**Group report**

**Executive Summary**

We present a boosted decision tree which correctly predicts if a flight will be delayed more by than 20 minutes with 84% accuracy. These predictions were made on hundreds of thousands of observations that had not been used to train the model. Using this model will massively improve your ability to avoid booking employees on flights which are subsequently delayed. This will save on costs, time and employee satisfaction.

**Introduction**

Dear *client*, in response to your request for a solution to the challenge of avoiding flight delays, we are delighted to present an accurate and easily interpretable model that correctly predicts delays with **84% accuracy**. There are always different ways to approach an answer, but we understood the importance to you of using a model that is simple and provides accurate results. Therefore, after testing many models, we present a boosted decision tree fits both of these criteria.

This model will bring many benefits to your business. Firstly, it will save administration time as it is quick and easy to use and provides an immediate answer on if the flight will be delayed or not. Also, time will be saved as employees will be unlikely to experiences delays when travelling. This also has the benefit of improving employee satisfaction, as studies have shown that business travel can have a detrimental effect on job satisfaction[[1]](#footnote-1)[[2]](#footnote-2). Therefore, it is important that this experience is as pleasant as possible. In addition to this, it has been reported that flight delays cost $billions and passengers have to foot the cost for much of this[[3]](#footnote-3).

**The Model**

Our model is a boosted decision tree which predicts if a flight will be delayed by 20 minutes or longer using the following attributes:

* Hour of scheduled departure
* Airline
* Distance of flight

By using only these variables as predictors, it allows you to make an accurate prediction with readily available information. Different decision trees were built and this one was selected as it was the most accurate. More importantly, it was the best model for correctly predicting flight delays, which we understand is more important than correctly predicting on-time departures (although it does this with great accuracy as well). The model was tested against unseen data and returned accuracy ratings as shown in the table below.

|  |  |  |
| --- | --- | --- |
| **Overall Accuracy** | **Delay Prediction Accuracy** | **On-time Prediction Accuracy** |
| 84.0% | 79.8% | 84.0% |

1. https://hub.gbta.org/groups/reviews/item/20/11/2385 (GBTA, 2016) [↑](#footnote-ref-1)
2. https://www.viewfrom36k.com/business-travel-employee-satisfaction [↑](#footnote-ref-2)
3. https://isr.umd.edu/NEXTOR/pubs/TDI\_Report\_Final\_10\_18\_10\_V3.pdf (Ball et al, 2010) [↑](#footnote-ref-3)