**LITERATURE SURVEY:**

### 1. TITLE: Machine Learning Model - based Prediction of Flight Delay

**AUTHOR:** N Lakshmi Kalyani, Jeshmitha G, Bindu Sri Sai U, Samanvitha M, Mahesh J, Dr.B.V.Kiranmayee Assistant Professor, Student, Student, Student, Student, Associate Professor

**YEAR: 2020**

**PAPER EXPLANATION:**

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.

### 2. TITLE: Applying Machine Learning to Aviation Big Data for Flight Delay Prediction

**AUTHOR:** Yushan Jiang, Yongxin Liu

**YEAR: 2019**

**PAPER EXPLANATION:**

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 OnTime 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.

**3. TITLE:** A Machine Learning Approach for Prediction of On-time Performance of Flights

**AUTHOR:** Balasubramanian Thiagarajanα, Lakshminarasimhan Srinivasanβ, Aditya Vikram Sharmaψ, Dinesh Sreekanthanγ, Vineeth Vijayaraghavanδ

**YEAR: 2019**

**PAPER EXPLANATION:**

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 ontime 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. Keywords—Flight Delay Prediction, Classification, Regression, Machine Learning, Deep Learning, Decision Support Tool

**4. TITLE: A Novel Integration Platform to Reduce Flight Delays in the National Airspace System**

**AUTHOR**: Chuyang Yang , Zachary A. Marshall

**YEAR**: 2019

**PAPER EXPLANATION:**

Flight delays in the U. S. National Airspace System (NAS) present a fundamental challenge to capacity growth under ever-increasing traffic volumes, and lead to significant financial burdens that reverberate across a multitude of aviation industry stakeholders. Roughly 20% of passengers’ total travel time is due to such delays, causing $35 billion annually in lost revenue and impacting not only the airline industry, but the retail, lodging, restaurant, and tourism industries, as well. The Federal Aviation Administration’s effort in aiding decision-making at airports is readily apparent in the Next Generation Air Traffic Control (NextGen) System’s System-Wide Information Management (SWIM) program, and in-flight delay information from the FAA Air Traffic Control System Command Center (ATCSCC). Academic researchers are concurrently developing various algorithms to predict flight delays that include advanced statistics, machine learning, and graph theory using various network topologies. Other stakeholders have initiated delay prediction methods to adjust their operational schedules. This suggests an opportunity to centralize, validate, and integrate the various delay prediction methods under development; furthermore, these methods are limited in scope with regard to geography, operators, and efficacy. The authors propose here a platform supporting the FAA’s Collaborative Decision-Making (CDM) process with the intent of reducing flight delays in the NAS. Building upon existing deep learning algorithms and utilizing the NextGen SWIM program, this research suggests a central delay prediction platform suited to the complex and dynamic needs of America’s airport infrastructure. assessments of risks and sustainability of the proposed platform are presented. The authors interviewed experts in industry and academic fields related to aviation and information technology, and used the information obtained to refine the model. It is anticipated that this model will accurately produce location-specific departure and arrival delay forecasts that can further be integrated into the CDM and Ground Delay Program (GDP) initiatives.

**5. TITLE:** Identification, Characterization, and Prediction of Traffic Flow Patterns in Multi-Airport Systems

**AUTHOR:** Mayara Condé Rocha Murça and Robert John Hansman

**YEAR: 2020**

**PAPER EXPLANATION:**

Efficient planning of airport capacity is key for the successful accomplishment of traffic flow management. Yet, the dynamic and uncertain behavior of capacity-determining factors makes it difficult to estimate flow rates precisely, especially for strategic planning horizons. Metroplex systems impose additional challenges in this decision-making process because of relevant operational interdependencies between the closely located airports. This paper presents a data-driven framework to identify, characterize, and predict traffic flow patterns in the terminal area of multi-airport systems toward improved capacity planning decision support in complex airspace. Through the identification and characterization of patterns in the terminal area traffic flows, we learn recurrent utilization patterns of runways and airspace as well as relevant decision factors, and use that knowledge to develop descriptive models for metroplex configuration prediction and capacity estimation. The framework is based on the application of machine learning methods on historical flight tracks, weather forecasts, and airport operational data. A multi-layer clustering analysis is first performed to mine spatial and temporal trends in flight trajectory data for identification of traffic flow patterns. Based on this knowledge, a multiway classification model is developed to generate probabilistic forecasts of the metroplex traffic flow structure for look-ahead times of up to eight hours. Finally, an empirical approach for arrival capacity estimation is proposed based on historical flow pattern behavior. The observed variability in throughput and terminal area delay performance emphasizes the importance of metroplex configuration predictability toward improved flow rate planning and ultimately better traffic regulation. Index Terms— Air traffic management, machine learning, multi-airport systems, traffic flow pattern.

**6. TITLE:** Sensor Dynamics in High Dimensional Phase Spaces via Nonlinear Transformations: Application to Helicopter Loads Monitoring

**AUTHOR:** Julio J. Valdes, Catherine Cheung

**YEAR: 2020**

**PAPER EXPLANATION:**

Accurately determining component loads on a helicopter is an important goal in the helicopter structural integrity field, with repercussions on safety, component damage, maintenance schedules and other operations. Measuring dynamic component loads directly is possible, but these measurement methods are costly and are difficult to maintain. While the ultimate goal is to estimate the loads from flight state and control system parameters available in most helicopters, a necessary step is understanding the behavior of the loads under different flight conditions. This paper explores the behavior of the main rotor normal bending loads in level flight, steady turn and rolling pullout flight conditions, as well as the potential of computational intelligence methods in understanding the dynamics. Time delay methods, residual variance analysis (gamma test) using genetic algorithms, unsupervised non-linear mapping and recurrence plot and quantification analysis were used. The results from this initial work demonstrate that there are important differences in the load behavior of the main rotor blade under the different flight conditions which must be taken into account when working with machine learning methods for developing prediction models.

**7. TITLE: A multi-index prediction method for flight delay based on long short-term memory network model**

**AUTHOR: yenpung jiang**

**YEAR: 2020**

**PAPER EXPLANATION:** . While the ultimate goal is to estimate the loads from flight state and control system parameters available in most helicopters, a necessary step is understanding the behavior of the loads under different flight conditions. This paper explores the behavior of the main rotor normal bending loads in level flight, steady turn and rolling pullout flight conditions, as well as the potential of computational intelligence methods in understanding the dynamics. Time delay methods, residual variance analysis (gamma test) using genetic algorithms, unsupervised non-linear mapping and recurrence plot and quantification analysis were used. The results from this initial work demonstrate that there are important differences in the load behavior of the main rotor blade under the different flight conditions which must be taken into account when working with machine learning methods for developing prediction models.

**8. TITLE:** A Model for Predicting Flight Delay and Delay Propagation Based on Parallel Cellular Automata

**AUTHOR:** Ding Jianli, Yu Yuecheng1

**YEAR: 2021**

**PAPER EXPLANATION:**

The extended cellular automata (ECA) is one of the extending application of cellular automata (CA) by extending the component of cell and the definition of cellular neighbors. Based on ECA, the definition of parallel cellular automata (PCA) composed of multi-ECA is proposed and the evolutions of all ECAs are parallel. Depending on the model of PCA, it becomes more easily to describe the synchronous evolution and the state propagation of a complicated system, which is composed of multi-subsystem. Flight delay and delay propagation among multi-aerodrome is one of the problems encountered by most airlines. The simulation based on the model of PCA shows that the model of PCA is a new approach to describe the prediction of flight delay and delay propagation.

**9. TITLE: An analysis method for flight delays based on Bayesian network**

**AUTHOR:** Li Qianya, Wang Lei, Fei Rong, Wang Bin, Hei Xinhong

**YEAR: 2021**

**PAPER EXPLANATION:**

With the rapid growth of air transportation, the number of flight delays is rising rapidly in China. In 2000, there are totally 14, 2700 delay flights, and it increases to 260000 in 2008. About 1/5 of the all visitors were suffering delays each year. According to the 190 million passengers in 2008, nearly 40 million persons delay. It leads to the economic losses arising in billions [1] to the civil aviation. So the flight delay obviously restricts the development of the civil aviation industry in China. Factors affecting the flight delays before taking-off include weather, air traffic control, passengers, airport and airline. Most of them are unpredictable. Whether the flight can be in normal navigation and landing, is similarly decided by many factors, including the weather during flying, air traffic control, idle degree of the runway, the stop idle and the flight takes off normally. However, parts of flight delays in flying are predictable.

**10. TITLE:** **Cloud based flight delay prediction using logistic regression**

**AUTHOR:** Rahul Nigam

**YEAR: 2022**

**PAPER EXPLANATION:**

In the modern world, airlines play a vital role for transporting people and goods on time. Any delay in the timings of these flights can adversely affect the work and business of thousands of people at any given moment. Forecasting these delays is very important during the planning process in commercial airlines. Several techniques have already been proposed for designing models to forecast the delay in departure time of aircraft. But because of the continuously increasing complexity of the airplane transportation and the amount of data related to it, designing accurate prediction methods has become very difficult. In this paper we utilise the method of logistic regression which is a supervised learning method to predict delay in departure times of aircraft. We utilise the Microsoft Azure Learning Studio platform which is an Integrated Development Environment for utilizing machine learning for training and testing the model on the cloud. We also join weather data such as temperature, humidity, precipitation, dew point along with the airport data to derive more accurate predictions as well as find out the effect of weather changes in flight delays. Our method was able to achieve about 80 percent accuracy in predicting whether a given aircraft would be delayed or not based on the training using past data.