PARSHVANATH CHARITABLE TRUST'S

A.P. Shah Institute of Technology

Thane, 400615

Academic Year: 2022-23
Department of Computer Engineering

CSL605 SKILL BASED LAB COURSE: CLOUD COMPUTING

Mini Project Report

- **Title of Project : Flight Price Prediction**
- **Year and Semester : T.E. Sem 6**
- **Group Members Name and Roll No. : Pranav Chopdekar**
  Soujanya Chavan
  Vedang Deshpande

Table of Contents

Problem Definition

Travelling through flights has become an integral part of today's lifestyle, as more and more people are opting for faster travelling options. The flight ticket prices increase or decrease every now and then depending on various factors like timing of the flights, destination, duration of flights. various occasions such as vacations or festive season. Therefore, having some basic idea of the flight fares before planning the trip will surely help many people save money and time. In practice, many travelers tend to purchase flight tickets as early as possible to avoid possible price hikes. However, this type of purchase behavior does not always lead to the most economical flight tickets. The problem is to predict flight prices for different airlines and routes using machine learning algorithms. The goal is to build a model that can accurately forecast flight prices for a given route and airline, based on historical data and other relevant features. The flight industry is highly dynamic and the prices can fluctuate rapidly, so accurate predictions are critical for both customers and airlines. A machine learning model that can provide accurate price predictions will help travelers plan their trips more efficiently and airlines to optimize their revenue. The dataset for this problem will consist of historical flight prices for different airlines, routes, and dates. The dataset may also include other features such as departure and arrival times, duration of the flight, number of stops, and other relevant information. The objective of this project is to develop a machine learning model that can accurately predict the prices of flights for a given airline and route, based on the provided dataset. The model should be able to handle the dynamic nature of the industry and be able to provide accurate predictions in real-time. The performance of the model will be evaluated using standard evaluation metrics such as mean absolute error, mean squared error, and R-squared score. The goal is to develop a model that can achieve high accuracy and low error rates, while also being scalable and efficient for large datasets.

Introduction

Flight price prediction is an important issue that has not been thoroughly studied. In our research, we attempt to apply machine learning techniques to this problem to improve prediction accuracy. At the point when the white-collar class of India is presented to air travel, buyers searching at modest costs. The rate of flight tickets at the least cost is continuously expanding. Flight price prediction is an important issue that has not been thoroughly studied. This project aims to develop an application which will predict the flight prices for various flights using machine learning model. The user will get the predicted values and with its reference the user can decide to book their tickets accordingly. In the current day scenario flight companies try to manipulate the flight ticket prices to maximize their profits. There are many people who travel regularly through flights and so they have an idea about the best time to book cheap tickets. But there are also many people who are inexperienced in booking tickets and end up falling in discount traps made by the companies where actually they end up spending more than they should have. The proposed system can help save millions of rupees of customers by proving them the information to book tickets at the right time.

Air travel has become an essential part of our daily life, and the demand for air travel is increasing every day. However, the prices of flights can be highly unpredictable, and this makes it difficult for travellers to plan their trips efficiently. To overcome this problem, many airlines and travel agencies use machine learning algorithms to predict flight prices accurately. The aim of this project is to develop a machine learning model that can accurately predict flight prices for a given airline and route. The model will use historical data to learn the patterns and trends in flight prices, which will enable it to make accurate predictions in real-time. The dataset for this project will consist of historical flight prices for different airlines, routes, and dates. The dataset may also include other relevant features such as departure and arrival times, duration of the flight, number of stops, and other relevant information. The project will involve several steps, including data pre-processing, feature engineering, model selection, and evaluation. The data pre-processing step will involve cleaning and transforming the dataset to make it suitable for machine learning.

Feature engineering will involve selecting and transforming the relevant features to improve the performance of the model. Model selection will involve evaluating different machine learning algorithms and selecting the best one for the problem. Finally, the model's performance will be evaluated using standard evaluation metrics such as mean absolute error, mean squared error, and R-squared score. Overall, the successful completion of this project will result in a machine learning model that can accurately predict flight prices for a given airline and route, which will be useful for both travellers and airlines in planning their trips and optimizing their revenue, respectively.

Description

The project involves developing a machine learning model that can accurately predict flight prices for a given airline and route using historical data. The dataset for this project will consist of historical flight prices for different airlines, routes, and dates. The dataset may also include other relevant features such as departure and arrival times, duration of the flight, number of stops, and other relevant information.

The project will involve several steps, including data preprocessing, feature engineering, model selection, and evaluation. The data preprocessing step will involve cleaning and transforming the dataset to make it suitable for machine learning. Feature engineering will involve selecting and transforming the relevant features to improve the performance of the model. Model selection will involve evaluating different machine learning algorithms and selecting the best one for the problem. Finally, the model's performance will be evaluated using standard evaluation metrics such as mean absolute error, mean squared error, and R-squared score.

Once the machine learning model is developed and trained, the next step is to deploy and host it on AWS cloud services. The deployment involves creating an EC2 instance, installing the necessary libraries and dependencies, and deploying the machine learning model as a RESTful API. The API will receive input data, predict flight prices using the machine learning model, and return the predicted prices as output.

The API will be exposed to the internet using API Gateway, and Lambda functions will be used to handle the incoming requests and responses. The API will be secured using appropriate authentication and encryption mechanisms. Finally, the deployed machine learning model will be tested by sending input data through the API and comparing the predicted prices with the actual prices.

The successful completion of this project will result in a machine learning model that can accurately predict flight prices for a given airline and route, deployed and hosted on AWS cloud services, which can be used for real-time predictions. This project will demonstrate your skills in machine learning, data science, and cloud computing, and will be a valuable addition to your portfolio.

Implementation screenshots :-

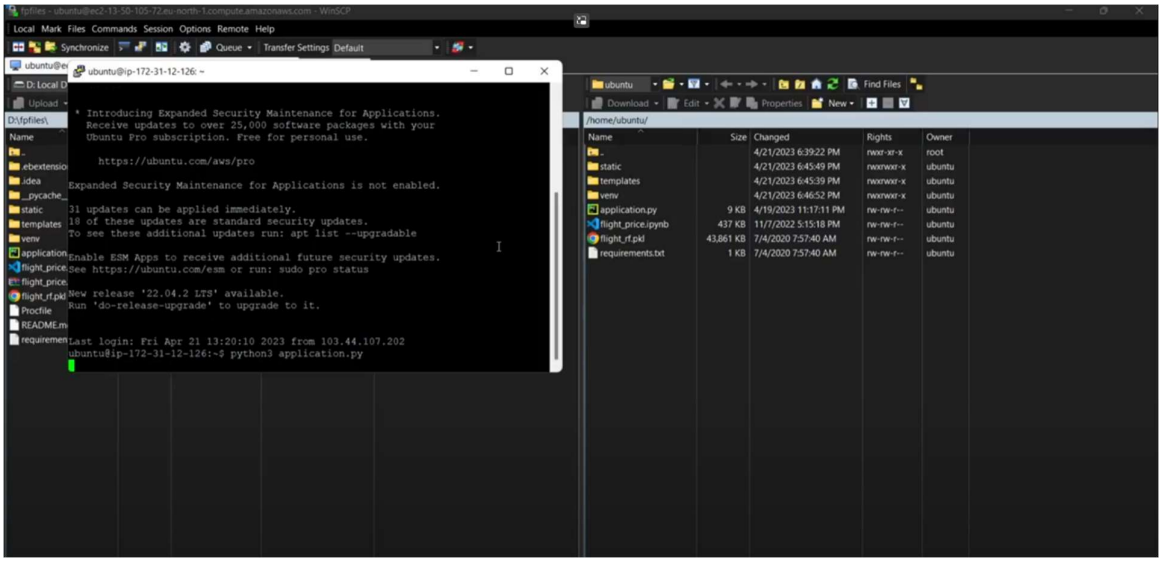

FIG 1:-LAUNCHING THE PROJECT

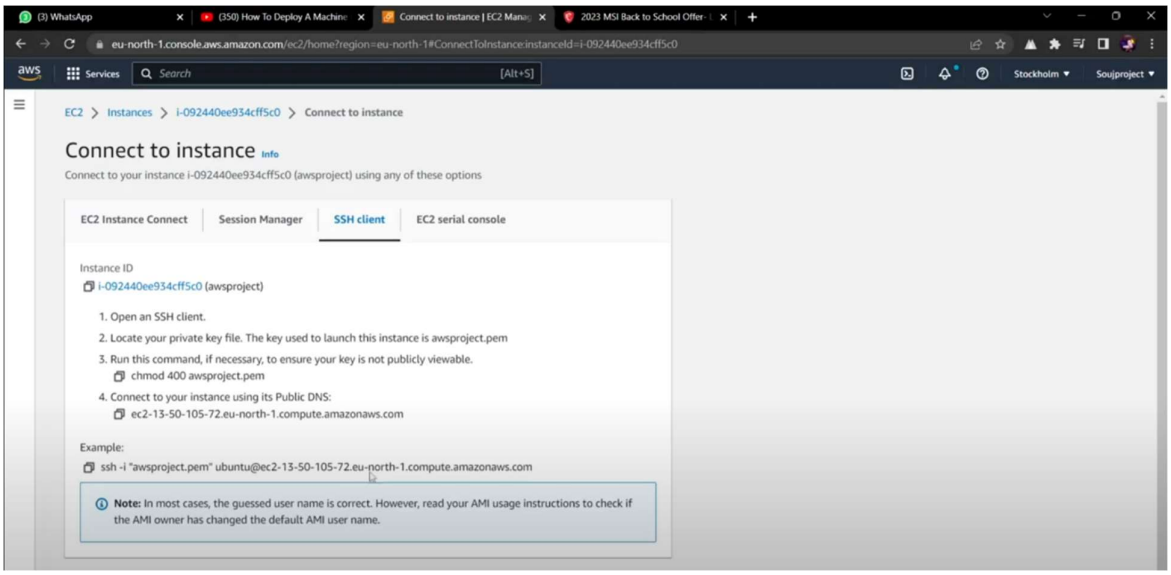

FIG 2 :- CONNECTING TO THE INSTANCE

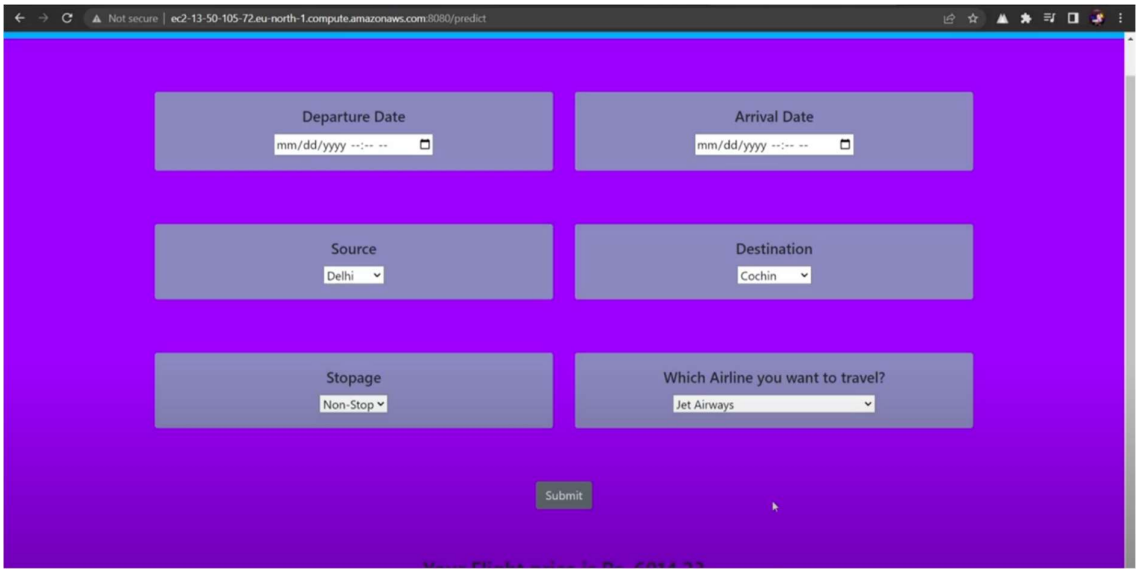

FIG 3:-PROJECT DEPLOYED

Learning Outcomes

1. Data preprocessing: You will learn how to preprocess and clean the raw data to make it suitable for machine learning.
2. Feature engineering: You will learn how to select and transform the relevant features to improve the performance of the model.
3. Model selection: You will learn how to evaluate different machine learning algorithms and select the best one for the problem.
4. Hyperparameter tuning: You will learn how to tune the hyperparameters of the selected model to achieve the best performance.
5. Evaluation metrics: You will learn how to use standard evaluation metrics such as mean absolute error, mean squared error, and R-squared score to evaluate the performance of the model.
6. Machine learning algorithms: You will gain knowledge of various machine learning algorithms such as Linear Regression, Decision Trees, Random Forest, and Gradient Boosting, and learn how to use them for regression problems.
7. Real-world application: You will learn how to apply machine learning algorithms to a real-world problem of predicting flight prices.
8. Python programming: You will gain proficiency in using Python programming language and its libraries such as Pandas, NumPy, Scikit-learn, and Matplotlib for data preprocessing, feature engineering, model selection, and evaluation.
9. AWS services: You will gain knowledge of various AWS cloud services such as EC2, S3, Lambda, and API Gateway and learn how to use them to deploy and host your machine learning model.
10. Model deployment: You will learn how to deploy your machine learning model on AWS cloud services and make it available for real-time predictions.
11. Scalability: You will learn how to make your machine learning model scalable to handle a large volume of requests and improve its performance.
12. Cost optimization: You will learn how to optimize the cost of deploying and hosting your machine learning model on AWS cloud services by selecting the appropriate instance type and managing the usage.
13. Security: You will learn how to secure your machine learning model by implementing appropriate security measures such as authentication and encryption.
14. Integration: You will learn how to integrate your machine learning model with other AWS services such as Lambda functions and API Gateway to create a complete application.
15. Real-world application: You will learn how to apply AWS cloud services to a real-world problem of predicting flight prices and make your machine learning model available for real-time predictions.