ROAD TRAFFIC ACCIDENT SEVERITY PREDICTION USING DATA MINING ALGORITHMS FOR SRI LANKA

N. K. Thilini Malika Wijesinghe

A dissertation submitted in partial fulfillment of the requirement for

Masters of Science degree in Business Analytics

Department of Computing Science and Digital Media

Informatics Institute of Technology, Sri Lanka

In collaboration with

Robert Gordon University, Aberdeen, UK

2022

Abstract

Transport sector is one of the major aspects that determines a country's economic situation. Therefore, ensuring road safety and administration of road accidents in order to establish a solid transport sector is crucial in the long run. Hence, authorities and government bodies such as Sri Lanka Police, National Council for Road Safety, Road Development Authority have been collectively working and formulating policies to ensure road safety and safe transportation within the country.

Currently data modelling using various machine learning algorithms or data mining algorithms has been explored by various researchers around the globe to expect the crash rates and to recognize the factors influencing higher rates of road traffic accidents. Major focus has been placed on this research arena by many countries in order to predict the severity classes of the expected accidents by many researchers worldwide mainly because it is crucial to put in place proper mechanisms and remedial measures to minimize the occurrences of those which would bring out major reductions in economic and social costs.

Using data mining algorithms has been pursued worldwide to determine the severity classes of road traffic accidents proved to give out best results from the previous researches undertaken. This research has focused on applying classification models to determine the traffic accident severity in Sri Lankan context given the fact that little focus has been given on broadening the road transport safety research arena on the aspect of using data mining algorithm based models to predict the severity classes of the probable accidents. This research has focused upon implementing several of the data mining algorithms such as multinomial logistic regression, KNN algorithm, decision trees, Naïve Bayes, Random Forest in developing its prediction models to determine the accident severity in Sri Lanka based on the existing accident records.

Key Words – Accident Severity Prediction, Classification Models, Data Mining Algorithms, Traffic Accidents, Road Safety, Sri Lanka