



SCHOOL OF COMPUTING SCIENCE AND DIGITAL MEDIA

MSc Business Analytics

Student Name: Sameera Prasad Alahakoonge Don	Matriculation Number: 1813679
Supervisor: Dr. Alroy Mascrenge	Second Marker:
Project Title: Predictive Model for Customer Satisfaction in Airline Industry	
	Start Date: 21 st April 2022
	Submission Date: 24 th August 2022

CONSENT

I agree
I do not agree

That the University shall be entitled to use any results, materials, or other outcomes arising from my project work for the purposes of non-commercial teaching and research, including collaboration.

DECLARATION

I confirm:

- That the work contained in this document has been composed solely by myself and that I have not made use of any unauthorised assistance.
- That the work has not been accepted in any previous application for a degree.
- All sources of information have been specifically acknowledged and all verbatim extracts are distinguished by quotation marks.

Student Signature: Sameera Prasad	Date Signed: 24 th August 2022
-----------------------------------	---





Abstract

Customer Satisfaction is one of the most crucial attributes of a service-providing company and specifically in the airline industry which falls under both travel and hospitality sectors, customer satisfaction plays a critical role. Having provided different services, airlines struggle to identify where customer satisfaction lies. Is it on the comfort they provide the passengers throughout a journey, is it preflight and post-flight services that the airline provides, or is it a combination of both facts. Budgets available to a company would be needed to be utilized in high optimization and identifying where exactly the investments should happen is key information that an airline company would need to have handy.

A dataset related to airline customer characteristics and customer satisfaction on each key service that an airline provides has been employed in this study. A descriptive statistical analysis, as well as predictive modeling using six different machine learning algorithms, are used to achieve the other objectives of the study. The algorithms are Logistic Regression, K-Nearest Neighbor Algorithm, Naïve Bayes, Random Forest, Support Vector Machine, and Extreme Gradient Boost algorithm. The secondary dataset which was obtained is divided into a training dataset as well as a testing dataset. The modeling has been done using the statistical language “R” with the assistance of “R Studio”.

The models built are compared against each other and the observations show that the Random Forest algorithm gives the best accuracy rate of 96.39%. Hence it has been selected as the best machine learning algorithm that makes predictions for the subject dataset in the context explained. Its noted that the type of travel, customer travel class and in-flight services have direct impacts on customer satisfaction and arrival and departure delays on flights do not have a very significant relationship on customer satisfaction. Based on the study results, recommendations are made to the company to keep satisfaction at an optimum level based on resource availability. The company is recommended to use the models to deploy their marketing strategies in to the model and predict customer satisfaction against different marketing strategies they have.