KALINGA INSTITUE OF INDUSTRIAL TECHNOLOGY, BHUBNESWAR, ODISHA



SOFTWARE REQUIREMENT SPECIFICATION

(SRS DOCUMENT)

FOR

AIRLINE PRICE PREDICTION USING MACHINE LEARNING

**SUPERVISOR**

**Dr. SAURABH BILGAIYAN**

**BY**

**SOMYA SINHA 21053365/ CSE - 37**

**Table of Contents**

**Table of Contents ii**

**1. Introduction 1**

1.1 Purpose 1

1.2 Product Scope 1

1.3 Definitions,acronyms and abbreviations 1

1.4 Technologies to be used 1

**2. Overall Description 2**

2.1 Product Perspective 2

2.2 User Classes and Characteristics 2

2.3 Operating Environment 2

2.4 User Documentation 2

2.5 Software interfaces 2

2.6 Hardware interfaces 2

2.7 External forces 3

2.7.1 Dependencies.........................................................................................................3

2.7.2 Constraints.............................................................................................................3

**3. System Features 3**

3.1 Functional requirements 3

3.2 Non-Functional requirements 4

**4. Design 4**

4.1 System Design(frontend design) 4

4.1.1 Usecase diagram 4

4.1.2 Class diagram 4

4.1.3 Sequence diagram 4

4.1.4 Activity diagram 4

4.2 Database Design 4

4.2,1 er-diagram................................................................................................................4

**5. GUI 5**

**6. FUTURE SCOPE 5**

**7. CONCLUSION 5**

**8. REFERENCES 6**

1. **Introduction**
   1. **Purpose:**

The purpose of predicting airline fare tickets is to help airlines maximize their profits by balancing individual ticket pricing and the capacity of their flights

Airlines use algorithms to determine the pricing of their tickets based on various factors such as passenger demand, seasonality, holidays, and the number of available airlines and flights. The algorithms are trained on the arcane rules of airfare, plus reams of historical data, and use that to hazard when customers should buy to get the best ticket price. The pricing of tickets is not fixed or controlled, and airlines seek to maximize profits made through ticket sales. The concept of ‘booking classes’ is used to determine the pricing and availability of fares. These are different from travel classes (economy, premium economy, and business class, for example) and are a series of letters that define the fare level paid. By predicting the optimal price to set for a ticket, airlines can maximize revenue while remaining competitive.

* 1. **Product Scope:**

- The purpose of predicting airline fare tickets is to help airlines maximize their profits by balancing individual ticket pricing and the capacity of their flights.

- The product scope for predicting airline fare tickets should include details on what problem the product is solving and why, what to build, and a detailed description of the solution in terms of features, user stories, use cases, and functionalities.

- The product scope should identify the characteristics and functions of the product or service, including physical features such as size and materials, as well as functional specifications.

- The product scope should focus on the result or the actual offering, which is the final product or service, and may also refer to a service or other item for customer use.

- The product scope for predicting airline fare tickets should consider how to evaluate whether the object is on track for completion and whether it meets the expected outcome.

- The product scope should summarize the different activities that are part of the production cycle, including the product development process, procurement and transportation of raw materials, parts and components, assembling process, quality control procedures, and distribution channels.

- The product scope for predicting airline fare tickets should be concise, attention-grabbing, and easy to scan for readers.

**1.3 Definition, Acronyms and Abbreviations:**

- API: Application Programming Interface. An interface that allows software applications to communicate with each other.

- AHELO: Assessment of Higher Education Learning Outcomes. A program developed by the OECD to assess the learning outcomes of higher education students.

- CAPI: Computer-Assisted Personal Interviews. A method of conducting surveys where the interviewer uses a computer to record responses.

- CD: Compact Disk. A digital optical disc used for storing data.

- FISIM: Financial Intermediation Services Implicitly Measured. A measure of the value of financial services provided by banks and other financial intermediaries.

- GDP: Gross Domestic Product. The total value of goods and services produced in a country in a given period.

- GUI: Graphical User Interface. A type of user interface that allows users to interact with electronic devices through graphical elements such as icons, buttons, and windows.

- HTML: Hypertext Markup Language. A standard markup language used to create web pages.

- JSON: JavaScript Object Notation. A lightweight data interchange format that is easy for humans to read and write and easy for machines to parse and generate.

- OECD: Organisation for Economic Co-operation and Development. An international organization that promotes economic development and cooperation among member countries.

- SRS: Software Requirements Specification. A document that describes the requirements for a software product or system.

- XML: Extensible Markup Language. A markup language that is used to encode documents in a format that is both human-readable and machine-readable.

**1.4 Technologies to be used:**

- Machine learning algorithms: Many companies, including Hopper, Kayak, Google Flights, Skyscanner, and others, use machine learning algorithms to predict flight prices. These algorithms are trained on historical data, such as price fluctuations, airline behavior, and economic conditions, to help travelers predict costsx.

- Data mining techniques: Data mining techniques like rule learning, reinforcement learning, time-series methods, and their combinations can be used to achieve greater accuracy in predicting the fare of flights.

- Artificial intelligence and deep learning techniques: Using artificial intelligence and deep learning techniques, an estimation of flight fares at a given time can be obtained within seconds.

- Flight data APIs: Google Flights provides a dedicated airline price prediction API to help customers save on airfare for any particular date.

- Price intelligence tools: Many businesses and price intelligence tools are based largely on data, crunching through a huge amount of historical data to help travelers predict costs.

- Fuel and oil price analysis: Fuel and oil prices are significant factors that impact airline ticket pricing.

- Ancillary revenue optimization: Airlines can use algorithms and forecasting to set base fares, plan seat availability and demand, and predict and optimize revenues from ancillary products and services.

1. **Overall Description**
   1. **Product Perspective**

The airline price prediction model serves as a valuable tool for users seeking the most cost-effective flight options. It operates as a standalone application with the primary objective of providing accurate forecasts of airline ticket prices based on comprehensive data analysis and advanced prediction algorithms.

**Key Features:**

* **Price Optimization:** The system's core functionality revolves around analyzing historical pricing data, market trends, and various influencing factors to predict the most opportune time to book a flight for the lowest fare.
* **Real-Time Updates:** The model provides users with timely updates on predicted price changes, allowing them to make informed decisions about their travel plans.
* **Intuitive User Interface:** The application offers an intuitive and user-friendly interface, ensuring that users can easily input their travel preferences and receive concise, actionable recommendations.
* **Integration Capabilities:** The model is designed to seamlessly integrate with existing online booking platforms and travel agency systems, enhancing the overall booking experience for users.
  1. **User Classes and Characteristics**

The users of the Airline Price Prediction system can be categorized into the following classes:

**2.2.1. End Users:**

**Characteristics:**

* Typically travelers or individuals seeking to estimate airfare prices.
* May not possess extensive knowledge of machine learning or data analysis techniques.
* Interact with the system through a user-friendly interface.

**Roles and Responsibilities:**

* Input relevant information such as departure and arrival destinations, travel dates, etc.
* View the predicted airfare prices generated by the system.

**2.2.2. Administrators :**

**Characteristics:**

* Have a deeper understanding of the system's functionality and technical aspects.
* Responsible for system maintenance, updates, and ensuring data integrity.
* Possess the ability to configure system parameters (if applicable).

**Roles and Responsibilities:**

* Monitor system performance and address any technical issues.
* Manage system configurations and data sources.

**2.2.3. Developers :**

**Characteristics:**

* Possess technical expertise in software development and machine learning.
* Involved in the creation, modification, and enhancement of the prediction model.
* Knowledgeable about programming languages and libraries used in the project.

**Roles and Responsibilities:**

* Develop, optimize, and fine-tune the prediction model.
* Implement new features or improvements based on feedback and requirements.

**2.2.4. Data Providers :**

**Characteristics:**

* Entities responsible for supplying data used in training and testing the prediction model.
* Can be airlines, travel agencies, or data providers offering relevant information.

**Roles and Responsibilities:**

* Supply accurate and up-to-date data for training purposes.
* Ensure data quality and compliance with privacy regulations.
  1. **Operating Environment**

The Airline Price Prediction system is designed to operate in a standard computing environment. It is compatible with the following operating systems and platforms:

**Operating Systems:**

* Windows 10 or later
* macOS Catalina (10.15) or later
* Linux distributions (e.g., Ubuntu 18.04 LTS or later)

**Web Browsers:**

* Google Chrome (latest version)
* Mozilla Firefox (latest version)
* Apple Safari (latest version)
* Microsoft Edge (latest version)

**Software Dependencies:**

* Python 3.7 or later
* Required Python libraries and frameworks ( NumPy, urllib, Pandas)

**Internet Connection:**

An active internet connection is required for fetching real-time data (if applicable) and accessing external data sources.

**Hardware Requirements:**

* Processor: Dual-core CPU or equivalent
* RAM: 4GB or higher
* Storage: 20GB free disk space (for model training and data storage)
* Display: Minimum 1366x768 resolution

**Development Environment (For Developers):**

* Integrated Development Environment (IDE) compatible with Python (Jupyter Notebook is being used here)
* Version control system for collaborative development.

**Data Sources:**

Access to airline data sources(live), APIs, for model training and prediction.

**Security Considerations:**

If handling sensitive user information, encryption protocols and secure connections should be implemented to ensure data privacy and security.

* 1. **User Documentation**

The documentation includes the following components:

**User Manual:**

A detailed guide on how to use the Airline Price Prediction system, including step-by-step instructions for various functionalities.

**Getting Started:**

An introduction to the system, outlining its purpose and benefits.

**System Overview:**

An overview of the features and capabilities of the application.

**User Interface Guide:**

Description of the user interface elements, navigation, and interaction.

**Input Data Guidelines:**

Instructions on how to provide input data (if applicable), including data format and requirements.

**Prediction Process:**

Explanation of how the prediction process works, including any advanced options or settings.

**Viewing Results:**

Instructions on how to interpret and utilize the prediction results generated by the system.

**Troubleshooting:**

Common issues and their resolutions, along with tips for optimal system performance.

**Frequently Asked Questions (FAQs):**

A compilation of frequently asked questions and their answers to assist users in resolving queries.

**Support and Contact Information:**

Information on how users can seek further assistance or report any issues encountered during system use.

**Appendices :**

Additional resources, glossary of terms, or supplementary information to enhance user understanding.

**Feedback Mechanism:**

Instructions on how users can provide feedback, suggestions, or report any concerns about the application.

* 1. **Software Interfaces**

The Airline Price Prediction system interfaces with several software components to ensure seamless operation. These interfaces include:

**Programming Language:**

The system is primarily developed using a Python and utilizes libraries, frameworks, and tools associated with it.

**Machine Learning Libraries:**

The system integrates machine learning libraries such as scikit-learn, TensorFlow, or PyTorch for building, training, and deploying predictive models.

**Database Management System (DBMS):**

The system involves data storage and retrieval, hence it interfaces with a DBMS (MySQL) to manage the database.

**Data Processing Tools:**

Tools for data preprocessing and feature engineering, such as pandas, NumPy, or specialized libraries for data manipulation.

**External APIs :**

Integration with external APIs for tasks like fetching real-time flight data, weather information, or other relevant data sources.

**Development Environment:**

The system is developed and tested in a specific integrated development environment (IDE) or code editor(jupyter).

**Version Control System :**

Use of version control to manage code repositories, track changes, and facilitate collaboration among developers.

**Deployment Platforms:**

The system may be deployed on specific platforms (e.g., cloud services like AWS, Google Cloud, or on-premises servers).

**Operating Systems**:

Compatible with specific operating systems (e.g., Windows, Linux, macOS) on which the system is designed to run.

**Frontend Technologies :**

If the system includes a user interface, it may interface with frontend technologies like HTML, CSS, JavaScript, and related frameworks.

**Visualization Libraries (if applicable):**

Integrated with visualization libraries (e.g., Matplotlib, Seaborn, D3.js) for generating charts, graphs, and visual representations of data

* 1. **Hardware Interfaces**

The term "hardware interface" refers to the interaction or connection points between software and hardware components. These interfaces are crucial for software systems to communicate with and control hardware devices. In the context of dynamic price prediction for airlines, here are the key hardware interfaces relevant to software engineering:

1. APIs (Application Programming Interfaces): APIs allow software to interact with external hardware and systems. For dynamic price prediction, software can use APIs to retrieve data from various sources like reservation systems, weather services, or financial data providers.

2. Device Drivers: Device drivers are software components that enable an operating system to communicate with and control hardware devices. For instance, software may need drivers to interact with sensors or data acquisition devices.

3. POS Terminal Integration: Software interfaces with point-of-sale terminals used at airline ticket counters and kiosks, ensuring that dynamic pricing updates are reflected in the user interface.

4. Payment Gateways:Software interacts with payment gateways through APIs to handle online ticket sales and payment processing securely.

5. Cloud Service Integrations:Software systems can interface with cloud services and platforms for data storage, computation, and scalability. Cloud APIs and SDKs are commonly used for this purpose.

6. Network Protocols: Dynamic pricing software communicates with hardware components over networks using various protocols, such as HTTP/HTTPS, FTP, or custom protocols for data exchange.

7. IoT Interfaces: For real-time data from IoT devices (e.g., sensors on aircraft or in airports), software engineers design and implement interfaces to collect and process this data.

Creating effective hardware interfaces is essential to ensure that software systems can interact with and control hardware components seamlessly. These interfaces enable the software to collect, process, and act upon data in real time, allowing airlines to optimize ticket prices based on dynamic factors while providing a smooth user experience.

**2.8. EXTERNAL FORCES**

Regulatory Changes in Airline Industry: Changes in regulations, such as new pricing policies or restrictions, may influence the accuracy and effectiveness of the price prediction model.

Technological Advancements: Advances in data analytics, machine learning techniques, or computing infrastructure may offer opportunities to enhance the accuracy and capabilities of the prediction algorithms.

Market Fluctuations: Economic factors, such as changes in fuel prices or geopolitical events, can have a significant impact on airline ticket prices, affecting the accuracy of predictions.

Emergence of New Airlines or Routes: New entrants in the airline industry or the introduction of new flight routes could introduce unpredictability in pricing patterns.

User Preferences and Behavior: Shifts in traveler preferences, such as changes in preferred airlines or travel times, may necessitate adjustments in the prediction model's algorithms.

Global Events and Crises: Events like natural disasters, political instability, or health crises (e.g., pandemics) can lead to sudden and drastic changes in airline ticket prices.

Integration Challenges with Booking Platforms: Changes or updates in the APIs or systems of online booking platforms may require adjustments in the model's integration capabilities.

Competitive Pricing Strategies: Changes in how airlines compete in terms of pricing (e.g., fare wars, discounts) can affect the accuracy of predictions.

Currency Exchange Rates: Fluctuations in currency exchange rates can impact the pricing of international flights, which in turn affects the accuracy of predictions.

Data Privacy and Security Regulations: Evolving data privacy laws may require adjustments in how the model handles and analyzes customer data.

Customer Feedback and Preferences: User feedback and changing customer expectations may drive the need for updates or additional features in the application's user interface.

Environmental Considerations: Environmental policies and initiatives within the airline industry may influence pricing strategies and, consequently, the accuracy of price predictions.

These external forces should be taken into account during the development and maintenance of the airline price prediction model to ensure its continued effectiveness and relevance in providing accurate forecasts for users seeking cost-effective flight options.

1. **System Features**
   1. **Functional Requirements**

R.1: User Registration

Description: Users must register to access the airline price prediction tool.

R.1.1: Select registration option

Input: "Register" option selection

Output: Prompt message to enter personal details.

R.1.2: Submit registration details

Input: Personal details (name, email, password, etc.)

Output: User has been registered

R.2: User Login

Description: Registered users can log in to access the airline price prediction tool.

R.2.1: Login option

Input: User ID and Password

Output: User logged in, or asked to enter valid account details if unsuccessful

R.2.2: OTP Verification

Input: OTP sent to user's registered email for confirmation

Output: User is logged in, or asked to re-enter OTP if unsuccessful

R.3: Search Flights

Description: Allows users to search for flights based on various criteria.

R.3.1: Select search option

Input: "Search" option selection

Output: User prompted to enter search criteria (departure city, destination, dates, etc.)

R.3.2: Search and display

Input: Search criteria

Output: User is displayed matching flight options

R.4: Price Prediction

Description: Provides accurate forecasts of airline ticket prices.

R.4.1: Select flight for prediction

Input: User selects a specific flight for price prediction

Output: Predicted price for the selected flight is displayed

R.4.2: Real-Time Updates

Input: User requests for real-time updates on predicted price changes

Output: User receives timely updates on predicted price changes

R.5: Integration with Booking Platforms

Description: The model integrates with existing online booking platforms and travel agency systems.

R.5.1: Seamless Integration

Input: User selects the option to book a flight

Output: User is directed to the integrated booking platform

R.6: User Preferences

Description: Users can input their travel preferences for more accurate predictions.

R.6.1: Input travel preferences

Input: User provides information such as preferred airlines, class, layovers, etc.

Output: System considers user preferences in price predictions

R.7: Logout

Description: Allows users to log out of the system.

R.7.1: Select logout option

Input: "Logout" option selection

Output: User has been logged out of the system

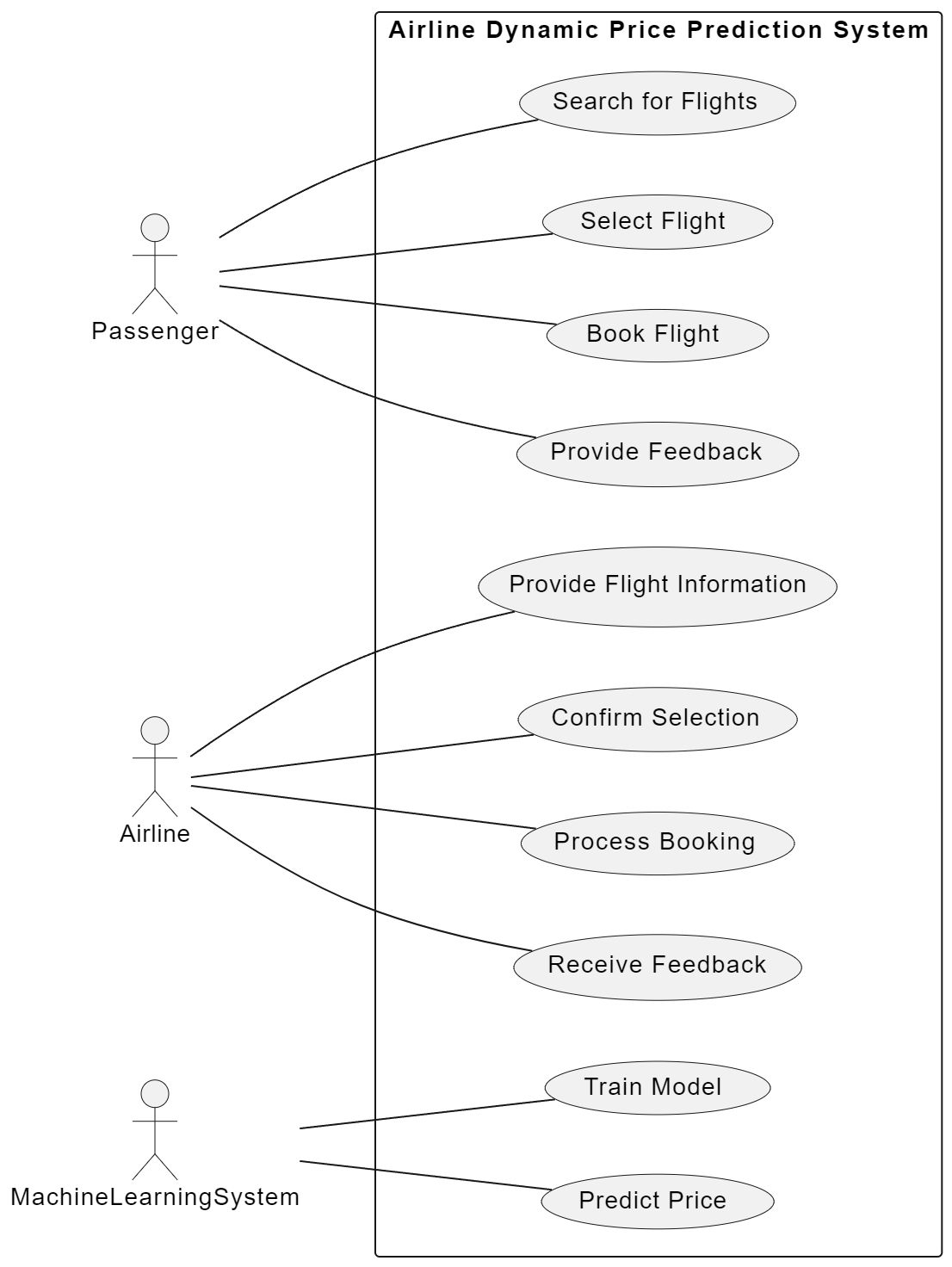
* 1. **Non- Functional Requirements:**

Non-functional requirements specify how a system should perform or behave rather than defining its specific functionality. Here are some non-functional requirements that are :

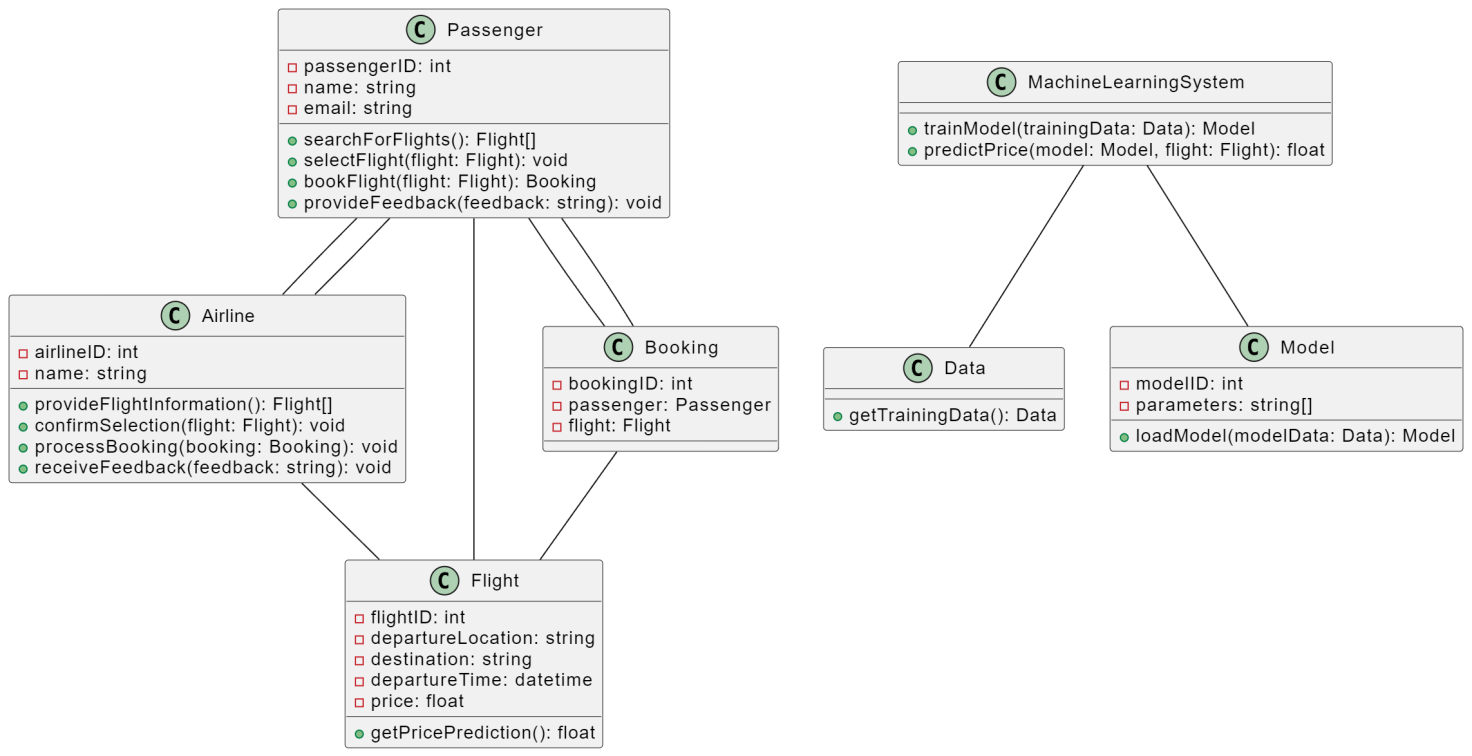
* Performance: The system should be able to provide fast and accurate predictions of airline ticket prices
* Reliability: The system should be reliable and available 24/7 to ensure that users can access it at any time
* Scalability: The system should be scalable to accommodate future growth and handle large amounts of data
* Usability: The system should be user-friendly and easy to use for both technical and non-technical users
* Security: The system should have robust security measures to protect the data and prevent unauthorized access
* Maintainability: The system should be easy to maintain and update to ensure that it continues to provide accurate predictions.
* Robustness: The system should be able to handle errors and exceptions gracefully and recover from failures quickly

1. **External Interface Requirements**
   1. **System Design(front-end design)**

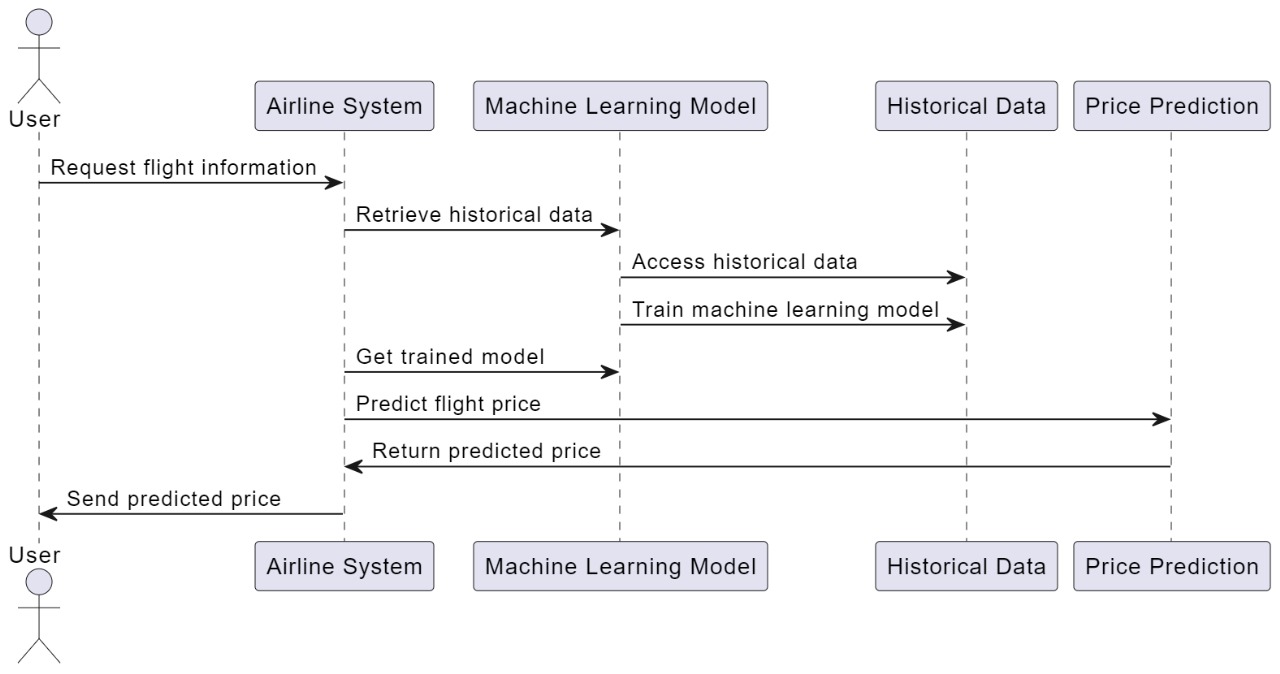
**4.1.1. Use Case Diagrams**



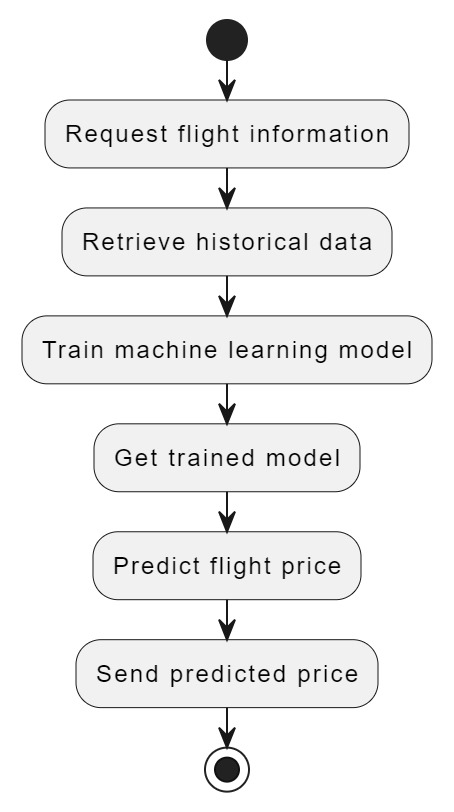
**4.1.2 Class Diagram:**



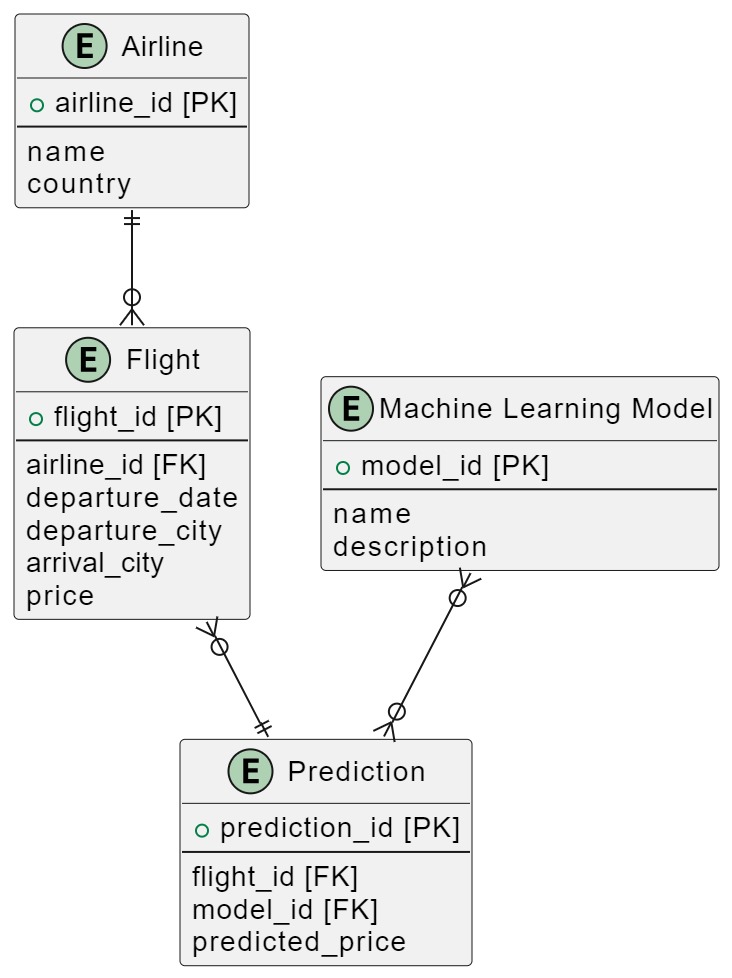
**4.1.3. Sequence Diagram**



**4.1.4. Activity Diagram**



* 1. **Database Design**
     1. **E-R Diagram**



1. **Future Scope**
2. **Integration of Additional Features:**

* Personalized recommendations
* Multi-city flight predictions
* Real-time fare updates

1. **Enhanced Data Sources:**

* Include more airlines and travel agencies
* Utilize up-to-date fare comparison platforms

1. **Advanced Machine Learning Models***:*

* *Explore latest algorithms for higher prediction accuracy*

1. **Multi-Platform Support:**

* *Develop mobile apps or responsive web designs*

1. **User Feedback and Ratings:**

* *Enable users to rate and review predictions*

1. **Historical Price Trends Analysis:**

* *View past price trends for informed decisions*

1. **Integration with Payment Gateways:**

* *Allow direct flight booking through the app*

1. **Notification and Alerts System:**

* *Alert users on significant fare changes or promotions*

1. **Multi-Lingual Support:**

* *Cater to non-English-speaking users*

1. **Social Media Integration:**

* *Share travel plans and predictions on social platforms*

1. **Integration with Travel Accommodations:**

* *Predict hotel, transportation, and package costs.*

1. **Geographical Expansion:**

* *Extend services to cover more regions and continents.*

1. **CONCLUSION:**

* To predict airline fare tickets, various technologies can be used, including machine learning algorithms, data mining techniques, artificial intelligence and deep learning techniques, flight data APIs, price intelligence tools, fuel and oil price analysis, and ancillary revenue optimization. The GUI for a flight fare prediction tool should be user-friendly, display the predicted fare based on inputted details, allow users to set alerts and compare fares across different airlines and flights, and have a responsive design that works well on both desktop and mobile devices.
* Fetching live data from a website can be achieved using APIs, web scraping tools, JavaScript, Python, polling, and push. It is important to choose the appropriate method based on the specific requirements of the project.
* In conclusion, writing an effective SRS document, using appropriate technologies, and fetching live data from a website are all critical components of developing a successful product that meets the needs of all stakeholders.

# REFERENCES

* Shukla, J., Srivastava, A., & Chauhan, A. Airline Price Prediction using Machine Learning.
* Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830
* Ren R., Yang Y., Yuan S. (2014). “Prediction of Airline Ticket Price”, University of Stanford
* S. Netessine and R. Shumsky, ‘‘Introduction to the theory and practice of yield management,’’ INFORMS Trans. Educ., vol. 3, no. 1, pp. 34–44 Sep. 2002
* M. V. Altaisky, N. E. Kaputkina, and V. A. Krylov, ‘‘Quantum neural networks: Current status and prospects for development,’’ Phys. Particles Nuclei, vol. 45, no. 6, pp. 1013–1032, Nov. 2014.