LITERATURE REVIEW

1)PAPER TITLE: Airport delay prediction using machine learning regression models as a tool for decision making process.

AUTHOR:

i. Edvin Simic Researcher and teaching assistant at University of Sarajevo.
ii. Muhamed Begovic Doctor of Engineering Faculty of Traffic and Communications.

ABSTRACT:

The expansion of air traffic capacity is very limited, primarily due to the sudden increase in the number of air operations as well as the lack of infrastructural capacity of the airport. The increase in the number of operations is not adequately accompanied by expanding the system capacity that would accept this increase without significant impacts on performance and qua lity of traffic service. Delay is one of the most important airport Key Performance Indicators (KPI) and it is necessary to establish a correlation with other traffic and surrounding airspace characteristics. Traffic management organizations at the European level collect a large amount of data related to the airport airside performance. For the analysis of such large databases, it is no longer practical to apply classical statistic methods. Therefore, the scientific community and industry are increasingly relying on the use of machine learning methods. The airport infrastructure in today's conditions is very difficult to expand, so to increase capacity and system efficiency existing operations should be optimized. To analyze the correlations of individual KPIs and their impact on output variables, several regression methods of machine learning were applied. This enabled additional validation of the proposed model and selection of the optimal approach for the decision-making process. It was observed how individual characteristics of the airport and related factors affect the overall delays. The correlations and predictions should enable decision-makers to find optimal solutions that would improve the system's capacity and efficiency.

DRAWBACKS:

i. If the number of observations is lesser than the number of features, Logistic Regression cannot not be used, otherwise, it may lead to overfitting.

2)PAPER TITLE: Flight delay prediction based on deep learning and Levenberg-Marquart algorithm.

AUTHOR:

i. Maryam Farshchian, Yazdi Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.
ii. Seyed Reza Kamel, Department of Computer Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
iii. Seyyed Javad Mahdavi Chabok, Department of Computer Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
iv. Maryam Kheirabadi, Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.

ABSTRACT:

Flight delay is inevitable and it plays an important role in both profts and loss of the airlines. An accurate estimation of fight delay is critical for airlines because the results can be applied to increase customer satisfaction and incomes of airline agencies. There have been many researches on modeling and predicting fight delays, where most of them have been trying to predict the delay through extracting important characteristics and most related features. However, most of the proposed methods are not accurate enough because of massive volume data, dependencies and extreme number of parameters. This paper proposes a model for predicting fight delay based on Deep Learning (DL). DL is one of the newest methods employed in solving problems with high level of complexity and massive amount of data. Moreover, DL is capable to automatically extract the important features from data. Furthermore, due to the fact that most of fight delay data are noisy, a technique based on stack denoising autoencoder is designed and added to the proposed model. Also, Levenberg-Marquart algorithm is applied to fnd weight and bias proper values, and fnally the output has been optimized to produce high accurate results. In order to study efect of stack denoising autoencoder and LM algorithm on the model structure, two other structures are also designed. First structure is based on autoencoder and LM algorithm (SAE-LM), and the second structure is based on denoising autoencoder only (SDA). To investigate the three models, we apply the proposed model on U.S fight dataset that it is imbalanced dataset. In order to create balance dataset, undersampling method are used. We measured precision, accuracy, sensitivity, recall and F-measure of the three models on two cases. Accuracy of the proposed prediction model analyzed and compared to previous prediction method.

DRAWBACKS:

i. As they are using deep learning concepts,they can perform well on benchmarked datasets may struggle when it is applied to real-world datasets.

3)PAPER TITLE: Applying Machine Learning to Aviation Big Data for Flight Delay Prediction.

AUTHOR:

i. Yushan Jiang,Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL.
ii. Yongxin Liu ,Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL.
iii. Dahai Liu ,College of Aviation, Embry-Riddle Aeronautical University, Daytona Beach, FL.
iv. Houbing Song,Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL.

ABSTRACT:

Flight delay has been a serious and widespread problem that needs to be solved. One promising solution is the flight delay prediction. Although big data analytics and machine learning have been applied successfully in many domains, their applications in aviation are limited. This paper presents a comprehensive study of flight delay spanning data pre-processing, data visualization and data mining, in which we develop several machine learning models to predict flight arrival delays. Two data sets were used, namely Airline On-Time Performance (AOTP) Data and Quality Controlled Local Climatological Data (QCLCD). This paper aims to recognize useful patterns of the flight delay from aviation data and perform accurate delay prediction. The best result for flight delay prediction (five classes) using machine learning models is 89.07% (Multilayer Perceptron). A Convolution neural network model is also built which is enlightened by the idea of pattern recognition and success of neural network method, showing a slightly better result with 89.32% prediction accuracy.

DRAWBACKS:

i. sensitive information are being collected for big data analytics. Those data need protection, and security risks can be demerits due to the lack of proper maintenance

4)PAPER TITLE: A machine learning approach for prediction of on-time performance of flights.

AUTHOR:

i. Balasubramanian Thiagarajan, Sri Venkateswara College of Engineering.
ii. Lakshminarasimhan Srinivasan ,SRM University.
iii. Aditya Vikram ,Sharma College of Engineering Guindy.
iv. Dinesh Sreekanthan ,SRM University.
v. Vineeth Vijayaraghavan, Solarillion Foundation.

ABSTRACT:

One of the major business problems that airlines face is the significant costs that are associated with flights being delayed due to natural occurrences and operational shortcomings, which is an expensive affair for the airlines, creating problems in scheduling and operations for the end-users thus causing bad reputation and customer dissatisfaction. In our paper, a two-stage predictive model was developed employing supervised machine learning algorithms for the prediction of flight on-time performance. The first stage of the model performs binary classification to predict the occurrence of flight delays and the second stage does regression to predict the value of the delay in minutes. The dataset used for evaluating the model was obtained from historical data which contains flight schedules and weather data for 5 years. It was observed that, in the classification stage, Gradient Boosting Classifier performed the best and in the regression stage, Extra-Trees Regressor performed the best. The performance of the other algorithms is also extensively documented in the paper. Furthermore, a real-time Decision Support Tool was built using the model which utilizes features that are readily available before the departure of an airplane and can inform passengers and airlines about flight delays in advance, helping them reduce possible monetary losses.

DRAWBACKS:

i. Requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality.

5)PAPER TITLE: Machine Learning Model - based Prediction of Flight Delay

AUTHOR:

i. N Lakshmi Kalyani ,Department of CSE, VNR VJIET, Hyderabad, India.
ii. G. Jeshmitha ,Department of CSE, VNR VJIET, Hyderabad, India.
iii. Bindu Sri Sai U ,Department of CSE, VNR VJIET, Hyderabad, India.
iv. M. Samanvitha,Department of CSE, VNR VJIET, Hyderabad, India.
v. J. Mahesh ,Department of CSE, VNR VJIET, Hyderabad, India.
vi. B.V. Kiranmayee,Department of CSE, VNR VJIET, Hyderabad, India.

ABSTRACT:

Prior prediction of flight arrival delays is necessary for both travelers and airlines because delays in flights not only trigger huge economic loss but also airlines end up losing their reputation that was built for several years and passengers lose their valuable time. Our paper aims at predicting the arrival delay of a scheduled individual flight at the destination airport by utilizing available data. The predictive model presented in this work is to foresee airline arrival delays by employing supervised machine learning algorithms. US domestic flight data along with the weather data from July 2019 to December 2019 were acquired and are used while training the predictive model. XGBoost and linear regression algorithms were applied to develop the predictive model that aims at predicting flight delays. The performance of each algorithm was analyzed. Flight data along with the weather data was given to the model. Using this data, binary classification was carried out by the XGBoost trained model to

predict whether there would be any arrival delay or not, and then linear regression model predicts the delay time of the flight.

DRAWBACKS:

i. It is often quite prone to noise and overfitting

6)PAPER TITLE: Flight Delay Prediction Using Machine Learning Algorithm XGBoost

AUTHOR:

i. K P Surya Teja ,Assistant Professor, Dept. of IT, Sreenidhi Institute of Science and Technology, (Autonomous), Hyderabad.
ii. Vigneswara Reddy, CVR College of Engineering.
iii. Subhani Shaik ,Sreenidhi Institute of Science & Technology.

ABSTRACT:

Growth in aviation industries has resulted in air-traffic jamming causing flight delays. Flight delays not only have economic impact but also injurious environmental properties. Air-traffic supervision is becoming increasingly challenging. Airlines delays make immense loss for business field as well as in budget loss for a country, there are so many reasons for impede in flights some of them are, some of them are due to security issues, mechanical problems, due to weather conditions, Airport congestion etc. we are proposing machine learning algorithms like XGBoost regressed, Linear regression Techniques. The aim of this research work is to predict Flight Delay, Which is highest economy producing field for many countries and among many transportation this one is fastest and comfort, so to identify and reduce flight delays, can dramatically reduce the flight delays to saves huge amount of turnovers, using machine-learning algorithms.

DRAWBACKS:

i. Training of unstructured data is not be possible because XGBoost does not perform so well on sparse and unstructured data.

7)PAPER TITLE: Flight Delay Prediciton at an Airport Using Machine Learning.

AUTHOR:

i. Bo Zhang ,School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China.
ii. Dandan Ma,School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China.

ABSTRACT:

The increasing demand for air transportation makes the problem of flight delays increasingly prominent. Accurate prediction of flight delays is of great significance to passengers' travel and the operation of airlines. In this paper, we establish a flight delay prediction model to predict the departure delay of Newark Liberty International Airport based on Catboost algorithm. The result shows that the prediction accuracy of our model can reach 0.77. In addition, we also analyze the contribution of each feature in the prediction by using SHAP.

DRAWBACKS:

i. Require lengthy batch training.
ii. Do not learn incrementally or interactively in real time.

8)PAPER TITLE: The Prediction of Flight Delay: Big Data-driven Machine Learning Approach.

AUTHOR:

i. Jiage Huo,Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China.
ii. K. L. Keung,Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China.
iii. C. K. M. Lee,Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China.
iv. Kam K. H. Ng,Interdisciplinary Division of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, China.
v. K.C. Li,Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China.

ABSTRACT:

Nowadays, Hong Kong International Airport faces the issues of saturation and overload. The difficulties of selecting taxiways and reducing the lead time at the runway holding position are the severe consequences that appeared from increasing the number of passengers and increased cargo movement to Hong Kong International Airport but without constructing a new runway. This paper is primarily about predicting flight delays by using machine learning methodologies. The prediction results of several machine learning approaches are compared and analyzed thoroughly by using real data from the Hong Kong International Airport. The findings and recommendations from this paper are

valuable to the aviation and insurance industries. Better planning of the airport system can be established through predicting flight delays.

DRAWBACKS:

i. Questionable data quality.
ii. Heightened security risks.
iii. Cost and infrastructure issues.

9)PAPER TITLE: Airline Flight Delay Prediction Using Machine Learning Models.

AUTHOR:

i. Yuemin Tang, University of Southern California, USA.

ABSTRACT:

Flight delays are gradually increasing and bring more financial difficulties and customer dissatisfaction to airline companies. To resolve this situation, supervised machine learning models were implemented to predict flight delays. The data set that records information of flights departing from JFK airport during one year was used for the prediction. Seven algorithms (Logistic Regression, K-Nearest Neighbor, Gaussian Naïve Bayes, Decision Tree, Support Vector Machine, Random Forest, and Gradient Boosted Tree) were trained and tested to complete the binary classification of flight delays. The evaluation of algorithms was fulfilled by comparing the values of four measures: accuracy, precision, recall, and f1-score. These measures were weighted to adjust the imbalance of the selected data set. The comparative analysis showed that the Decision Tree algorithm has the best performance with an accuracy of 0.9777, and the KNN algorithm has the worst performance with an f1-score of 0.8039. Tree-based ensemble classifiers generally have better performance over other base classifiers.

DRAWBACKS:

i. Time-consuming and more resources required.

10)PAPER TITLE: Analyzing flight delay prediction under concept drift.

AUTHOR:

i. Lucas Giusti ,Federal Center for Technological Education of Rio de Janeiro, CEFET/RJ, Rio de Janeiro, Brazil.
ii. Leonardo Carvalho ,Petroleo Brasileiro S.A.
iii. Antonio Tadeu Azevedo Gomes Laboratorio Nacional de Computaçao Científica Rafaelli de Carvalho Coutinho Centro Federal de Educacao Tecnológica Celso Suckow da Fonseca (CEFET/RJ).
iv. Jorge De Abreu Soares Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ).
v. Eduardo Ogasawara ,Centro Federal de Educaçao Tec.

ABSTRACT:

Flight delays impose challenges that impact any flight transportation system-predicting when they will occur in a meaningful way to mitigate this issue. However, the distribution of the flight delay system variables changes over time. This phenomenon is known in predictive analytics as concept drift. This paper investigates the prediction performance of different drift handling strategies in aviation under different scales (models trained from flights related to a single airport or the entire flight system). Specifically, two research questions were proposed and answered: (1) how do drift handling strategies influence the prediction performance of delays? (2) Do different scales change the results of drift handling strategies? In our analysis, drift handling strategies are relevant, and their impacts vary according to scale and machine learning models.

DRAWBACKS:

i. Absence of updates to the model after real concept drift, the model will no longer correctly describe the full target concept space.