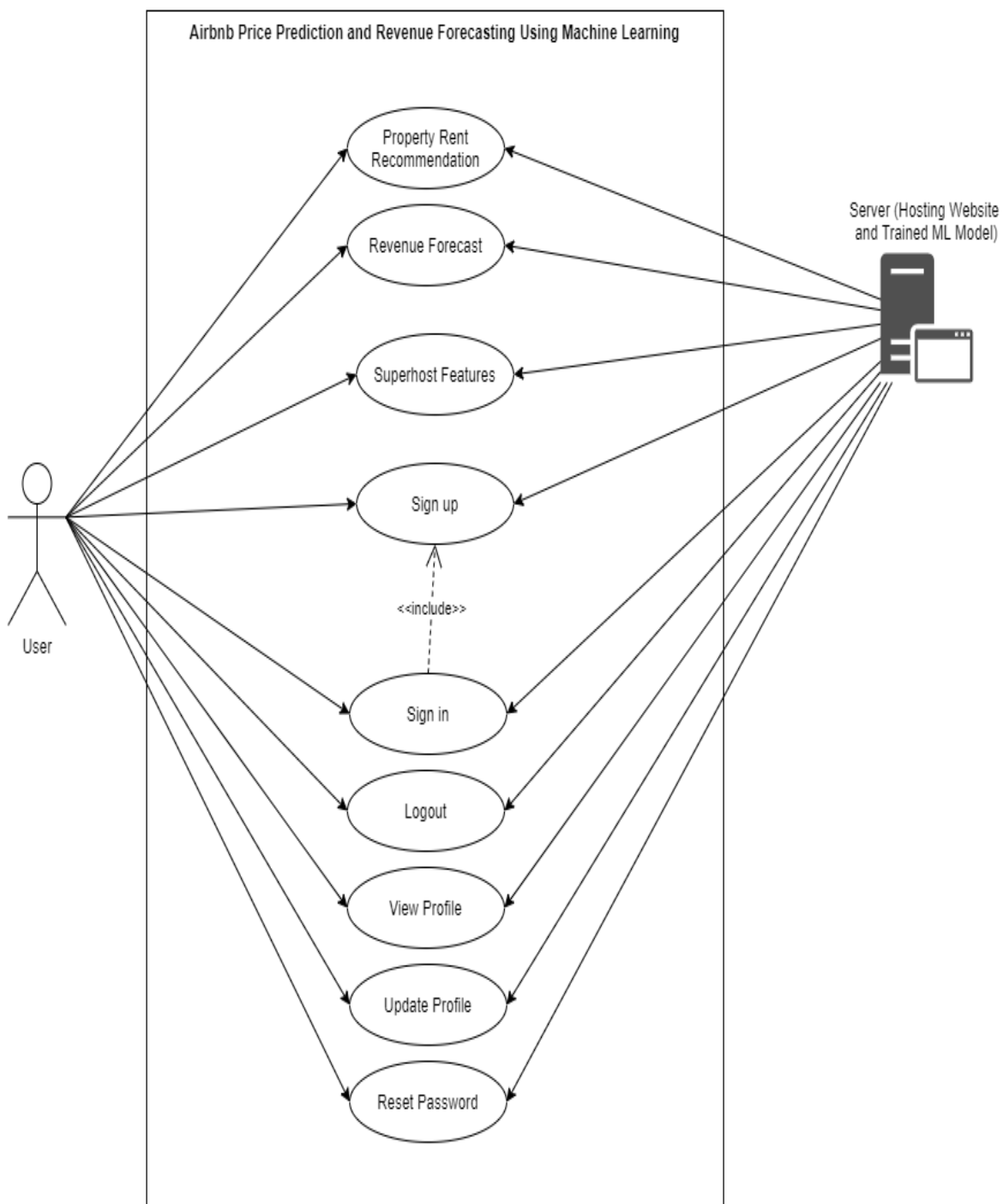


Figure 3.3:User Use case

3.6.3 Admin Use Case Diagram

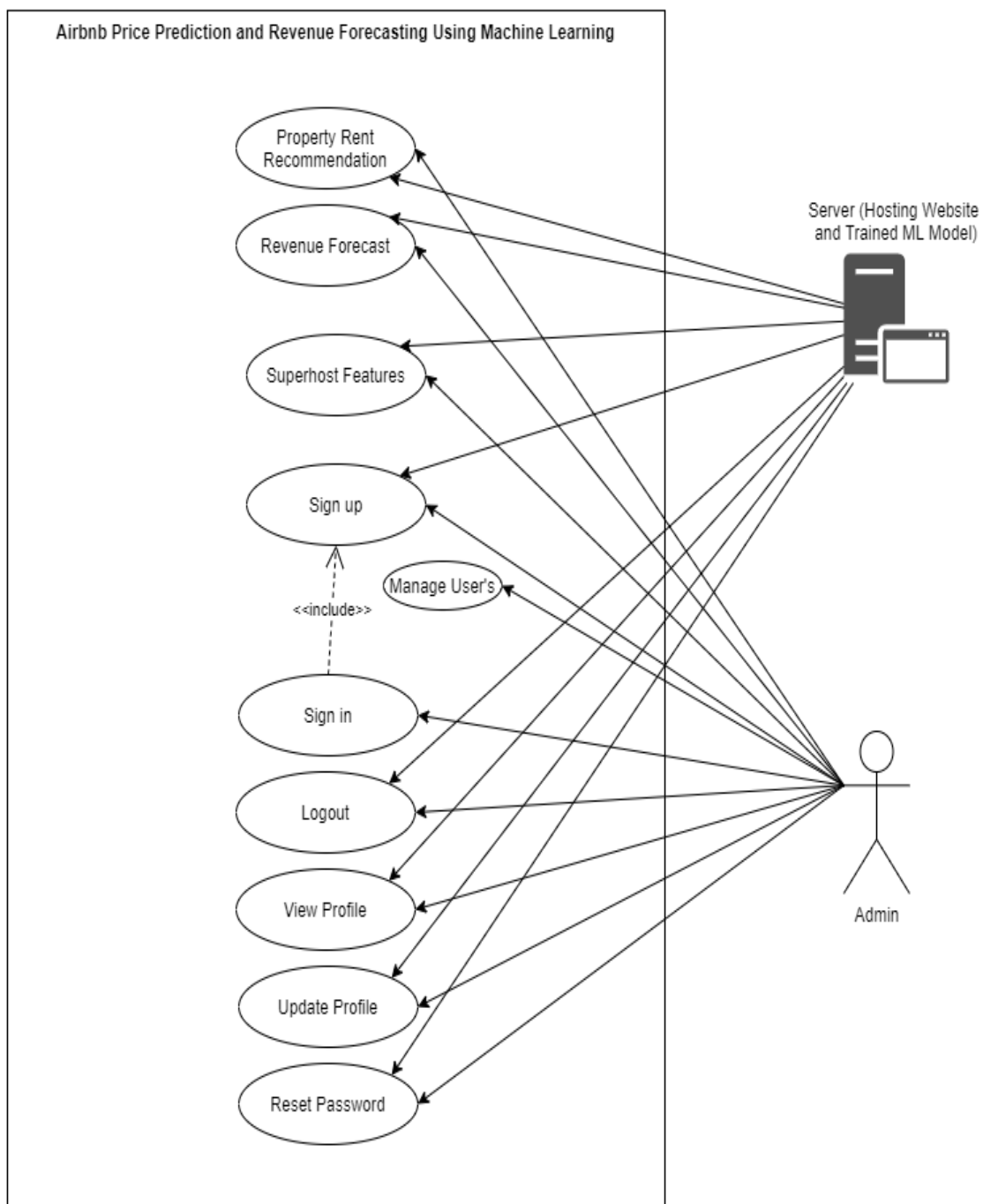


Figure 3.4:Admin Use case

3.7 Domain Model

Domain model is a conceptual model of the domain that incorporates both behaviour and data.

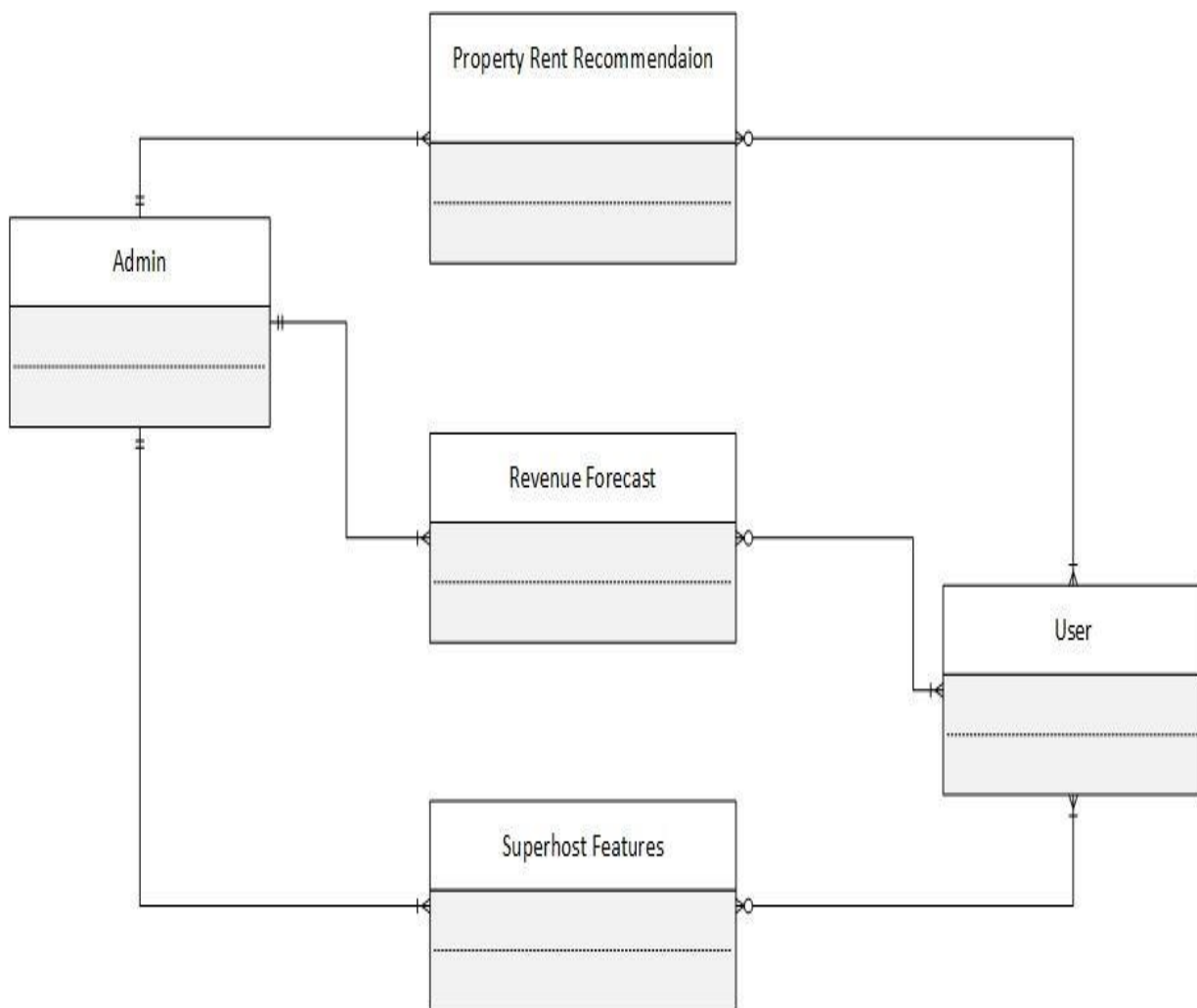


Figure 3.5: Domain Model

3.8 Sequence Diagram

3.8.1 Sign In

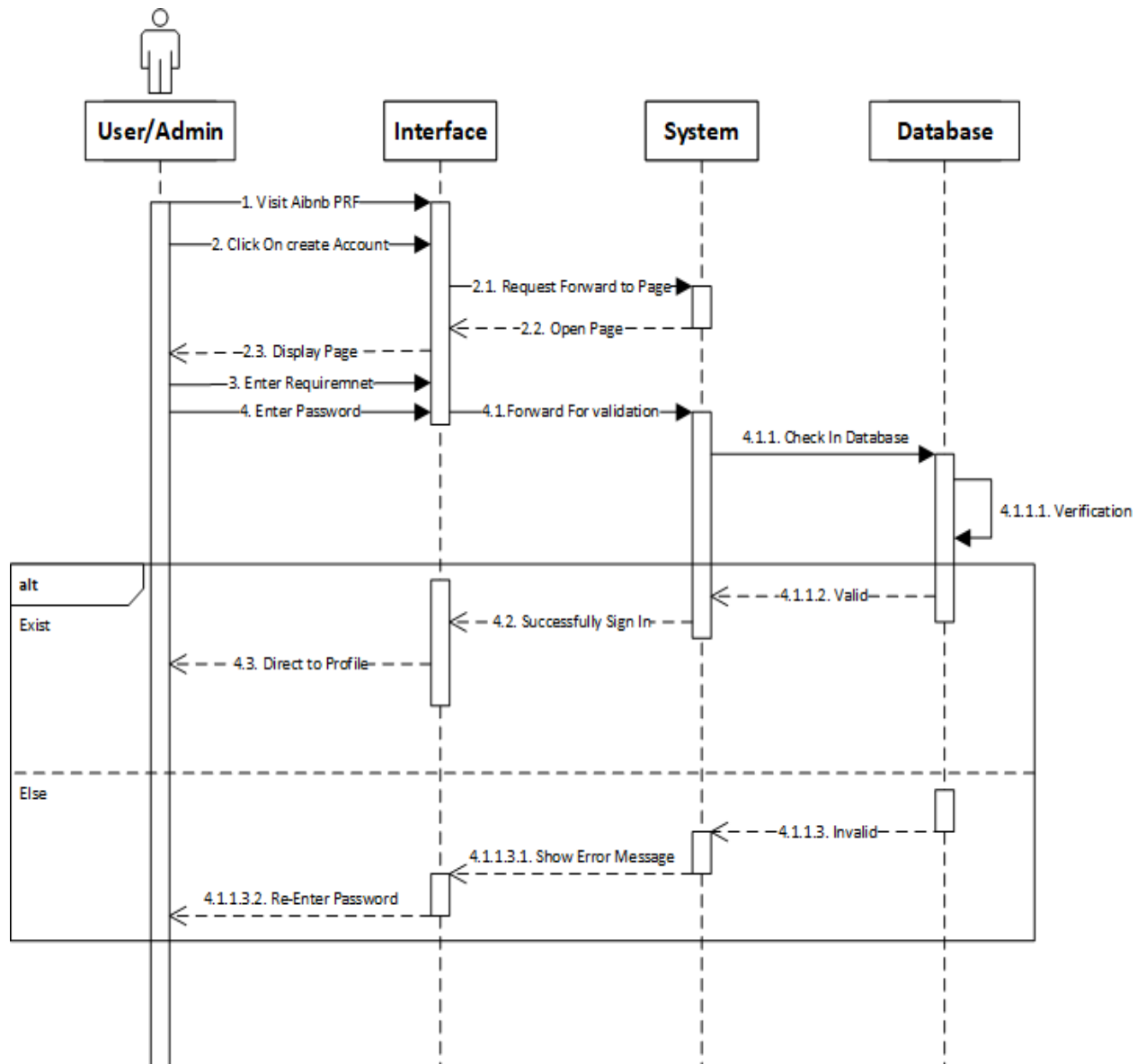


Figure 3.6: Sign In

3.8.2 Sign Up

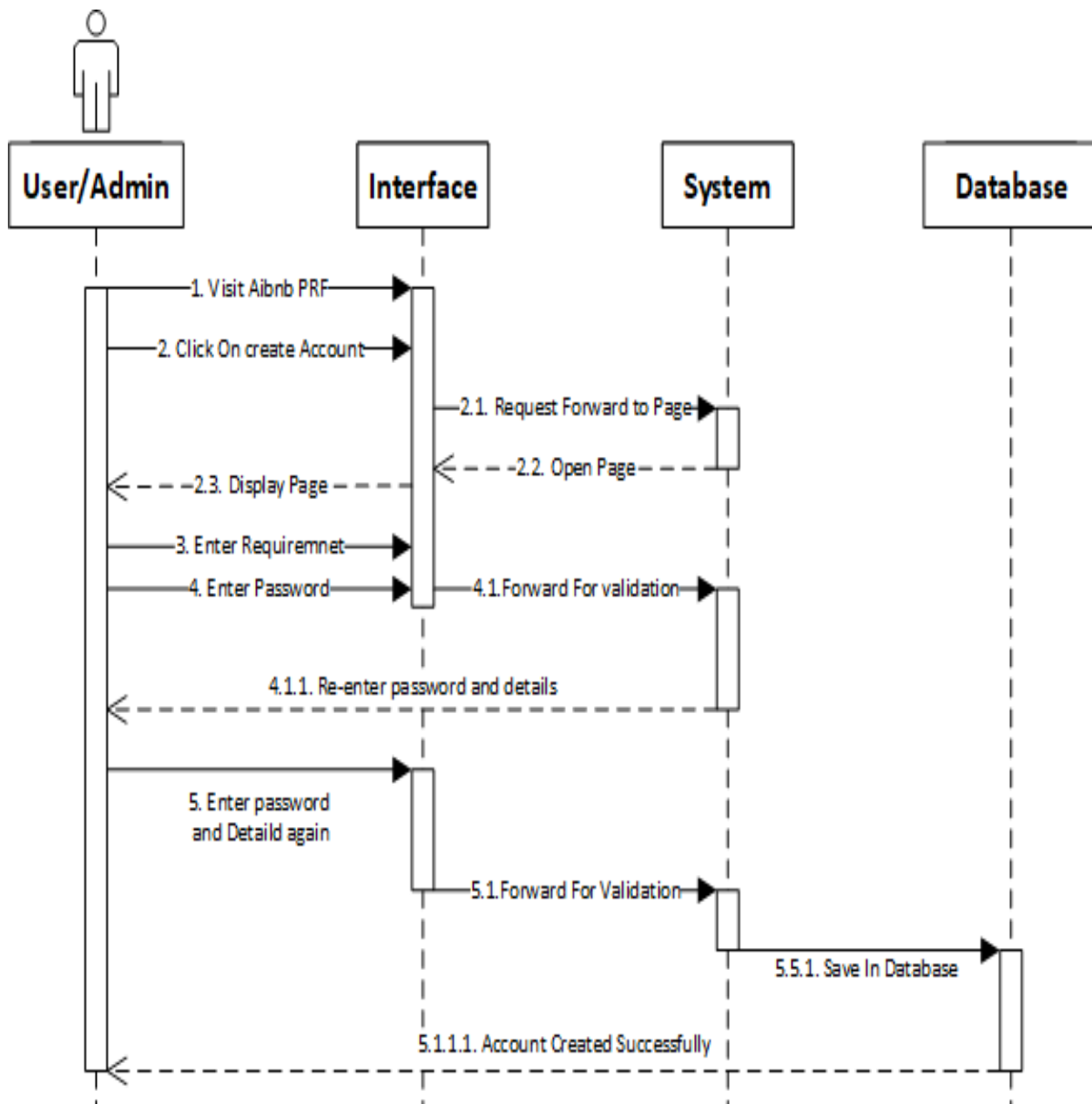


Figure 3.7: Sign Up

3.8.3 Log out

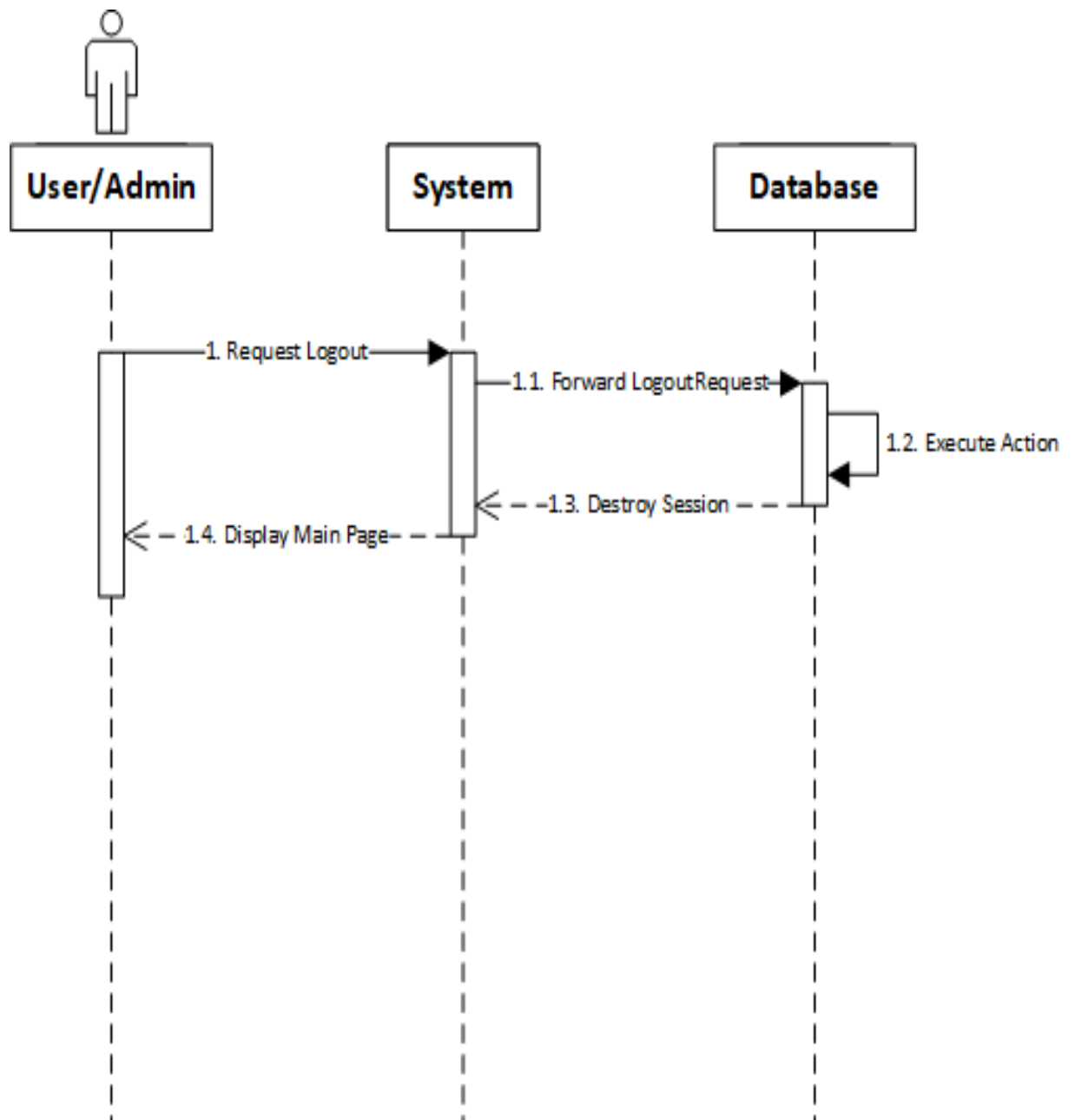


Figure 3.8:Log out

3.8.4 Reset Password

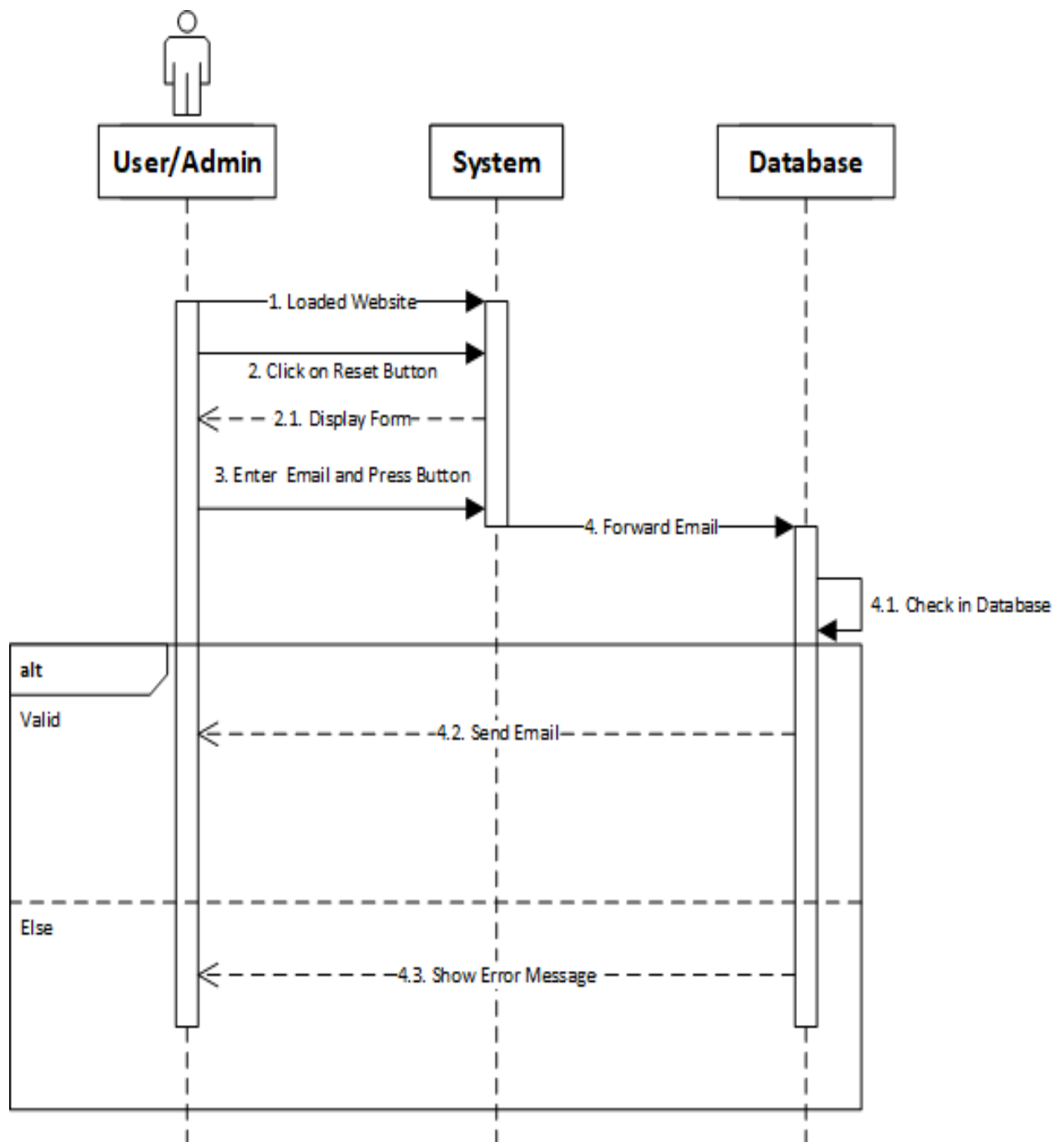


Figure 3.9:Reset Password

3.8.5 Update

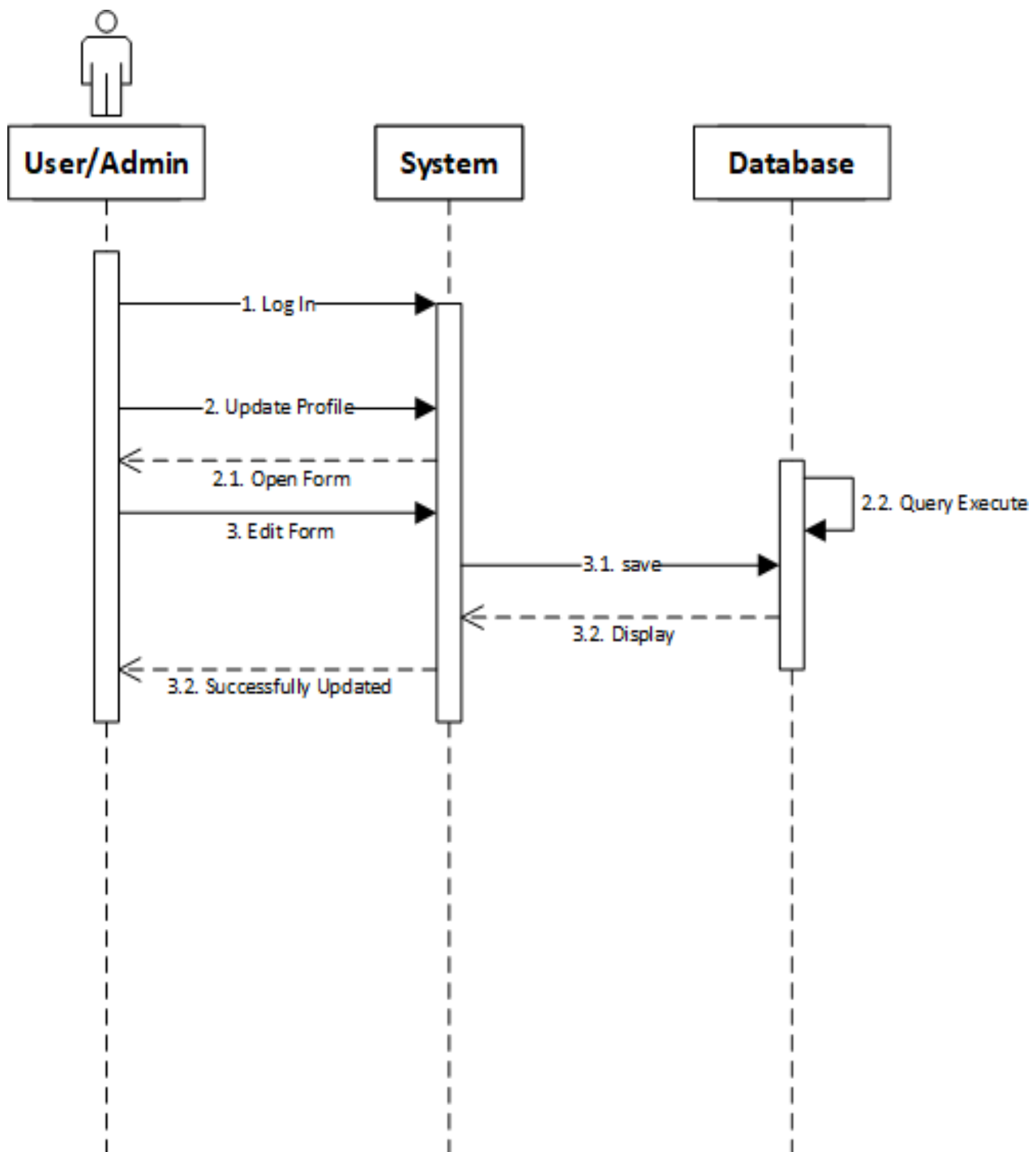


Figure 3.10:Update

3.8.6 View Profile

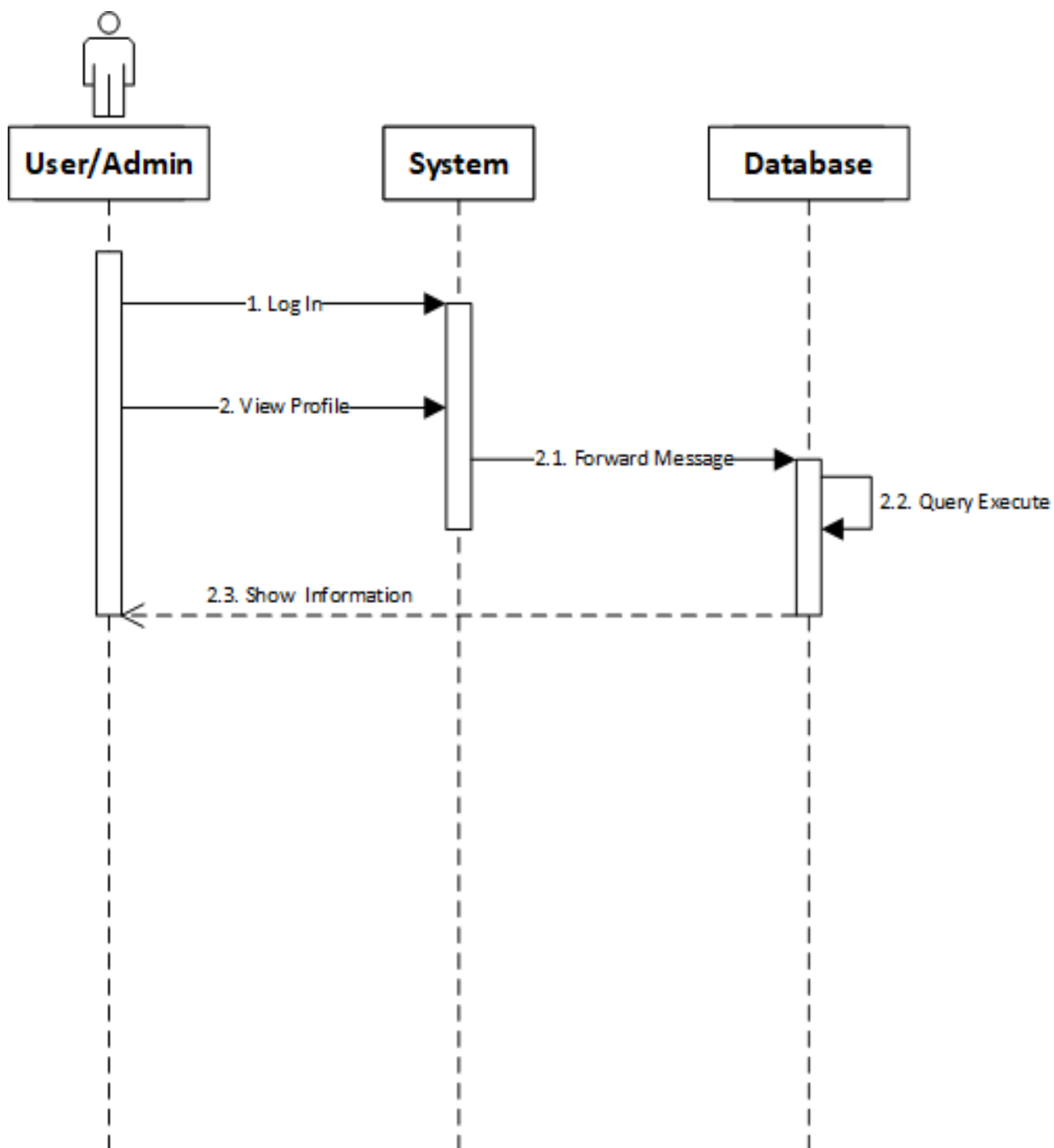


Figure 3.11:View Profile

3.8.7 Property Rent Recommendation

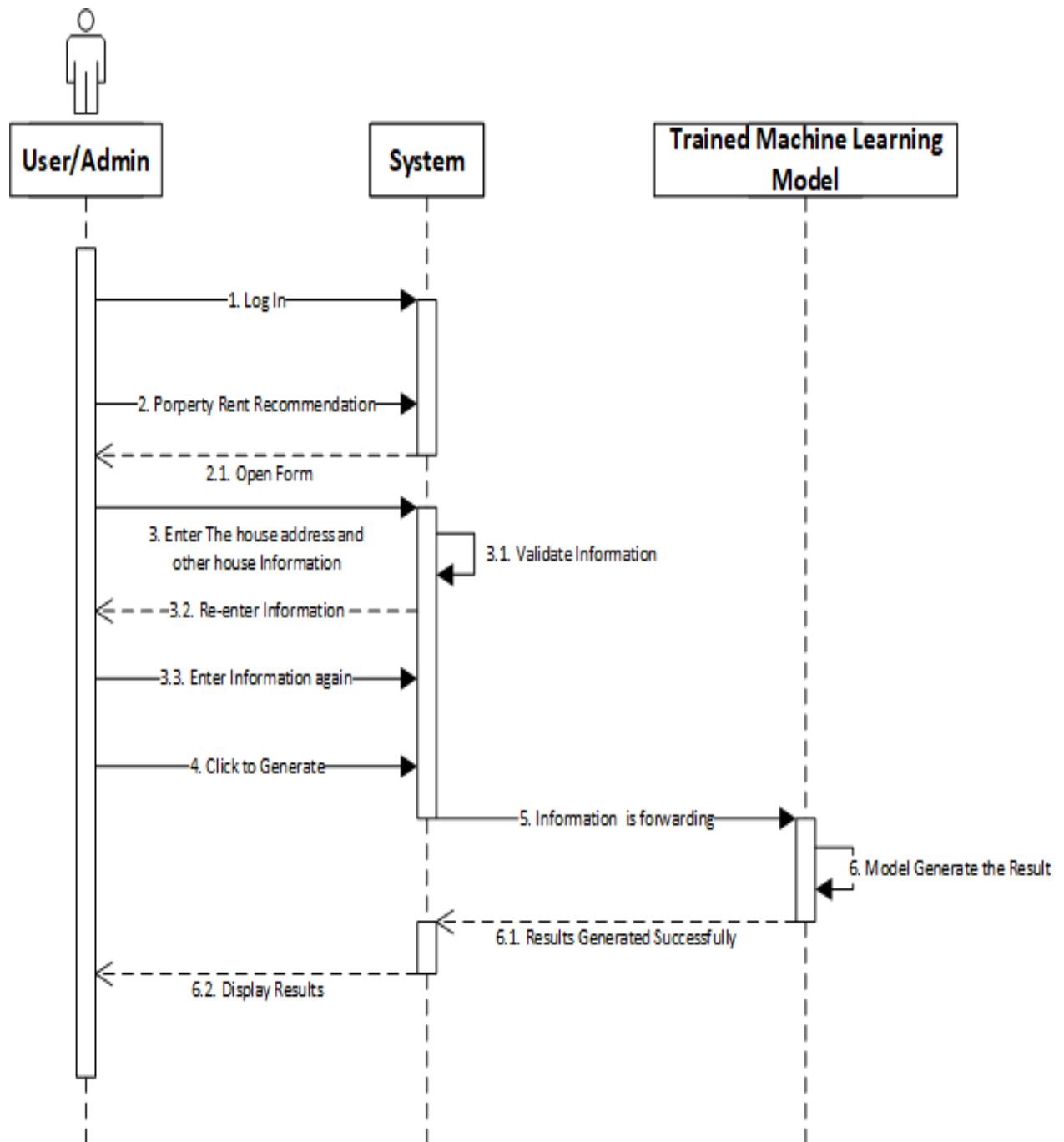


Figure 3.12:Property Rent Recommendation

3.8.8 Revenue Forecast

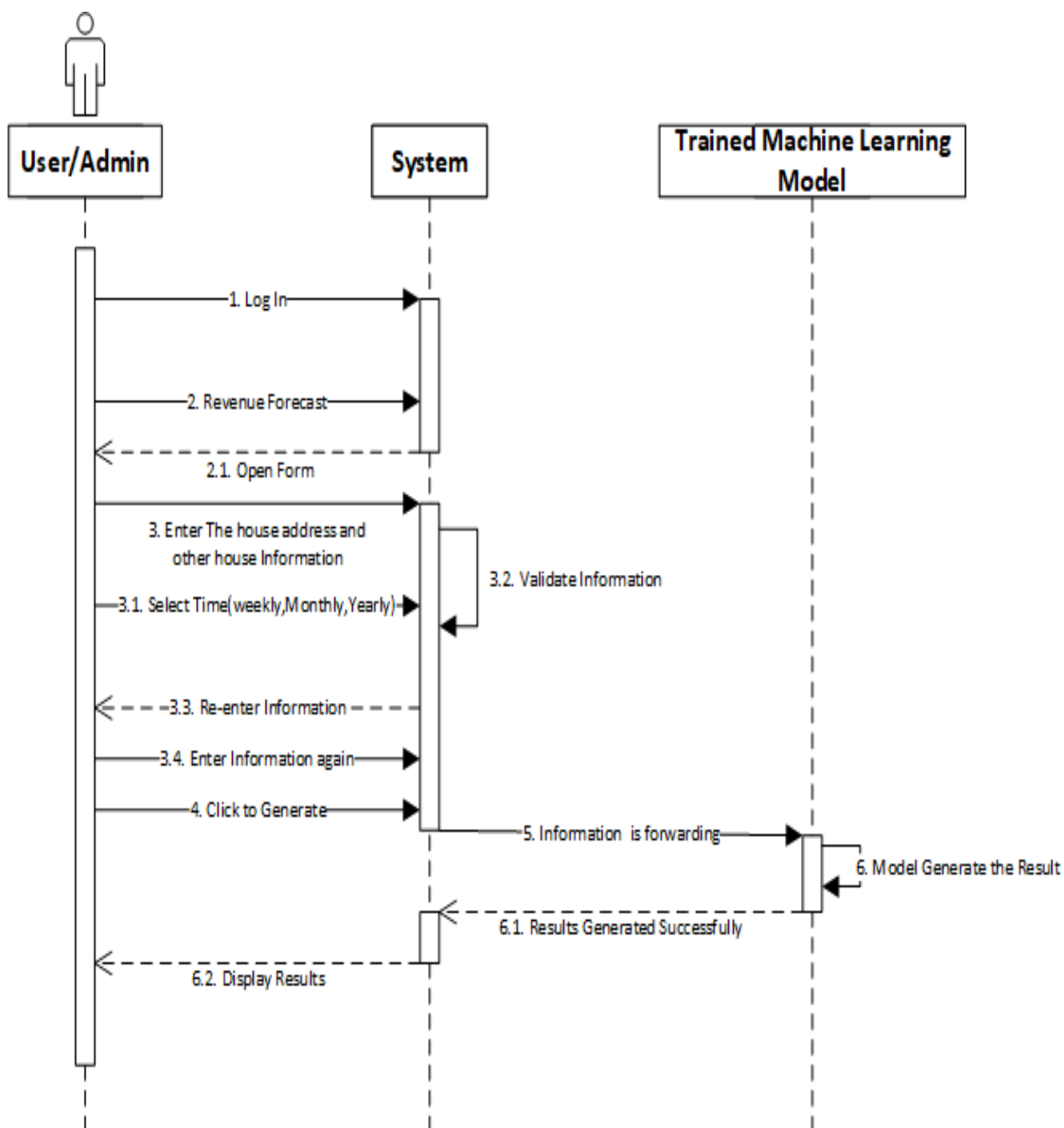


Figure 3.13:Revenue Forecast

3.8.9 Superhost Feature

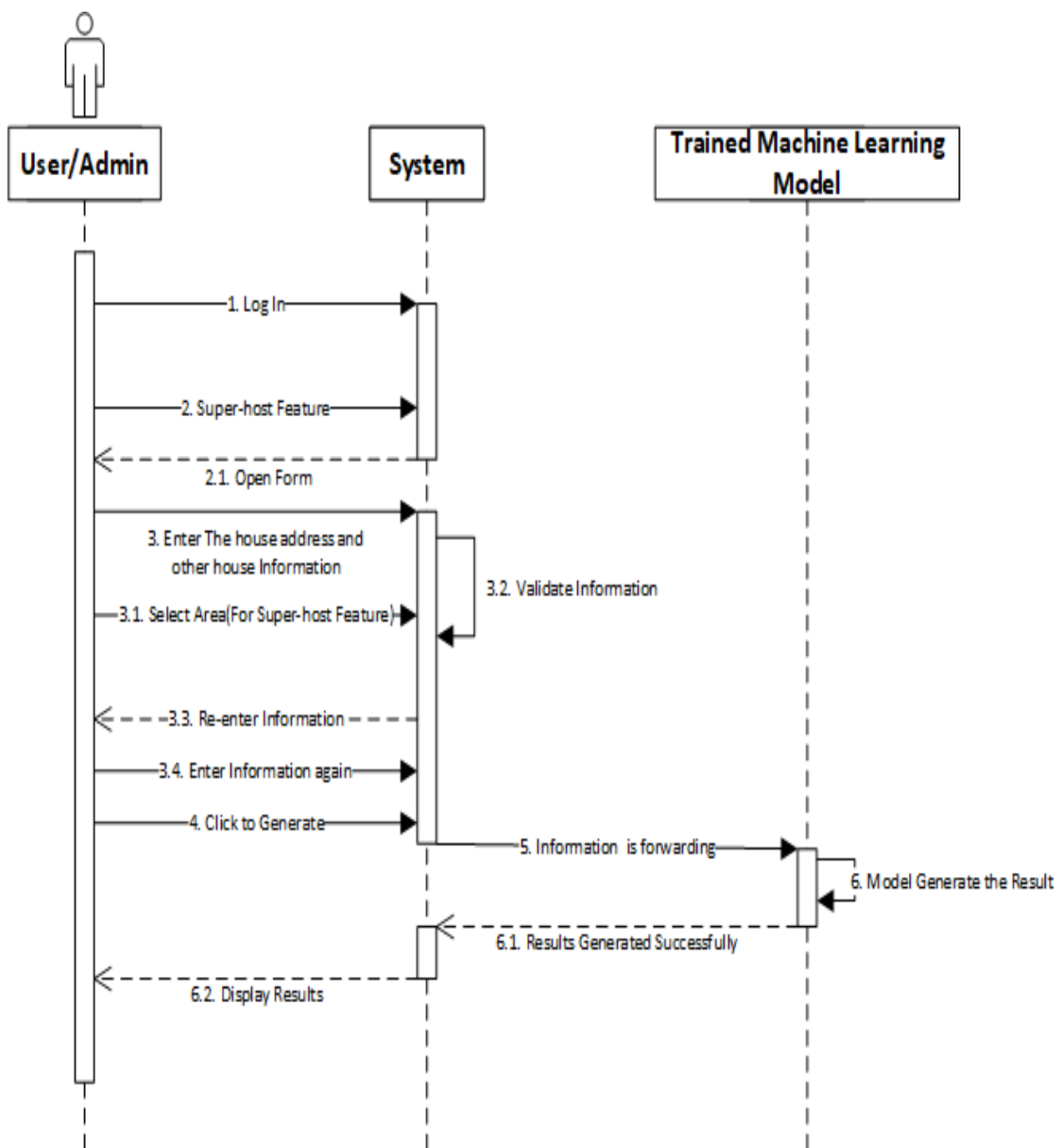


Figure 3.14: Superhost Feature

3.8.10 User Management

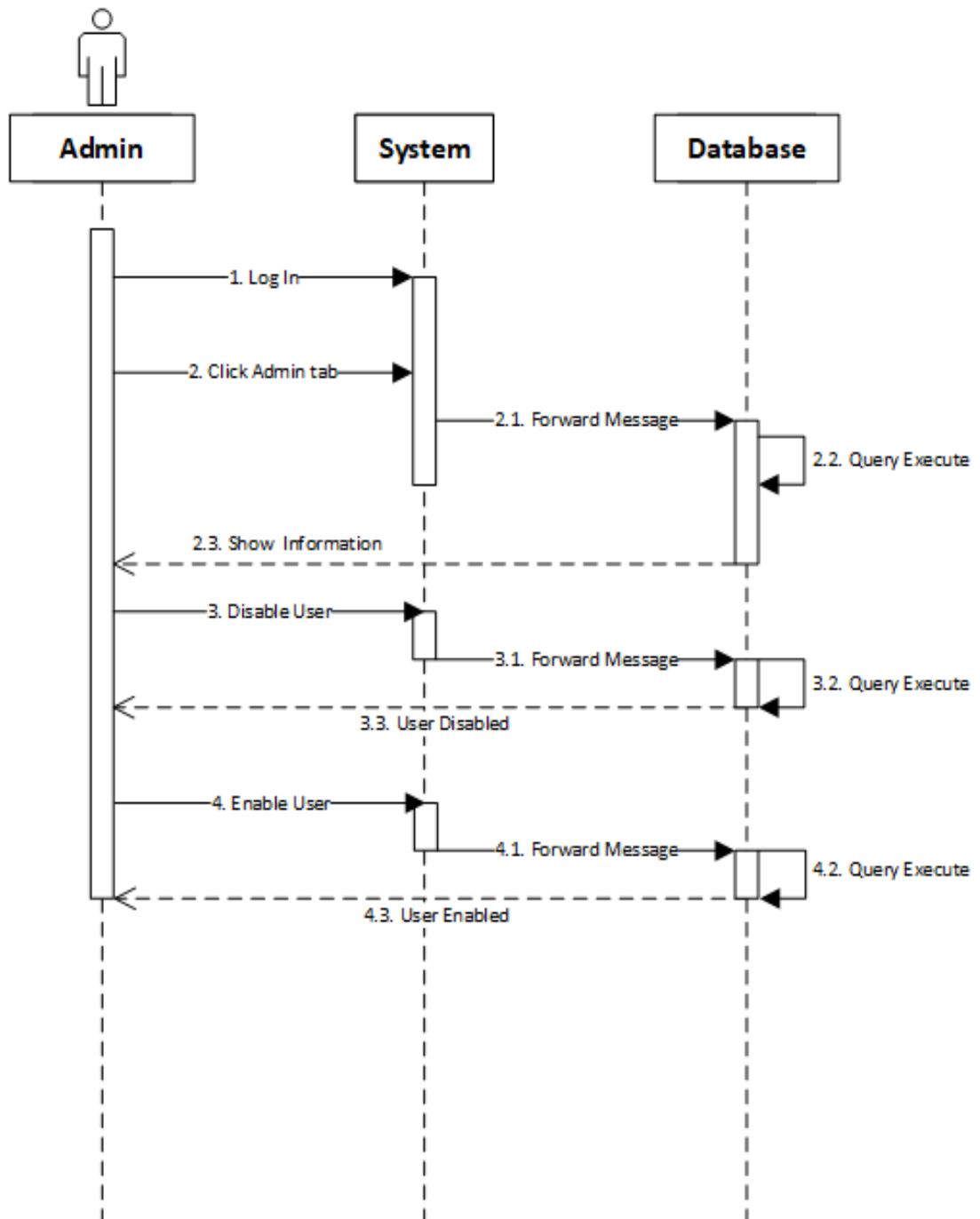


Figure 3.15:User Management

3.9 Collaboration Diagram

Collaboration diagram, also known as a communication diagram, is an illustration of their relationships and interactions among software objects in the Unified Modeling Language (UML).

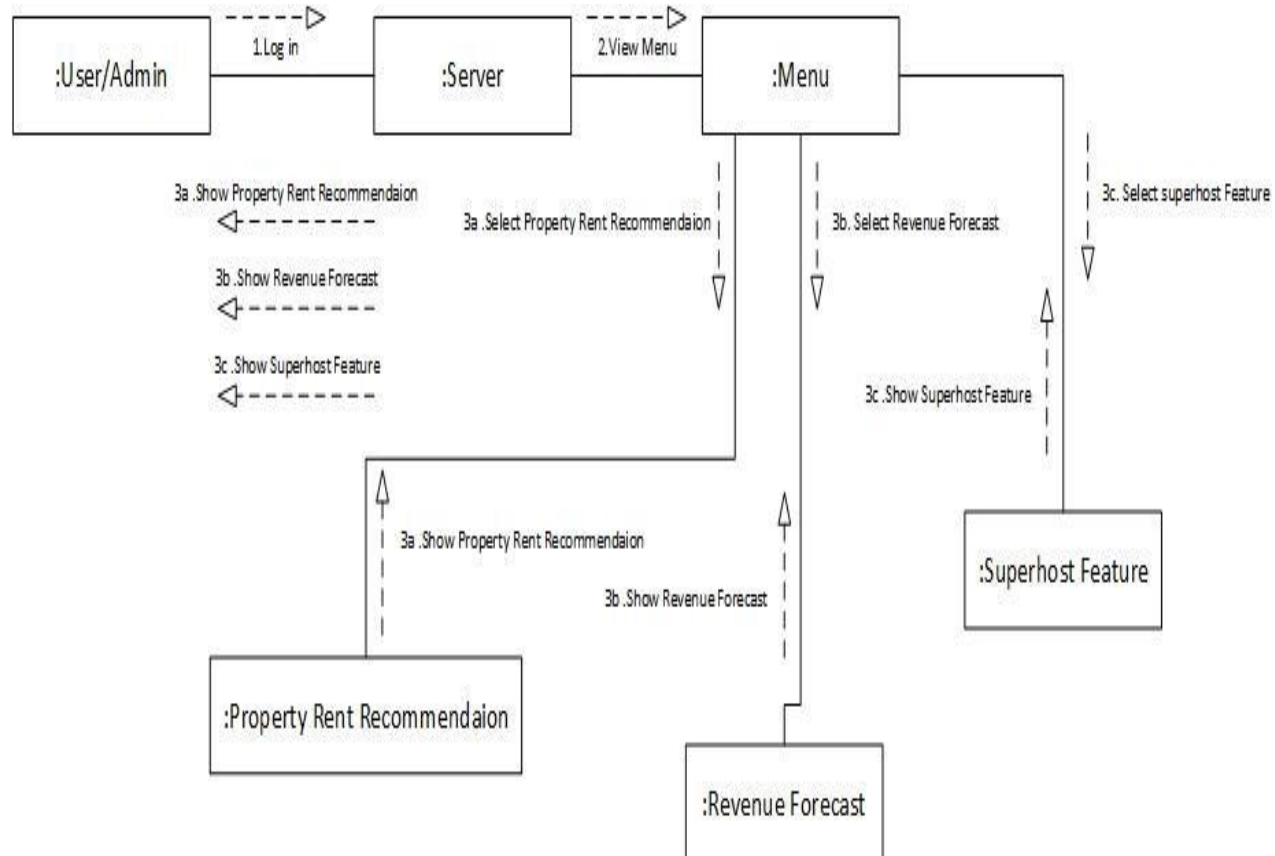


Figure 3.16:Collaboration Diagram

3.10 Design Class Diagram

Classes are the workhorses of the design effort—they actually perform the real work of the system. The other design elements—subsystems, packages and collaborations simply describe how classes are grouped or how they interoperate. Active classes are design classes, which coordinate and drive the behaviour of the passive classes - an

active class is a class whose instances are active objects, owning their own thread of control

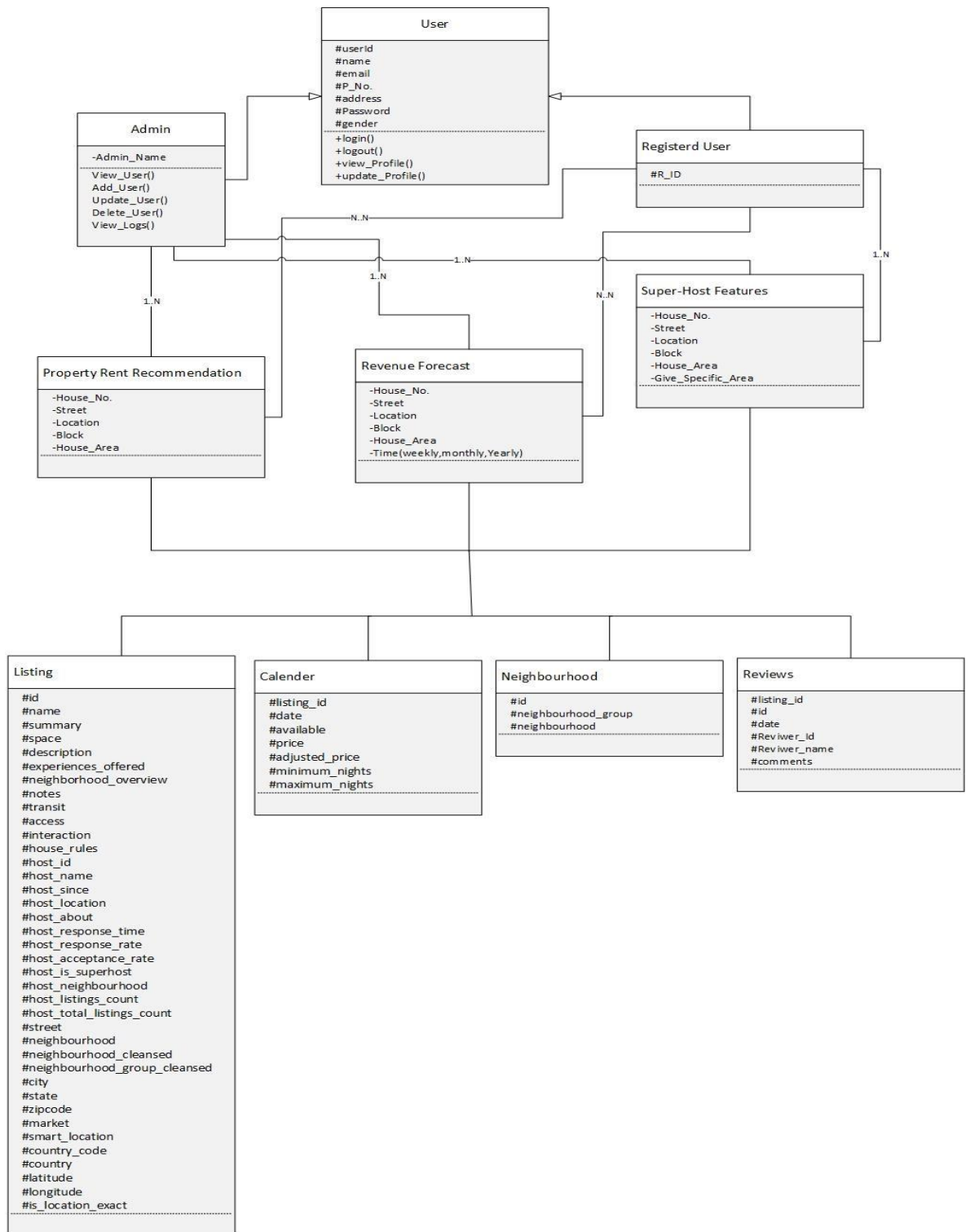


Figure 3.17:Class Diagram

3.11 Data Model

The data model is a subset of the implementation model, which describes the logical and physical representation of persistent data in the system.

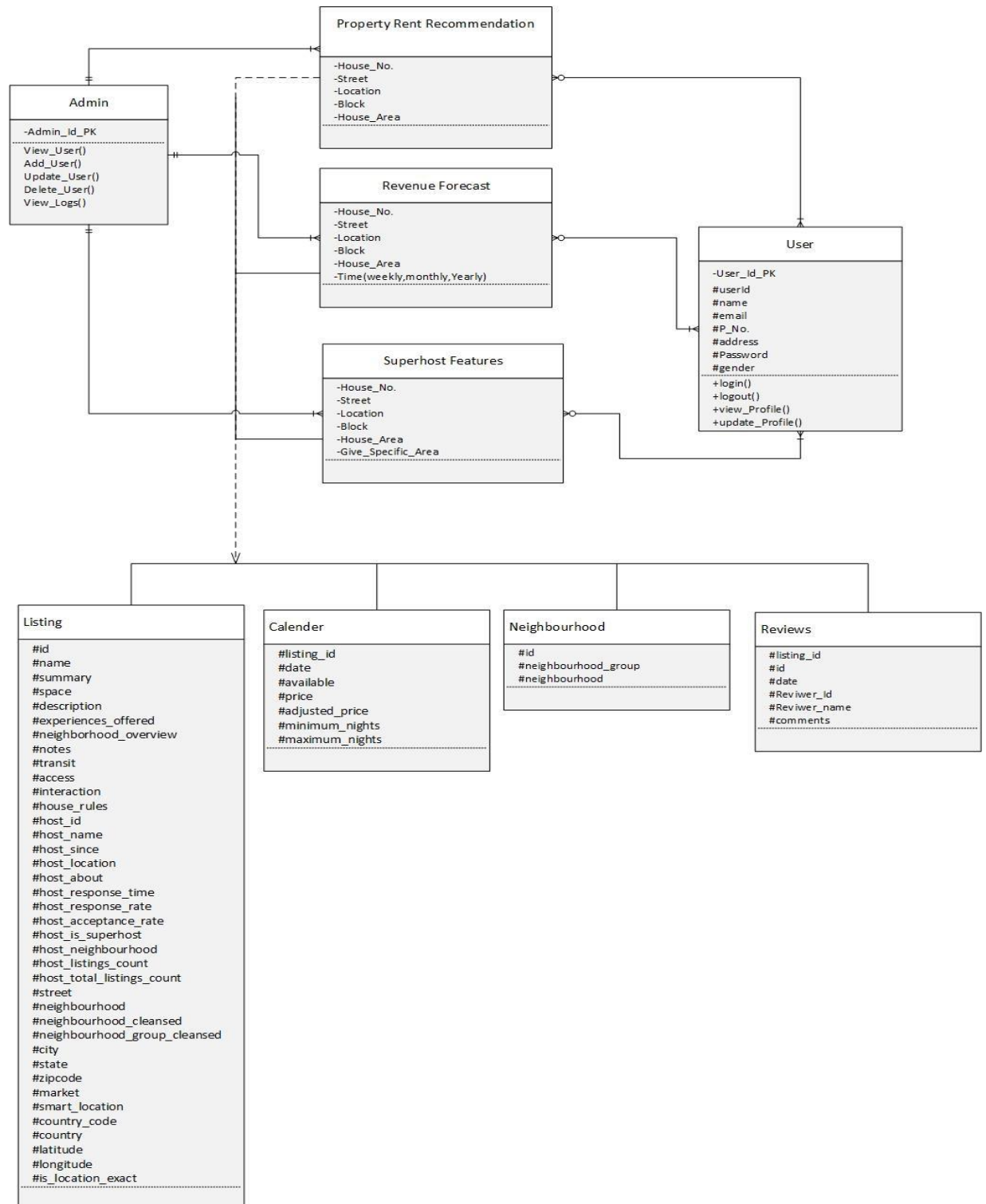


Figure 3.18:Data Model

CHAPTER 4

DATA AND EXPERIMENTS

4.1 Languages Used

Following are the tools and technologies which will be used in his project

4.1.1 HTML

HTML stands for Hypertext Mark-up Language. It is the standard mark-up Language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript, jQuery, and jQuery UI

Version: HTML5

4.1.2 JavaScript

It is used to control the website behaviour programmatically. JavaScript is used in our project for both frontend and backend extensively.

4.1.3 Nodejs

Node.js is an open-source, cross-platform, asynchronous JavaScript runtime environment. It is designed to build scalable network applications. It executes JavaScript code outside of a browser.

4.1.4 Pug

Pug is a template engine that enables you to use static template files in your application. At runtime, the template engine replaces variables in a template file with actual values and transforms the template into an HTML file sent to the client. This approach makes it easier to design an HTML page.

4.1.5 Python

Python is an interpreted, high level and general-purpose programming language. We use python for developing and training machine learning models and for inference server. For the training model, we use the following python Libraries. Pandas, NumPy, matplotlib, sklearn, XGBoost, IPython. Keras, date, geopanda.

4.2 Framework

4.2.1 Express.js

Express is a minimal and flexible Node.js web application framework that provides a robust set of features for web and mobile applications. which is specifically designed for building single-page, multi-page, and hybrid web applications.

4.2.2 Bootstrap

Bootstrap is the most popular HTML, CSS, and JavaScript Framework for developing responsive, mobile-first websites.

4.3 Tools

- Visual Studio Code
- Jupyter Notebook
- Mongo DB Compass

4.4 Algorithm

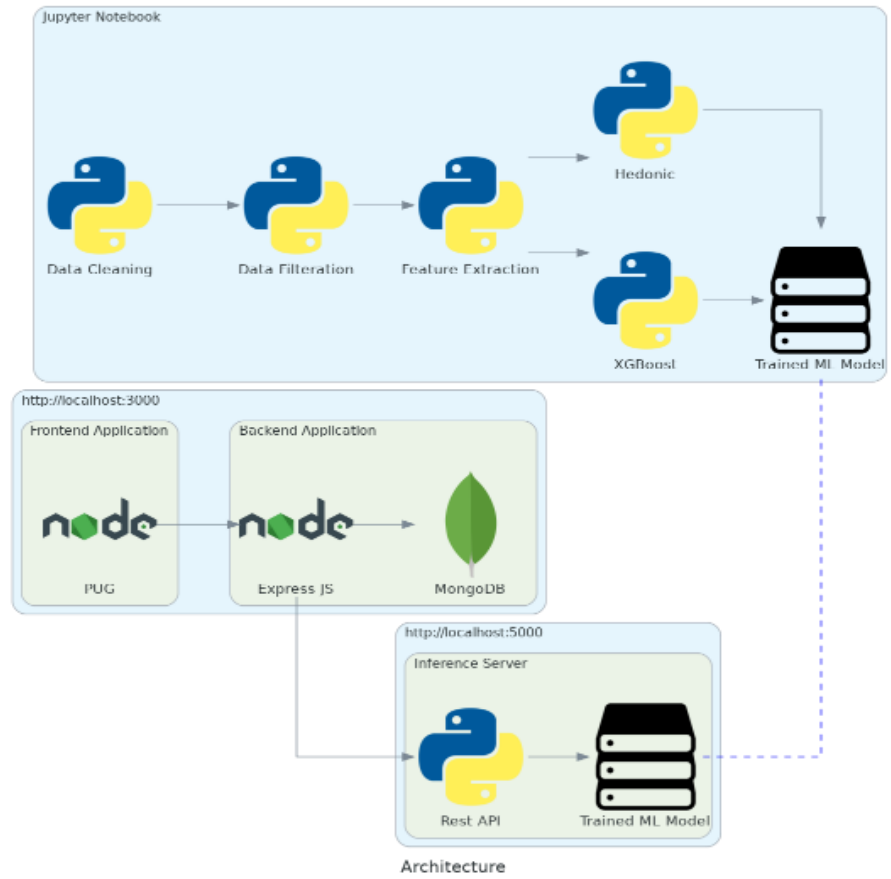
This Rent prediction and revenue forecast experiment algorithm use the XGBoost library to train the model. This process works in two phases: 1) the training phase, 2) the testing phase.

In training phases, all features are prepared from data and split into 80-20 ratio. And then input to the machine learning model. The objective function (loss function and regularization) at iteration t that we need to minimize is the following:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i(t-1) + f_t(x_i)) + \Omega(f_t)$$

In testing phases, user input is transformed into feature and input to the training model for inference.

4.5 Project Workflow



CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 USER MANUAL

5.1.1 Main Page

To open Airbnb Rent Prediction type URL <http://localhost:3000/> in the browser and the user will automatically redirect to the main page.

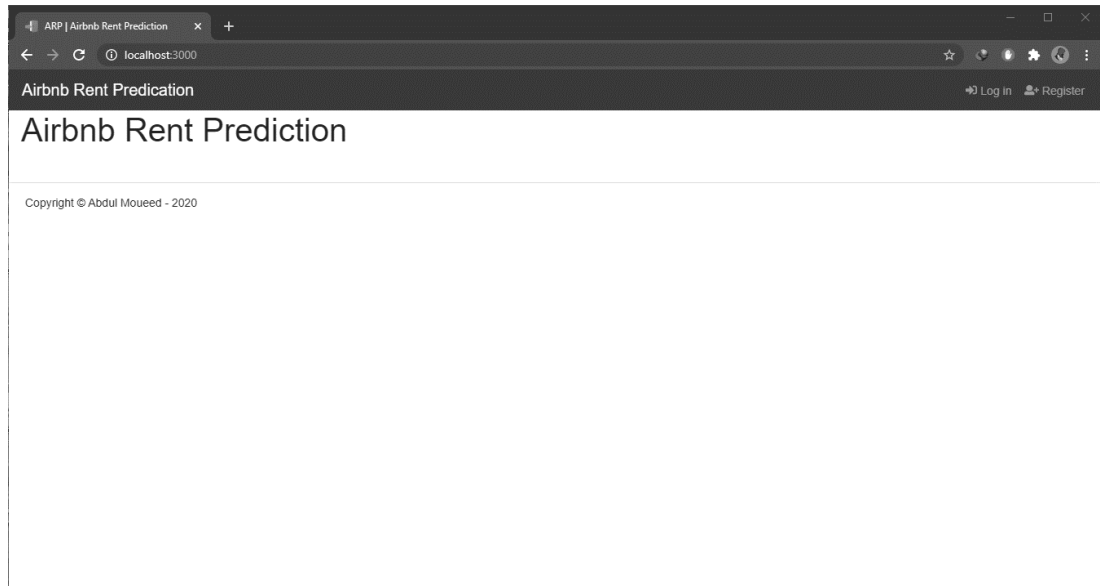
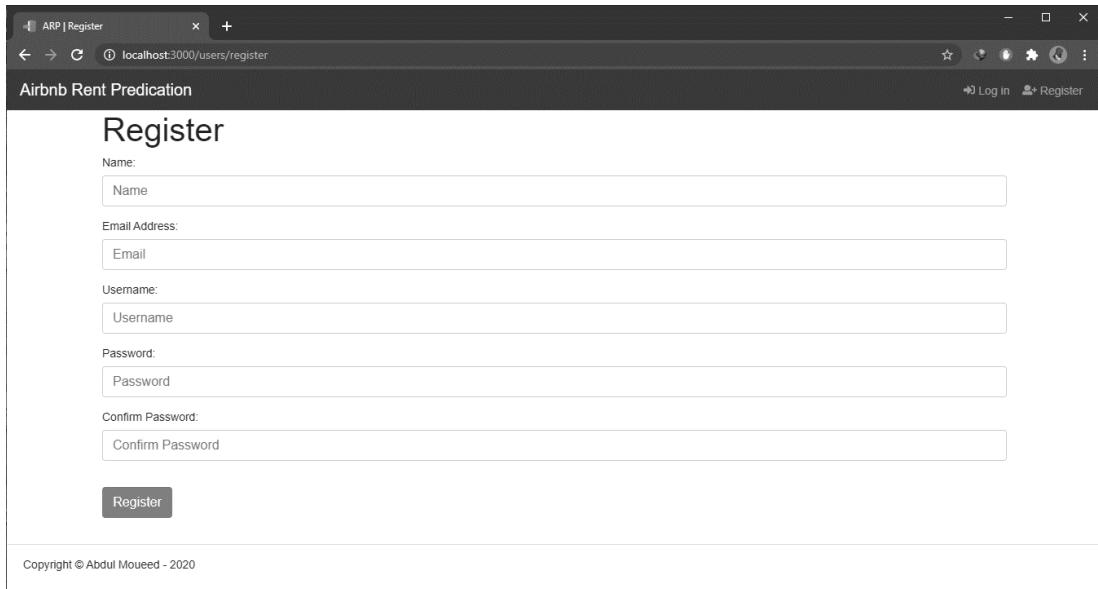


Figure 5.1: Main Page

5.1.2 Register

When a user clicks to register on the top right corner of the nav-bar a registration page will open as shown.



The screenshot shows a web browser window with the URL `localhost:3000/users/register`. The page title is "Airbnb Rent Predication". The main heading is "Register". The form contains the following fields:

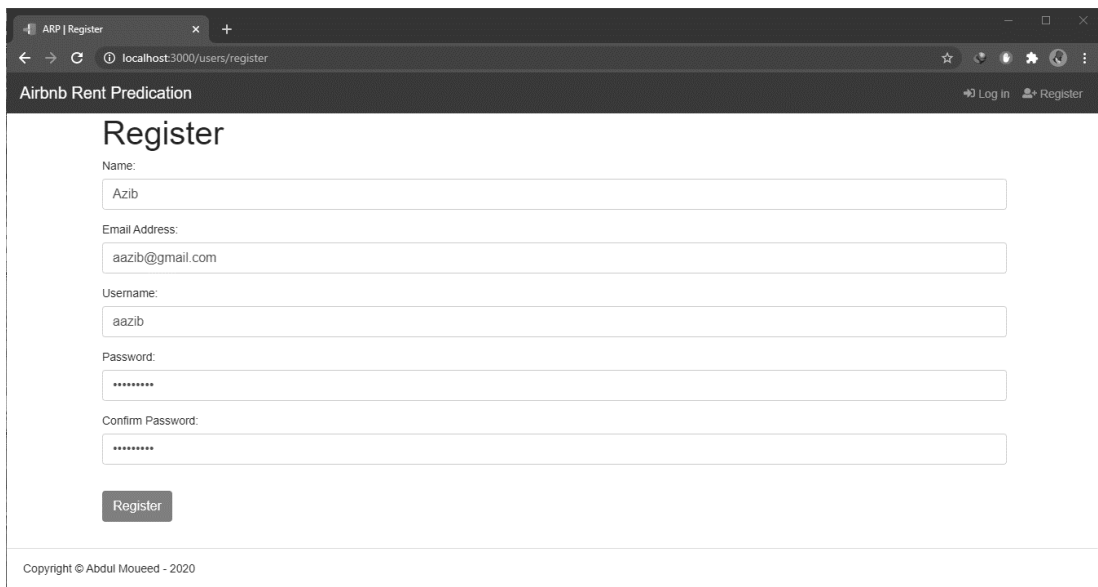
- Name:
- Email Address:
- Username:
- Password:
- Confirm Password:

A "Register" button is located below the form. The footer text reads "Copyright © Abdul Moueed - 2020".

Figure 5.2: Register

5.1.3 User Registration

In the shown figure a process is presented that a user can register to the website.



The screenshot shows the same web browser window as Figure 5.2, but with the registration form filled out. The fields contain the following data:

- Name:
- Email Address:
- Username:
- Password:
- Confirm Password:

The "Register" button is still visible below the form. The footer text remains "Copyright © Abdul Moueed - 2020".

Figure 5.3: User Register

5.1.4 Login

When a user is registered, then the user will simply enter the credential and login to the website.

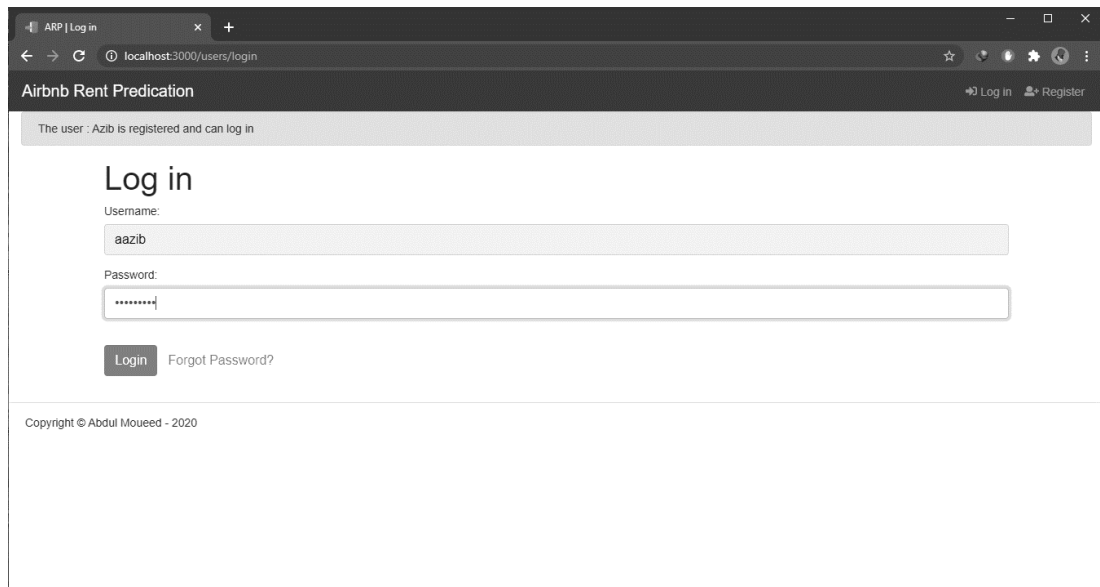


Figure 5.4: Login

5.1.5 Dashboard

After Login user is directed to the Dashboard of the website where the user will enter the housing details and forecast the price and get revenue generation details.

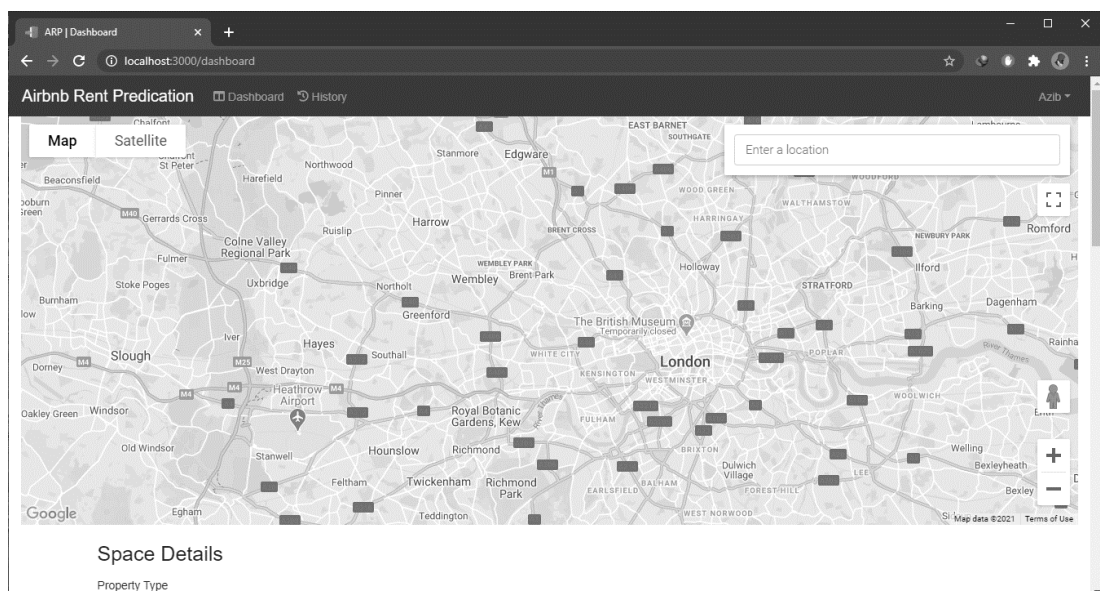


Figure 5.5: Dashboard 1



ARP | Dashboard

localhost:3000/dashboard

Property Type
House

Room Type
Entire House/Apt

Accommodates
1

Bathrooms
1

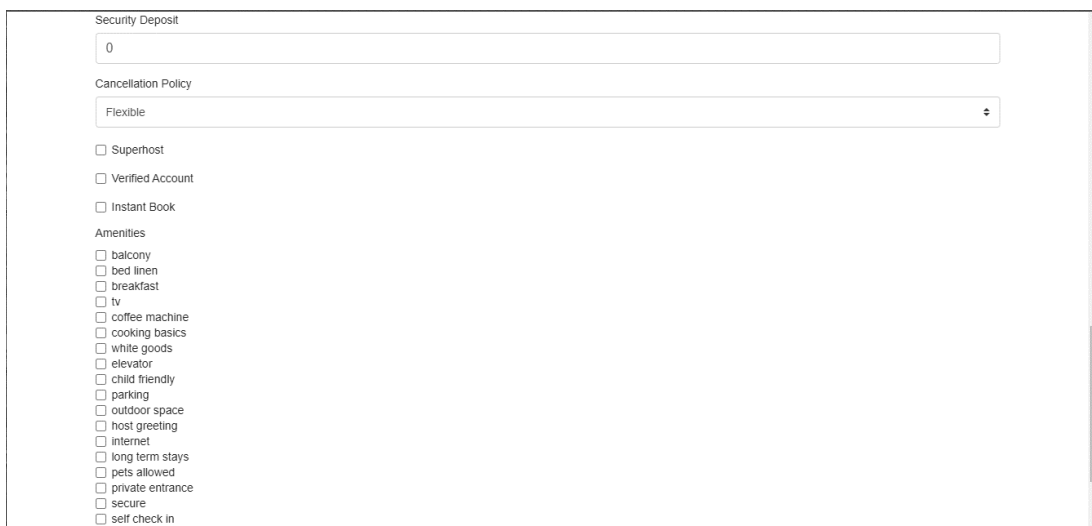
Cleaning Fees
0

Extra People
0

Maximum Nights
1

Minimum Nights
1

Figure 5.6: Dashboard 2



Security Deposit
0

Cancellation Policy
Flexible

Superhost

Verified Account

Instant Book

Amenities

- balcony
- bed linen
- breakfast
- tv
- coffee machine
- cooking basics
- white goods
- elevator
- child friendly
- parking
- outdoor space
- host greeting
- internet
- long term stays
- pets allowed
- private entrance
- secure
- self check in

Figure 5.7: Dashboard 3



Get Estimate

Copyright © Abdul Moueed - 2020

Figure 5.8: Dashboard 4

5.1.6 Profile

Here user can see his details and can also update the profile and change the password.

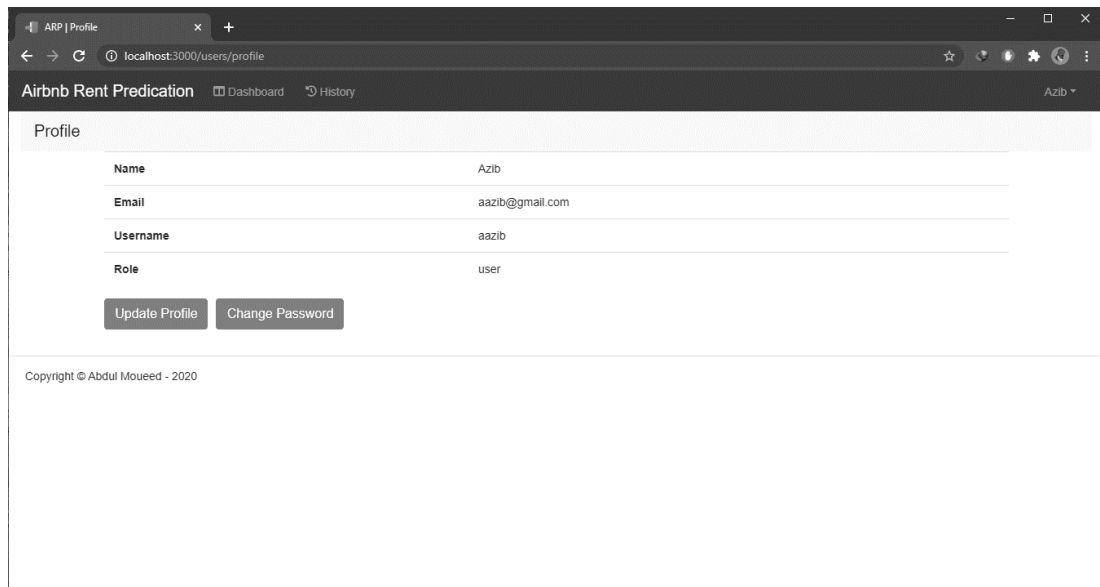


Figure 5.9: Profile

5.1.7 Update User Information

Here the user is updating his profile. User can update Name, username, Email.

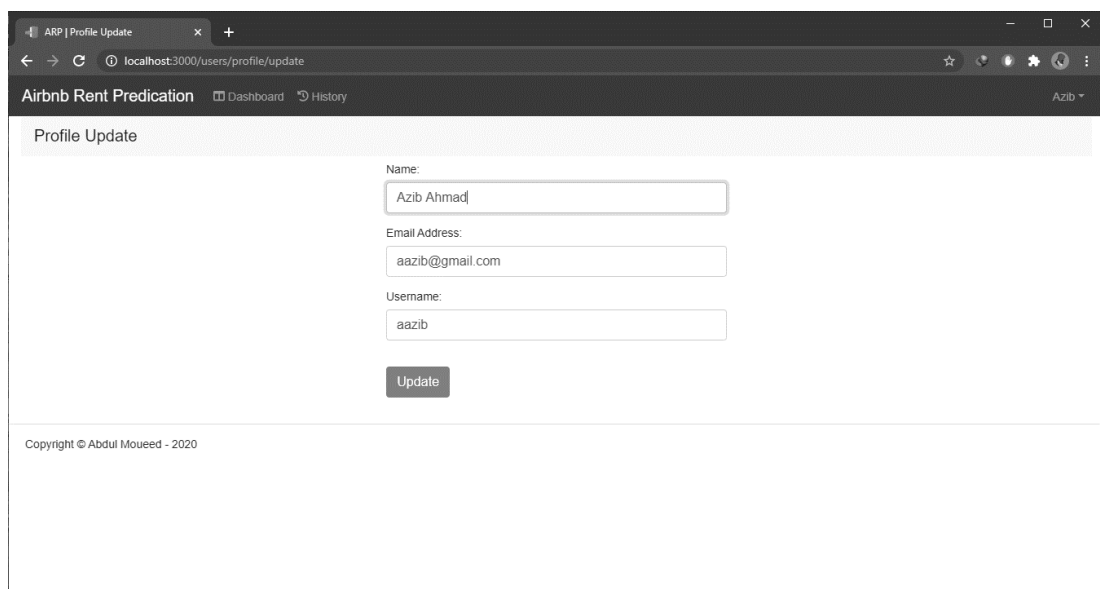


Figure 5.10: Update User Information

In the shown figure user updated their profile.

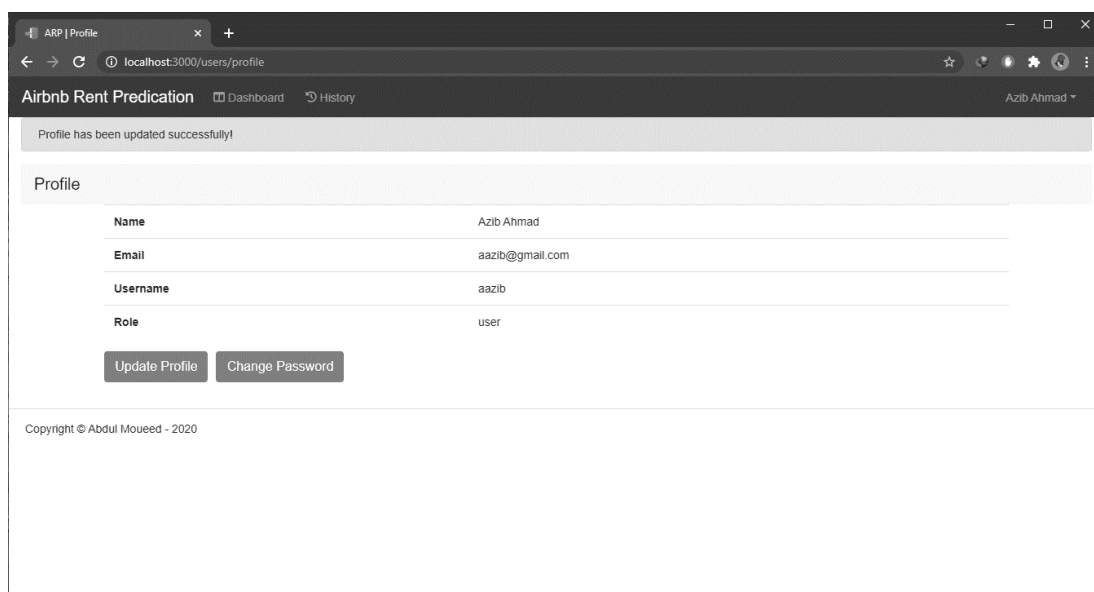


Figure 5.11: Updated User Information

5.1.8 Change Password

Here the user changed their password as shown in the figure.

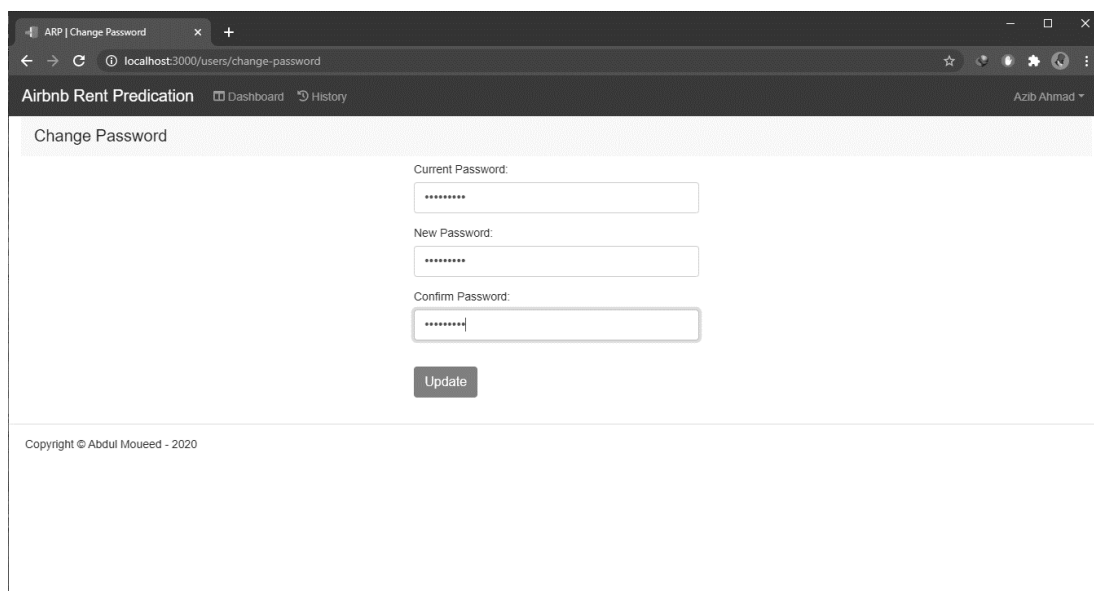


Figure 5.12: Change Password

5.1.9 Forecasting and Revenue

Here are the results of the prediction that the user enter information for his house rent recommendation.

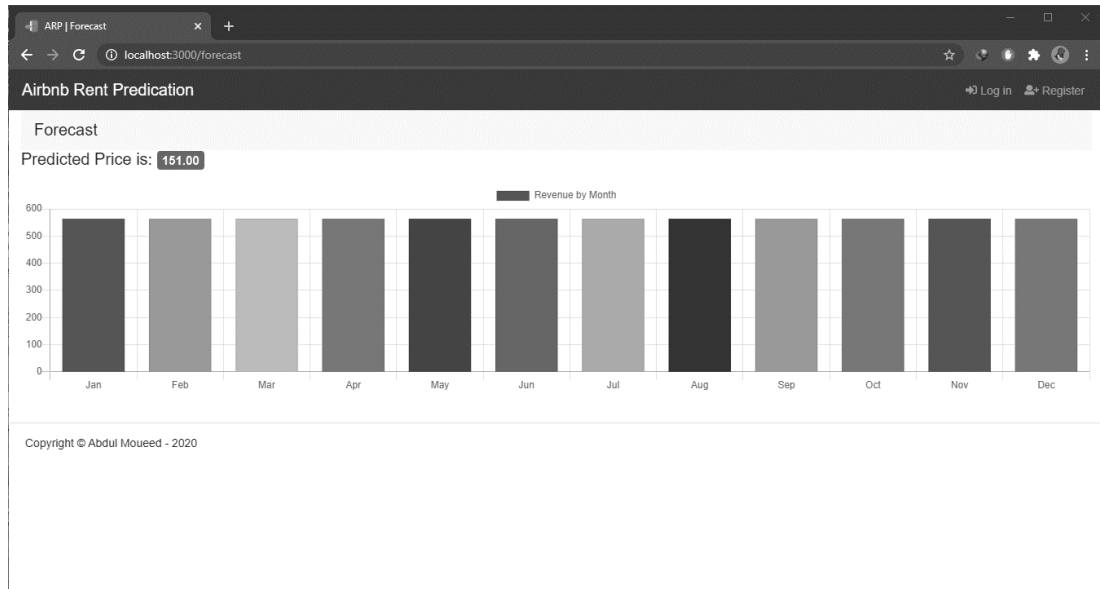


Figure 5.13: Revenue and Forecasting

5.1.10 History

Here is the history of predictions that the user had made in the past.

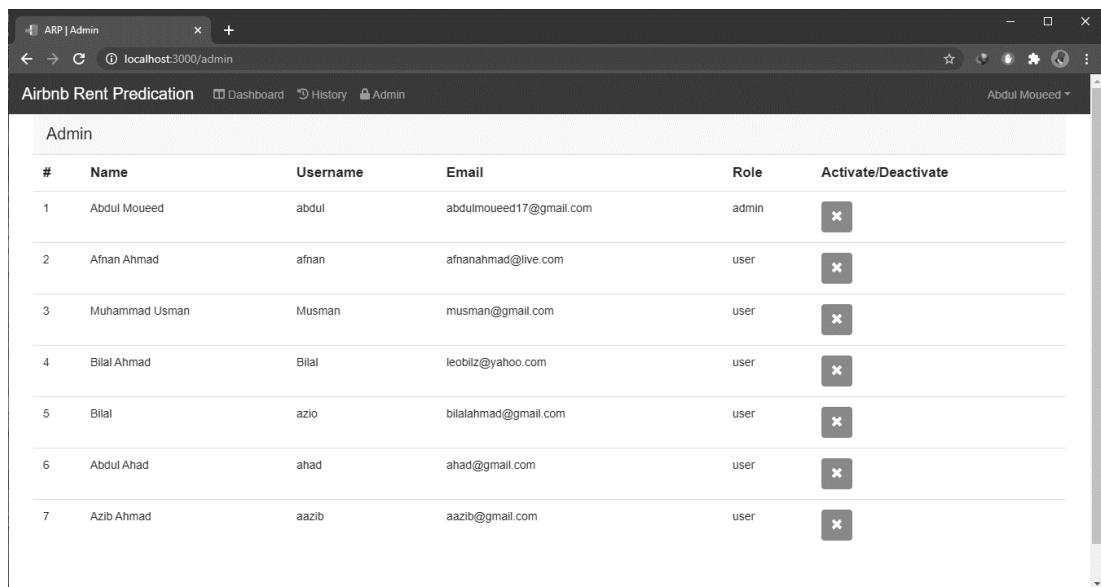
The screenshot shows a web browser window with the URL localhost:3000/history. The page title is 'Airbnb Rent Prediction'. Below the title, there is a 'History' section with a table of past predictions. The table has four columns: '#', 'Address', 'Prediction', and 'Date'. There are 5 rows of data. Below the table, there is a copyright notice: 'Copyright © Abdul Moueed - 2020'.

#	Address	Prediction	Date
1	221B Baker St Marylebone	216.99	Wed Dec 30 2020 16:33:34 GMT+0500 (Pakistan Standard Time)
2	GB Feltham Greater London	218.67	Thu Dec 31 2020 15:19:23 GMT+0500 (Pakistan Standard Time)
3	Feltham Railway Station Station Approach High St Feltham	316.60	Thu Dec 31 2020 15:20:26 GMT+0500 (Pakistan Standard Time)
4	221 Erith Rd Bexleyheath	203.87	Tue Jan 05 2021 02:17:21 GMT+0500 (Pakistan Standard Time)
5	221b Harrison Rd Leicester	205.45	Tue Jan 05 2021 02:18:57 GMT+0500 (Pakistan Standard Time)

Figure 5.14: History

5.1.11 Admin Panel

This is the admin Panel where Admin can block and unblock users. By clicking the Red Button User is disabled and Green Button appears when the admin click the green button user is enabled and can access the website.



#	Name	Username	Email	Role	Activate/Deactivate
1	Abdul Moueed	abdul	abdulmoueed17@gmail.com	admin	X
2	Afnan Ahmad	afnan	afnanahmad@live.com	user	X
3	Muhammad Usman	Musman	musman@gmail.com	user	X
4	Bilal Ahmad	Bilal	teobiliz@yahoo.com	user	X
5	Bilal	azio	bilalahmad@gmail.com	user	X
6	Abdul Ahad	ahad	ahad@gmail.com	user	X
7	Azib Ahmad	aazib	aazib@gmail.com	user	X

Figure 5.15: Admin Panel

5.2 Model Results

5.2.1 Numerical Features

Distribution of prices from 0 to 1000

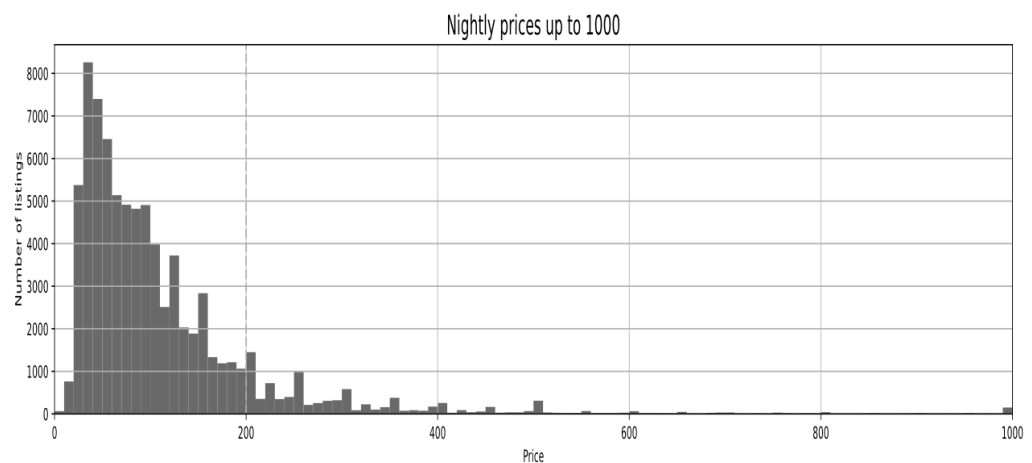


Figure 5.16: Numerical Features

Distribution of prices from 200 upwards

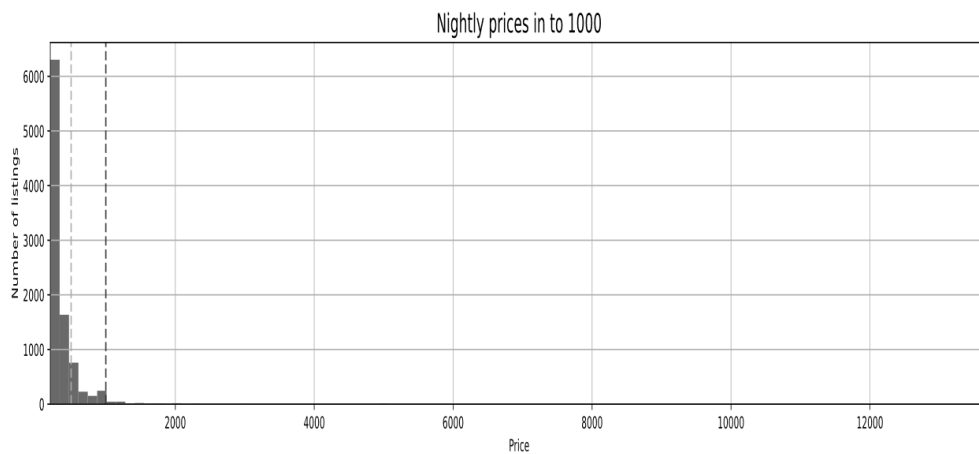


Figure 5.17: Numerical Features 1

Number of people accommodated bathrooms, bedrooms, and beds

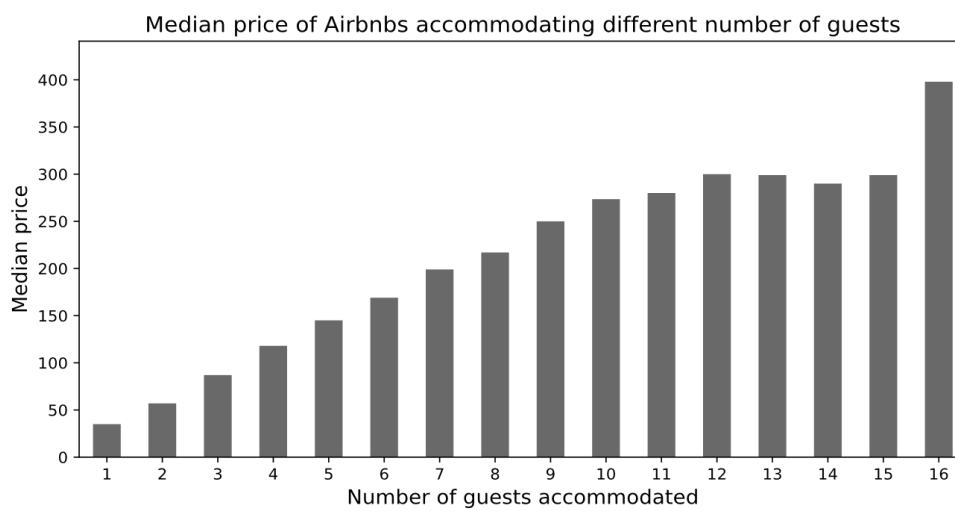


Figure 5.18: Numerical Features 2

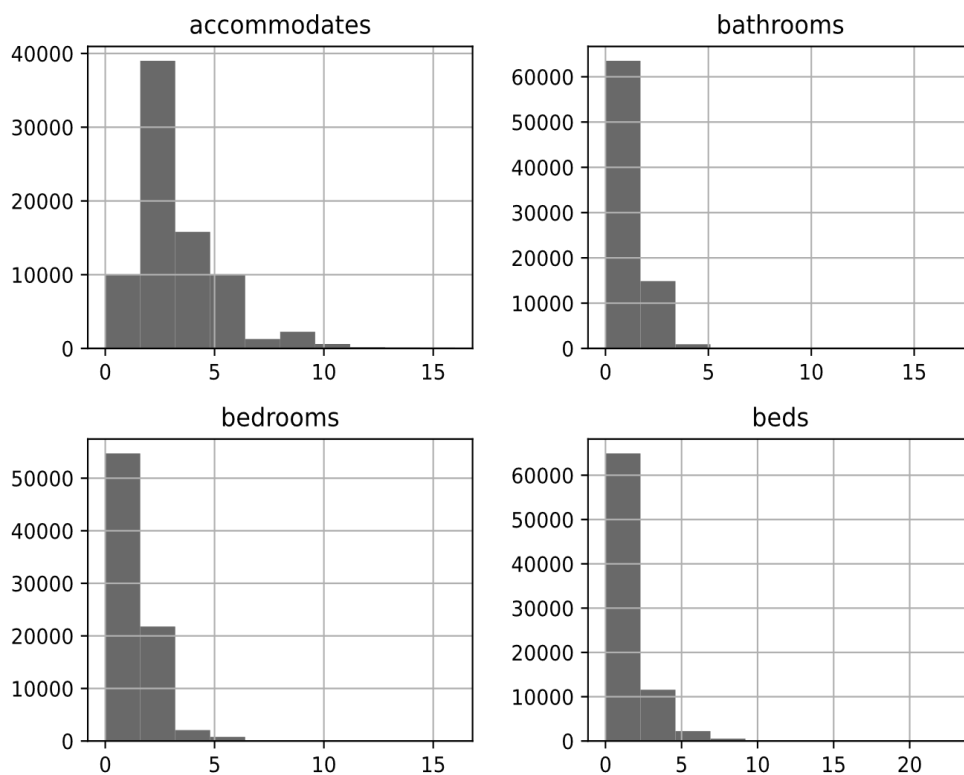


Figure 5.19: Numerical Features 3

5.2.2 Categorical features

Property and room types

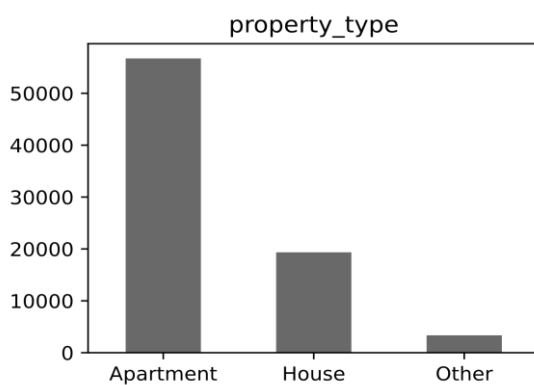


Figure 5.20: Property Type

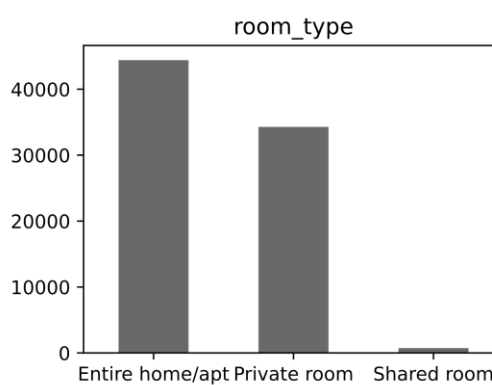


Figure 5.21: Room Type

5.2.3 Boolean Features

Here are the Boolean Features.

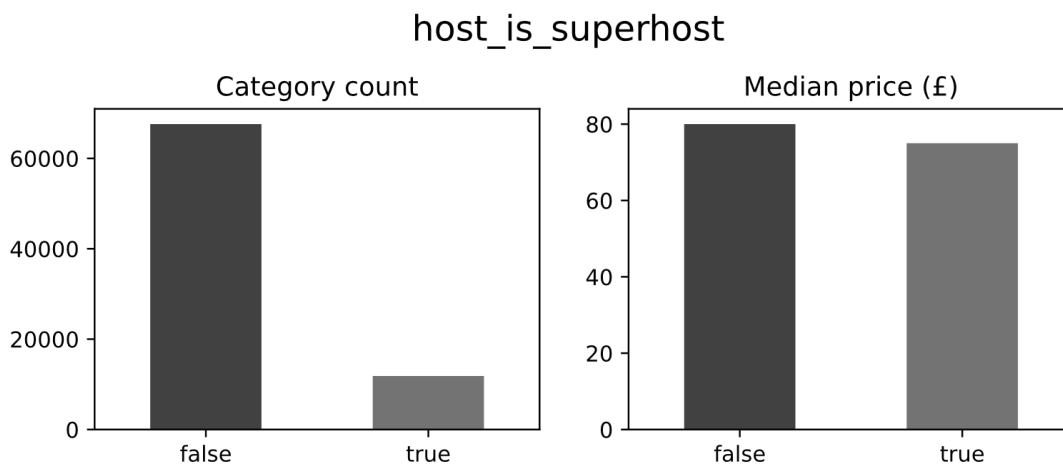


Figure 5.22: Host is Superhost

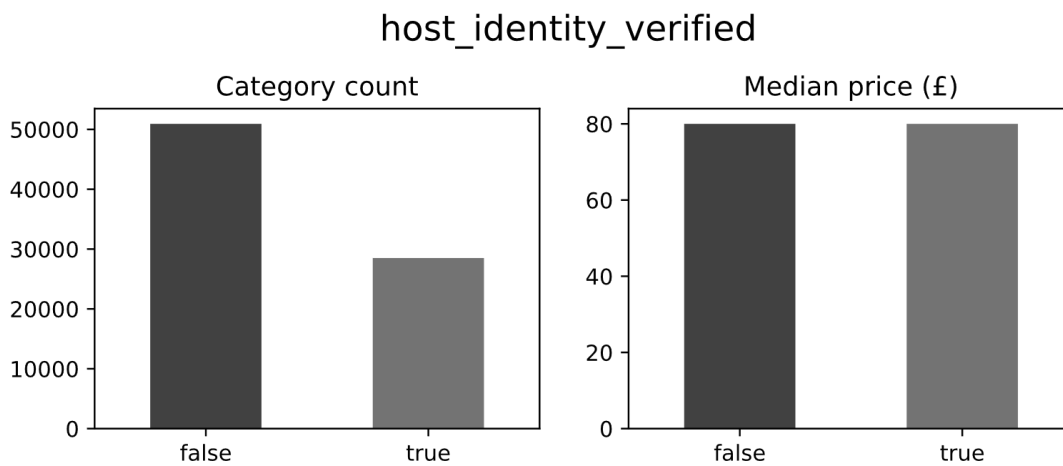


Figure 5.23: Host Identity Verified

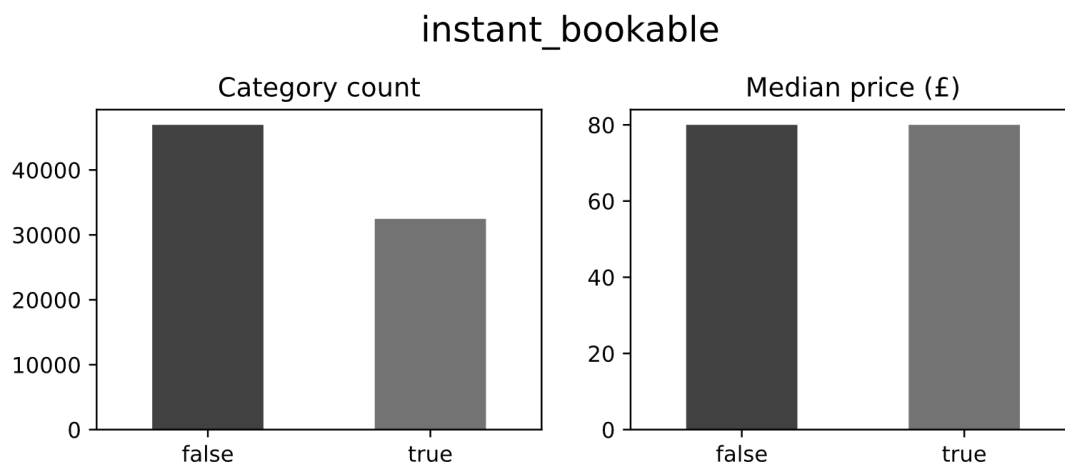


Figure 5.24: Instant Bookable

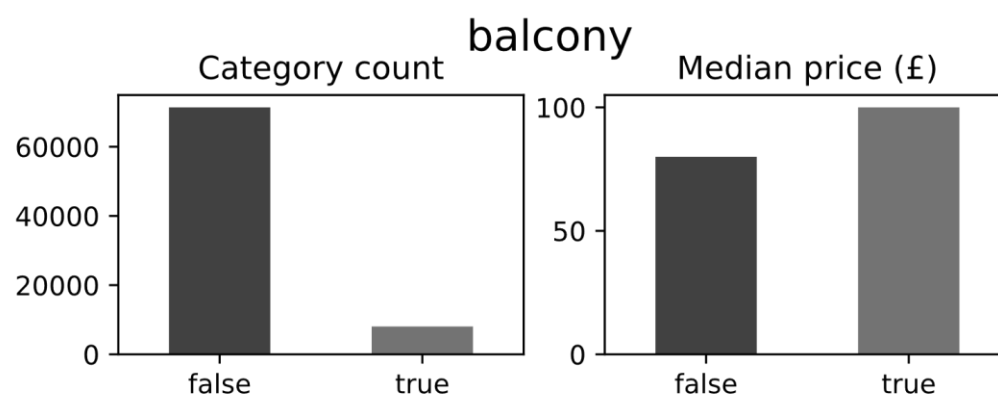


Figure 5.25: Balcony

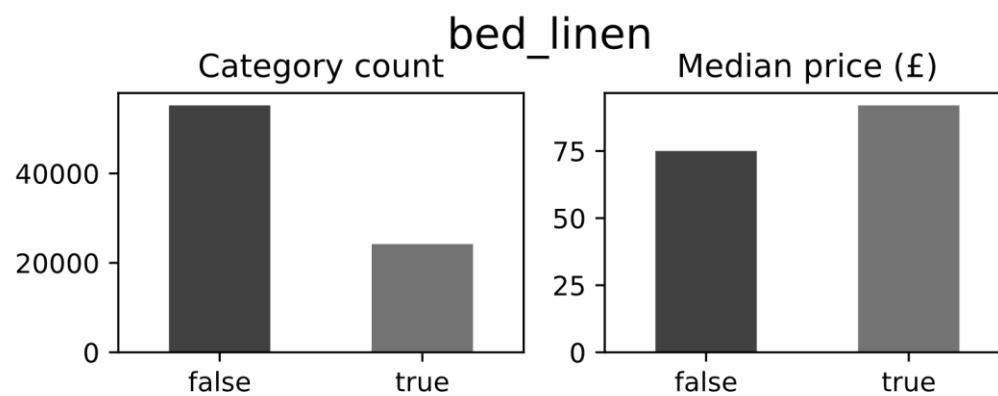


Figure 5.26: Bed Linen

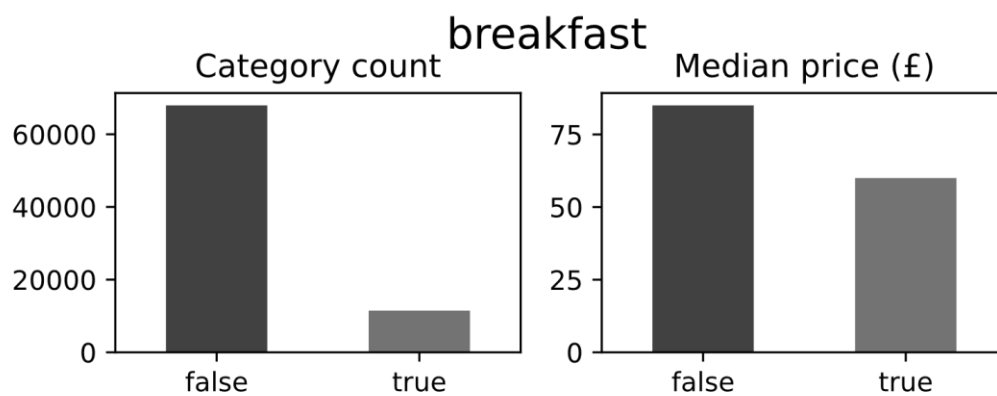


Figure 5.27: Breakfast

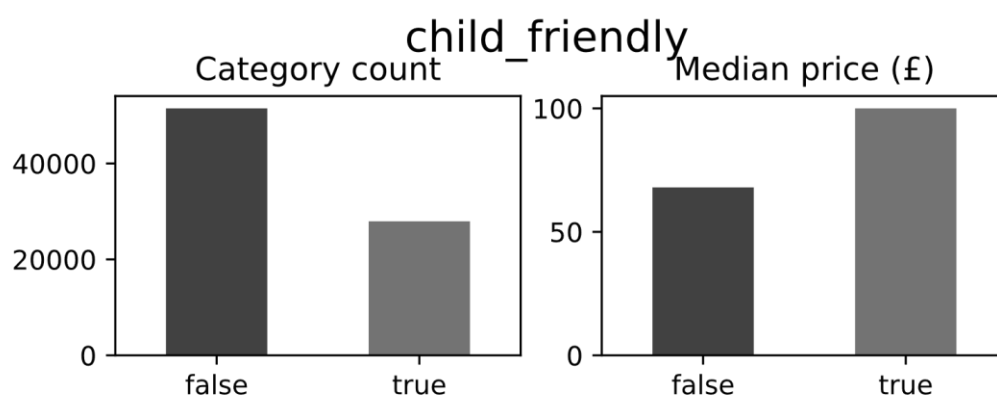


Figure 5.28: Child Friendly

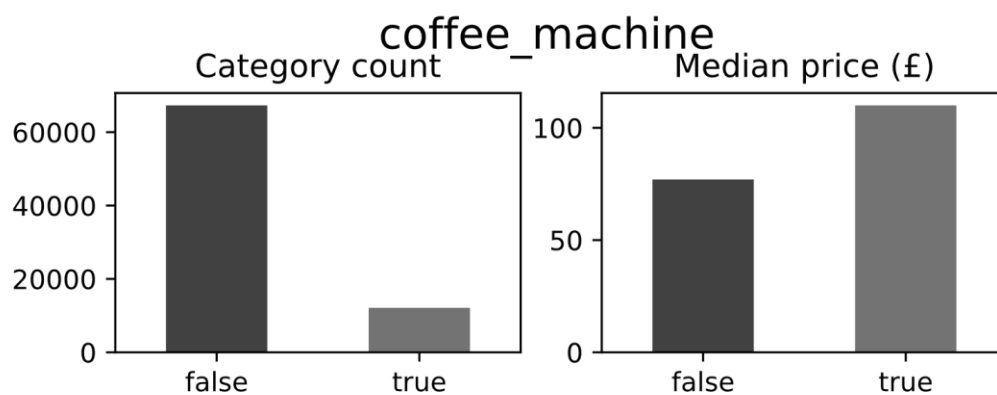


Figure 5.29: Coffee Machine

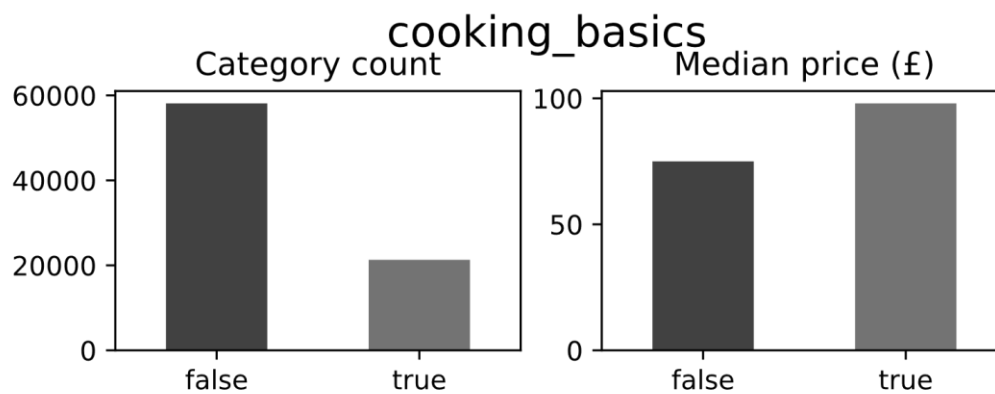


Figure 5.30: Cooking Basics

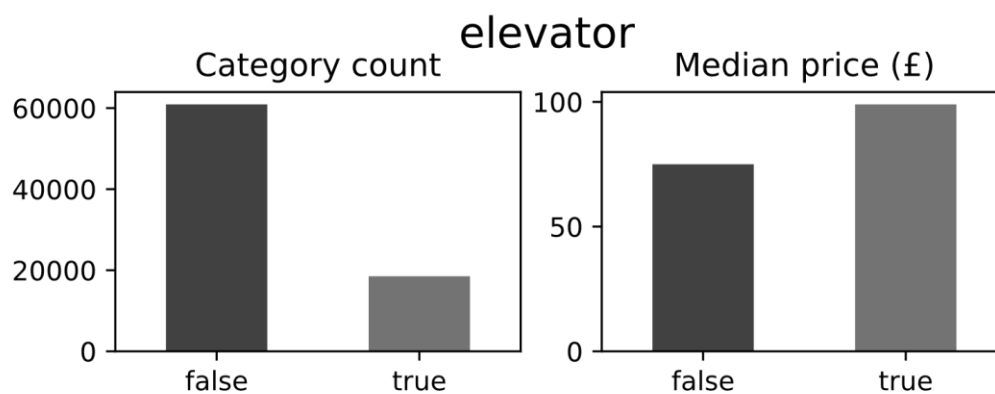


Figure 5.31: Elevator

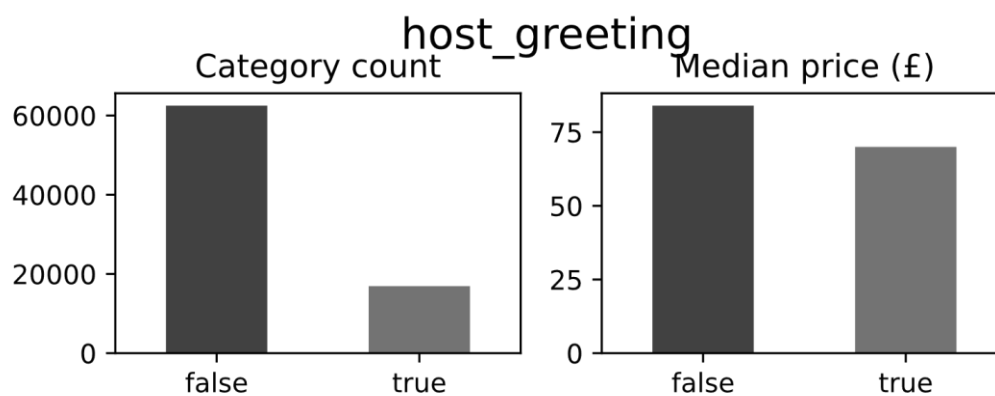


Figure 5.32: Host Greeting

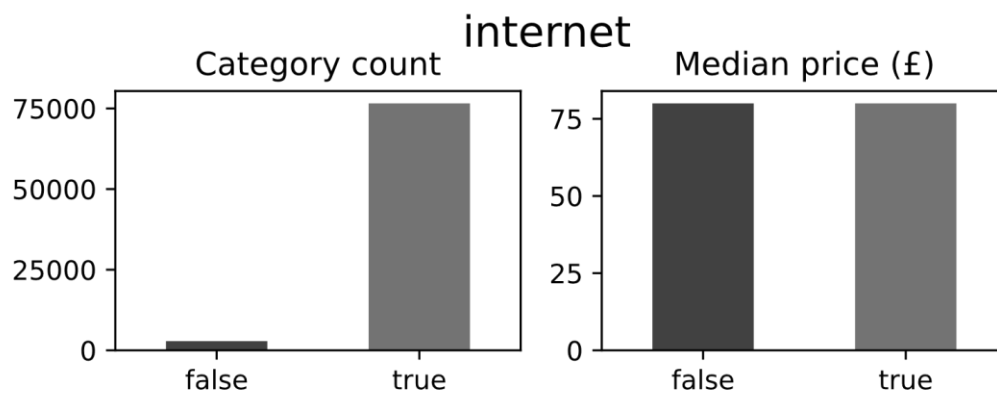


Figure 5.33: Internet

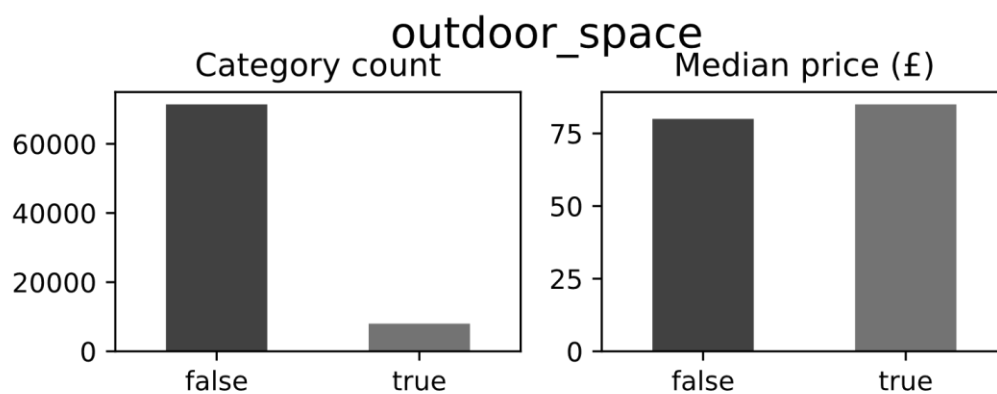


Figure 5.34: Outdoor Space

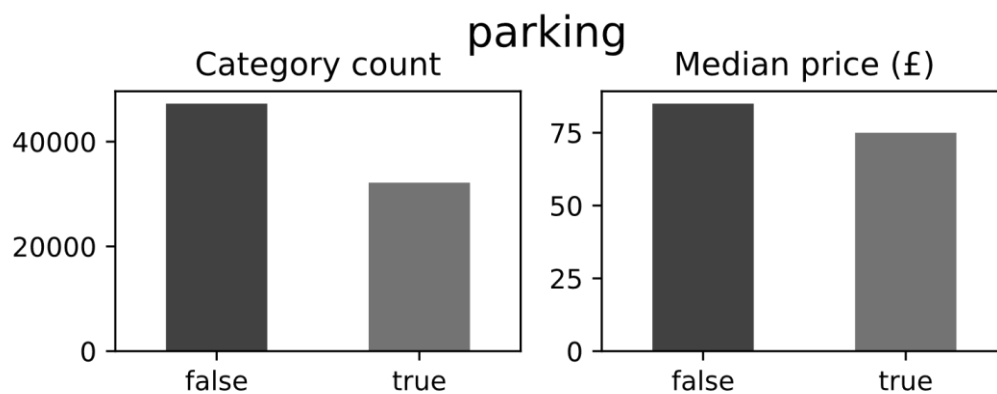


Figure 5.35: Parking

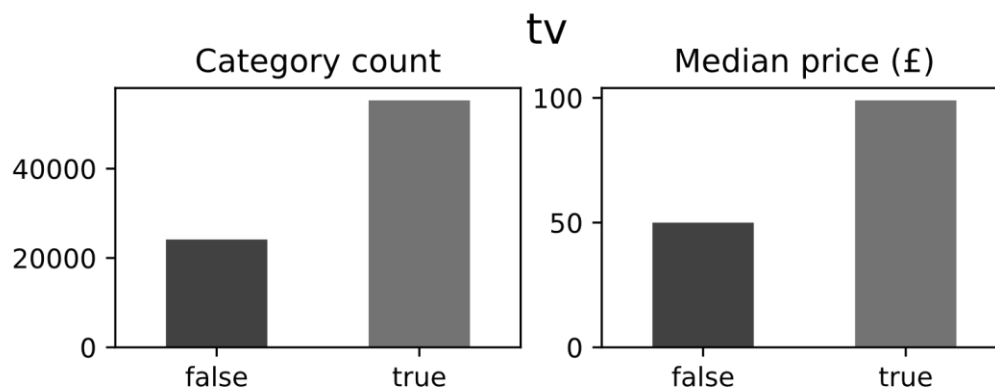


Figure 5.36: TV

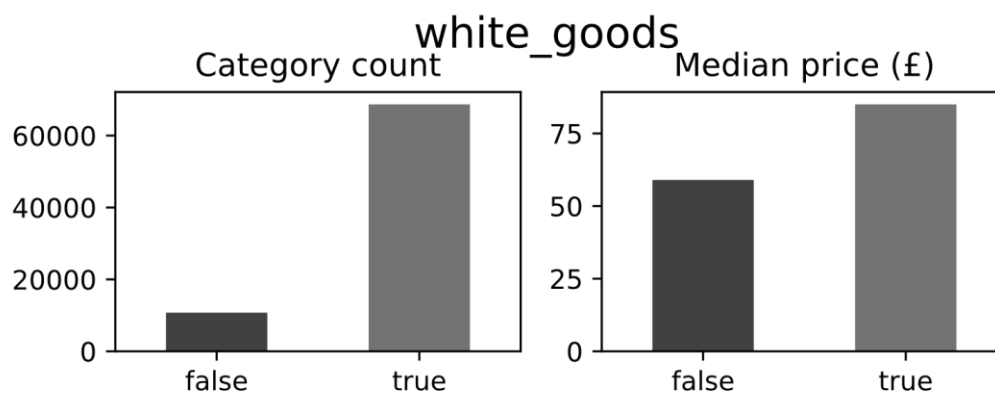


Figure 5.37: white Goods

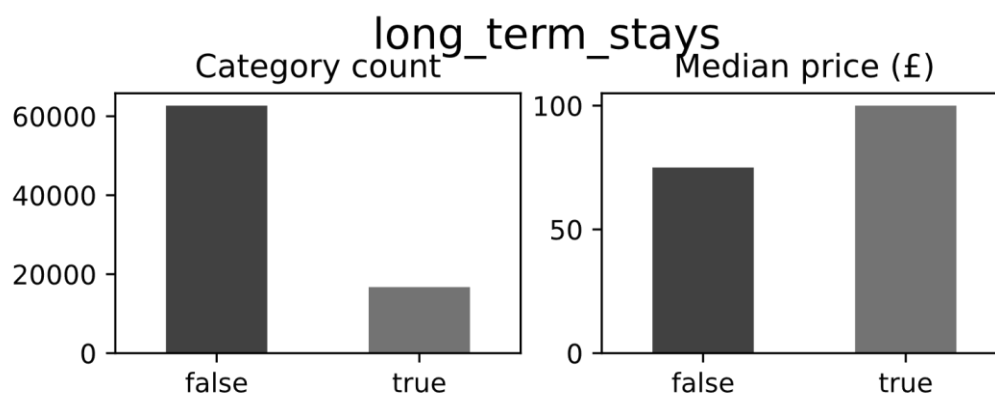


Figure 5.38: Long-term Stays

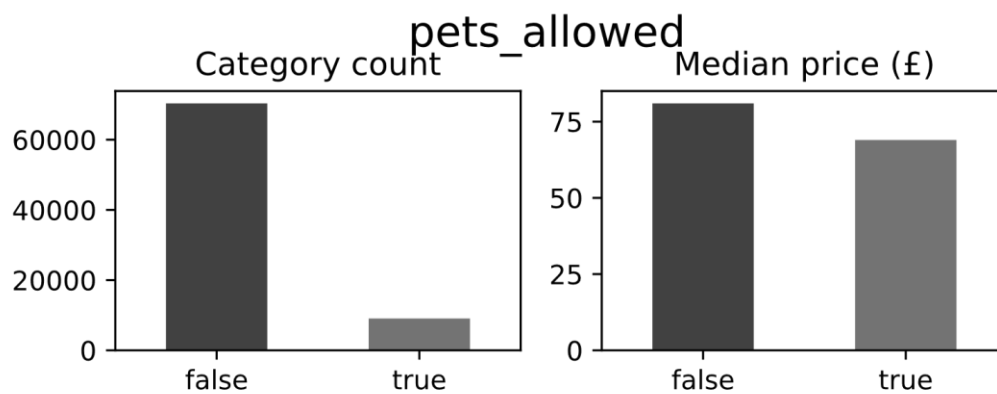


Figure 5.39: Pets Allowed

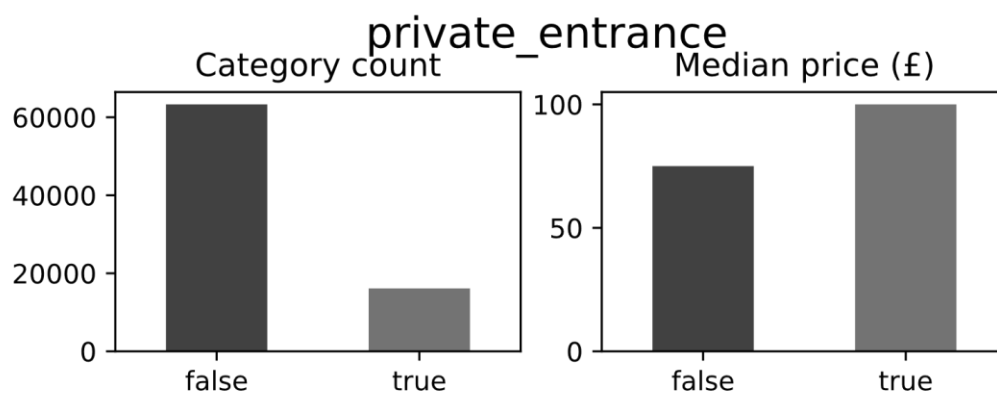


Figure 5.40: Private Entrance

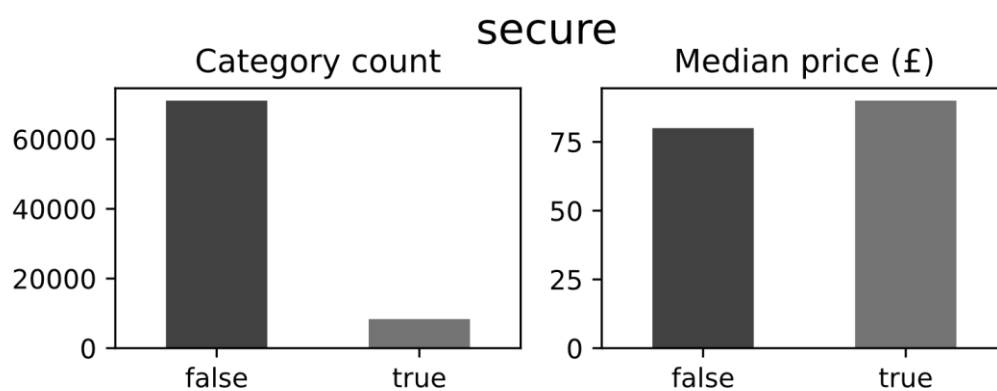


Figure 5.41: Secure

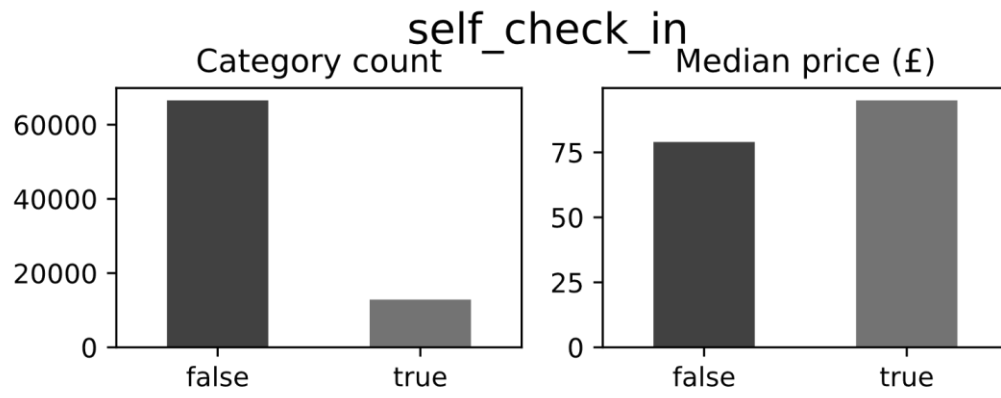


Figure 5.42: Self Check in

5.2.4 Important Features



Figure 5.43: Important Feature

5.2.5 Results

Table 5.1: Results

Model	Phase	RMSE	r²
Hedonic	Training	0.1946	0.669
	Testing	0.1896	0.671
	Cross-Validation		0.67
XGBoost	Training	0.0425	0.928
	Testing	0.1399	0.757
	Cross-Validation		0.74

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

This project aims to predict the price of Airbnb rental properties based on its features and amenities available. Using this tool, a user can sign up on the portal and can easily input property information to get a prediction result. We can see from the results that XGBoost performs way better than the Hedonic model. The model can predict 74% of the variation in price with an RMSE value of 0.14.

The current model is only trained with a dataset gathered from one city. The accuracy can be improved by scraping/collecting datasets for more cities and training models again. Also, we can improve the accuracy of the model by doing sentiment analysis of the review and assigning a score. This can be included as a feature during the training model.

Accessibility is another important feature that can be included in the training model. The current dataset lacks this information, this is because Airbnb doesn't allow downloading/scraping dataset. There are a few paid Airbnb Data collection companies which offer high-quality dataset which can be acquired through future funding to allow more development on this project.

REFERENCES

- [1] *Airbnb Data Collection*. (2019, 9 16). Retrieved from Airbnb:
<https://github.com/tomslee/airbnb-data-collection>
- [2] *Inside Airbnb*. (2020, 1 31). Retrieved from <http://insideairbnb.com/get-the-data.html>
- [3] *Spatial Regression*. (2020, 01 29). Retrieved from Geographic Data Science with PySAL and the pydata stack:
http://darribas.org/gds_scipy16/ipynb_md/08_spatial_regression.html
- [4] *Sopranzetti, B.* (2010). Hedonic Regression Analysis in Real Estate Markets: A Primer. *Handbook of Quantitative Finance and Risk Management*, pp.1201-1207.
- [5] *Cobián Álvarez, J. and Resosudarmo, B.* (2019). The cost of floods in developing countries' megacities: a hedonic price analysis of the Jakarta housing market, Indonesia. *Environmental Economics and Policy Studies*, 21(4), pp.555-577.
- [6] *Neloy, A., Haque, H. and Ul Islam, M.* (2019). Ensemble Learning Based Rental Apartment Price Prediction Model by Categorical Features Factoring. *Proceedings of the 2019 11th International Conference on Machine Learning and Computing - ICMLC '19*.