AIRLINE ON TIME PERFORMNCE

PHASE 1 REPORT

Submitted by

R.G.JANANE (RCAS2021MDB037)

in partial fulfillment for the award of the degree of

MASTER OF SCIENCE
SPECIALIZATION IN
DATA SCIENCE AND BUSINESS ANALYSIS

DEPARTMENT OF COMPUTER SCIENCE

RATHINAM COLLEGE OF ARTS AND SCIENCE

(AUTONOMOUS)

COIMBATORE - 641021 (INDIA)

DECEMBER-2022

RATHINAM COLLEGE OF ARTS AND SCIENCE
(AUTONOMOUS)
COIMBATORE - 641021

BONAFIDE CERTIFICATE

This is to certify that the phase-1 report entitled **Airline on time performance** submitted by **R.G.Janane**, for the award of the Degree of Master in Computer Science specialization in **"DATA SCIENCE AND BUSINESS ANALYSIS"** is a bonafide record of the work carried out by him/her under my guidance and supervision at Rathinam College of Arts and Science, Coimbatore

Mr.P.Sivaprakash
Supervisor

Mr.P.Sivaprakash
Mentor

Submitted for the University Examination held on 02.12.2022

INTERNAL EXAMINER

EXTERNAL EXAMINER

RATHINAM COLLEGE OF ARTS AND SCIENCE

(AUTONOMOUS)

COIMBATORE - 641021

DECLARATION

I, **R.G.Janane**, hereby declare that this Phase-1 Report entitled **"Airline On Time Performance",** is the record of the original work done by me under the guidance of **Mr.P.Sivaprakash**, Faculty Rathinam college of arts and science, Coimbatore. To the best of my knowledge this work has not formed the basis for the award of any degree/diploma/ associateship/fellowship/or a similar award to any candidate in any University.

Signature of the Student:

R.G.Janane

Place: Coimbatore

Date: 01.12.2022

COUNTERSIGNED

Mr.P.Sivaprakash

Supervisor

Contents

Acknowledgement

On successful completion for project look back to thank who made in possible. First and foremost, thank **"THE ALMIGHTY"** for this blessing on me without which I could have not successfully our project.I am extremely grateful to **Dr.Madan.A. Sendhil, M.S., Ph.D.,** Chairman, Rathinam Group of Institutions, Coimbatore and **Dr. R.Manickam MCA., M.Phil., Ph.D.,** Secretary, Rathinam Group of Institutions, Coimbatore for giving me opportunity to study in this college.I am extremely grateful to **Dr.R.Muralidharan, M.Sc., M.Phil., M.C.A., Ph.D.,** Principal Rathinam College of Arts and Science(Autonomous), Coimbatore.Extend deep sense of valuation to **Mr.A.Uthiramoorthy, M.C.A., M.Phil., (Ph.D),** Rathinam College of Arts and Science (Autonomous) who has permitted to undergo the project.

Unequally I thank **Mr.P.Sivaprakash, M.E., (Ph.D).,** Mentor and **Dr.Mohamed Mallick, M.E., Ph.D.,** Project Coordinator, and all the Faculty members of the Department - iNurture Education Solution pvt ltd for their constructive suggestions, advice during the course of study.I convey special thanks, to the supervisor **Mr.P.Sivaprakash, M.E., (Ph.D).,** who offered their inestimable support, guidance, valuable suggestion, motivations, helps given for the completion of the project.

I dedicated sincere respect to my parents for their moral motivation in completing the project.

List of Figures

List of Abbreviations

OTP	ON Time Performance
AOTP	Airline On Time Performance
LSTM	Long Short Term Memory

Abstract

On-time performance is a significant problem for the airline industry that puts financial strain on the owners. About 20 percentage of flights are either delayed or cancelled in any given year. Both airlines and passengers are impacted by it. Airline on-time performance is the percentage of times an airline delivers on its schedule. Among these include the weather, scheduling difficulties, delays in and arrivals of passengers, and other factors. The aim of this project is to extract data about airline on-time performance using Python programming. With the help of this article, airline owners will be able to get some advantages, and passengers will be able to determine which airline is the easiest to use based on on-time performance.

Chapter 1

Introduction

The public values on-time performance (OTP) highly and frequently uses it to determine which airline to travel on. OTP affects an airline's reputation for punctuality as well as the general public's opinion and passengers' contentment. Customers place a great importance on schedule reliability. Customers will favour an airline with a proven track record of dependability, other factors being equal.

Flight delays are possible in some airports, despite the efforts of some airports and/or airlines to manage their operations to minimise potential delays. In fact, many of the factors that influence flight delays are independent. The causes of flight delays are very varied. The capabilities of the state's ATC, airport congestion, weather conditions on airports or/and routes, and the airline's own flight scheduling are all internal issues that might cause delays. The timeline could also be considerably impacted by people arriving late and by delays at immigration, security checks, and quarantine procedures.

1.1 Objective of the project

This work is the public values on-time performance (OTP) highly and frequently uses it to determine which airline to travel on. OTP affects an airline's reputation for punctuality as well as the general public's opinion and passengers' contentment.Customers value schedule reliability. Customers will favour an airline with a consistent track record of dependability, other things being equal. Many business travellers place a high value on being on time. Be aware that the OTP definition can vary depending on whether you're in the passenger or aviation sectors. The aviation industry follows what is known as "the 15-minute rule," which differs from how passengers typically perceive arrival delays, which are defined as any time beyond the scheduled time of arrival. Within 15 minutes of the scheduled arrival time, the flight is still regarded as having arrived "on-time."Regardless of how little time is involved, delays take away time that the passenger could be spending at their destination. This time has some value to the traveller. Long delays may cause the customer to miss appointments or connections, among other outcomes. Typically, the severity of the effects is inversely correlated with the duration of the delay Due to the nature of their trips, this is particularly true for business travellers. They frequently have busy schedules and cannot afford to be late and ruin their plans. Because they simply do not want to spend their time at the airport and are aware that arriving earlier will not hasten their departure, business travellers typically are familiar with the passenger handling procedures at airports and tend to arrive there 45 minutes or less before their flight takes off. They depend heavily on the

dependability of airlines because of these factors.

Neural Network:

Neural network is a sequence of algorithms that try to build the relationships between the data and the human brain. Neural networks adapt the changing input so, the network generates the best possible output without needs to redesign again.

1.2 Scope of the Project

In recent years, flight delays have cost the aviation industry millions of dollars and have become into a major issue. Thus, it's crucial to comprehend how flight delays behave. This essay focuses on how operational, temporal, and meteorological factors cause flight delays. To study departure and arrival delay, many econometric models are developed. According to the findings, arrival delays are more likely than departure delays to be impacted by earlier delays and the buffer effect. In terms of reducing flight delays, block buffer has a seven-times bigger impact than turnaround buffer. Convective weather causes more flight delays for departures than for arrivals. Convective weather has a bigger effect on flight delays at the destination airport than it does at the departure airport. Aside from that

1.3 Module Description

TensorFlow: It is an end-to-end Open-Source platform for Machine Learning. It has a different tools, libraries and community resources.

Keras: It is a deep learning API written in Python language, running on the top of

the machine learning platform i.e., Tensor flow. It is used to create layers in Neural Network.

1.4 Existing System

schedule buffer:

Schedule buffer is a calculating the data set and finding the delay timings.And identifying by using scatter plot.

Hypothesis:

Hypotheses Departure (arrival) delay is the difference between the actual departure (arrival) time and the scheduled departure (arrival) time. Based on the Figure 1, the departure and arrival delay can be established as Equations (1) and (2): Departure delay = Actual departure time Scheduled departure time, (1) Arrival delay = Actual arrival time Scheduled arrival time. (2)

Chapter 2

Literature Survey

1) Airline on-time performance management-2020 The issue of airline on-time performance is discussed in this essay, with a focus on the timetable cushion that airlines build into their itineraries. Numerous studies look at how the scheduling buffer affects expenses, daily aircraft use, and other operational issues. This essay provides information on the rationale behind schedule buffer usage by airlines. Also included are explanations of the schedule buffer's affecting elements and a process for determining the right airline schedule buffer and its analysis. In order to clearly show how the computation serves its intended function, fictitious data are used in the study. Also discussed are the advantages of reducing (or regulating) the airline scheduling cushion. The study concludes by suggesting new applications for the computation. 2) The Cause And Effect Of Commercial Flight OnTime Performance-2020 The demand for air transportation is growing daily. The airline industry developed full service, medium service, and low cost carrier airline concepts to suit demand. However, there may be significant risks associated with airline concepts, most notably the safety risk. Time is just as important to an airline's operations as safety is. Due to various factors, airlines occasionally need

to adjust their schedules. However, the airline does occasionally have to postpone the schedule, which has an impact on on-time performance (OTP). Exploratory descriptive research using the cause and effect diagram approach was utilised in this study to determine the causes and effects of OTP in Citilink Airlines. In specifically, the categories of causes employed by 5M with reference to labour, method, and mother 3) A Study on the Factors that Affect the On-Time Performances of Airline Companies-2021 When compared to other airlines, on-time performance (OTP), a measure of an airline's dependability, can be a key indicator. By comparing data from domestic K airlines for the year before and after March 2020, when the number of flights was significantly reduced owing to the corona virus, this study examined the effects and degree of delay causes on airline punctuality. Additionally, the impact of seasonal weather effects on delays was examined, as well as the impact of the current situation's dramatic decline in passenger numbers on delay causes. Regarding the delay factors, the K airline's delay codes—which the IATA added to as well—were used. Considering the outcomes. 4) Development of a predictive model for on-time arrival flight of airliner by discovering correlation between flight and weather data of year 2019 - Increasing client happiness is a key component of the airline company. Flights are delayed and cause consumer displeasure due to inclement weather, a mechanical issue, and the delayed arrival of the aircraft at the place of departure. With the aid of weather and flight data, a predictive model for flights arriving on time is put forth. The investigation of the relationship between flight data and weather data is the main research topic in this work. It is discovered that the sea-level pressures of three weather observation spots,

namely Wakkanai as the most northern spot, Minami-Torishima as the most eastern spot, and Yonagunijima as the most western spot, can classify the relationship between pressure pattern and flight data of Peach Aviation, an LCC (low-cost carrier) in Japan. 5)A performance improvement case study in aircraft maintenance and its implications for hazard identification of year 2010- The industry of aircraft maintenance is highly regulated, sophisticated, and competitive. To address process efficiency without compromising safety and quality, creative solutions must be developed. The argument made in this paper is that in order to improve a system as complex as aircraft maintenance, a complete and ecologically sound model of the operational system must be created. This model must show not only what is intended to happen, but also what actually does. This model then offers the framework for changing or bettering the system. Data about anything that 'blocks' task or check performance is gathered in a performance report built specifically for aeroplane maintenance called the Blocker Report. To remove obstacles, a blocker resolution process was created. 6) Potential On-Time Performance Improvements at the Lufthansa Station at Frankfurt Airport - 2006 - In recent years, departure punctuality has drawn more and more attention. This change resulted from the recognition of the significance of this issue to the economics of the airline sector. Being on time is more than just a good performance. sign, but also enables airlines to set themselves apart from rivals. In Additionally, there is a large chance for cost savings with this issue. Due to this, some airlines have launched special initiatives to enhance Lufthansa also performed well in terms of punctuality. Nevertheless, despite this rising Outbound punctuality levels still need improvement in terms of attention.

In order to aid in this process, this study sought to identify potential off- improvements to the Lufthansa station's timing at Frankfurt Airport.

2.1 Delays

The airline reported an all-causes delay: The key kinds of delays, and the primary cause of delays in Europe, are responses to prior delays. Such a delay is brought about by a flight's delayed arrival and affects subsequent flights or the entire network of the airline. The scheduling buffer's primary function is to handle this kind of delay. The second most common reason for delays is airline-related, which is followed by reactionary delays.

2.2 5 Ms

1. Machine (machine or technology), 2. Method (method or process), 3. Material (include raw material, consumption, and information), 4. Man Power (labor or physical work) / Mind Power (labor: suggestion, etc.) 5. Measurement/ Management (measurement or inspection) 6. Milieu / Mother Nature (environment).

2.3 Proposed System

The proposed system concentrates on analyzing airline data history to provide the important and interesting statistics related to airline on time performance. The proposed system architecture is shown in the figure.

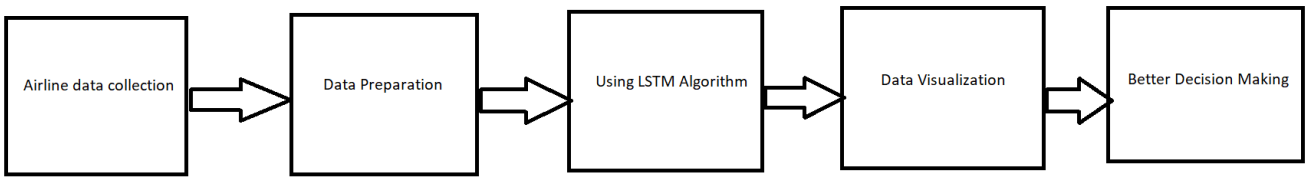

Figure 2.1: flow chart

Chapter 3

Methodology

3.1 Long Short-Term Memory - LSTM

. LSTM basically poss ed with memory blocks, simply it creates an input memory block which reduces the smaller gradient effect, and previous inputs which are fed into the input layer is controlled by forget gate and make LSTM ahead from the vanishing gradient issue, forget gate used to determine the state of input whether its remembered or forgot-ted. Totally LSTM has three gates input gate, output gate and forgot gate.

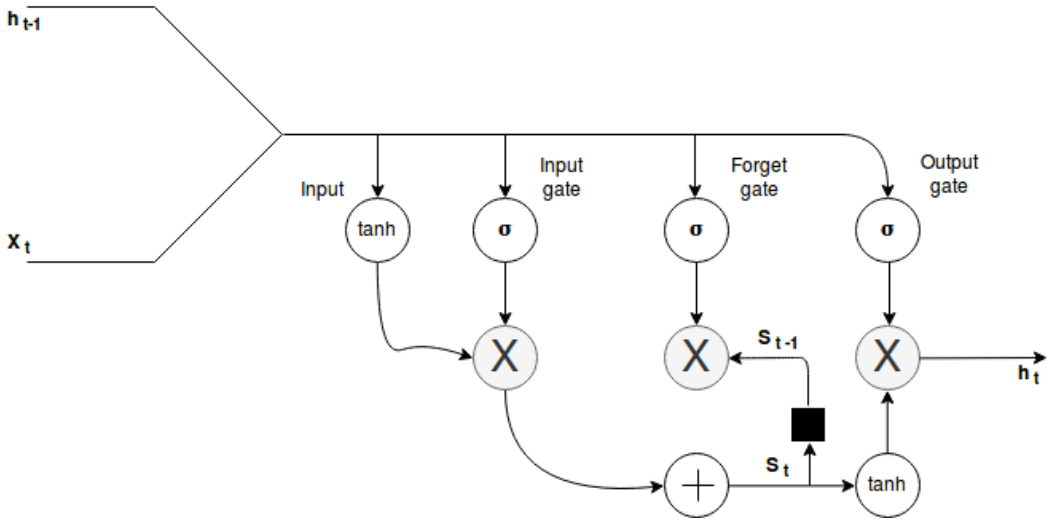

Figure 3.1: Long short term memory

The entire flow of LSTM is been shown in fig, x_t as input, h_{t-1} as a output cell. In this activation function $tanh$ is take and distinguishing values are from -1 to 1. The following theroy is expressed as

$$g = tanh(b^g + x_t R^g + h_{t-1} J^g) \tag{3.1}$$

Where weight from the previous cell input and output is denoted by R^g and J^g. Bias is denoted by b^g, g is the input weight of the neuron.

The output of the input gate undergoes element wise multiplication, with sigmoid nodes with x_t of weights and input value ofh_{t-1} . The input gate of LSTM is expressed as

$$k = \sigma(b^k + X_t U^k + h_{t-1} V^k \tag{3.2}$$

The output of the LSTM cell is expressed as where $\circ$ is multiplication operator element wise.

$$g \circ k \tag{3.3}$$

In the forget gate a new state is introduce S_t . This inner state S_t provides a internal loop with time step adding to the $g \circ k$ input state. Forget gate usually set a node for activation function which is S_{t-1}. It Remembers the information from the previous input state, this helps LSTM to learn the exact context fed into the network. The Forget gate is expressed as

$$l = \sigma(b^1 + X_t U^1 + h_{t-1} V^1) \tag{3.4}$$

S_{t-1} is output of the forget gate and each time the inputs added to the this gate, all inputs are filtered without multiplication it is mixed along with sigmod function and weights. This helps in eliminating vanishing gradient problem. Finally the output gate posses with gating function from each cell and produce the output.

3.2 LSTM via attention-based approach for flight delay

The LSTM is a significant RNN category that was first introduced in 1997 [31]. Text, video, and audio processing have all benefited from the usage of RNN, and it is gradually being applied to additional disciplines. However, the conventional RNN struggles to save previous data far from the time of decision when the input sequence data is excessively long [32]. The long-term dependence of serial data can be resolved naturally using LSTM [33]. Each LSTM unit filters the input data while incorporating prior historical data to preserve the long-term memory of sequence data. The intrinsic links between sequences can be effectively maintained using LSTM, aiding in the preservation of distant information. From the airport's perspective, it is envisaged that the ability to forcast.

3.3 Flow chart

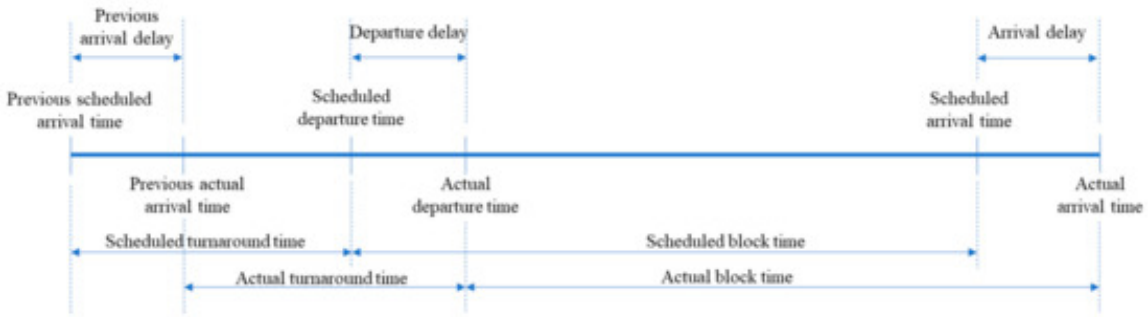

Figure 3.2: flow chart

Chapter 4

Experimental Setup

4.1 Data Analysis

Due to the constraints of the available data, a thorough analysis was done using the general present situation. After the epidemic, on-site support personnel lacked enough people to complete actual job owing to unpaid leave or leave of absence, therefore expenditures had to be cut. This is said to have delayed things. Before takeoff in snowy weather, aeroplanes must undergo de-icing procedures, which typically take 20 to 40 minutes every trip. Additionally, loading and unloading of freight are delayed when it is snowing heavily. Unloading and loading are particularly difficult and time-consuming tasks, especially when the temperature is too low and the wind is blowing hard at the same time. Workers must take precautions against slipping, which slows down job speed, and stay warm to avoid frostbite.

4.2 Data set

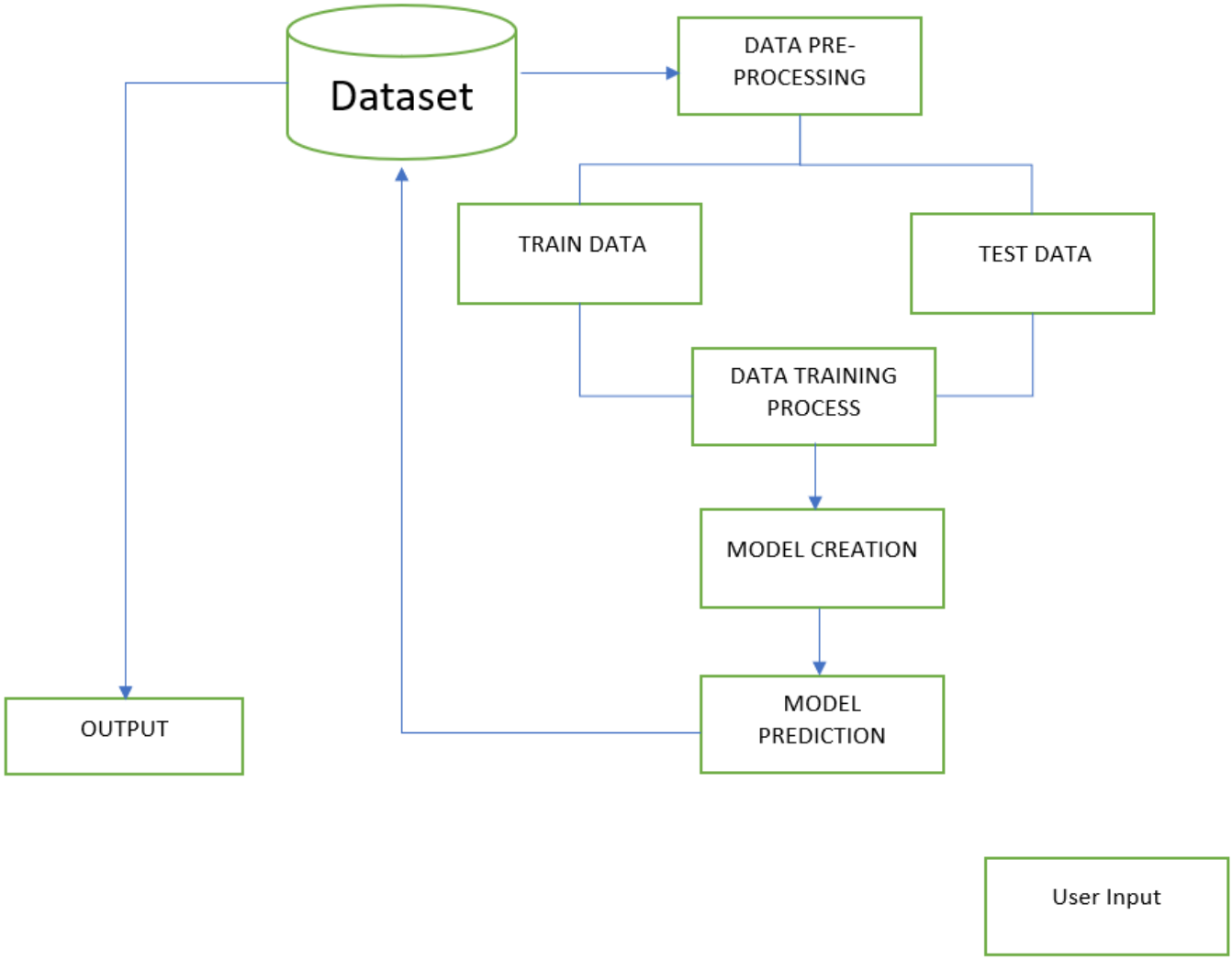

Figure 4.1: flow chart

This is the basic overview of data pre-processing model. In data classification, there will be two sets of data as Training data and Testing data. Once the model is created, if the user is asking any queries the chatbot will response according to the given input.

4.3 Data Processing

Define Neural Network architecture for proposed model and for that use the Sequential model class of Keras. Keras is an open source, high level library for developing neural network model.

The steps for creating a Keras (Sequential) model are the following:

Step 1: Define neural network model (i.e., Sequential model). The network is defined as a sequence of layers, each with its own customizable size and activation function. In this sequential model, the first layer is the input layer, which define the size of the input, which is feed to neural network. After this more and more layers (hidden layers) can be added and customized until reached to the final output layer

Step 2: Define the optimization algorithm that will be used to train the sequential model, for that use SGD optimizer and also choose the loss function. Then, compile the neural network, which transforms the simple sequence of layers into a complex group of matrix operations that describes the behavior of the network.

Step 3: Train or fit the neural network and save it in chatbot$_m odel.pfile$.

Step 4: Now, model/network is trained, an input can be used to make prediction

Chapter 5

Results and Discussions

The findings of the methodologies chosen and used on the dataset to create two predictive models are presented in this part, along with an explanation of how the results of the two predictive models can be used for Peach flights.

5.1 Results & Discussion

In this experiment, the model's performance is assessed using fresh flight and meteorological data. New flight and weather data that were gathered between January 1, 2019, and January 7, 2019, and between February 8, 2019, and February 14, 2019, are used to evaluate the experiment. The accuracy and F-score of Gradient Boosting are identical to those of Random for models using weather data. Nevertheless, it is possible to demonstrate that the Random Forest model outperforms the Gradient Boosting model in terms of accuracy, precision, recall, and F-score.

Chapter 6

Conclusion

This essay offers information on a number of aviation schedule buffer-related topics. This paper provides a definition and a description of the scheduling buffer for the first time. The majority of the information is theoretical, however a calculating method has also been established. The exact size of the buffer cannot be determined with this computation, but we can determine the delay that has the most impact on it and analyse the effects of adding or removing additional buffer time from the schedule. Although the formula is very straightforward, it can easily modified by adding variables to get more complex findings and a larger set of data to produce more accurate results. The 20 flights included in this package are only used for demonstration purposes.

6.1 Future Works

Future analyses of airline schedules and operational models may benefit from the concepts offered in this work. The necessity of using a scheduling buffer is emphasised in the paper. It is simple to modify the computation formula to account for additional scheduling buffer impacting elements, such as holding time, ATFM slots, alternate

routes, etc. By including these variables, the formula's outputs are made possible and more room is provided for data interpretation. Because the authors were unable to take into account all of these criteria in this article, several choices are left open for later analyses. A few of the potential outcomes include: - Different daytime flights - Winter season analysis

References

1. ACI International, 2020. Slot Sanctions [online].Berdy, P., Gershkoff, I., 2000. Improving On-Time Performance and Operational Dependability. In Handbook of Airline Operation. New York: Aviation week, 2000. ISBN 007 -982386-6. pp. 310

2. Coda Eurocontrol, 2019. All-causes delay and cancellations to air transport in Europe SUPPORTING EUROPEAN AVIATION Q2 2019.

3. Danesi, A. 2006. Measuring airline hub timetable co-ordination and connectivity: definition of a new index and application to a sample of European hubs. In European Transport Trasporti Europei [online]. 2006, issue 34.

   , K., Balint, V., Novák, A., 2017. A number of conflicts at route intersections - Rectangular model. Communications - Scientific Letters of the University of Zilina 19(2), 145-147.

4. Kazda, A., Badánik, B., Tomová, A., Laplace, I., Lenoir, N., 2013. Future airports development strategies. In: Communications : scientific letters of the University of Žilina. - ISSN 1335-4205. - Vol. 15, no. 2, 19-24.

5. Kohl, N. et al. 2007. Airline disruption management – Perspectives, experiences and outlook. In Journal of Air Transport Management [online]. 2007, vol. 13, issue 3.

6. Lazar, T., Novák Sedláčková, A., Bréda, R., 2015. Regression in personal air transport of passengers evolution at selected airport time series method. Naše More 62, 228-232. doi:10.17818/NM/2015/SI26.

, M., Novak, A., 2008. Models of processes as components of air passenger flow model. Komunikacie 10(2), 50-54

7. Novák, A., Havel, K., Adamko, P., 2019. Number of conflicts at the route intersection – Minimum distance model. Aviation 23(1), 1-6

8. OAG, 2020. On-time performance is becoming increasingly important to an airlines and airports

9. Qatar Airways, 2011. Available online at: https://sky4all.wordpress.com/2011/10/30/qatar-airways-ultimate-operational-analysis

10. Tomová, A., Materna, M., 2017. The Directions of On-going Air Carriers' Hybridization: Towards Peerless Business Models? Procedia Engineering 192, 569-573.

11. Tomová, A., Materna, M., 2018. Miscellaneous "modi vivendi" of regional and network airlines: The tracks for future research. Transportation Research Procedia 35, 305-314

12. Wu, C.- L., Caves, R. E., 2000. Aircraft operational costs and turnaround efficiency at airports. In Journal of Air Transport Management [online]. 2000, vol. 6, issue 4.