Flight delay Prediction and Visualization

[1]Darshan Parab, [2]Vinod Yadav, [3]Akash Patil, [4]Prof.Reena chaudhari

[1]B.E. Student, [2]B.E. Student, [3]B.E. Student, [4]Professor

Department of Computer Engineering

Shivajirao S. Jondhale College of Engineering

Dombivli East, Thane, Maharashtra 421204

Abstract: *Flight delay has been one of the serious issues in the airline company. . A study by Frankfurt based consulting company 'Aviation Experts', presented that costs of $25 billion were incurred in 2014 due to flight delays* worldwide *Flight delays not only have economic impact but also harmful environments effects. In this project we are predicting the flight delay customer can choose flight with least delay.*

Keyword*– Machine learning, Regression, Data manipulation, Data visualization*

I. INTRODUCTION

Flight delay may be a serious and widespread drawback within the u. s. Flight traveling has become more in style among travelers over the past twenty years.Mainly due to its speed and luxury in sure things. This has junction rectifier to important growth in traffic and also the ground. The rise in traffic growth has additionally junction rectifier to serious delays in craft and also the ground. These delays have caused Brobdingnagian economic and environmental losses. In line with reports, the most reason for the 2007 taxi operation within the u. s. was four, 000 loads of hydrocarbons, 8,000 loads of gas oxides and forty five, 000 loads of CO emissions. Additionally, the economic impact of flight delays still exists.

It is calculable that USA domestic flights offer airlines with over $19 billion in votes annually, and annual national fares exceed $4.1 billion. In response to growing issues concerning fuel emissions and their negative health effects, the aviation trade is actively finding out the way to accurately predict flight delays to optimize flight operations and minimize delays. The input to our algorithmic rule is that the row of the feature vector, like the departure date, the departure delay, the space between the 2 airports, the scheduled point in time, and so on. Then we tend to use the choice tree classifier to predict if the flight arrival are delayed. within the u. s.,

The USA Federal Aviation Administration believes that once the distinction between the scheduled point in time and also the actual point in time exceeds quarter-hour, the flight are delayed.

Delay is outlined in many alternative ways that, relying upon the context. scheduled departure and arrival delay is however late a flight departs or arrives compared to an airline's schedule. Flights will incur delays whereas mobile or on the bottom, for instance as craft taxi between the runway AND gate. Delay propagates throughout the NAS attributable to the interdependencies between totally different scheduled flights. for instance, a late arrival of 1 flight might cause a late departure of following flight on the itinerary of the aircraft.

II. LITERATURE REVIEW

1 Karthik Gopalkrishnan, hamsa Balakrishnan "A comparative Analysis delays in Air traffic network" It compared the performance of several algorithms for delay prediction (ANN, MJLS, CART, LR). Temporal (timeof-day, day-of-week, season), local (airport delay, OD-pair delay) and network (type-of-hour, type-of-day) factors were used to make these predictions.

2.Masoud Yaghini*, Mohammad M. Khoshraftar and Masoud Seyedabadi" Railway passenger train delay prediction via neural network model" In this paper, for being able to predict passenger train delays in Iranian Railways, a neural network model with high accuracy is presented. To predict passenger train delays, the registered data of Iranian Railways from year 2005 to the end of year 2009 were used. To evaluate the quality of the results, we took advantage of decision tree and multinomial logistic regression models.

3.Navaneeth Jamadagni, Nathalie Kuhn," Application of Machine Learning Algorithms to Predict Flight Arrival Delays" There are several work in the literature that focus on air-traffic management and optimization a queuing model that computes delays at individual airports, and a delay propagation algorithm. In response to the local delays calculated by the queuing model, the delay propagation algorithm continuously updates flight schedules and demand-rates at all airports in the network. Such technique is unique in the area and more research using such techniques could be very beneficial to the aviation industry in terms of practical applications.

III. PROPOSED SYSTEM

Growth in air-line Company has triggered in flight traffic. Growing population peoples for traveling uses air transportation, because of its speed, time and comfort. In this system we used machine learning algorithm like linear regression and polynomial regression to predict if flight arrival will be delayed or not.

Fig 1- Architecture diagram of the system

.

The input to our system is parameters such as YEAR, MONTH, DAY,DAY_OF_WEEK, AIRLINE, FLIGHT_NUMBER, TAIL_NUMBER,ORIGIN_AIRPORT, DESTINATION_AIRPORT, SCHEDULED_DEPARTURE,DEPARTURE_TIME, DEPARTURE_DELAY, TAXI_OUT,WHEELS_OFF, SCHEDULED_TIME, ELAPSED_TIME,AIR_TIME, DISTANCE, WHEELS_ON, TAXI_IN, SCHEDULED_ARRIVAL, ARRIVAL_TIME, ARRIVAL_DELAY, DIVERTED, CANCELLED, CANCELLATION_REASON, AIR_SYSTEM_DELAY, SECURITY_DELAY,AIRLINE_DELAY, LATE_AIRCRAFT_DELAY, WEATHER_DELAY

STEPS:-

1. Cleaning

- 1.1 Dates and times
- 1.2 Filling factor

2. Comparing airlines

- 2.1 Basic statistical description of airlines
- 2.2 Delays distribution: establishing the ranking of airlines

3. Delays: take-off or landing?

4. Relation between the origin airport and delays

- 4.1 Geographical area covered by airlines
- 4.2 How the origin airport impact delays 4.3 Flights with usual delays?

5. Predicting flight delays

1. Cleaning

In the initial data frame, dates are coded according to 4 variables: **YEAR, MONTH, DAY**, and **DAY_OF_WEEK**. In fact, python offers the **date time** format which is really convenient to work with dates and times clean the data frame BY throwing the variables that won't be used and re-organized to read the columns easily.

2. Comparing airlines

The **AIRLINE** variable contains the airline abbreviations. Their full names can be retrieved from the "*airlines*" file: As a first step, we consider all the flights from all carriers. Here, the aim is to classify the airlines with respect to their punctuality and for that purpose, we compute a few basic statistical parameters: we examine more in details the distribution of delays for every airlines:

3. Delays: take-off or landing?

These delays differ somewhat from the delays recorded at arrival.

4. Relation between the origin airport and delays

we will now try to define if there is a correlation between the delays registered and the airport of origin. Here, we have a quick look at the number of destination airports for each airline: In this section, we will have a look at the variations of the delays with respect to the origin airport and for every airline. The first step thus consists in determining the mean delays per airport In the previous section, it has been seen that there is variability in delays when considering the different airlines and the different airports of origin. We are now going to add a level of granularity by focusing not just on the original airports but on flights: origin →→ destination. The objective here is to see if some flights are systematically delayed or if, on the contrary, there are flights that would always be on time.

5. Predicting flight delays

- 5.1 one airline, one airport
  - 5.1.1 Pitfalls
  - 5.1.2 Polynomial degree: splitting the dataset
  - 5.1.3 Model test: prediction of end-January delays
- 6.2 one airline, all airports
  - 5.2.1 Linear regression
  - 5.2.2 Polynomial regression
  - 5.2.3 Setting the free parameters
  - 5.2.4 Model test: prediction of end-January delays
- 5.3 Accounting for destinations
  - 5.3.1 Choice of the free parameters
  - 5.3.2 Model test: prediction of end-January delays

IV. RESULT

Figure 4.1 Number of destination airports for each airline

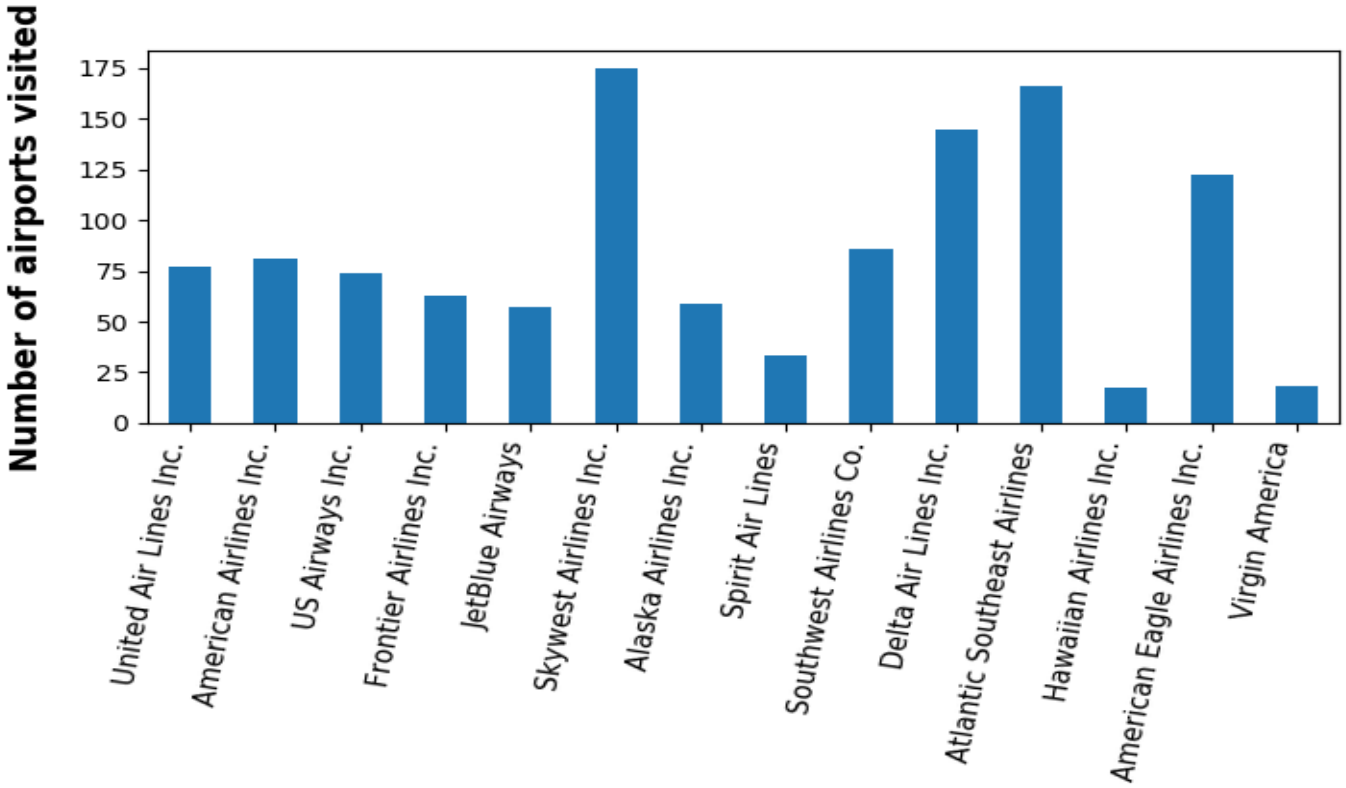
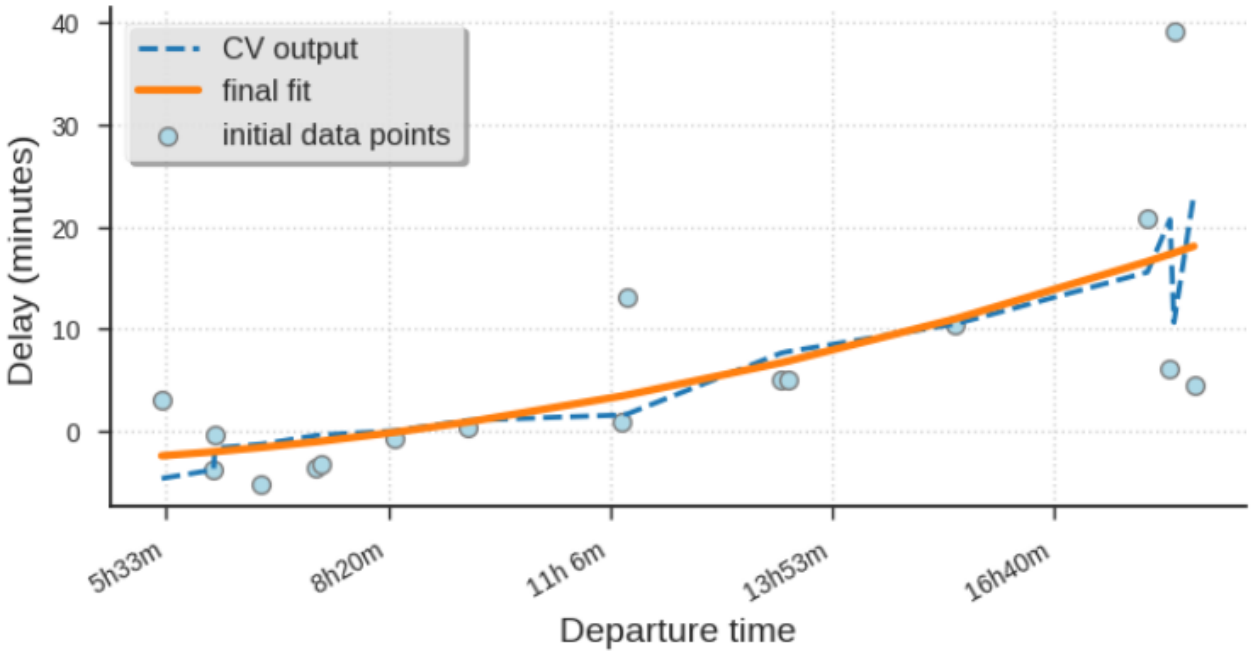


Figure 4.2 Delay of Sky west Airline



V. MAINTENANCE

Maintenance is the phase that brings more revenue to the system developers then the development phase. As time goes by or based on the level or size of the society there might occur a situation where the system has much more amount of user then it can usually handle and might lead to a crash down of the system. So for such thing to be avoided the systems will be in a need of maintainability after a certain finite amount of time. The task of maintenance in generally divided on basis of the module that might have a certain issue, but basically it is divided into hardware and software.

VI. SYSTEM REQUIREMENT

a) Software requirements:

- Windows XP, 7, 8.1 or higher
- IOS devices

b) Hardware Requirements:

- Processor: Intel i3 or higher
- Hard disc space: 5Gb minimum
- MEMORY: 1Gb RAM

VII. FEATURES

1. Information regarding flight(day, day of week , airline, flight range, tail number)
2. Data regarding origin and destination(origin airport, destination airport)
3.Data regarding the departure (scheduled departure, time of departure, departure delay, taxi-out, wheels-of
4. Information concerning the flight-journey (scheduled time, period of time, air time, distance)
5. Data concerning the arrival (wheels-on, taxi-in, regular arrival, point in time, arrival delay)
6. Data concerning diversion, cancellation and reason of delay (air system delay, security delay, airline delay, late craft delay, weather delay)

VIII. CONCLUSION

Finally we would like to conclude that while working on this paper we have learned many new technologies ,concepts and have also about working in a team.

In this project ,we tend to were able to with success apply machine learning algorithms rule to visualize flight Arrival delay,using this model customer can chose flight with least delay we show simple linear and polynomial models.for further work we will improve our models our models and work with more dataset.

IX. REFERENCE

[1] Karthik Gopalkrishnan, hamsa Balakrishnan “A comparative Analysis delays in Air traffic network”2016
[2] Masoud Yaghini*, Mohammad M. Khoshraftar and Masoud Seyedabadi” Railway passenger train delay prediction via neural network model” 2009
[3] Navaneeth Jamadagni, Nathalie Kuhn,” Application of Machine Learning Algorithms to Predict Flight Arrival Delays”, 2017
[4] Young Jin Kim, Sun Choi Briceno“A Deep Learning Approach to Flight Delay Prediction” 2016
[5] Anish M. Kalliguddi, Area K. Leboulluec “Predictive Modeling of Aircraft Flight Delay”
2017