HOTEL SCRAPPER

A PROJECT REPORT

Submitted by

JINDAL CHANDRAKANT PRIYADARSHI [20BECE30215]

PRIYA PRAVINBHAI SUTARIYA [20BECE30256]

SURYAPRATAPSINH BALVANTSINH THAKOR [20BECE30264]

KRISHN BHARATBHAI THAKRAR [20BECE30265]

In fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

Computer Engineering

LDRP Institute of Technology and Research, Gandhinagar

Kadi Sarva Vishwavidyalaya

April, 2023

LDRP INSTITUTE OF TECHNOLOGY AND RESEARCH GANDHINAGAR

CE-IT Department

CERTIFICATE

This is to certify that the Project Work entitled **"Hotel Scrapper"** has been carried out by **Jindal Chandrakant Priyadarshi (20BECE30215)** under my guidance in fulfilment of the degree of Bachelor of Engineering in Computer Engineering Semester-6 of Kadi Sarva Vishwavidyalaya University during the academic year 2022-2023.

Prof. Hitesh Patel	Dr. Sandip Modha
Internal Guide	**Head of the Department**
LDRP ITR	**LDRP ITR**

LDRP INSTITUTE OF TECHNOLOGY AND RESEARCH GANDHINAGAR

CE-IT Department

CERTIFICATE

This is to certify that the Project Work entitled **"Hotel Scrapper"** has been carried out by **Priya Pravinbhai Sutariya (20BECE30256)** under my guidance in fulfilment of the degree of Bachelor of Engineering in Computer Engineering Semester-6 of Kadi Sarva Vishwavidyalaya University during the academic year 2022-2023.

Prof. Hitesh Patel	Dr. Sandip Modha
Internal Guide	**Head of the Department**
LDRP ITR	**LDRP ITR**

LDRP INSTITUTE OF TECHNOLOGY AND RESEARCH GANDHINAGAR

CE-IT Department

CERTIFICATE

This is to certify that the Project Work entitled **"Hotel Scrapper"** has been carried out by **Suryapratapsinh Balvantsinh Thakor (20BECE30264)** under my guidance in fulfilment of the degree of Bachelor of Engineering in Computer Engineering Semester-6 of Kadi Sarva Vishwavidyalaya University during the academic year 2022-2023.

Prof. Hitesh Patel	Dr. Sandip Modha
Internal Guide	**Head of the Department**
LDRP ITR	**LDRP ITR**

LDRP INSTITUTE OF TECHNOLOGY AND RESEARCH GANDHINAGAR

CE-IT Department

CERTIFICATE

This is to certify that the Project Work entitled **"Hotel Scrapper"** has been carried out by **Krishn Bharatbhai Thakrar (20BECE30265)** under my guidance in fulfilment of the degree of Bachelor of Engineering in Computer Engineering Semester-6 of Kadi Sarva Vishwavidyalaya University during the academic year 2022-2023.

Prof. Hitesh Patel	Dr. Sandip Modha
Internal Guide	**Head of the Department**
LDRP ITR	**LDRP ITR**

Second Presentation for Project-II

1. Name & Signature of Internal Guide	**Prof. Hitesh Patel**
2. Comments from Panel Members	
3. Name & Signature of Panel Members	

ACKNOWLEDGEMENT

I take this opportunity to humbly express our thankfulness to all those concerned with my project.

First of all I am thankful to **LDRP ITR** for undertaking this project. I am sincerely indebted to **Prof. Hitesh Patel** for giving me the opportunity to work on this project. Her continuous guidance and help have proved to be a key to my success in overcoming the challenges that I have faced during my project work. Her support made the project a pleasantly memorable one. Without her help at all stages in spite of her own work load; the completion of the project would not have been possible.

I express my sincere gratitude to **Prof. Hitesh Patel** for his valuable guidance and positive feedback.

There are so many persons without whose help I would never have conceived and learnt, to whom I would like to express my gratitude – my friends, colleagues, and of course CE & IT Department of LDRP ITR.

Last but not least I am thankful to almighty GOD and my PARENTS for giving me such a good atmosphere to work hard and to succeed.

JINDAL PRIYADARSHI (20BECE30215)

PRIYA SUTARIYA (20BECE30256)

SURAYAPRATAPSINH THAKOR (20BECE30264)

KRISHN THAKRAR (20BECE30265)

ABSTRACT

Web Scraping is a technique to extract large amount of data from several websites. The term "scraping" refers to obtaining the information from another source and saving it into a local file. so here we are working on project based on hotels where you require the name of hotel and price of hotel rooms make comparison between the different hotels. If you collect these details by checking various sites, it will take much time. In that case web scraping plays important role where by writing few lines of code you can get the desired results.

TABLE OF CONTENTS

CHAPTER NAME	PAGE NO
Acknowledgement	**I**
Abstract	**II**
Table of Contents	**III**
List of Figures	**IV**
List of Tables	**V**
1. Introduction 1.1 Introduction 1.2 Aims and Objective of the work 1.3 Problem definition 1.4 Plan of the work	**1**
2. Technology and Literature Review	**2**
3. System Requirements Study 3.1 User Characteristics 3.2 Hardware and Software Requirements 3.3 Assumptions and dependencies	**4**
4. System Diagrams	**5**
5. Data Dictionary	**13**
6. Result, Discussion and Conclusion	**14**
7. References	**17**

LIST OF FIGURES

NO	NAME	PAGE NO
1	Gantt Chart	5
2	Class Diagram	6
3	Use Case Diagram	7
4	Sequence Diagram	8
5	Entity Relationship Diagram	9
6	Context Level Data Flow Diagram	10
7	First Level Data Flow Diagram	10
8	Second Level Data Flow Diagram	11
9	Activity Diagram	12

LIST OF TABLES

NO	NAME	PAGE NO
1	Admin Table	13
2	User Table	13

1. Introduction

1.1 Introduction

- As a result of scaping we can use data from different website and from take some useful insights from it. So here we use web scraping for our personal use and get data of different hotels.

- What we do in this project that we fetch the price and name of hotels from different hotel booking sites and make comparison between price of hotels.

1.2 Aims and Objective of the work

- So through this Hotel Scrapper website we are trying to show comparison of price of hotels from different hotel booking sites in one particular city so they can choose hotel according to their requirement.

1.3 Problem Definition

- The problem is that for checking the price of one particular hotel in a city we have to go through different hotel booking websites and that is much time taking process so here we are trying to solve that problem and make available that data on one website.

1.4 Plan of the work

- Here first when user open the website then they have to first create new account and they have to give some required details and if they are already registered user then then can do direct login and can show the main page of website.

- After that, main page of website is open and in that page hotel scrapping data is available and they can show of comparison of prices of more than one hotels of one particular city.

2. Technology and Literature Review

Technology:

html:

- HTML stands for Hyper Text Markup Language. It is the standard markup language for creating Web pages. It describes the structure of a Web page and consists of a series of elements. Also it contains elements that tells the browser how to display the content.

CSS:

- CSS stands for Cascading Style Sheets and is used to style web pages. For example, to alter the font, color, size, and spacing of your content, split it into multiple columns, or add animations and other decorative features.

js:

- JavaScript is used to create dynamic and interactive web content like applications and browsers and used as a client-side programming language.

React.js:

- It is an open-source front end JavaScript framework for building user interfaces and single-page applications.

Django:

- Django is a high-level Python web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source.

MongoDB:

- MongoDB is a document database used to build highly available and scalable internet applications. Instead of storing data in tables of rows or columns like SQL databases, each record in a MongoDB database is a document described in BSON, a binary representation of the data. Applications can then retrieve this information in a JSON format.

MySQL:

- MySQL is a relational database management system.

Python's Beautiful soup Library :

- Beautiful Soup is a python package which allows us to pull a python package which allows us to pull data out of HTML and XML documents.

Express api:

- Express is a popular framework for Node.js, intended for use in building web applications and APIs.

Literature Review:

- In recent years, the digital transformation has fundamentally changed numerous industries. Among the examples are media (news portals, streaming services), retailing (mail ordering and online groceries), tourism (online travel agencies, sharing platforms), and banking (online banking, smart payments).

- Similar changes may be observed in the hospitality sector, a part of the larger tourism industry that comprises the hotel industry. Over the past years, this field has seen the rise of digital innovations, increased competition and, most recently, it has experienced the decisive event of the Corona pandemic.

3. System Requirement Study

3.1 User Characteristics

- Register page display information of registration process
- Login page display details to be logged in.
- Main page display data of hotels from different hotel booking sites.

3.2 Hardware and Software requirements

Hardware Requirements:

- Any browser supported device
- Dedicated server
- Internet

Software Requirements:

- Latest browsers with v8 engine
- 2gb RAM

3.3 Assumptions and Dependencies

- User have to create new account on our website.
- After that they can log in.
- And then they can show the available data in websites.

4 System Diagrams

Gantt Chart

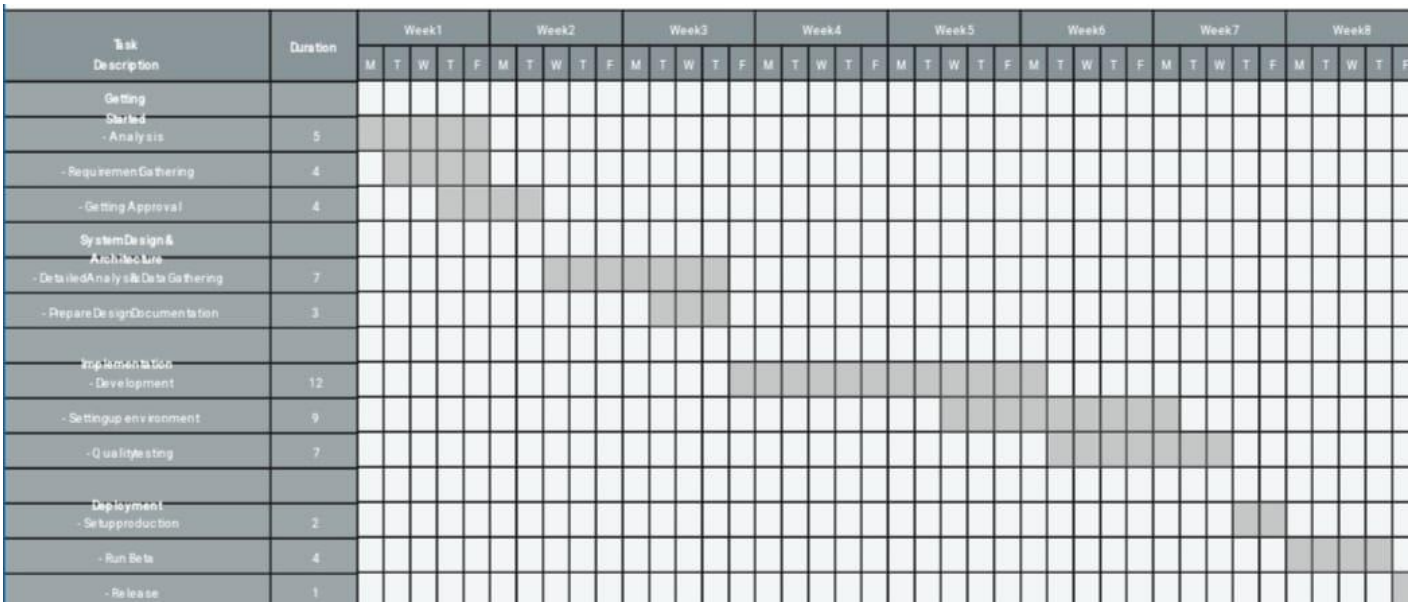

Fig – 4.1 : Gantt Chart

Class Diagram

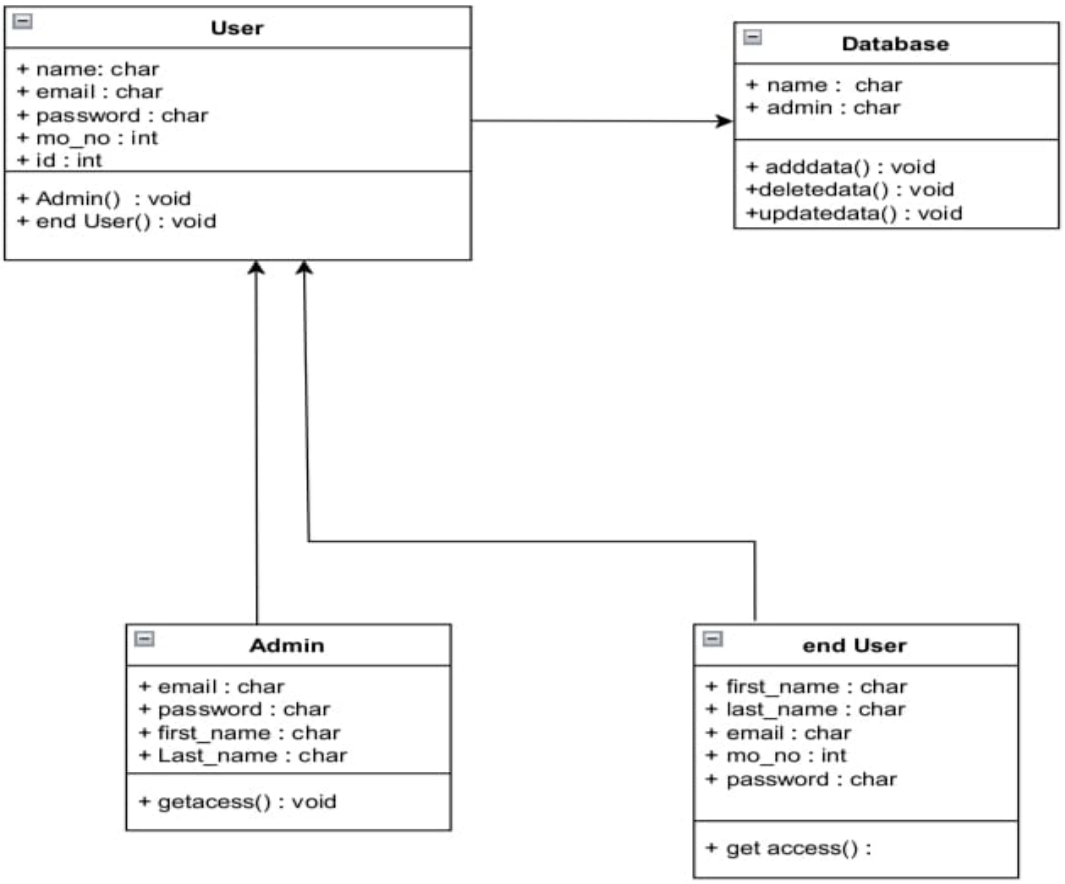

Fig – 4.2 : Class Diagram

Use Case Diagram

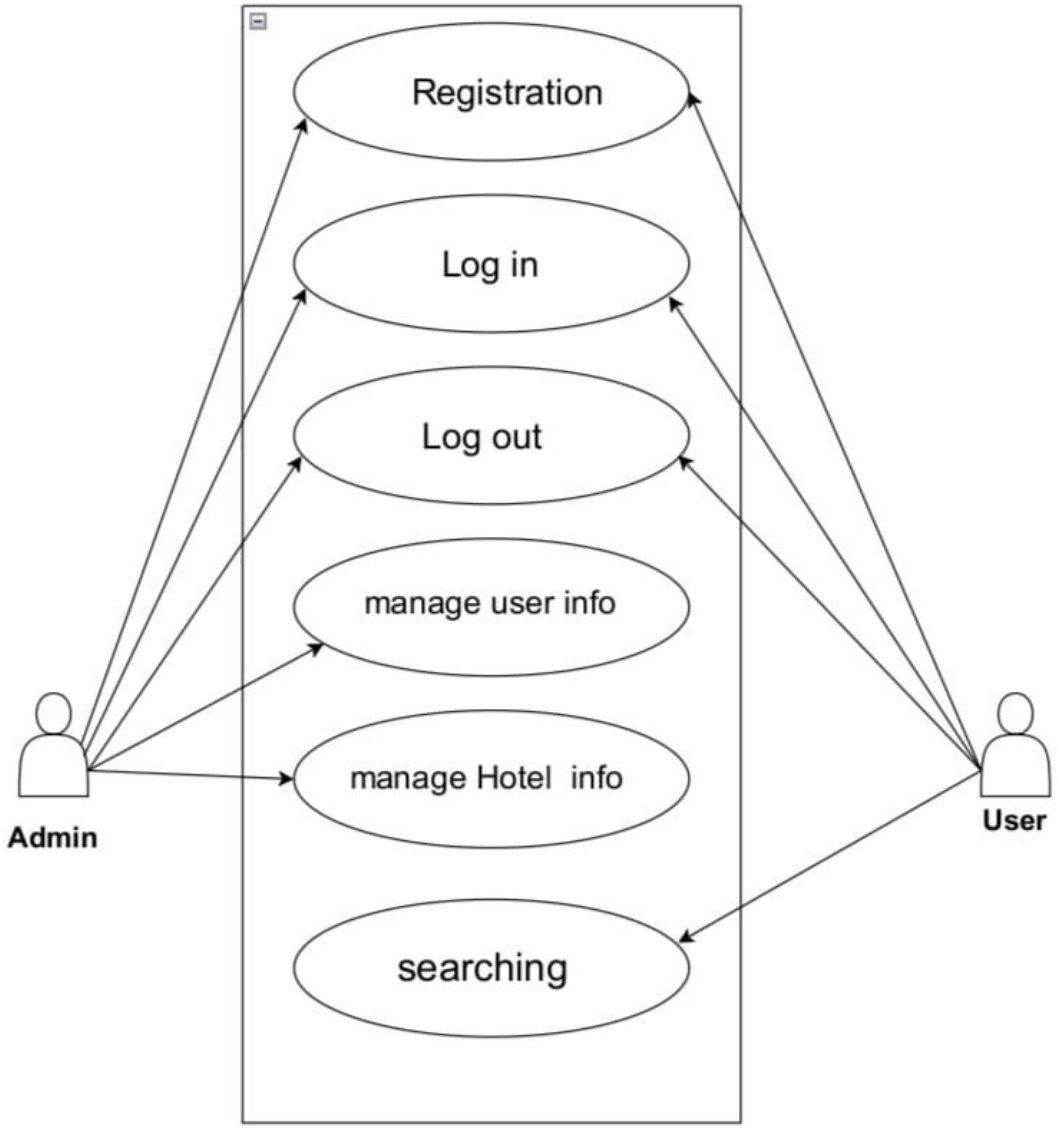

Fig – 4.3 : Use Case Diagram

Sequence Diagram

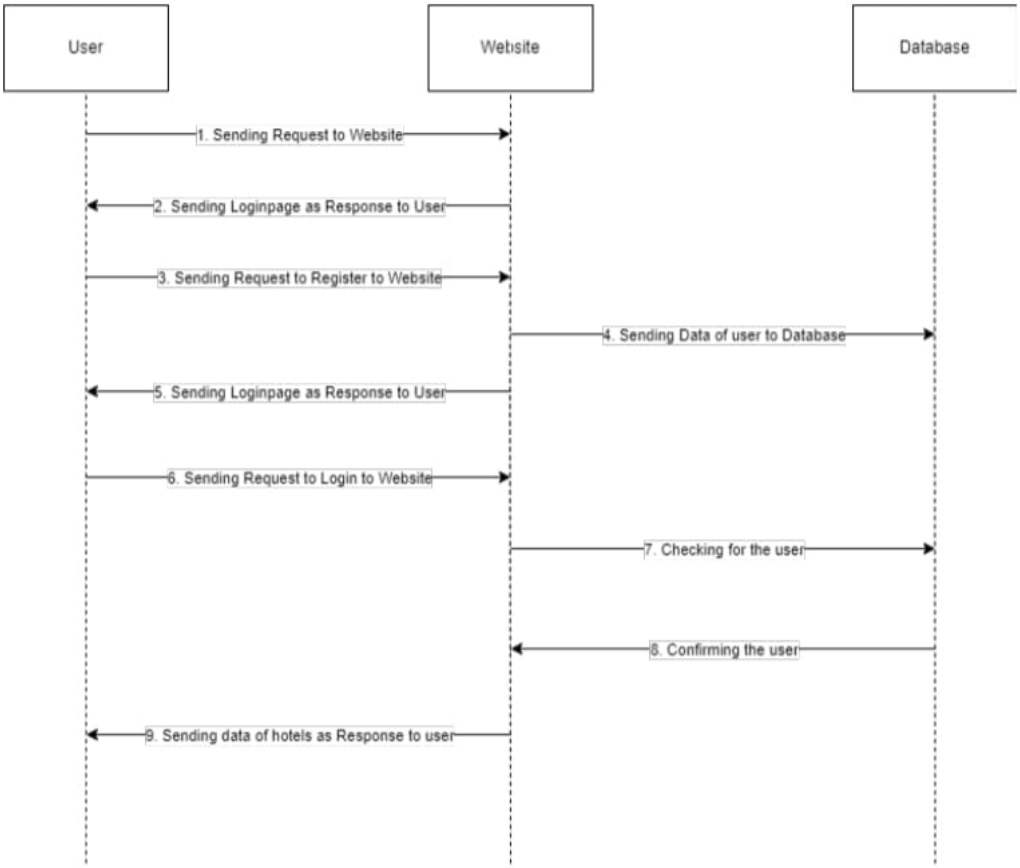

Fig – 4.4 : Sequence Diagram

Entity Relationship Diagram

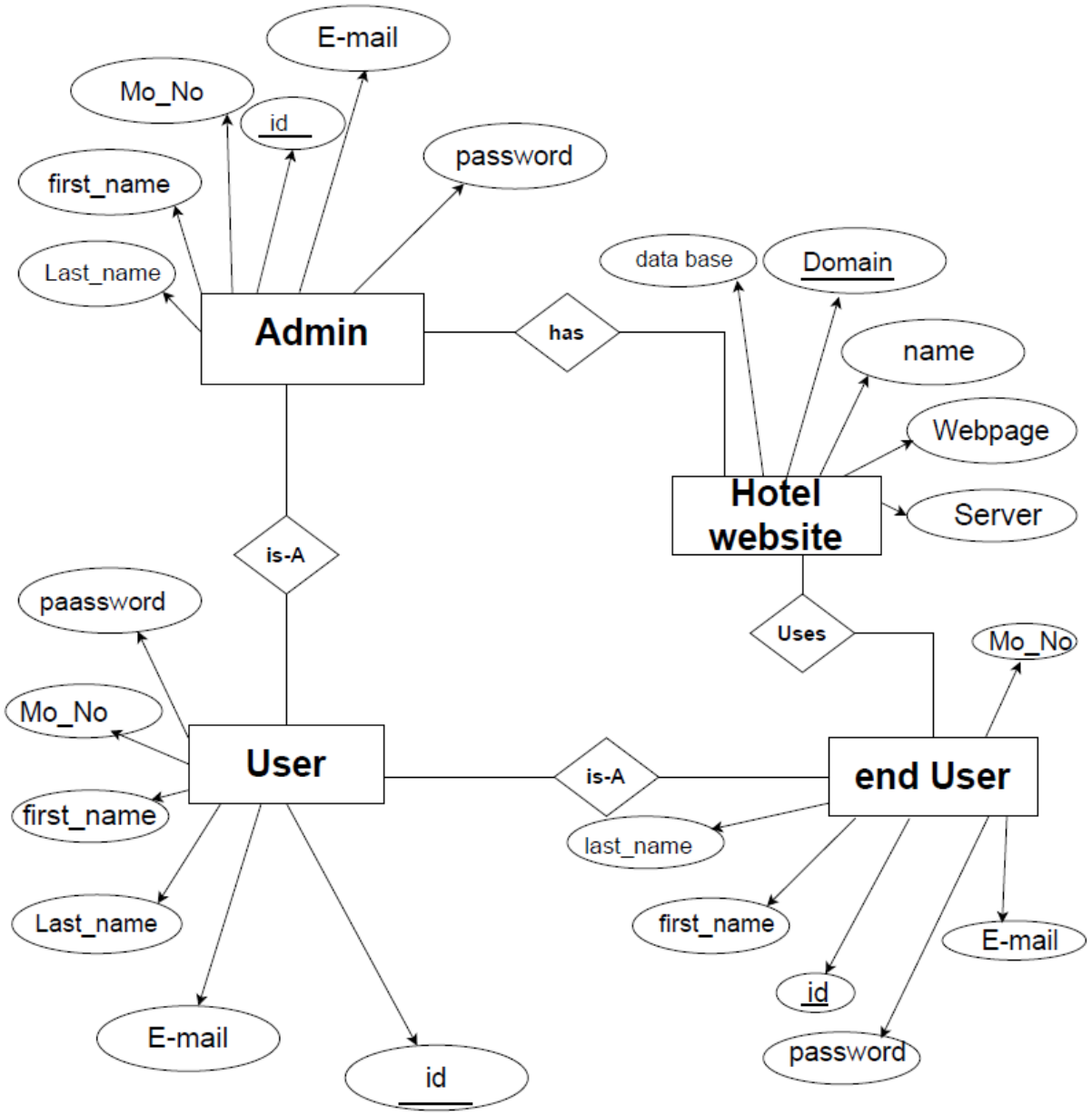

Fig – 4.5 : Entity Relationship Diagram

Context Level Data Flow Diagram

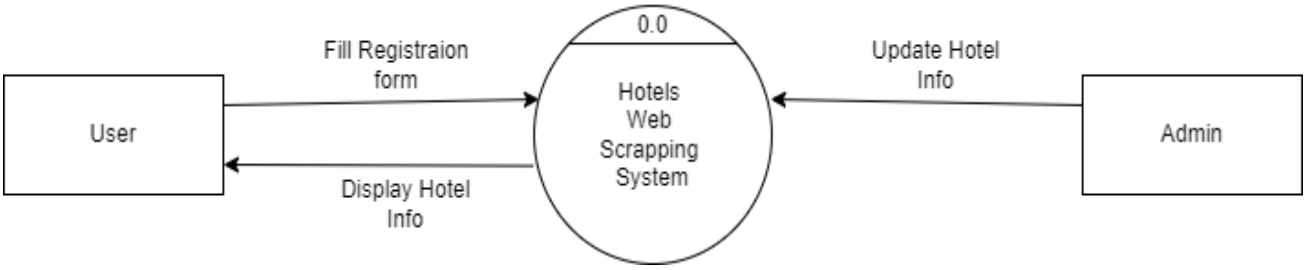

Fig – 4.6.1 : Context Level Data Flow Diagram

First Level Data Flow Diagram

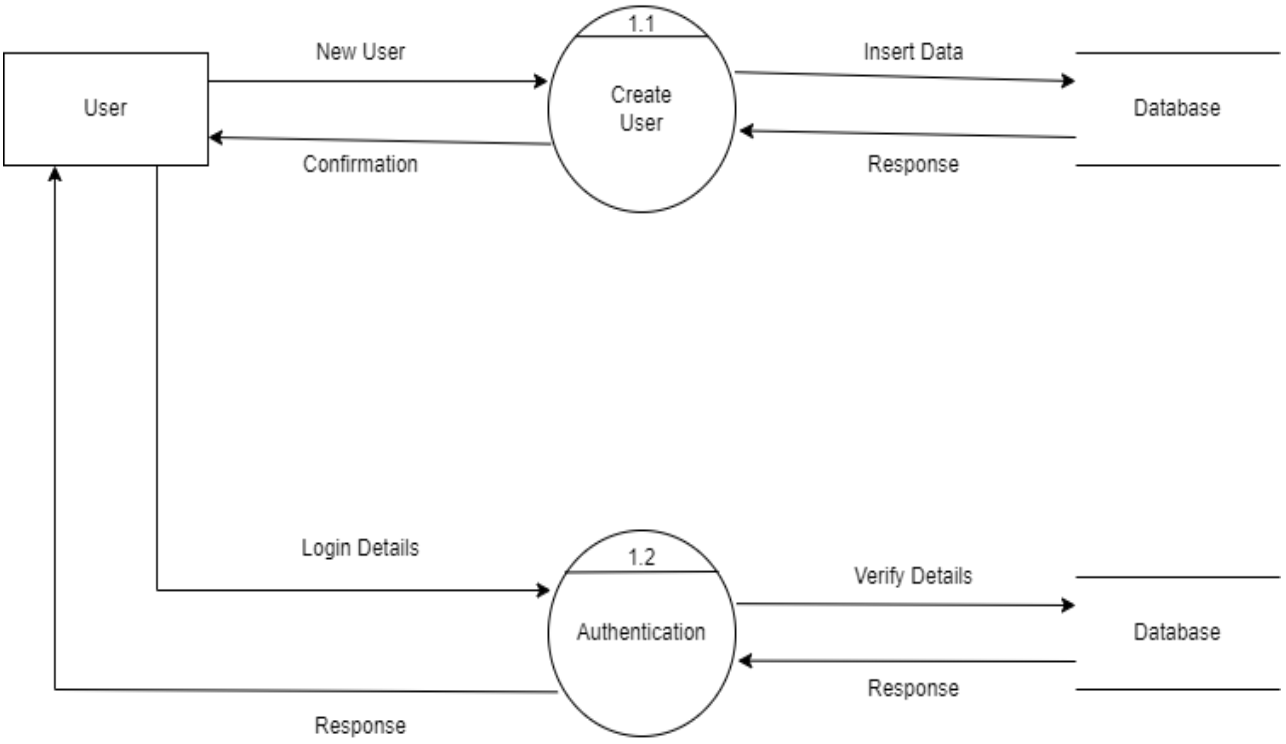

Fig – 4.6.2 : First Level Data Flow Diagram

Second Level Data Flow Diagram

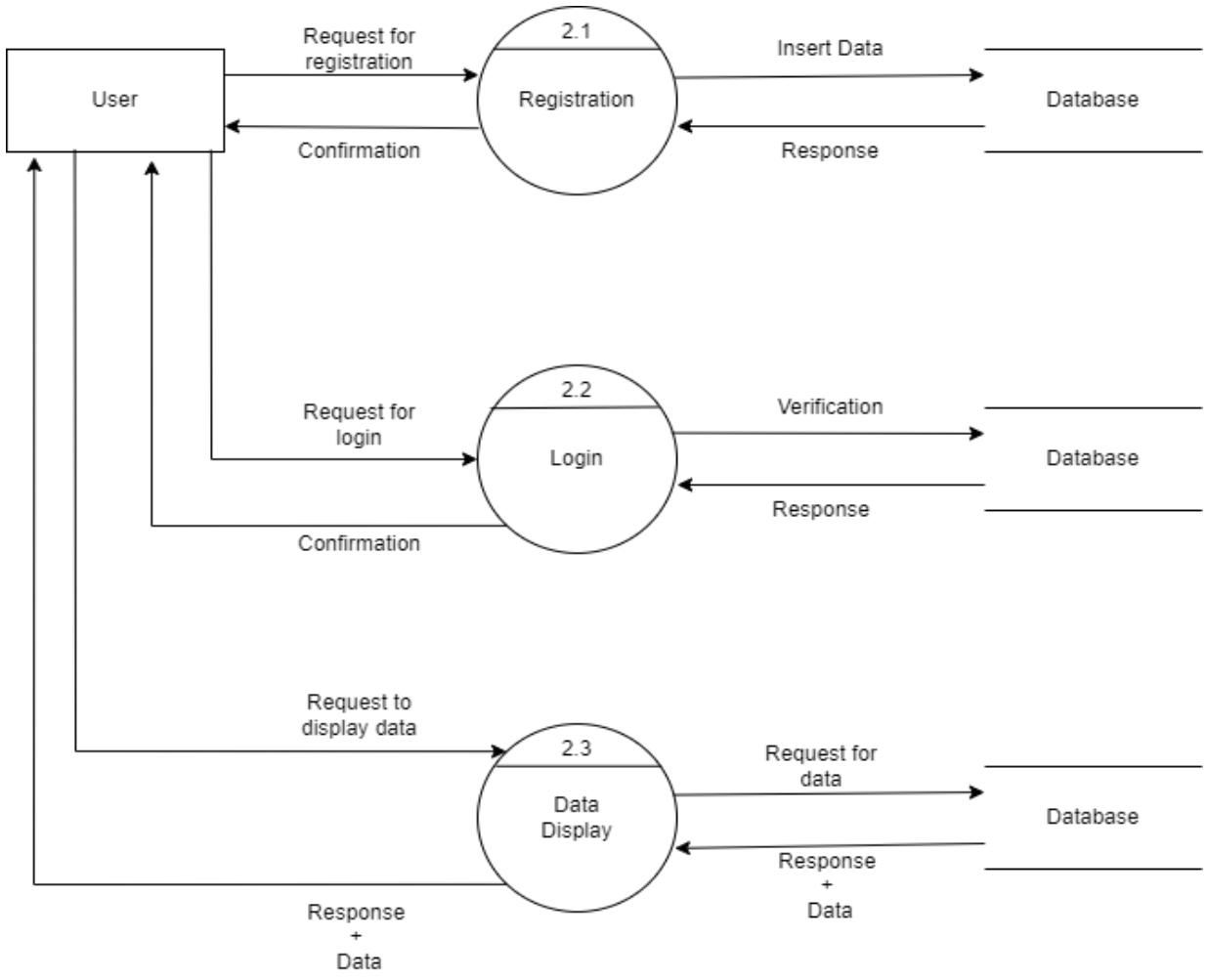

Fig – 4.6.2 : Second Level Data Flow Diagram

Activity Diagram

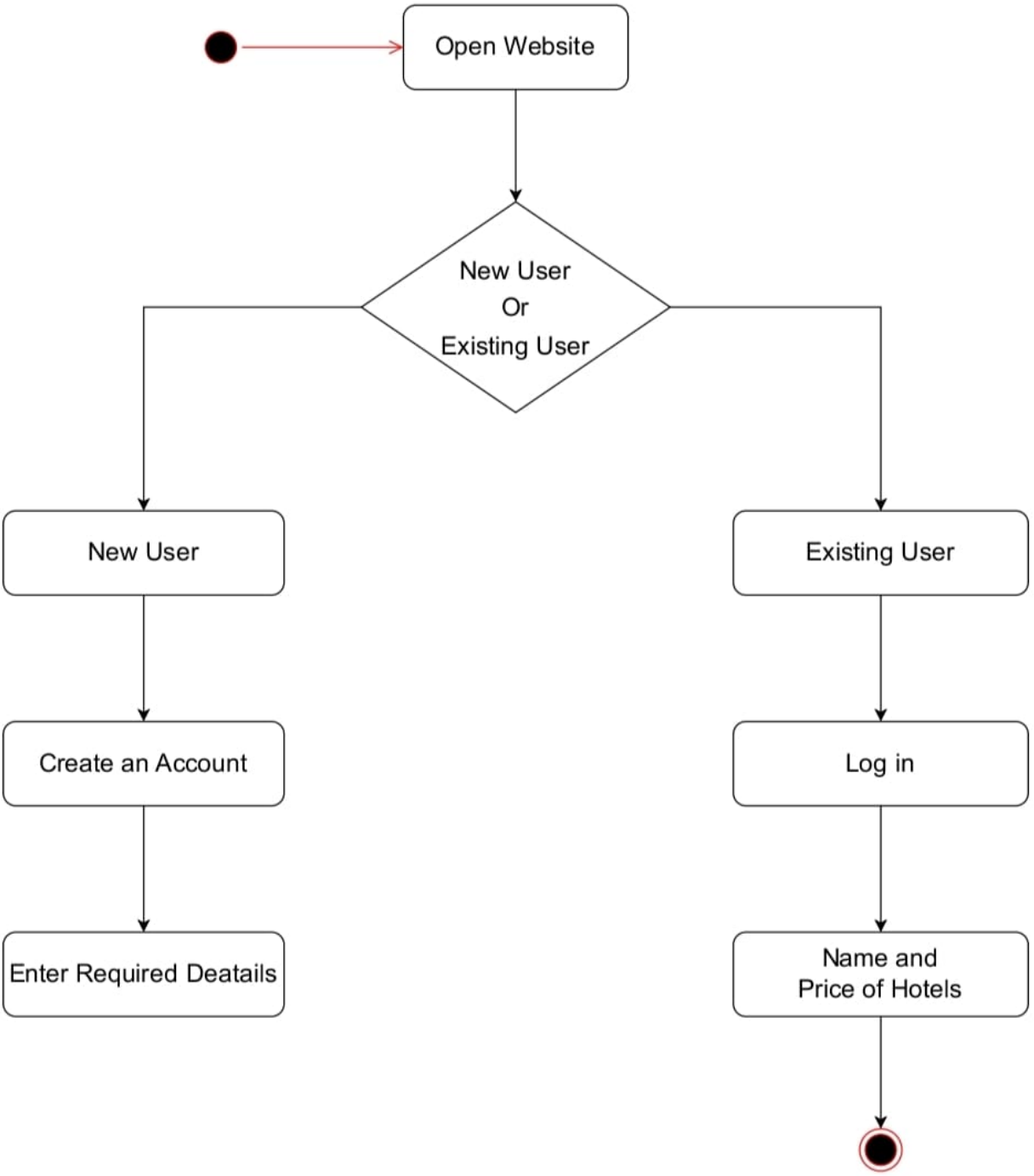

Fig – 4.7 : Activity Diagram

5 Data Dictionary

Admin Table:

Name of the Attribute	Datatype
ID	int (10)
Email	varchar (10)
Password	varchar (10)

User Table:

Name of the Attribute	Datatype
ID	int (10)
First Name	varchar (10)
Last Name	varchar (10)
Email	varchar (20)
Mobile No	int (10)
Password	varchar (20)

6 Result, Discussion and Conclusion

Result:

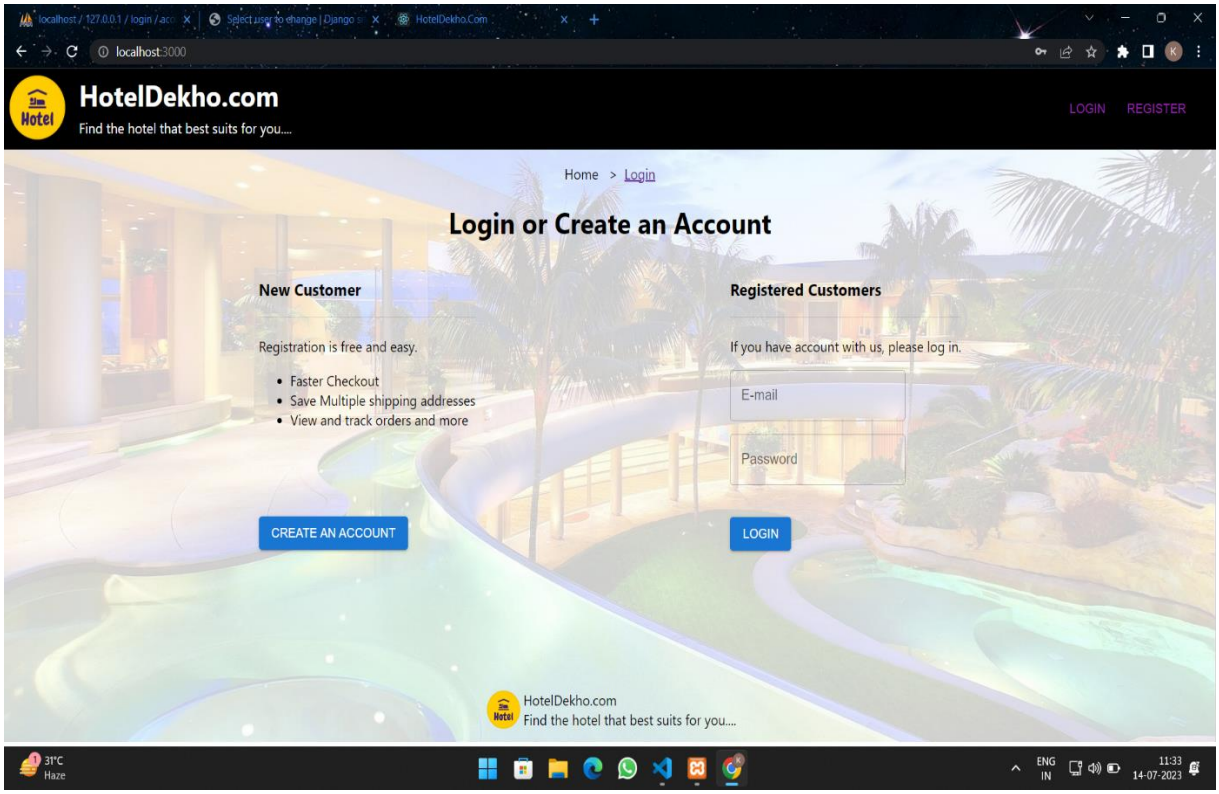

Fig – 6.1

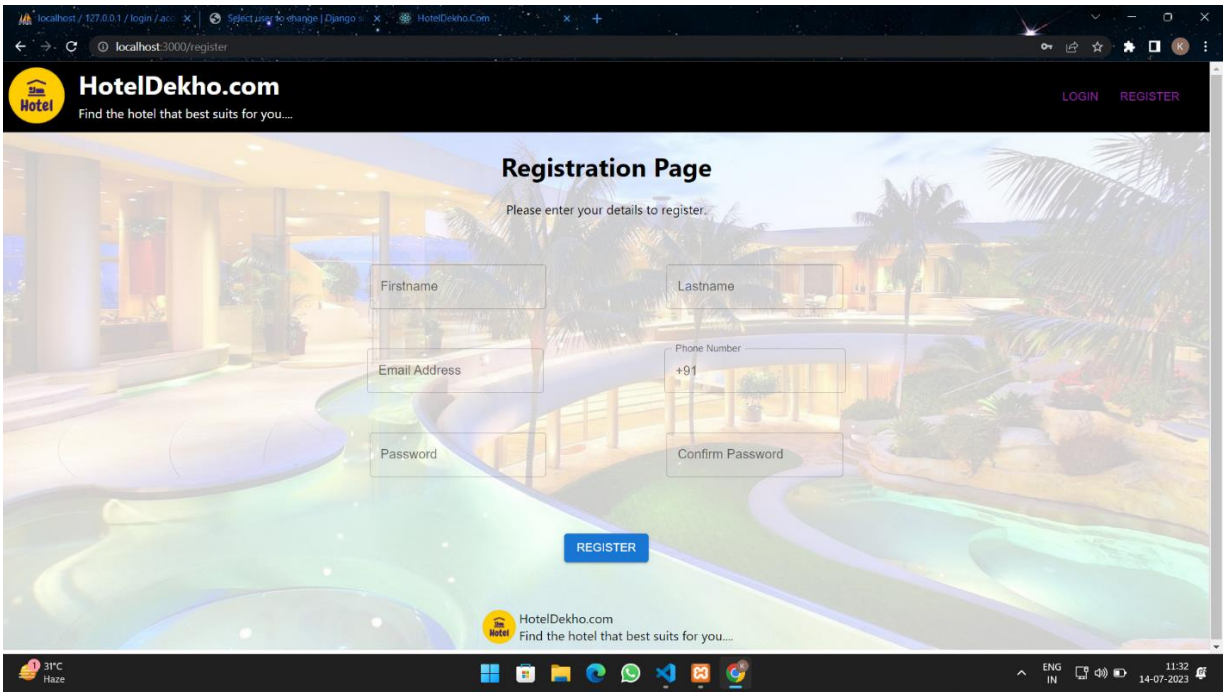

Fig – 6.2

Hotel Name	Booking Price	Hotel Ebooking Price
The Leela Gandhinagar	₹ 9,500	₹ 10,000
Hotel Shiv Inn	₹ 2,374	₹ 2,849
Grand Mercure Gandhinagar GIFT City - An Accor Hotels Brand	₹ 5,499	₹ 10,000
Gift City Club	₹ 6,554	₹ 6,899
Hotel Lemon	₹ 1,699	₹ 2,199
Hotel Ashooka Inn	₹ 2,000	₹ 3,000
Fortune Inn Haveli, Gandhinagar - Member ITC's Hotel Group	₹ 5,400	₹ 6,480
Taj Gandhinagar Resort and Spa	₹ 7,000	₹ 8,000
FabHotel Siddharth Corporate	₹ 2,281	₹ 2,281
Siara Styles Amba Suites, Gandhinagar	₹ 15,699	₹ 4,599
Hotel German Palace Near to Airport	₹ 2,520	₹ 3,780
Hotel Asopalav	₹ 1,188	₹ 1,188
Hotel Royal Mansion, Gandhinagar	₹ 1,231	₹ 1,693
Lovely Ac Condo with Terrace view	₹ 1,960	₹ 1,820
Hotel Prime	₹ 2,399	₹ 3,679

Fig – 6.3

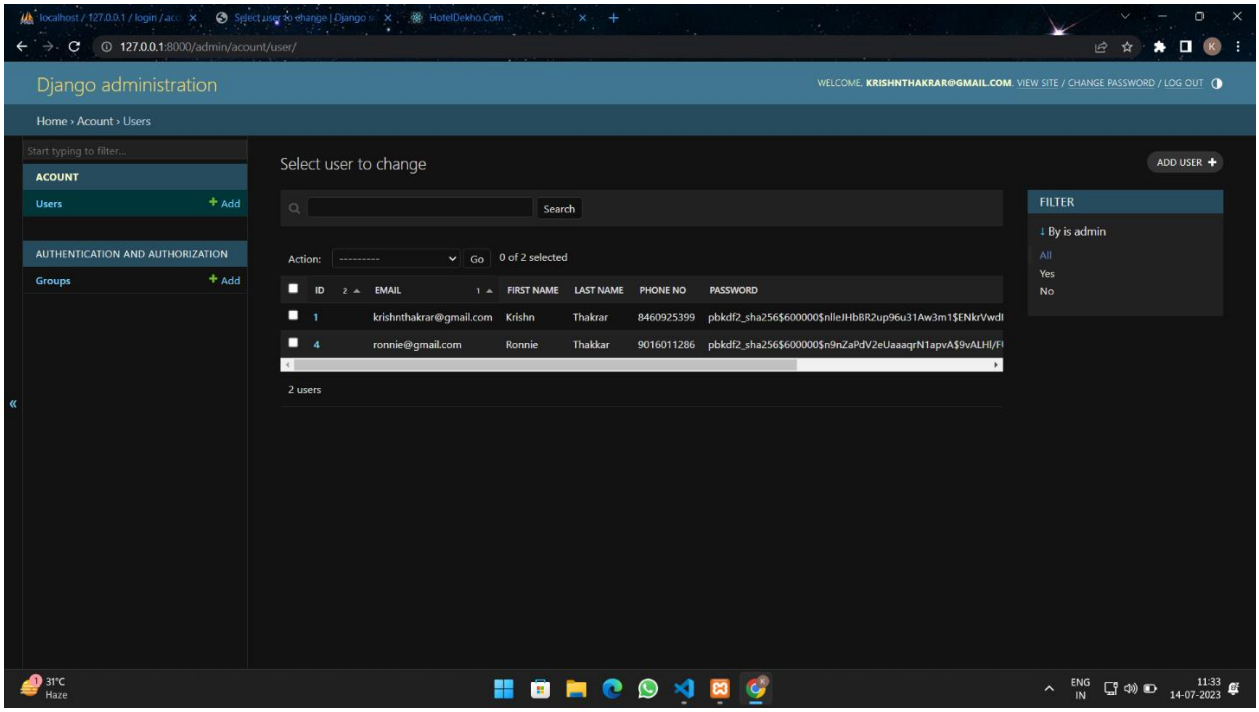

Fig – 6.4

Discussion:

- Web scraping involves using a tool to surf the internet and collect data from different websites. Without it, you would have to physically visit hundreds of competitor websites and travel booking sites, enter your search parameters, and record your data in a spreadsheet. Using a web scraper saves you valuable time and effort so you can focus on other parts of your business.

Conclusion:

- So here we can conclude that we can display the data of comparison of hotel prices to find best suitable hotel according to your requirement.

7 References

- https://www.geeksforgeeks.org/python-web-scraping-tutorial/amp/
- https://towardsdatascience.com/a-step-by-step-guide-to-web-scraping-in-python-5c4d9cef76e8
- https://ieeexplore.ieee.org/document/8822022
- https://www.analyticsvidhya.com/blog/2021/10/a-detailed-guide-on-web-scraping-using-python-framework/