# FLIGHT DEPARTURE DELAY PREDICTION USING MACHINE LEARNING TECHNIQUES WHILE IDENTIFYING THE IMPACT FROM COVID19

**Tharanga Lakmali**

A dissertation submitted in partial fulfilment of the requirement for

Master of Science degree in Business Analytics

**Department of Computing**

**Informatics Institute of Technology, Sri Lanka**

**in collaboration with**

**Robert Gordon University, UK**

**2022**

**Abstract**

Flight delays are a major concern when it comes to the airline industry hence the delays end up in numerous circumstances causing passenger dissatisfaction and financial loses. As per the statistics, the percentage of flights that had delayed during year 2021 in the United States is around 10.93% only in departures. When it comes to both arrival delays and departure delays, the total percentage is above 20% which is very high. If the underlying reasons can be identified and if the delay can be predicted in advance, it will be beneficial for all related stakeholders. Therefore, the research has focused on identifying the major factors affecting flight departure delays considering the actual flight operational data of the busiest airport in the United States, Hartsfield-Jackson Atlanta International Airport along with the weather data of years 2019 and 2020.

Prediction models are developed using different classification algorithms to identify the best model to predict flight departure delays in identified four classes as 'No delay', 'delay between 15-60 minutes', '61- 120 minutes' and 'above 120 minutes'. While identifying the best model, the research tried to identify whether the covid19 pandemic had made any impact in delaying flight departures in the United States. To identify this, three combinations of training and testing datasets was used during the implementation and testing stages.

As the outcome, the model which resulted with the highest value for macro average of F1 value was identified as the best model for flight departure delay prediction.