NAAC Accredited with ‘A” Grade
Recognised by UGC under sections 2(f)
and 12(b) & Affiliated to Bangalore
University

DEPARTMENT OF COMPUTER APPLICATIONS

T JOHN COLLEGE

(Affiliated to Bangalore University & Approved to AICTE)
Gottigere, Bengaluru – 560 083

AIRFARE PRICE PREDICTION

A project report submitted to the Bangalore University in the partial fulfillment

of the requirements for the award of the degree of

MASTER OF COMPUTER APPLICATIONS

Submitted by

MANOJ KUMAR N

(P03ML21S0059)

Under the guidance of

Mr. GOWREESWARAN DHANDAYUTHABANI

(Assistant Professor, Department of Computer Applications)

NOVEMBER 2023

NAAC Accredited with ‘A” Grade
Recognised by UGC under sections 2(f)
and 12(b) & Affiliated to Bangalore
University

DEPARTMENT OF COMPUTER APPLICATIONS

T JOHN COLLEGE

(Affiliated to Bangalore University & Approved to AICTE)

Gottigere, Bengaluru – 560 083

PROJECT WORK

AIRFARE PRICE PREDICTION

Bonafide Work Done by

MANOJ KUMAR N

(P03ML21S0059)

The project submitted in partial fulfillment of the requirements for the award of Master of Computer Applications, of Bangalore University, Bengaluru.

GUIDE **HEAD OF THE DEPARTMENT**

Submitted for the Viva-Voce Examination held on ______________

INTERNAL EXAMINER **EXTERNAL EXAMINER**

DECLARATION

I hereby declare that the project entitled **“AIRFARE PRICE PREDICTION”**, submitted to the Department of Computer Applications, T John College, Bangalore in the partial fulfillment of the requirements for the award of Degree in Master of Computer Applications is a record of original work done by me under the guidance of **Mr. GOWREESWARAN DHANDAYUTHABANI**, Assistant Professor in MCA Department during my period of study in T John College, Bangalore.

SIGNATURE OF THE CANDIDATE

Place: Bengaluru

Date:

ACKNOWLEDGEMENT

The project titled **AIRFARE PRICE PREDICTION** has been successfully completed my project with the help of my faculty. I thank her for providing great tips that I have incorporated in this project.

I would express my sincere thanks to **Sr. Rose Mary Balappa**, Principal, T. John College for her valuable advice, co-operation and support and for providing all the faculties for the construction of the project.

I am thankful to **Dr. S. Dinakaran**, Head, Department of Computer Applications department for his valuable advice and co-operation.

I am thankful to various resources that provide requirements for my projects, because requirements are back bone of every project. I am also thankful to my teachers, friends, family members, for their support and prayer for me to complete my project.

ABSTRACT

Traveling via flights has become an integral part of today's lifestyle, as more and more people choose faster travel options. Airfare prices increase or decrease every day here and there depending on various factors such as flight timings, destination and duration of flights, various occasions such as holidays or holiday season. As a result, many people will save time and money by having a basic understanding of flights before making travel arrangements.

A predictive model will be created in the proposed system by application of machine learning algorithms to collected historical data. This system will give people an idea of the trends the prices follow and also provide the predicted value of the price they can check before booking flights to save money. This kind system or service can be provided to customers through flight booking companies to help customer book tickets.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 ABOUT THE PROJECT

The ticketing system is to buy the ticket many days before the flight takes off to avoid the effects of the most extreme fees. Air routes usually do not agree with this procedure. Airlines can reduce costs at times when they need to build a market and when tickets are less affordable. They can maximize costs. So the price may depend on various factors. In order to predict the costs, this business uses AI to show the ways of the tickets after some time. All groups are free to adjust the price of their tickets at any moment

An explorer can save cash by booking the lowest cost flight. People who have traveled frequently by plane are aware of the fluctuations in prices. Airlines use comprehensive Revenue Management principles to implement distinctive rating systems. As a result, the evaluation system changes the fee depending on the time, season and holidays to change the header or footer on the following pages.

The ultimate goal of airways is to make a profit while the customer seeks the minimum rate. Customers usually try to buy a ticket well in advance of the departure date to avoid the increase in ticket prices as the date approaches. But in reality it is not. A customer may end up paying more than they should for the same location.

1.2 OBJECTIVE AND SCOPE OF THE PROJECT

This project aims to predict flight prices for different flights using the machine learning model. The user receives the expected values, and using these as a guide, they can choose whether to purchase tickets.

In the current scenario, airlines are trying to manipulate ticket prices to maximize their profits. Many people frequently take flights, so they are aware of the optimum times to get affordable tickets. But there are also many people who have no experience while booking tickets and they end up falling into the trap of discounts from the companies where they actually end up spending more than they should have. By providing clients with the information they need to order tickets at the proper moment, the suggested system can help them save millions of rupees

Definition of the Problem:

The problem addressed in this project is the fluctuation of airfare prices in the airline industry. Air travel has become an essential part of modern life due to its speed and convenience. However, airfare prices are highly dynamic and can change daily, depending on a multitude of factors such as flight schedules, destinations, flight durations, and special occasions like holidays. This constant price variation poses a challenge for travelers who seek to make cost-effective travel arrangements.

The primary issue is that travelers lack the necessary insights into how airfare prices fluctuate over time and the factors that influence these changes. Without this knowledge, travelers may miss out on potential cost savings when booking flights. Therefore, there is a need for a solution that provides travelers with a better understanding of airfare price trends and offers predictive pricing information. This solution will empower travelers to make more informed decisions, save both time and money, and ultimately enhance their travel experience.

1.3 PROJECT PLAN:

The project plan involves the development of a predictive model that utilizes machine learning algorithms. The primary goal is to create a system that can analyze historical airfare data and generate predictions regarding future price trends. Here are the key components of the project plan: Data Collection: Gather historical airfare data from various sources, including airlines, travel agencies, and online booking platforms. This data should cover a wide range of flight routes, airlines, and time periods.

Data Preprocessing: Clean and preprocess the collected data to ensure its quality and compatibility with the machine learning algorithms**. This step may involve data cleaning, feature engineering, and data transformation.**

Gantt Chart:

The Gantt chart is an invaluable project management tool used for various purposes. It aids in tracking and reporting project progress and visually represents the project schedule. A Gantt chart includes components such as scheduling, budgeting, and resource planning. It provides a quick overview of:

The project's various activities.

- Start and end dates for each activity.
- Duration of each scheduled activity.
- Overlapping activities and their durations.
- The project's start and end dates.
- This chart is a crucial part of project management, allowing stakeholders to understand the project's progress and status at a glance.

CHAPTER 2

LITERATURE REVIEW

2.1 LITERATURE REVIEW

1. A Prediction of Flight Fare Using K-Nearest Neighbors

Year: 2022

Authors: S. Naveen Prasath; Sathish Kumar M; Sherin Eliyas

Someone who buys plane tickets on a regular basis can be able to predict the optimum moment to buy an airline ticket in order to afford the greatest price. For revenue management, many airlines modify ticket pricing. when the demand for the ticket is high, the airline may hike the pricing. To calculate the minimal airfare, data for a given air route was gathered over a period of time, including parameters such as departure time, arrival time, and airways. To use Machine Learning (ML) models, features are retrieved from the obtained data. The cost of a plane ticket is determined by several distinct strands. The main purpose of the research is to find out the factors that impacts drive. airfare price fluctuations and how they connect to price changes. Then, based on this data, create a mechanism to assist purchasers in deciding whether or not to purchase a ticket. The KNN technique is used in this paper to estimate prices at a given time using machine learning regression approaches.

2. Forecasting of airfare prices using time series

Year:2019

Authors: Anastasiia Gordiievych; Igor Shubin

Airline ticket purchase timing is a strategic problem that requires both historical data and domain knowledge to be solved consistently. Even with some historical information (often a feature of modern travel reservation web sites), it is difficult for consumers to make true costminimizing decisions. As product prices become increasingly available on the World Wide Web, consumers attempt to understand how corporations vary these prices over time. However, corporations change the prices based on proprietary algorithms and hidden variables. This work is devoted to the analysis of time series forecasting methods on the example of air tickets forecasting prices. Then, that information it is planned to build a system that will help customers to make purchasing decisions by forecasting how air ticket prices will evolve in the future.

3. OTPS: A decision support service for optimal airfare Ticket Purchase

Year:2018

Authors: Yuchang Xu; Jian Cao

Due to the fast development and wide application of Web technology, online purchases have become the main method to buy airfare tickets. Because information about pricing factors is unavailable, it is difficult for customers to buy airfare tickets at the lowest cost or even at a relatively lower cost. Although several approaches have been proposed to suggest the optimal timing for purchasing airfare tickets, these methods are based on predictions of tomorrow's price information. Clearly, local optimal decisions which rely solely on tomorrow's price prediction can be misleading or incorrect, resulting in missed opportunities to buy tickets at a lower cost in the future. In this paper, a novel Optimal airfare Ticket Purchase decision-support Service (OTPS) is proposed, which makes continuous recommendations on the best purchase time before the departure date. OTPS is based on the strategy of the Dynamic ratio of numbers of Potential days with Lower Price (DPLP), which takes the fluctuation tendency of airfare prices for a period into consideration. Specifically, parameter values of the model are set in a route-specific manner and updated periodically to enhance the generalization ability and reliability of OTPS. Extensive experiments are conducted on the real-world ticket price dataset of multiple routes. The experimental results prove that OTPS shows superiority over other state-of-the-art solutions.

4. Predicting Airfare Price Using Machine Learning Techniques: A Case Study for Turkish Touristic Cities

Year:2022

Authors: Yekta Said Can; Koray Büyükoğuz; Efe Batur Giritli; Mustafa Şişik

Airline ticket price is influenced by several elements, such as flight distance, purchasing time, number of transfers, etc. Furthermore, every carrier has its own proprietary rules and techniques to determine the ticket price accordingly. With recent improvements in Machine Learning (ML), these rules could be inferred and the price variation could be modeled. In this study, we first created the first dataset containing flight prices for Turkey. The flight price dataset consists of over 1000 domestic flights towards the touristic cities in Turkey. We then use machine learning algorithms to model the ticket price based on different origin and destination pairs. We achieved promising results for predicting the flight ticket price on our dataset

5. Joint decision making about price and duration of discount airfares

Year:2019

Authors: Yanli Fang; Yan Chen; Xin Li

Dynamic pricing is a common mechanism used in the market to enhance revenue generation from different segments of customers. In this paper, we study a joint decision making problem, faced by an airline, about the optimal price and duration of the discounted airfares to maximize its revenue. Customers have heterogeneous valuations about the air ticket. Furthermore, customers' valuations increase over the sales horizon as customers learn more about their travel schedule. In this study, the interaction between the airline and customers is modeled as a Stackelberg game, where airline acts as the leader and customers act as followers. The optimal joint decision is derived to maximize airline's total expected revenue. At last, a numerical example is conducted to illustrate the strategic pricing decision of the airline.

CHAPTER 3
SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

Air ticket price prediction is a challenging task since the factors involved in pricing dynamically change overtime and make the price fluctuate. In the last decade, researchers have incorporated machine learning algorithms and data mining strategies to better model observed prices. Among them, regression models, such as Linear Regression(LR), Support Vector Machines (SVMs), Random Forests(RF), are frequently used in predicting accurate airfare price. Early work also considered using classification models top redict the trends of the itineraries.

3.2 PROPOSED SYSTEM

Forecasting the price of an airline ticket is a very challenging task because many factors depend on the price of an airline ticket. Many researchers used various machine learning algorithms to obtain a model with higher prediction accuracy from the ticket price. Researchers use various regression models such as support vector machines (SVMs), Linear regression (LR), decision tree, random forests, etc. to predict the exact price of a flight.

3.3 MODULE DESCRIPTION

Admin: It is Django Admin by default its will there in for view details of login

User Module: It is module which the consist of login and register

- Predict Module- it consist of test data for predicting price
- Feedback module – users can give feedback about the airfare

3.4 SYSTEM ARCHITECTURE

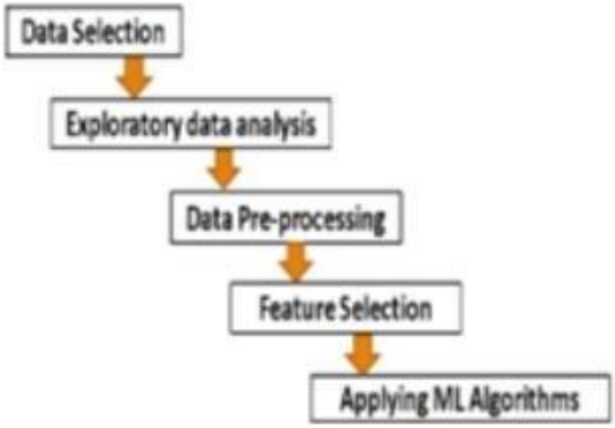


Fig 3.1.System architecture

3.5 METHODOLOGIES

Step1: Downloading Dataset from Kaggle:

Kaggle lets in customers to discover and put up data sets, discover and construct fashions in a web-primarily based totally factstechnology data science environment, paintings with different facts scientists and device mastering engineers, and input competitions to resolve data science challenges.

Step2: Data Processing:

Data preprocessing is a data mining technique.It is used to convert the raw records in a beneficial and useful format. In the dataset, many attributes include the identical information. Directly merging the tables creates many replica fields. Also, the records pronounced via way of means of the airways can also additionally consist of faulty values because of human error, foreign money conversion error, etc. Hence, as it should be designed records preprocessing workflow is vital to generate correct enter records for you to build the machine learning model.

Step3: Feature Extraction:

Feature Extraction goals to lessen the wide variety of functions or features in a dataset through developing new functions from the prevailing ones (after which discarding the unique functions). These new decreased set of functions must then be capable of summarize maximum of the statistics contained within side the unique set of functions.

Step4:Feature Selection:

Feature selection is the system of decreasing the wide variety of input variables while making a predictive model. It is ideal to lessen the wide variety of input variables to each lessen the computational value of modeling and, in a few cases, to enhance the overall performance of the model.

Step5:Predicting model:

Algorithm used in our model is Random forest. Random forest is supervised machine learning algorithm. It is a collection of multiple decision trees whose results are aggregated into one final result. As the name suggests, "Random Forest is a classifier that contains a multiple number of decision trees on various subsets of the given dataset and takes the average to improve the predictive or final accuracy of that dataset.

Step 6:

User can give rating and reviews according to the predicted airfare price outcome.s

- **NON-FUNCTIONAL REQUIREMENTS**

NON-FUNCTIONAL REQUIREMENT (NFR) specifies the quality attribute of a software system. They judge the software system based on Responsiveness, Usability, Security, Portability and other non-functional standards that are critical to the success of the software system. Example of nonfunctional requirement, *"how fast does the website load?"* Failing to meet non-functional requirements can result in systems that fail to satisfy user needs. Non-functional Requirements allows you to impose constraints or restrictions on the design of the system across the various agile backlogs. Example, the site should load in 3 seconds when the number of simultaneous users are > 10000. Description of non-functional requirements is just as critical as a functional requirement.

- Usability requirement
- Serviceability requirement
- Manageability requirement
- Recoverability requirement
- Security requirement
- Data Integrity requirement
- Capacity requirement
- Availability requirement
- Scalability requirement
- Interoperability requirement

- Reliability requirement
- Maintainability requirement
- Regulatory requirement
- Environmental requirement

3.6 FEASIBILITY STUDY

All projects are feasible given unlimited resources and infinite time. Unfortunately the development of computer-based system in many cases is more likely to be plagued by scarcity of resources and delivery date. Hence, we have made use the concept of reusability that is what Object Oriented Programming (OOPS) is all about. The feasibility report of the project holds the advantages and flexibility of the project. This is divided into three sections: ¬ Economical Feasibility ¬ Technical Feasibility ¬ Behavioral Feasibility

- **3.6.1 ECONOMIC FEASIBILITY:**

Economic analysis is the most frequently used method for evaluating the effectiveness of the candidate system. More commonly known as cost/benefit analysis, the procedure is to be determining the benefits and savings that are expected from a candidate and compare them with costs. If benefits outweigh costs, then the decision is made to design and implement the system. A systems financial benefit must exceed the cost of developing that system. i.e. a new system being developed should be a good investment for the organization. Economic feasibility considers the following i. The cost to conduct a full system investigation ii. The cost of hardware and software for the class of application. iii. The benefits in the form of reduced cost or fewer costly errors. iv. The cost if nothing changes (i.e. the proposed system is not developed). The proposed Movie Recommendation System is economically feasible because the system requires very less time factors. ii. The system will provide fast and efficient automated environment instead of slow and error prone manual system, thus reducing both time and man power spent in running the system. iii. The system will have GUI interface and very less user-training is required to learn it. iv. The system will provide service to view various information for proper managerial decision making.

- **3.6.2 TECHNICAL FEASIBILITY:**

Technical feasibility centers around the existing computer system (Hardware and Software etc.) and to what extend it support the proposed addition. For example, if the current computer is operating at 80 percent capacity - an arbitrary ceiling - then running another application could overload the system or require additional Hardware. This involves financial considerations to accommodate technical enhancements. If the budget is a serious constraint, then the project is judged not feasible. In this project, all the necessary cautions have been taken care to make it technically feasible. Using a key the display of text/object is very fast. Also, the tools, operating system and programming language used in this localization process is compatible with the existing one.

- **3.6.3 BEHAVIORAL FEASIBILITY:**

People are inherently resistant to change, and computers have been known to facilitate change. An estimate should be made of how strong a reaction the user staff is likely to have toward the development of a computerized system. Therefore it is understandable that the introduction of a candidate system requires special efforts to educate and train the staff. The software that is being developed is user friendly and easy to learn. In this way, the developed software is truly efficient and can work on any circumstances, tradition, locales. Behavioral study strives on ensuring that the equilibrium of the organization and status quo in the organization are not disturbed and changes are readily accepted by the users.

CHAPTER 4

SYSTEM REQUIREMENTS

AND

SPECIFICATIONS

4.1 Hardware Requirements

Processor	:	Intel i3 7th Gen 2.4 GHz
Hard Disk	:	Min 40 GB
RAM	:	4GB or above
External	:	Mouse,Keyboard

4.2 Software Requirements

Operating system	:	Windows 10
Front End	:	HTML5, CSS3, JavaScript, Bootstraps
Coding Language	:	Python
Database	:	postgresql/PgAdmin
Tool	:	Pycharm IDE
Framework	:	Django web frameworks
Browsers	:	Chrome, Mozilla firefox

4.3 TOOLS AND TECHNOLOGIES

Tensorflow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.

TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

It is the fundamental package for scientific computing with Python. It contains various features including these important ones:

- A powerful N-dimensional array object
- Sophisticated (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using Numpy which allows Numpy to seamlessly and speedily integrate with a wide variety of databases.

Pandas

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

Matplotlib

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

Scikit – learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use.

Python

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace.

Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

- Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- Python is Interactive − you can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

Python also acknowledges that speed of development is important. Readable and terse code is part of this, and so is access to powerful constructs that avoid tedious repetition of code. Maintainability also ties into this may be an all but useless metric, but it does say something about how much code you have to scan, read and/or understand to troubleshoot problems or tweak behaviors. This speed of development, the ease with which a programmer of other languages can pick up basic Python skills and the huge standard library is key to another area where Python excels. All its tools have been quick to implement, saved a lot of time, and several of them have later been patched and updated by people with no Python background - without breaking.

CHAPTER 5

SYSTEM DESIGN

5.1 TABLE DESIGN

1. User Registration

FIELD NAME	DATATYPE	SIZE	DESCRIPTION
ID	INTEGER	20	Primary key
First Name	VARCHAR	100	First Name
Last Name	VARCHAR	100	Last Name
User Name	VARCHAR	100	User Name
Email ID	VARCHAR	100	Email ID
Password	VARCHAR	500	Password
Confirm Password	VARCHAR	500	Confirm Password

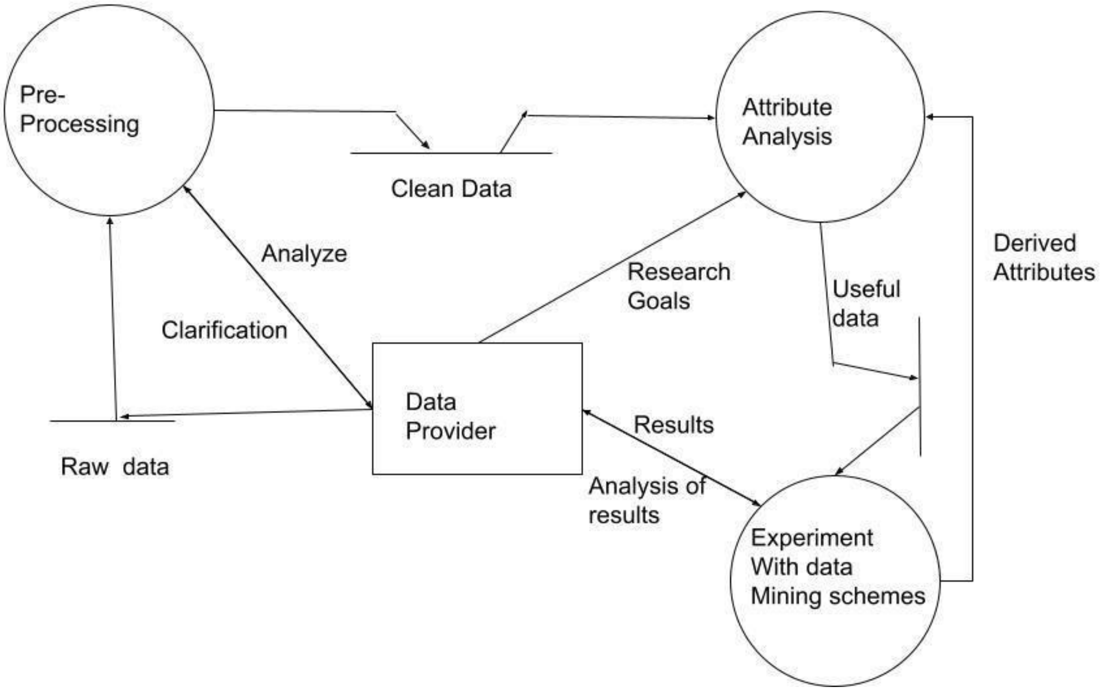
2. User Login

FIELD NAME	DATATYPE	SIZE	DESCRIPTION
ID	INTEGER	**20**	Primary key
User Name	VARCHAR	100	User Name
Password	VARCHAR	500	Password

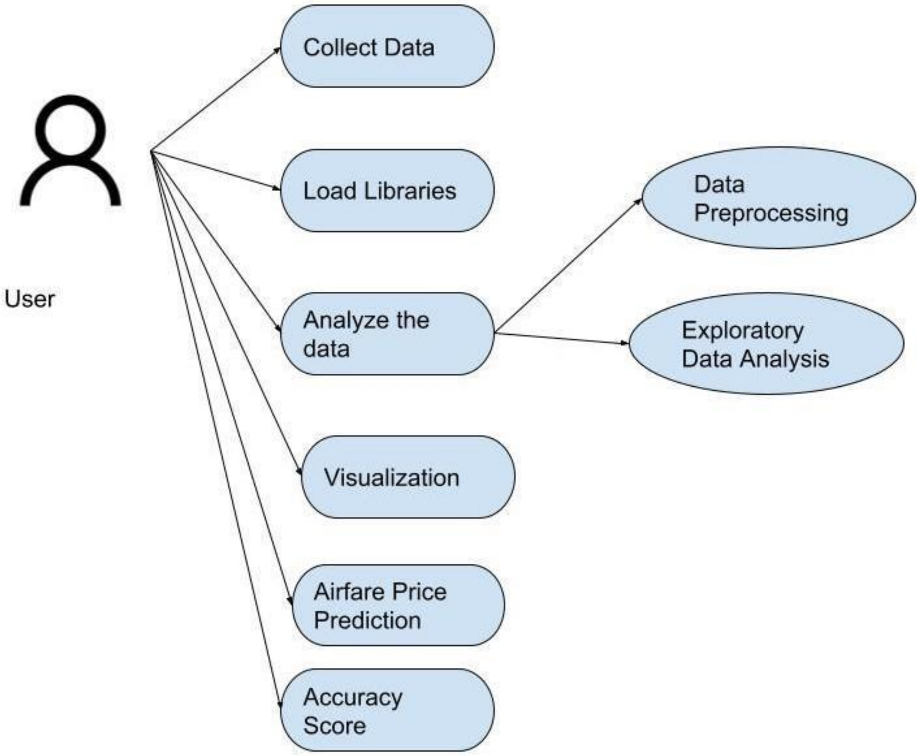
3. Feedback Table

FIELD NAME	DATATYPE	SIZE	DESCRIPTION
ID	INTEGER	20	Primary key
Name	VARCHAR	100	User Name
Email	VARCHAR	100	Email ID
Your Message	VARCHAR	500	Password

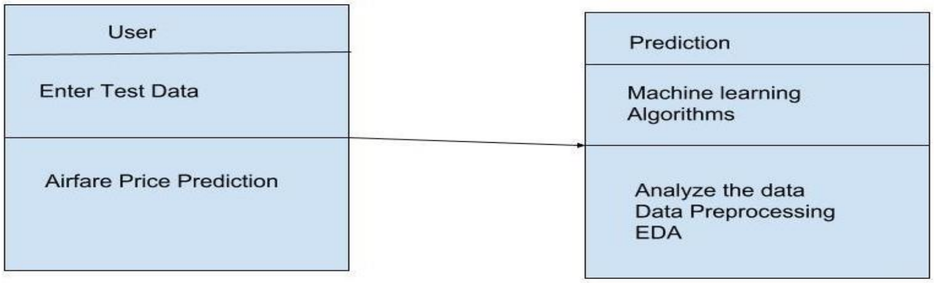
5.2 DATAFLOW DIAGRAM



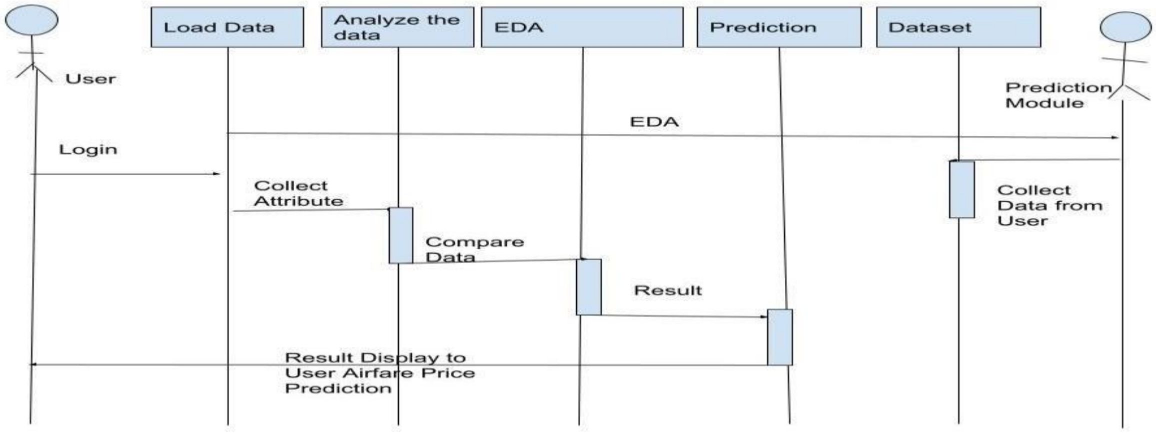
5.3 USECASE DIAGRAM



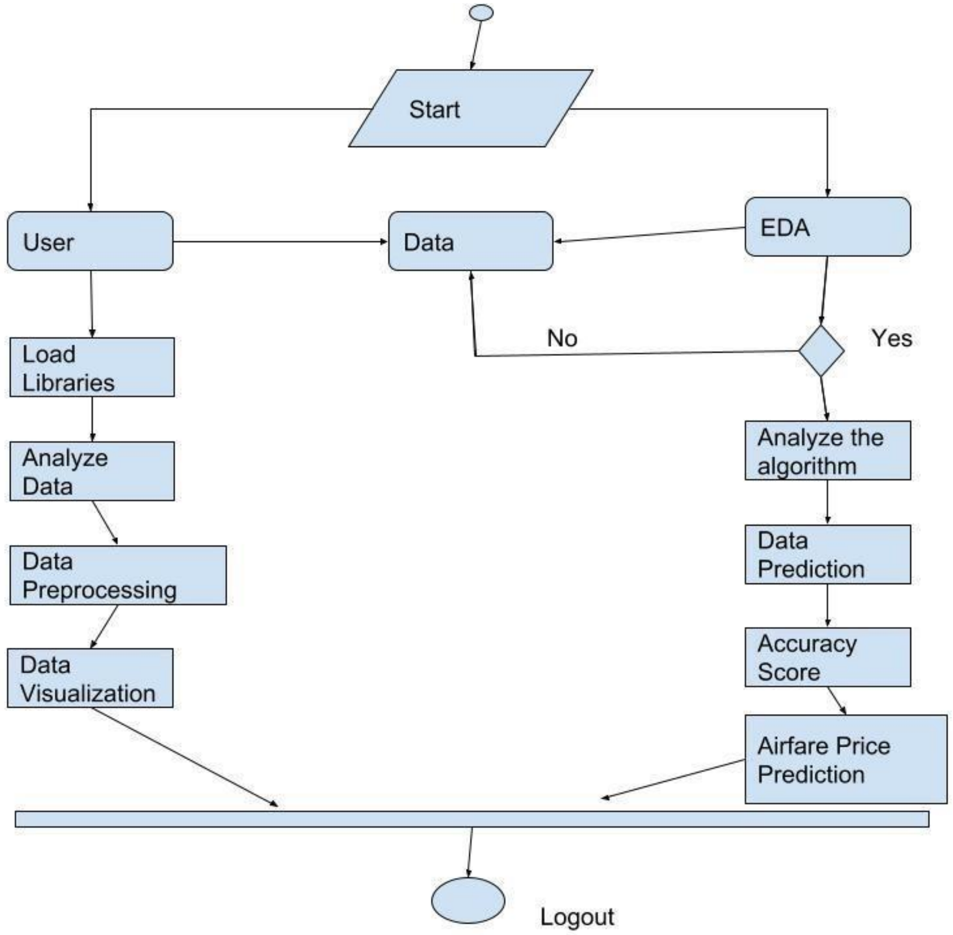
5.4 CLASS DIAGRAM



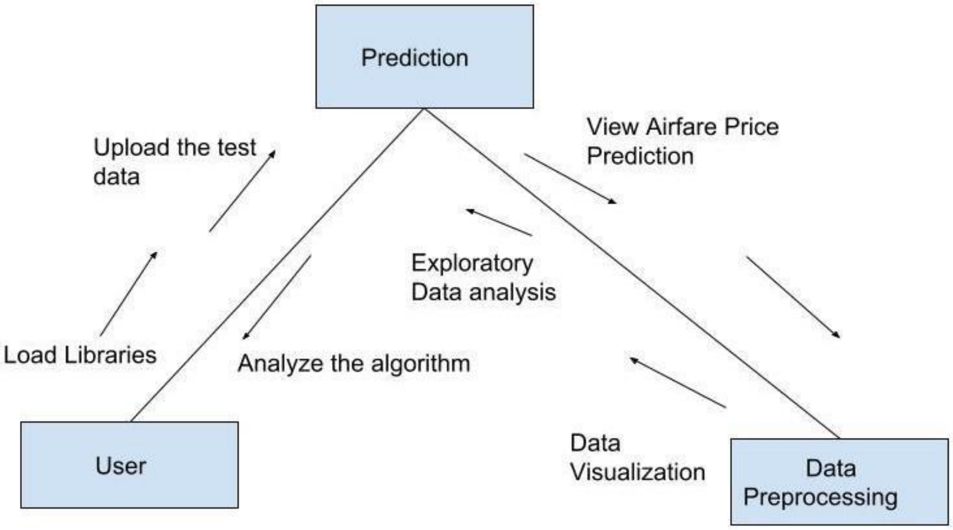
5.5 SEQUENCE DIAGRAM



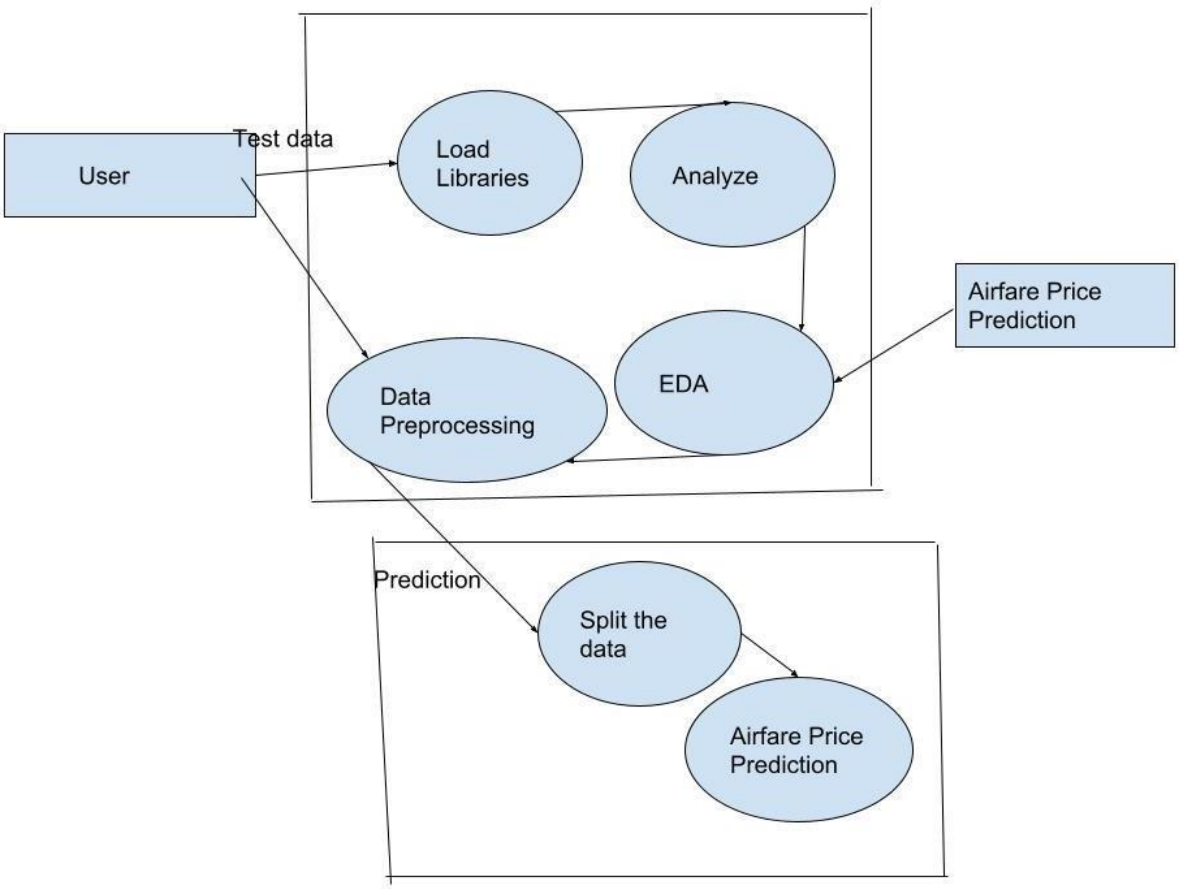
5.6 ACTIVITY DIAGRAM



5.7 COLLABORATION DIAGRAM



5.8 CONTEXT DIAGRAM



CHAPTER 6

SYSTEM IMPLEMENTATION, TESTING & MAINTENANCE

6.1 CAAD Model:

The airfare price prediction system is based on a predictive model that employs a Convolutional Neural Network (CNN) architecture. CNNs are widely used for image analysis. The model is trained to learn from historical airfare data to understand the typical price distribution. Once it has learned what is considered a "normal" airfare for various flight routes and dates, it can be used to detect anomalies or deviations from this learned normal representation. Here's how the CAAD model works for airfare price prediction:

1. Data Preprocessing: The input data, including flight details, departure dates, and other relevant factors, is preprocessed. This may involve standardizing the format of data, encoding categorical variables, and normalizing numerical features to ensure they are in a suitable format for the model.

2. Training the CNN Model: The system trains a CNN model on a large dataset of historical airfare prices. The CNN model is trained to extract relevant features from the input data and classify them as either normal or abnormal. It learns the patterns and variations in airfare prices.

3. Anomaly Detection Module: To detect anomalies in airfare prices, the model is fed with input data for a specific flight. If the predicted airfare is similar to the learned normal representation, the case is considered normal. However, if the predicted airfare deviates significantly from the normal distribution, it is classified as an anomaly. This is done using a threshold-based approach, where the difference between the predicted price and the expected price is compared to a predefined threshold.

4. Threshold Setting: During the training phase, the system determines the threshold for anomaly detection based on the distribution of differences between predicted and actual airfares in the training dataset. The threshold value is crucial for classifying anomalies, and it can be adjusted to control the sensitivity of the model.

6.2 SYSTEM IMPLEMENTATION:

The system is implemented using a combination of programming languages and technologies, including:

- **HTML**: HTML is used to create the user interface and web pages for interacting with the system.
- **CSS:** Cascading Style Sheets are used for styling and formatting the web pages to make them visually appealing.
- **JavaScript:** JavaScript is used for client-side scripting, making web pages interactive and enhancing user experience.
- **Python:** Python is employed as the backend language for implementing the CAAD model and handling data processing and predictions.

6.3 SYSTEM TESTING:

Thorough testing is conducted to ensure the accuracy and reliability of the system. Various types of testing are performed:

- Black Box Testing: This type of testing focuses on the system's functionality and user interactions, evaluating user registration, login procedures, and the functionality of different modules. It also checks the CAAD model's accuracy in anomaly detection and various user interface aspects.

- White Box Testing: White box testing examines the internal code structure of the system. It includes code coverage analysis, control flow testing, boundary value analysis, and other structural assessments to ensure the code is robust and efficient.

6.4 TESTING STRATEGIES

6.4.1 UNIT TESTING

Unit testing, is a testing technique using which individual modules are tested to determine if there are issues by the developer himself.. it is concerned with functional correctness of the standalone modules. The main aim is to isolate each unit of the system to identify, analyze and fix the defects.

Unit Testing Techniques:

Black Box Testing - Using which the user interface, input and output are tested.

White Box Testing –Used to test each one of those functions behavior is tested.

6.4.2 DATA FLOW TESTING

Data flow testing is a family of testing strategies based on selecting paths through the program’s control flow in order to explore sequence of events related to the status of Variables or data object. Dataflow Testing focuses on the points at which variables receive and the points at which these values are used.

6.4.3 INTEGRATION TESTING

Integration Testing done upon completion of unit testing, the units or modules are to be integrated which gives raise too integration testing. The purpose of integration testing is to verify the functional, performance, and reliability between the modules that are integrated.

6.4.4 BIG BANG INTEGRATION TESTING

Big Bang Integration Testing is an integration testing Strategy wherein all units are linked at once, resulting in a complete system. When this type of testing strategy is adopted, it is difficult to isolate any errors found, because attention is not paid to verifying the interfaces across individual units.

6.4.5 USER INTERFACE TESTING

User interface testing, a testing technique used to identify the presence of defects is a product/software under test by Graphical User interface [GUI].

6.5 TEST CASES:

S.NO	INPUT	If available	If not available
1	Dataset upload & analysis	Dataset loaded	There is no process
2	Dataset Processing & Analytical Methods	encode data with integer ID and then split dataset into train and test	There is no process
3	Run Regression Model	trained Regression algorithm	There is no process
4	Performance Graph	Plot Price Prediction graph	There is no process
5	Predict Airfare Price	predict price test data contains any airfare data	There is no process

6.6 Validation Testing:

Validation testing checks the accuracy of user inputs and ensures that the system handles data correctly. It validates that data provided by users is accurately processed and that the system's behavior aligns with specified requirements.

6.7 Test Report:

A test report is generated during the testing phase, documenting the results of various tests and their outcomes. It provides an overview of the system's performance and any issues encountered during testing.

6.8 System Maintenance:

System maintenance is an essential part of software development. It involves creating a maintenance plan that addresses how users can request modifications or report problems, resource and cost estimates, and decision-making for new system features and quality objectives.

CHAPTER 7

CONCLUSION

AND

FUTURE SCOPE

7.1 CONCLUSION

This paper reported on a preliminary study in “airfare prices prediction”. We gathered airfare data from a Kaggle website and showed that it is feasible to predict prices for flights based on historical fare data. The experimental results show that ML models are a satisfactory tool for predicting airfare prices. Other important factors in airfare prediction are the data collection and feature selection from which we drew some useful conclusions. From the experiments we concluded which features influence the airfare prediction at most.

7.2 FUTURE SCOPE

More routes can be added and the same analysis can be extended to major airports and travel routes in India. More of data points and historical data should be taken into account for analysis. This will train the model better and provide better accuracy and more savings. Additional rules may be added to rule-based learning based on our understanding of the industry, including offer periods provided by airlines. Development of this project more user-friendly interface for different routes giving users more flexibility.

Currently, there are many fields where prediction services are used such as stock price prediction tools used by stockbrokers and services like estimate that provide estimated value of house prices. That's why there exists the demand for services like this in the aviation industry which can assist customers in booking tickets. There are many of them that examine the work that has been done on it using various techniques and further research is needed to improve prediction accuracy using different algorithms. Results can be more accurately obtained by using data that is more accurate and has better features.

CHAPTER 8

BIBLIOGRAPHY

REFERENCES

[1] Joshi, Achyut & Sikaria, Himanshu & Devireddy, Tarun. (2017). Predicting Flight Prices in India.

[2] Wang, Tianyi & Pouyanfar, Samira & Tian, Haiman & Tao, Yudong & Alonso, Miguel & Luis, Steven & Chen, Shu-Ching. (2019). A Framework for Airfare Price Prediction: A Machine Learning Approach.200- 207.10.1109/IRI.2019.00041.

[3] Supriya Rajankar, Neha Sakharkar, Omprakash Rajankar, “Predicting The Price Of A Flight Ticket With The Use Of Machine Learning Algorithms”, INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 12, ISSN 2277-8616, DECEMBER 2019.

[4] Neel Bhosale, Pranav Gole, Hrutuja Handore, Priti Lakde, Gajanan Arsalwad, “Flight Fare Prediction Using Machine Learning”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653, Volume 10, Issue V, May 2022.

[5] Jaywrat Singh Champawat, Udhhav Arora, Dr. K. Vijaya, “INDIAN FLIGHT FARE PREDICTION: A PROPOSAL”, International Journal of Advanced Technology in Engineering and Science, Vol. No.09, Issue No.03, ISSN 2348-7550, March 2021.

[6] Zhichao Zhao, Jinguo You, Guoyu Gan, Xiaowu Li, Jiaman Ding, “Civil airline fare prediction with a multi-attribute dual-stage attention mechanism”, Appl Intell 52, 5047-5062 (2022), 03 August 2021.

[7] Pavithra Maria K, Anitha K L, “Flight Price Prediction for Users by Machine Learning Techniques”, International Advanced Research Journal in Science, Engineering and Technology, Vol.8, Issue 3, DOI: 10.17148/IARJSET.2021.8321, March 2021.

[8] Juhar Ahmed Abdella, NM Zaki, Khaled Shuaib, Fahad Khan, “Airline ticket price and demand prediction: A survey”, Journal of King Saud University-Computer and Information Sciences, Volume 33, Issue 4, May 2021.

[9] Rutuja Konde, Rutuja Somvanshi, Pratiksha Khaire, Prachi Zende, Kamlesh Patil, “Airfare Price Prediction System”, International Research Journal of Engineering and Technology (IRJET), Volume: 09 Issue: 05, May 2022.

[10]Vivekanand P. Thakare, Ankita Sanjay Murraya, Roshani Bandu Gawade, Mrunali Mukundrao Sawarkar, Trupti Khemraj Shende, Ujjwala Kamlesh Badole, “Flight Ticket Price Predictor Using Python”, Advancement and Research in Instrumentation Engineering Volume 5 Issue 1 e-ISSN: 2582-4341.

[11]Kunal Khandelwal, Atharva Sawarkar, Dr. Swati Hira, “A Novel Approach for Fare Prediction Using Machine Learning Techniques”, International Journal of Next-Generation Computing, VOLUME 12, SPECIAL ISSUE 5, Online ISSN: 0976-5034, NOVEMBER 2021.

[12]Prof. Ms. Archana Dirgule, Shubham Agarwal, Ram Agrawal, Neha Singh, Kiran Adsul, “Flight Fare Prediction using Random Forest Algorithm”, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), Volume 2, Issue 3, ISSN (Online) 2581- 9429, May 2022.

[13]Tadvi Shabana, Khan Mariya, Shaikh Afifa, Sayyed Naziya Begum, “A NOVEL DATA SCIENCE AND ML APPROACH TO PREDICT AIRFARE”, JETIR, Volume 6, Issue 4, ISSN-2349-5162, April 2019.

[14]Abhijit Boruah, Kamal Baruah, Biman Das, Manash Jyoti Das, Niranjan Borpatra Gohain, “A Bayesian Approach for Flight Fare Prediction Based on Kalman Filter”, Progress in Advanced Computing and Intelligent Engineering, Volume 2, ICACIE 2017.

CHAPTER 9

SAMPLE CODE

Sample Code:

```
from django.shortcuts import render, redirect
from django.contrib import messages
from django.contrib.auth.models import User, auth

# Create your views here.
def login(request):
    if request.method == 'POST':
        username = request.POST*'username'+
        password = request.POST*'password'+

        user = auth.authenticate(username=username, password=password)

        if user is not None:
            auth.login(request, user)
            return redirect('/')
        else:
            messages.info(request, 'invalid credentials...')
            return redirect('login')
    else:
        return render(request, 'login.html')

def register(request):
    if request.method == 'POST':
        first_name = request.POST*'first_name'+
        last_name = request.POST*'last_name'+
        username = request.POST*'username'+
        email = request.POST*'email'+
        password1 = request.POST*'password1'+
        password2 = request.POST*'password2'+

        if password1==password2:
            if User.objects.filter(username=username).exists():
                messages.info(request, "username taken...")
                return redirect('register')
            elif User.objects.filter(email=email).exists():
                messages.info(request, "email taken...")
                return redirect('register')
            else:
                user = User.objects.create_user(username=username, password=password1, email=email,
first_name=first_name, last_name=last_name)
                user.save()
                messages.info(request, "user created...")
                return redirect('login')
        else:
            messages.info(request, "password not matched...")
            return redirect('register')
        return redirect('/')
    else:
        return render(request, 'register.html')


def logout(request):
    auth.logout(request)
    return redirect('/')


def contact(request):
    return render(request, 'contact.html')
```

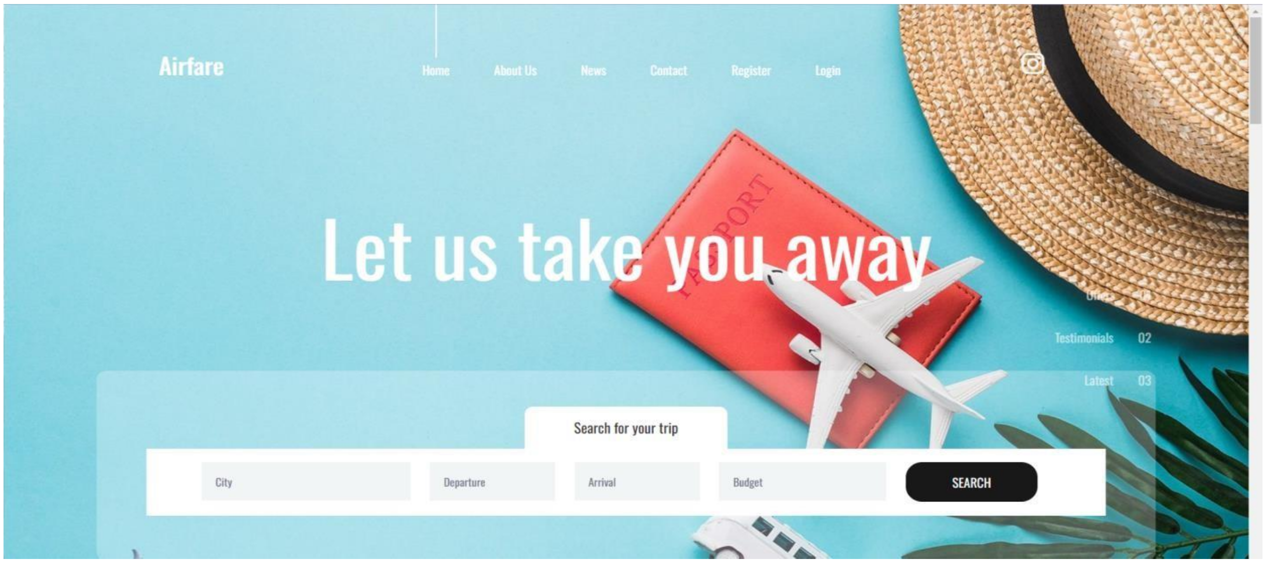
```
def about(request):
    return render(request, 'about.html')

def news(request):
    return render(request, 'news.html')

def predict(request):
    return render(request,"predict.html")


def data(request):
    if (request.method == 'POST'):
        airline = request.POST*'airline'+
        flight = request.POST*'flight'+
        source_city = request.POST*'source_city'+
        departure_time = request.POST*'departure_time'+
        stops = request.POST*'stops'+
        arrival_time = request.POST*'arrival_time'+
        destination_city  =  request.POST*'destination_city'+

        duration=float(request.POST*'duration'+)
        days_left=int(request.POST*'days_left'+)
        from sklearn.preprocessing import LabelEncoder
        l=LabelEncoder()
        l.fit_transform(*airline,flight,    source_city,departure_time,stops,arrival_time,destination_city+)

        air = l.fit_transform(*airline+)
        fl = l.fit_transform(*flight+)
```

```
src = l.fit_transform(*source_city+)
dept  =  l.fit_transform(*departure_time+)
st = l.fit_transform(*stops+)
arr  = l.fit_transform(*arrival_time+)
dest = l.fit_transform(*destination_city+)
#dur = l.fit_transform(*duration+)
#days  = l.fit_transform(*days_left+)
import pandas as pd
df = pd.read_csv("static/Dataset/Clean_Dataset.csv")
df.dropna(inplace=True)
df.isnull().sum()

airline1 = l.fit_transform(df*"airline"+)
flight1 = l.fit_transform(df*"flight"+)
source_city1 = l.fit_transform(df*"source_city"+)
departure_time1 = l.fit_transform(df*"departure_time"+)
stops1 = l.fit_transform(df*"stops"+)
arrival_time1 = l.fit_transform(df*"arrival_time"+)
destination_city1 =  l.fit_transform(df*"destination_city"+)
#duration1 = l.fit_transform(df*"duaration"+)
#days_left1  =  l.fit_transform(df*"days_left"+)

df*"airline"+=airline1
df*"flight"+=flight1
df*"source_city"+=source_city1
df*"departure_time"+=departure_time1
df*"stops"+=stops1
df*"arrival_time"+=arrival_time1
df*"destination_city"+=destination_city1
```

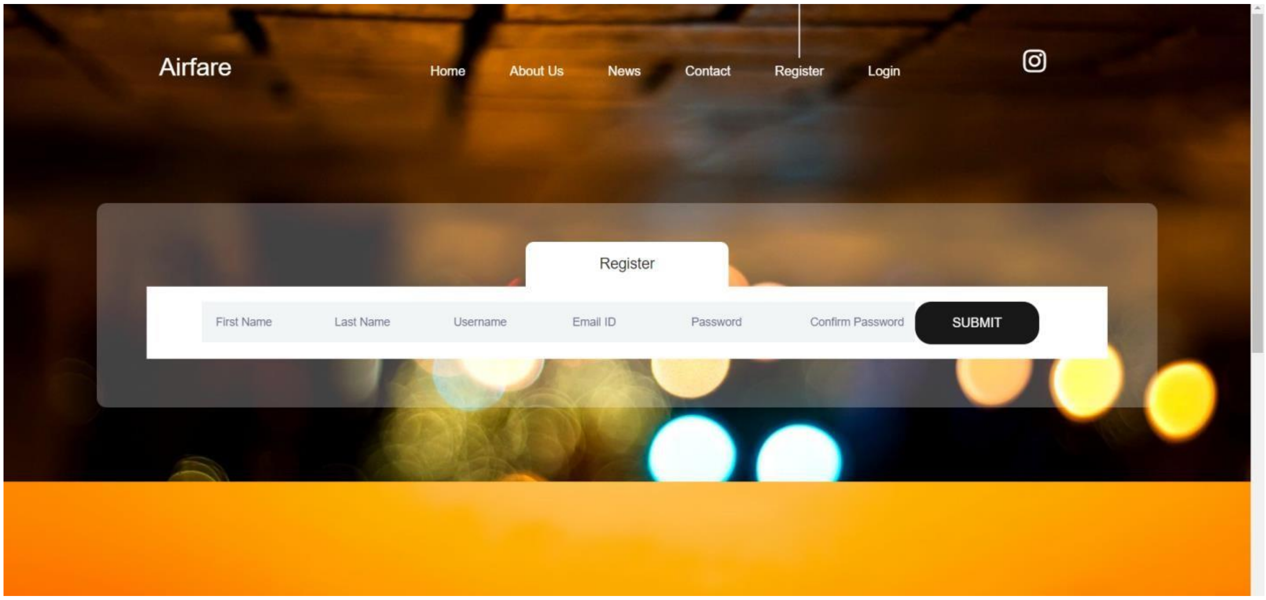
CHAPTER 10

SCREENSHOTS

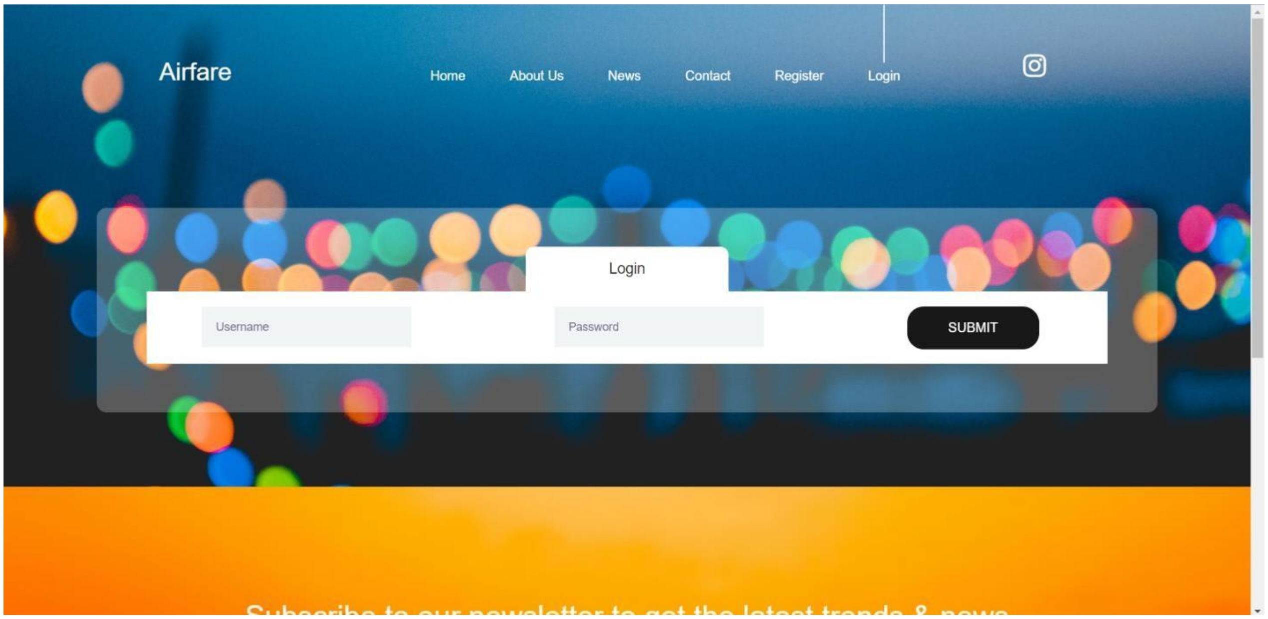
Home page:



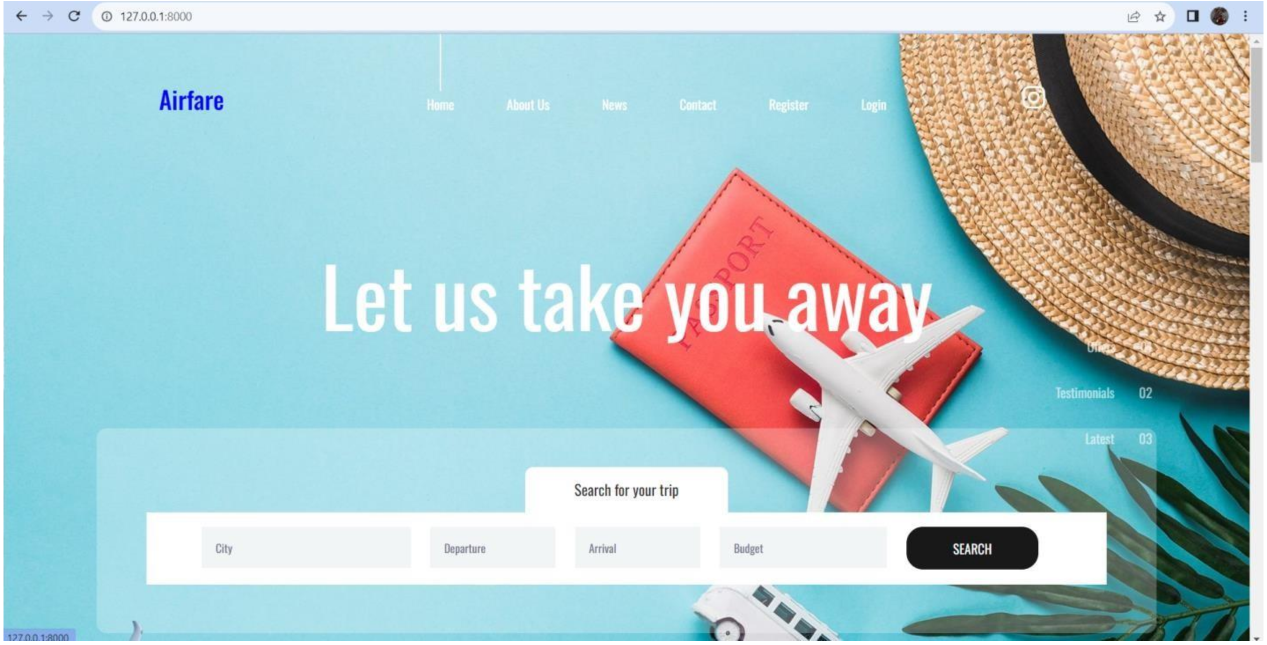
Register Page:



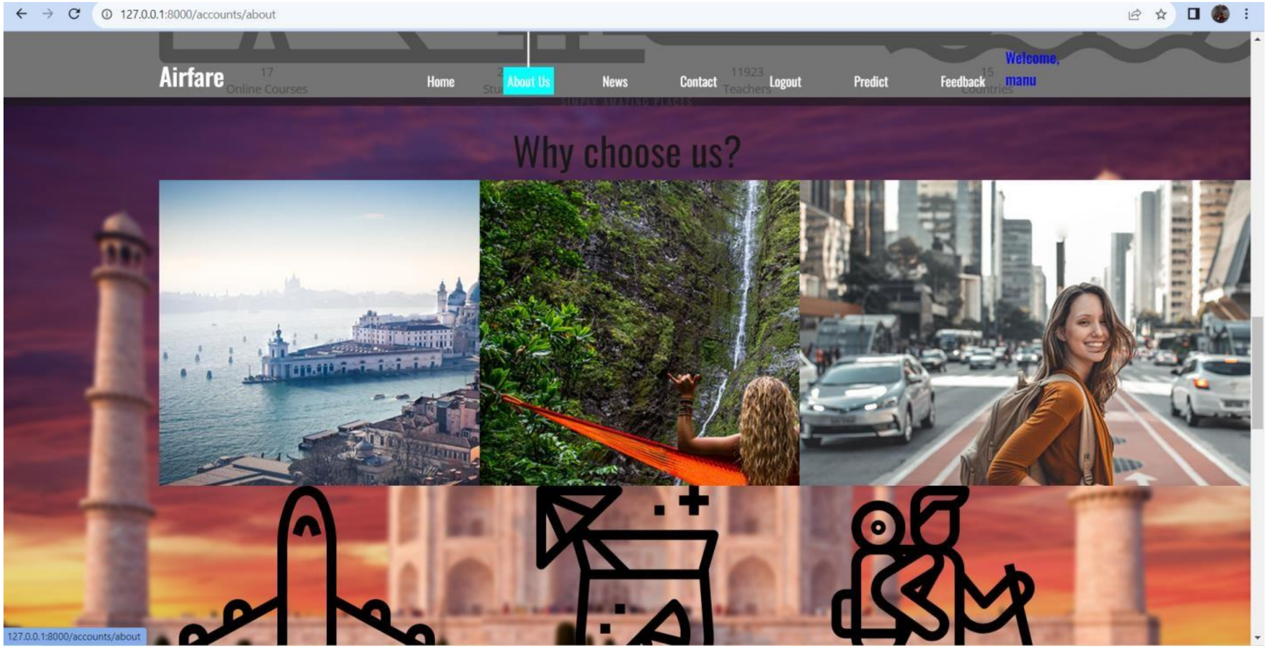
Login Page:



Index Page:

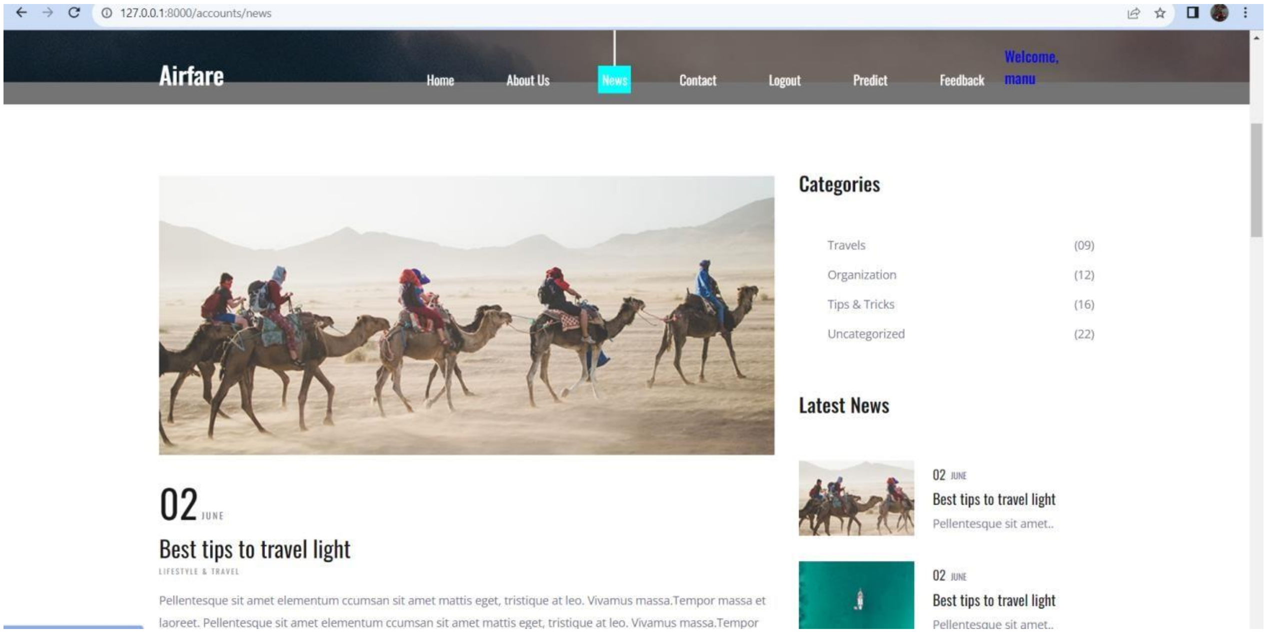


About Page:



Why choose us?

News Page:



Categories

Travels	(09)
Organization	(12)
Tips & Tricks	(16)
Uncategorized	(22)

Latest News

02 JUNE
Best tips to travel light
Pellentesque sit amet..

02 JUNE
Best tips to travel light
Pellentesque sit amet..

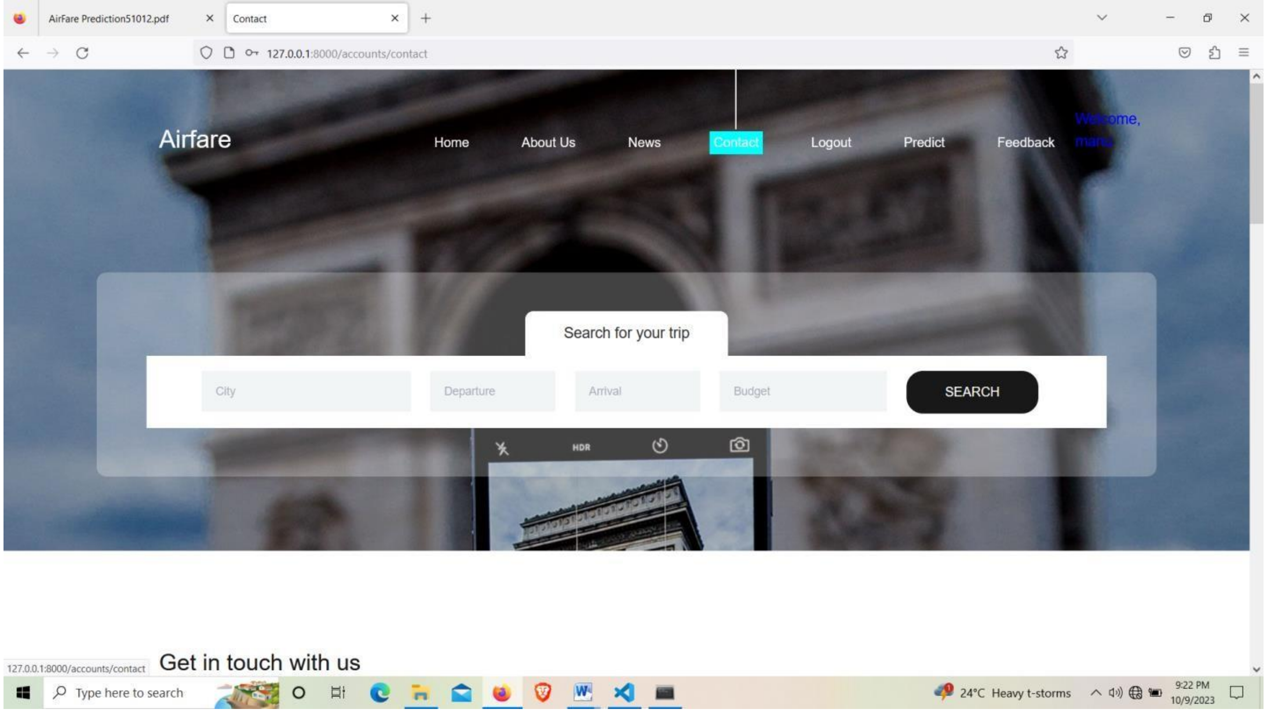
02 JUNE

Best tips to travel light

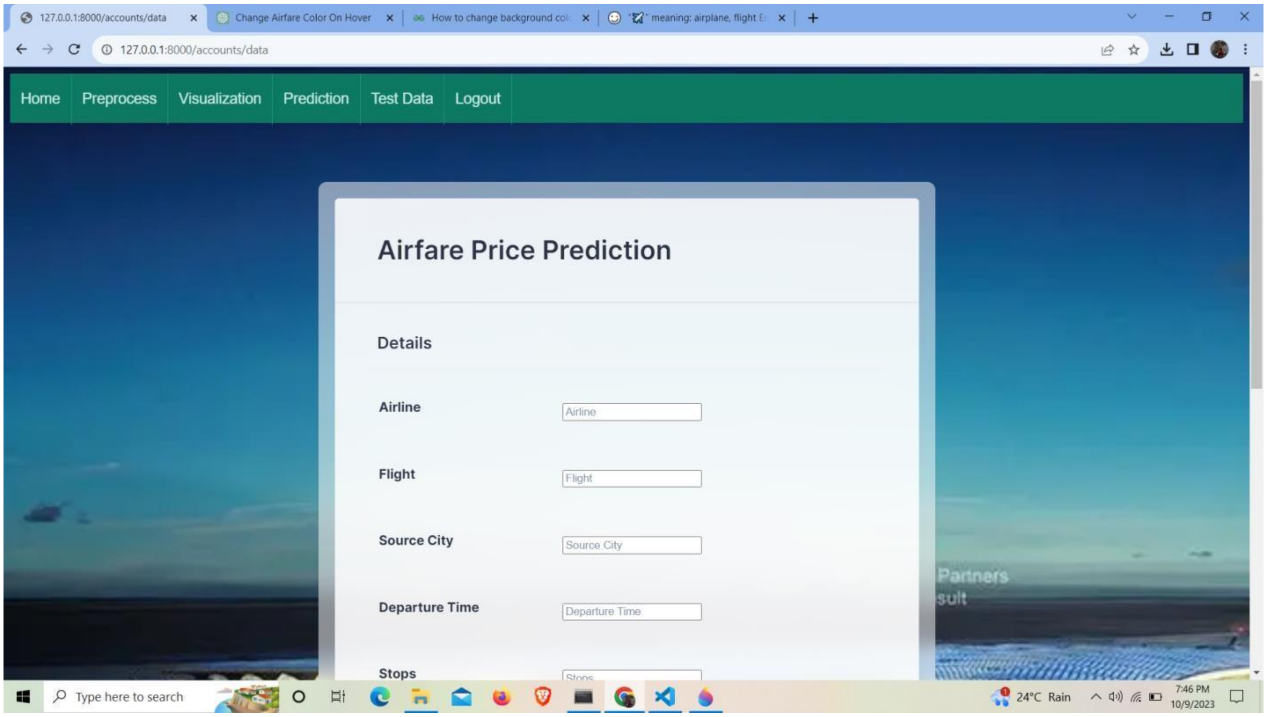
LIFESTYLE & TRAVEL

Pellentesque sit amet elementum ccumsan sit amet mattis eget, tristique at leo. Vivamus massa.Tempor massa et laoreet. Pellentesque sit amet elementum ccumsan sit amet mattis eget, tristique at leo. Vivamus massa.Tempor

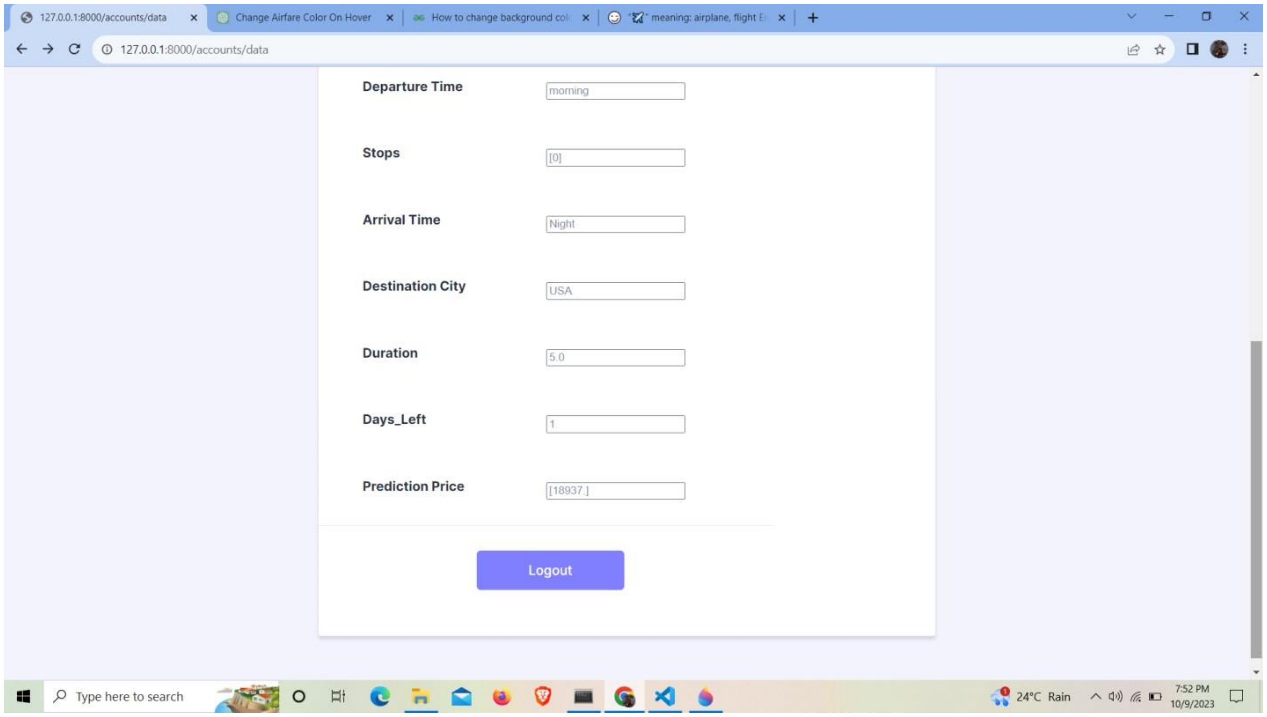
Contact Page:



Predict Data Page:



Predict Page:


127.0.0.1:8000/accounts/data
Change Airfare Color On Hover
How to change background col
meaning: airplane, flight E
Type here to search
24°C Rain
7:52 PM
10/9/2023
Logout

Departure Time: morning

Stops: [0]

Arrival Time: Night

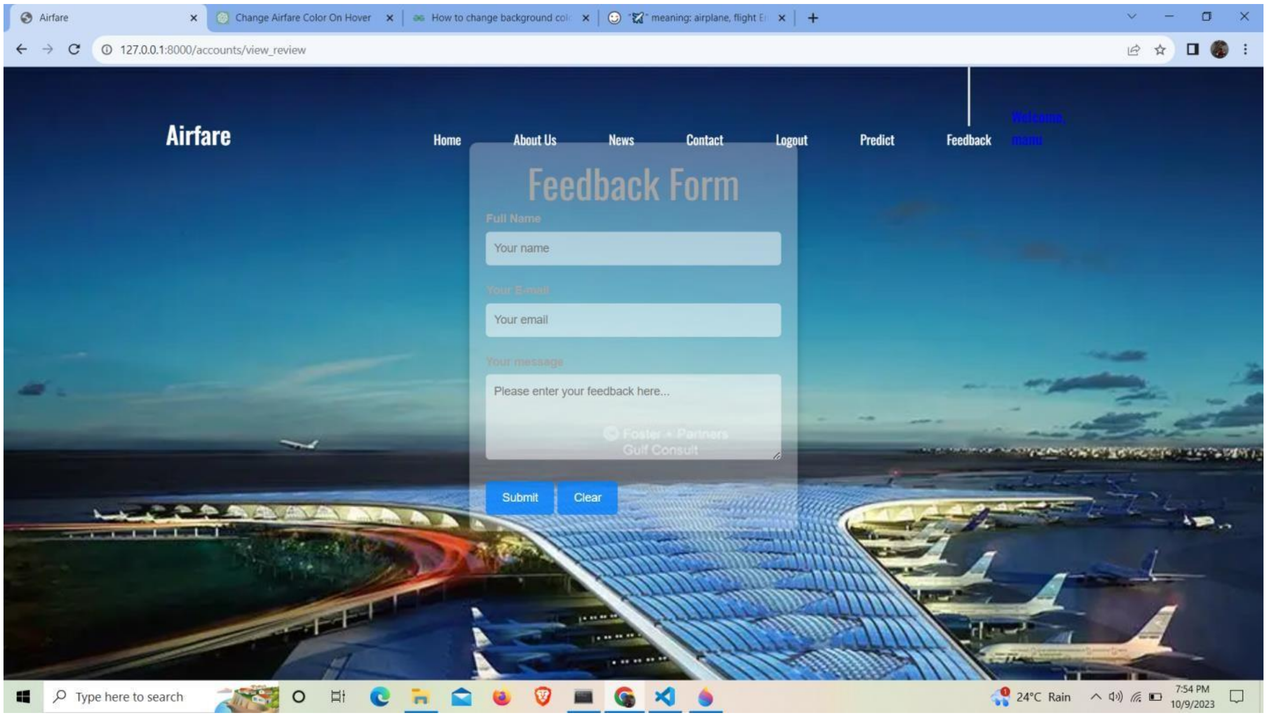
Destination City: USA

Duration: 5.0

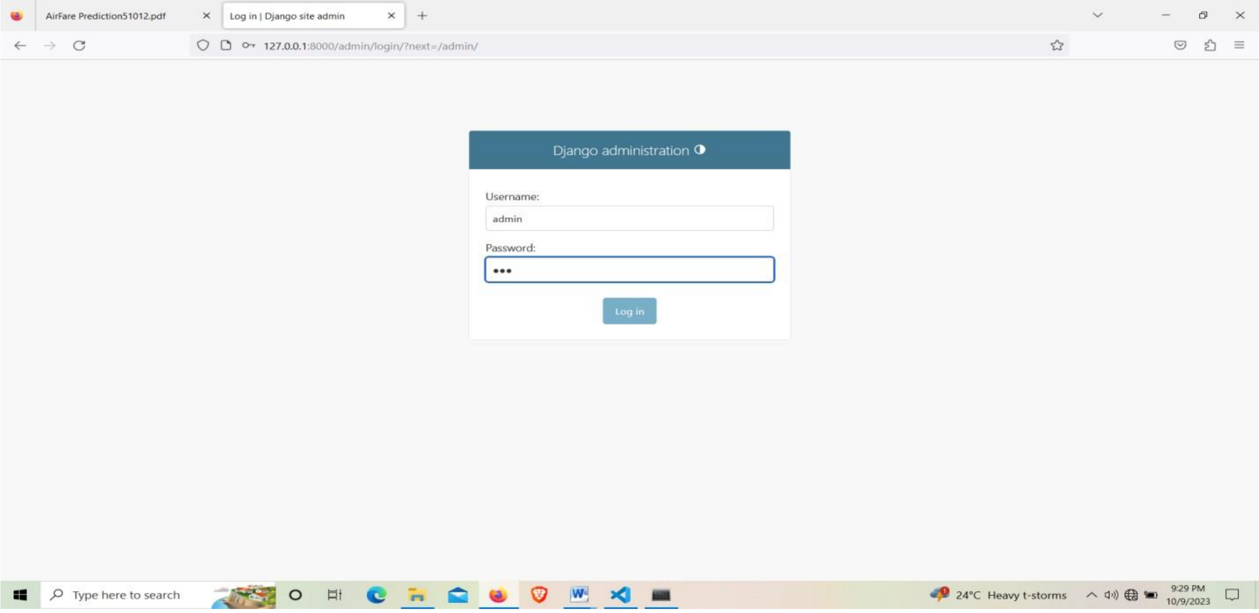
Days_Left: 1

Prediction Price: [18937.]

Feedback Page:



Admin:



AirFare Prediction51012.pdf
Log in | Django site admin
127.0.0.1:8000/admin/login/?next=/admin/
Type here to search
24°C Heavy t-storms
9:29 PM
10/9/2023

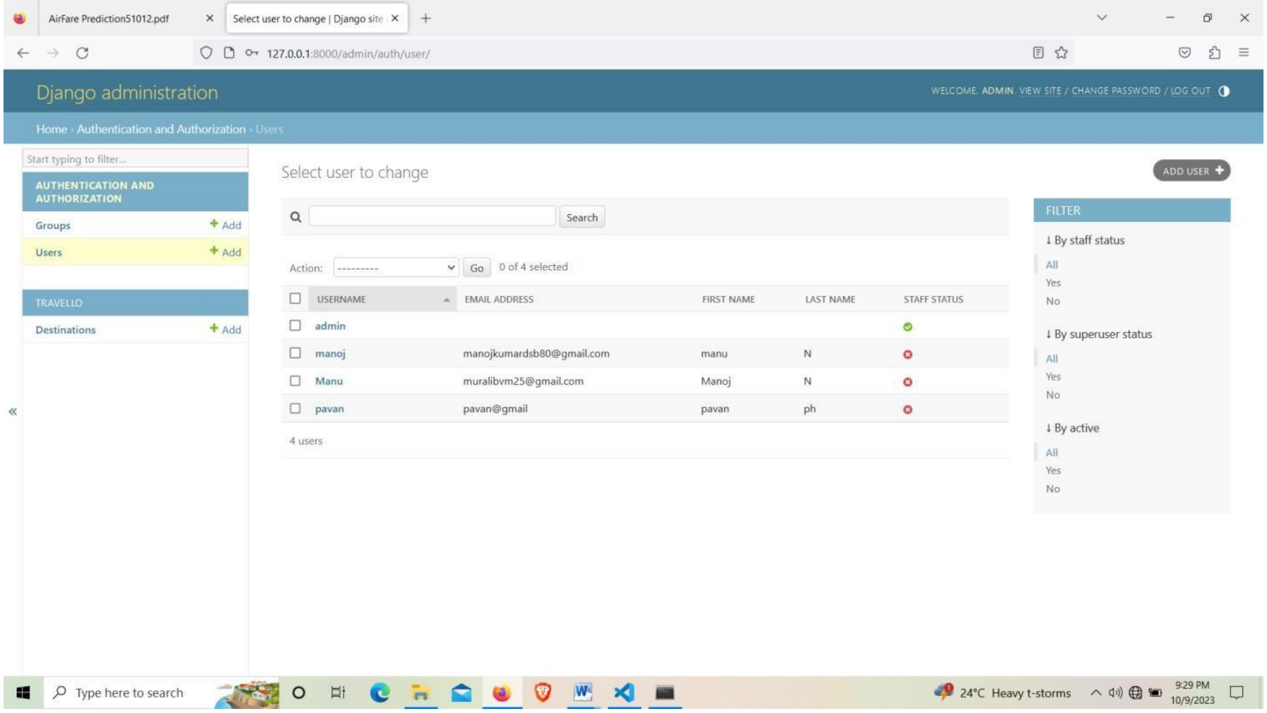
Django administration

Username: admin

Password: •••

Log in

Admin View:



AirFare Prediction51012.pdf
Select user to change | Django site
127.0.0.1:8000/admin/auth/user/
WELCOME, ADMIN. VIEW SITE / CHANGE PASSWORD / LOG OUT
Home › Authentication and Authorization › Users
Start typing to filter...
AUTHENTICATION AND AUTHORIZATION
Groups + Add
Users + Add
TRAVELLO
Destinations + Add
ADD USER +
Search
Action: ---------
Go
FILTER
↓ By staff status
All
Yes
No
↓ By superuser status
All
Yes
No
↓ By active
All
Yes
No
Type here to search
24°C Heavy t-storms
9:29 PM
10/9/2023

Django administration

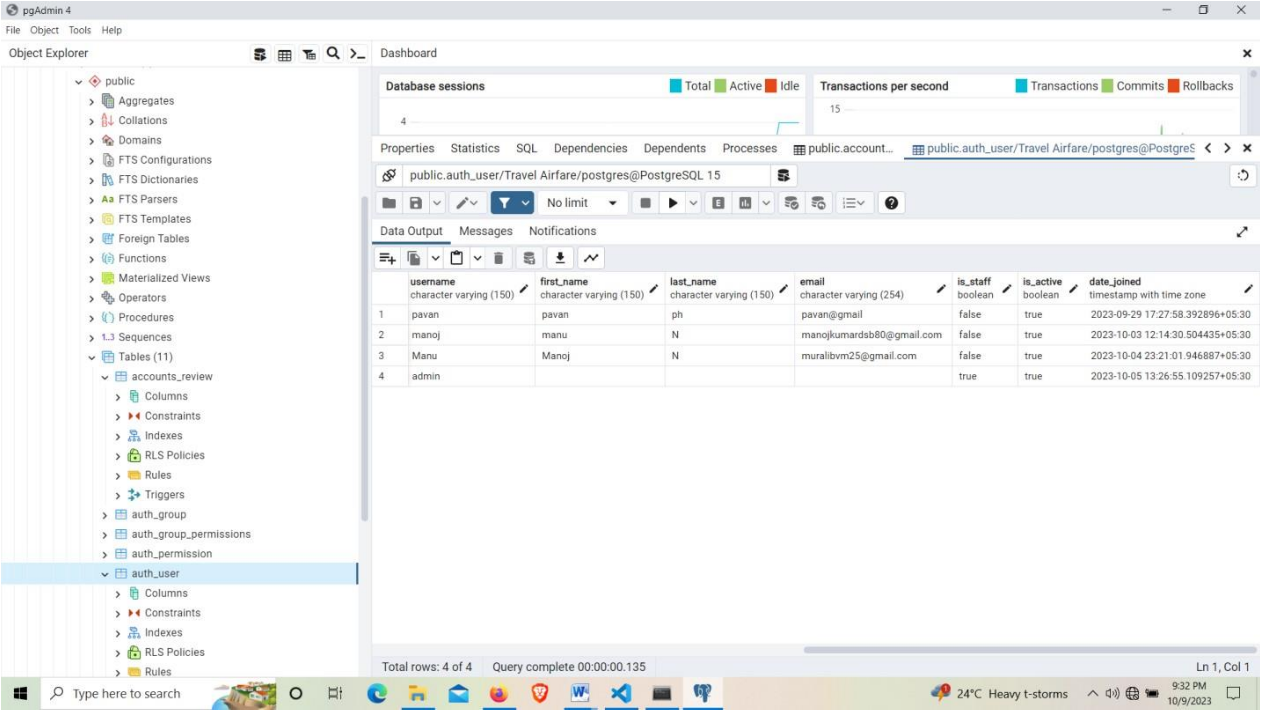
Select user to change

0 of 4 selected

□	USERNAME	EMAIL ADDRESS	FIRST NAME	LAST NAME	STAFF STATUS
□	admin				✓
□	manoj	manojkumardsb80@gmail.com	manu	N	✗
□	Manu	muralibvm25@gmail.com	Manoj	N	✗
□	pavan	pavan@gmail	pavan	ph	✗

4 users

Database View: Login details

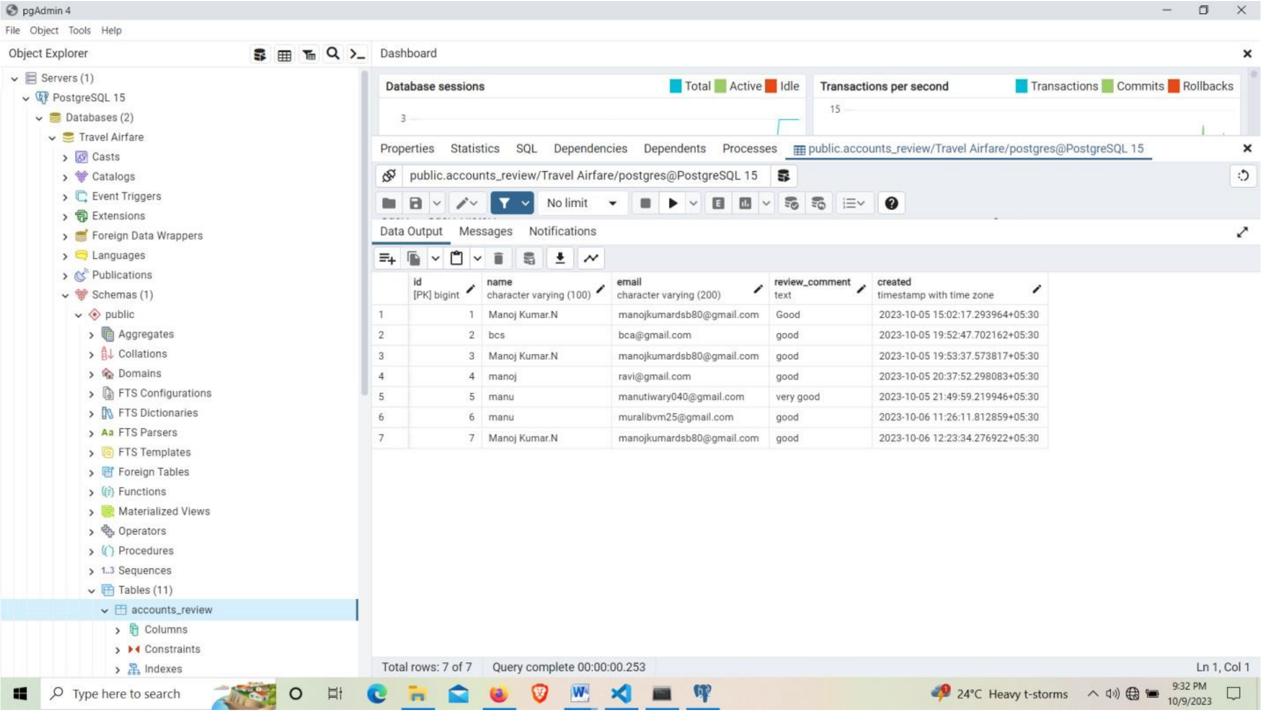


public.auth_user/Travel Airfare/postgres@PostgreSQL 15

	username character varying (150)	first_name character varying (150)	last_name character varying (150)	email character varying (254)	is_staff boolean	is_active boolean	date_joined timestamp with time zone
1	pavan	pavan	ph	pavan@gmail	false	true	2023-09-29 17:27:58.392896+05:30
2	manoj	manu	N	manojkumardsb80@gmail.com	false	true	2023-10-03 12:14:30.504435+05:30
3	Manu	Manoj	N	muralibvm25@gmail.com	false	true	2023-10-04 23:21:01.946887+05:30
4	admin				true	true	2023-10-05 13:26:55.109257+05:30

Total rows: 4 of 4 Query complete 00:00:00.135 Ln 1, Col 1

Feedback Details:



public.accounts_review/Travel Airfare/postgres@PostgreSQL 15

	id [PK] bigint	name character varying (100)	email character varying (200)	review_comment text	created timestamp with time zone
1	1	Manoj Kumar.N	manojkumardsb80@gmail.com	Good	2023-10-05 15:02:17.293964+05:30
2	2	bcs	bca@gmail.com	good	2023-10-05 19:52:47.702162+05:30
3	3	Manoj Kumar.N	manojkumardsb80@gmail.com	good	2023-10-05 19:53:37.573817+05:30
4	4	manoj	ravi@gmail.com	good	2023-10-05 20:37:52.298083+05:30
5	5	manu	manutiwary040@gmail.com	very good	2023-10-05 21:49:59.219946+05:30
6	6	manu	muralibvm25@gmail.com	good	2023-10-06 11:26:11.812859+05:30
7	7	Manoj Kumar.N	manojkumardsb80@gmail.com	good	2023-10-06 12:23:34.276922+05:30

Total rows: 7 of 7 Query complete 00:00:00.253 Ln 1, Col 1