Flight Delay Prediction for Aviation Industry usingMachine Learning

TEAM ID: **NM2023TMID14898**

TEAM SIZE: 4

TEAM LEADER: **DEVAYANI M**(CB20S203192)

TEAM MEMBERS: **RAMYA P**(CB20S203210)
PARAMESHWARI A(CB20S203208)
KAVIYA K (CB20S203198)

FLIGHT DELAY PREDICTION FOR AVIATION INDUSTRY

1. INTRODUCTION:

In the present world, the major components of any transportation system include passenger airline, cargo airline, and air traffic control system. With the passage of time, nations around the world have tried

to evolve numerous techniques of improving the airline transportation system. This has brought drastic change in the airline operations. Flight delays occasionally cause inconvenience to the modern passengers. Every year approximately 20% of airline flights are cancelled or delayed, costing passengers more than 20 billion dollars in money and their time.

Overview:

Nowadays, the aviation industry plays a crucial role in the world's transportation sector, and a lot of businesses rely on various airlines to connect them with other parts of the world. But, extreme weather conditions may directly affect the airline services by means of flight delays. To solve this issue, accurately predicting these flight delays allows passengers to be well prepared for the deterrent caused to their journey and enables airlines to respond to the potential causes of the flight delays in advance to diminish the negative impact. VI The purpose of this project is to look at the approaches used to build models for predicting flight delays that occur due to bad weather conditions. In the first part of the project, we look at using Python based Logistic Regression along with Support Vector Machine and then plugging the dataset into our classifier for results. In the second part of the project, we primarily focus on gathering a dataset from Twitter, breaking the

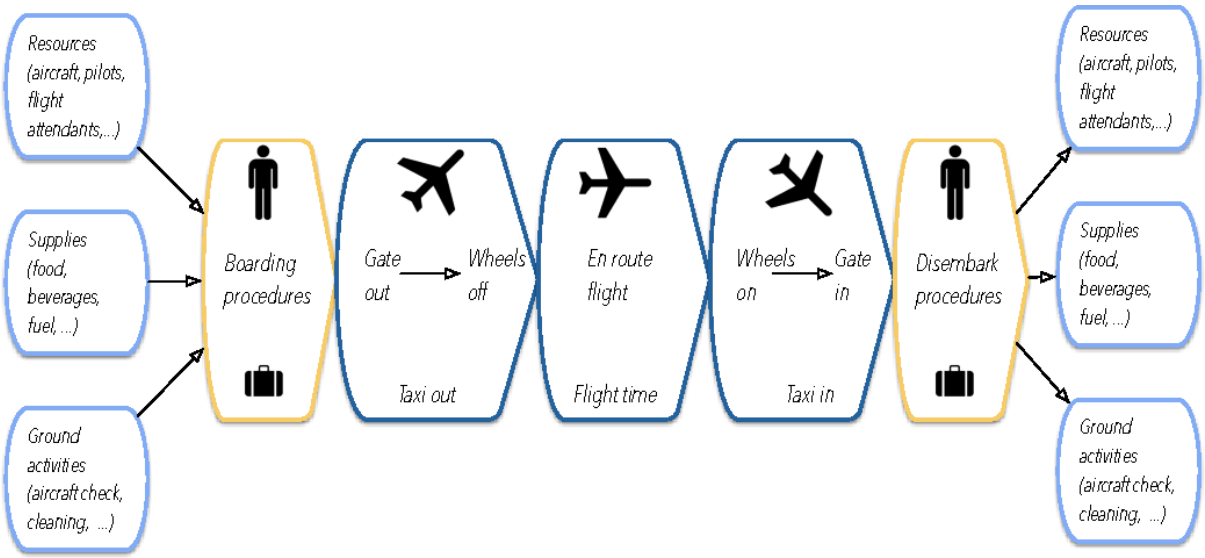
dataset down and identifying relevant attributes. Upon examining the results, we compare the results with other models such as Random Forest Classifier and derive the best classifier to solve the problem.

Purpose:

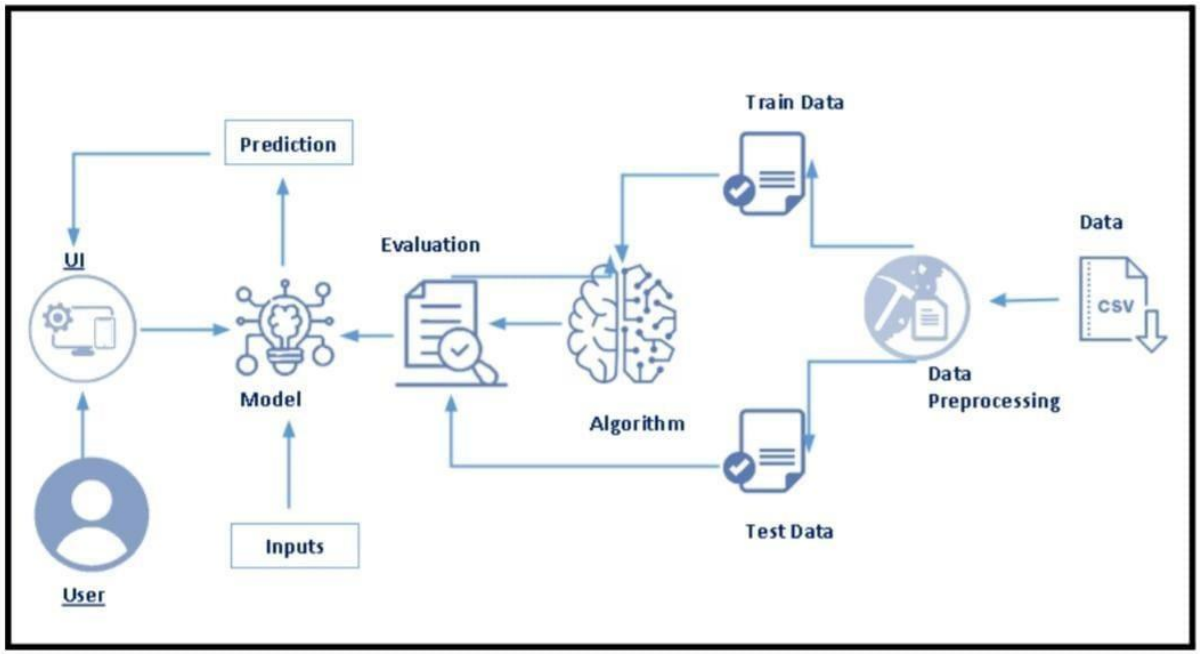
Using a machine learning model, we can predict flight arrival delays. The input to our algorithm is rows of feature vector like departure date, departure delay, distance between the two airports, scheduled arrival time etc. We then use decision tree classifier to predict if the flight arrival will be delayed or not. A flight is delayed when difference between scheduled and actual arrival times is greater than 15 minutes. Furthermore, we compare decision tree classifier with logistic regression and a simple neural network for various figures of merit. Finally, it will be integrated to web based application

2. PROBLEM DEFINITION & DESIGN THINKING:

EMPATHY MAP:



IDEATION AND BRAINSTORMING MAP:


Train Data
Prediction
Data
UI
Evaluation
CSV
Model
Algorithm
Data
Preprocessing
Test Data
Inputs
User

3. RESULT:


Prediction of Flight Delay
Enter the Flight Number : 1399
Month : 2
Day of Month : 4
Day of Week : 5
origin JFK
destination SEA
Scheduled Departure Time : 1
Scheduled Arrival Time : 22
Actual Departure Time : 1
SUBMIT

4. ADVANTAGES**:**

A. From a social perspective:

1. Flight delay prediction can help improve the travel experience for passengers.

2. By providing accurate and timely predictions of flight delays, passengers can make more informed decisions about their travel plans and potentially avoid delays or missed connections.

3. This can lead to a reduction in travel-related stress and inconvenience.

B. From a business perspective:

1. Flight delay prediction can help airlines and airports improve their operations and reduce costs.

2. By identifying and addressing the factors that contribute to flight delays, airlines and airports can take proactive measures to mitigate the impact of delays. This can lead to improved on-time performance, which can help airlines and airports attract and retain customers and increase revenue.

3. Additionally, flight delay prediction can help airlines and airports optimize their staffing and resource allocation, resulting in cost savings.

4. Flight delay prediction can also contribute to safety by allowing airlines to proactively address any safety concerns that may arise due to weather, Mechanical issues, or other factors.

4. DISADVANTAGES:

1. Flight delays irritate air passengers.

2. Flight delays disrupt their schedules but also cause a decrease in efficiency, an increase in capital costs, reallocation of flight crews and aircraft, and additional crew expense.

3. The impact of flight delay can be a risk and this risk represents financial losses.

4. The dissatisfaction of passengers, time losses, loss of reputation and bad business relations.

5. Privacy concerns: There may be concerns about the collection and use of personal data in implementing a flight delay detection system.

5. APPLICATION:

I develop a model aimed at predicting flight delays at take-off. The purpose is not to obtain the best possible prediction but rather to emphasize on the various steps needed to build such a model. Along this path, I then put in evidence some basic but important concepts. Among then, I comment on the importance of the separation of the dataset during the training stage and how cross-validation helps in determine accurate model parameters. I show how to build linear and polynomial models for univariate or multivariate regressions and also, I give some insight on the reason
why regularisation helps us in developing models that generalize well.

From a ***technical point of view***, the main aspects of python covered throughout the notebook are:

- **visualization**: matplolib , seaborne , baseman
- **data manipulation**: pandas, jumpy
- **modelling**: sclera, spicy
- **class definition**: regression, figures

During the EDA, I intended to create good quality figures from which the information would be easily accessible at a first glance. An important aspect of the data scientist job consists in divulgating its findings to people who do not necessarily have knowledge in the technical aspects data scientists' master. Graphics are surely the most powerful tool to achieve that goal, and mastering visualization techniques thus seems important.

Also, as soon as an action is repeated (mostly at identical) a few times, I tend to write classes or functions and eventually embed them in loops. Doing so is sometimes longer than a simple *copy-paste-edit* process but, on the one hand, this improves the readability of the code and most importantly, this reduces the number of lines of code (and so, the number of opportunities to introduce mistakes!!). In the current notebook, I defined classes in the modelling part in order to perform regressions. I also defined a class to wrap the making of figures. This allows creating stylish figures, by tuning the matplotlib parameters that can be subsequently re-used thanks to that template. I feel that this could be useful to create nice looking graphics and then use them extensively once you are satisfied with the tuning. Moreover, this helps to keep some homogeneity in your plots.

6. CONCLUSION:

In this project, we were able to successfully apply machine learning algorithms to predict flight arrival-delay and show simple classifiers like decision tree and logistic regression can predict if a flight's arrival will be delayed or not fairly accurately. For further work we like to further improve our models, perhaps with more training-data or deeper neural network, or both. Taxi-delay prediction is a natural progression to this work, considering

amount of fuel wasted while taxiing. Accurate taxi-delay prediction requires taking airport runway and taxiway configurations in to consideration where very little work exists.

7. FUTURE SCOPE:

The future of flight delay prediction is promising, thanks to the advancements in technology and data analytics. Here are some potential developments in the field:

1. Increased accuracy: With the help of machine learning and artificial intelligence, flight delay prediction algorithms can become more accurate over time. The algorithms can learn from historical flight data and real-time data to identify patterns and make predictions with greater precision.

2. Real-time updates: Flight delay prediction systems can provide real-time updates to passengers and airlines, allowing them to make informed decisions about travel plans. This can help minimize disruptions and improve customer satisfaction.

3. Integration with other systems: Flight delay prediction systems can be integrated with other systems, such as airline reservation systems and airport operations systems, to provide a more comprehensive view of flight status and improve overall operational efficiency.

4. Use of Iota: The Internet of Things (Iota) can be used to gather data from various sources, such as weather sensors, aircraft sensors, and passenger devices, to improve flight delay prediction accuracy.

5. Predictive maintenance: Predictive maintenance systems can be used to identify potential mechanical issues that can cause flight delays. By addressing these issues proactively, airlines can prevent delays and improve operational efficiency.

Overall, the future of flight delay prediction is bright, with the potential to improve customer experience, reduce costs, and increase operational efficiency for airlines. As technology continues

to evolve, we can expect to see even more advanced systems and capabilities in the field of flight delay prediction.

8. Appendix