**LITERATURE SURVEY**

**[1]. (2020) Flight Delay Prediction System. International Journal Of Engineering Research & Technology (IJERT) Volume 09, Issue 03 (March 2020).**

**Authors: Borse, Y., Jain, D., Sharma, S., Vora, V., and Zaveri, A**

Flight Planning is one of the challenges in industrial world which faces many uncertain conditions. One such condition is delay occurrence, which stems from various factors and imposes considerable costs on airlines, operators, and travelers. Delays in departure can occur due to bad weather conditions, seasonal and holiday demands, airline policies, technical issue such as problems in airport facilities, luggage handling and mechanical apparatus, and accumulation of delays from preceding flights. Here in flight delay prediction system based on the weather parameters which can result in delays. The system considers the temperature, humidity, rain in mm, visibility and month number as important parameters for prediction of delay.

**[2] Data mining applications in civil aviation sector: State-of-art review. In CEUR Workshop Proc (Vol. 1852, pp. 18-25).**

**Authors: Akpinar, M.T. and Karabacak, M.E. (2017).**

Artificial Intelligence (AI) and its related disciplines (Machine Learning, Data Mining, Big Data…) offer opportunities whose practical implementation pose complex challenges. Their fast evolution evidences potential but has caused a gap between academia and certain areas of the industry – which seem to lack the required agility to implement such technologies. This study aims to suggest some recommendations and a roadmap aligning both communities through a comprehensive quantitative meta-analysis and visualization of the existing literature. Although four modes of transport are initially compared, the focus is placed on AI within air transport (273 works since 1987) and its relationship with organizational areas. Results show that the most popular topics are Machine Learning and neural networks. Nevertheless, as many documents only mention one AI-related term, visibility is hindered in specific-keyword searches. Operations seem to be thoroughly explored while there is room for research in Strategy and Resourcing.

**[3] Mining Aviation Data to Understand Impacts of Severe Weather. In Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC.02) pp. 518-523**.

**Authors: Nazeri, Z. and Zhang, J. (2017).**

This paper describes our latest experiment with application of data mining to analyzing severe weather impacts on National Airspace System (NAS) performance. We show the importance of data preparation and feature extraction in our work. Two types of data - weather and air traffic data - were used in this experiment. Weather data are represented as binary images. A severe-weather day for air traffic is represented as a set of severe-weather regions, each with a set of weather- and traffic-related features. The set of severe-weather regions for each day was first converted into a vector of attribute values, and then classification, regression and clustering were applied to the data. Initial results were encouraging, while later results were improved and impressive. Meaningful classification rules were generated and the clusters generated for weather-traffic days were clearly correlated with NAS performance.

**[4]. Predicting Ground Delay Program at an airport based on meteorological conditions. In 14th AIAA Aviation Technology, Integration, and Operations Conference (pp. 2713-2718).**

**Authors: Mukherjee, A., Grabbe, S. R., and Sridhar, B. (2014).**

In this paper, we present two supervised-learning models, logistic regression and decision tree, to predict occurrence of a ground delay program at an airport based on meteorological conditions and scheduled traffic demand. Predicting the occurrence of ground delay programs can help the Federal Aviation Administration traffic managers and airline dispatchers prepare mitigation strategies to reduce impact of adverse weather. The models are developed for two major U.S. airports: Newark Liberty and San Francisco International airports. The logistic regression model estimates the probability that a ground delay program will occur during a given hour. The decision tree model, on the other hand, classifies whether or not a ground delay program is likely during an hour based on the input variables. Results indicate both models perform significantly better than a purely random prediction of ground delay program occurrence at the two airports. The degree to which various input variables impact the probability of ground delay program vary between the two airports. While the enroute convective weather is a dominant factor causing ground delay programs at Newark Liberty Intl. airport, poor visibility and low cloud ceiling caused by marine stratus are major drivers of ground delay program occurrence at San Francisco Intl. airport.

**[5]. "A Novel Approach: Airline Delay Prediction Using Machine Learning," 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA pp. 1081-1086, doi: 10.1109/CSCI46756.2018.00210**

**Authors: Natarajan, V., Meenakshisundaram, S., Balasubramanian, G. and Sinha, S. (2018)**

Flight delays often create exasperation in airports when not properly mobilized. Increasing development of machine learning models provoke the researchers and scientists to harness the modern research problems. Due to an increase in customer satisfaction in the air transportation system, there needs a proper decision-making process to mobilize the air-traffic with minimum delay. It is recorded that 19% of United States domestic flights reach their destination with an average delay of 15 minutes. Moreover, the complexity of the air transportation system limits the availability of accurate prediction models. Due to the stochastic nature of delays, this research investigates the qualitative prediction of airline delays to implement necessary changes and provide better customer experience. Collection of historical weather data and operational data during departure and arrival at airports serves as the source for building prediction models. A logistic regression model is used to get the status of the delay which is further contrasted to a decision tree model for evaluating the performance of the delay. The proposed research empirically evaluates the effectiveness of the decision tree algorithm over logistic regression. The results of this simulation indicate the potential delays in major airports including the time, day, weather, etc., and hence the volume of delay shall be minimum based on the constructed model.