A Decision Support System to Predict Highway Accident Alerts in Sri Lanka

K. A. D. S. H. Rodrigo
179476A

Faculty of Information Technology
University of Moratuwa
2020

A Decision Support System to Predict Highway Accident Alerts in Sri Lanka

K. A. D. S. H. Rodrigo
179476A

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the fulfillment of the requirements of Degree of Master of Science in Information Technology.
2020

Declaration

I declare that this is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student

K. A. D. S. H. Rodrigo

Signature of Student

……....

Date: ……...

Supervised By

Name of Supervisor

S. C. Premarathne

Signature of Supervisor

…….

Date: ……

Acknowledgement

Firstly, I would like to express my sincere gratitude to my supervisor Mr. Saminda Premarathne for the continuous support of my dissertation study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my research study.

Besides my supervisor, I would like to thank Dr. Mohamed Firdhous who taught Research Methodology and Literature Review and Thesis Writing modules which helped me to widen my research from various perspectives.

I would not forget to remember the Sri Lanka Police for their timely support by providing all the accident records that have been required for the success of this research.

In addition, I would like to thank all the staff of Paragon Software Lanka for their insightful comments and encouragement which helped me lot to enhance my knowledge from various perspectives.

Furthermore, I would also like to thank all the batch mates of the M.Sc. in IT degree program who gave their valuable feedbacks to improve the results of the research, my family for the support they provided me throughout my entire life and in particular.

Abstract

Road traffic and rapidly increasing road accidents become a vast problem not only for Sri Lankans but for all human beings who are living in this planet. According to the Sri Lanka police highway reports, there have been a number of deaths recorded due to fatal accidents in highways and also affected significant amount of people for non-fatal accidents as well. Several key factors vastly contributed directly to cause an accident, such as environmental factors, human factors and road condition factors., etc. In this research, machine learning techniques and methods have been applied to the southern highway accident records retrieved by Police Highway Control Center for the period of 2015 to June 2019 in order to establish a model which enables to forecast reason for the accidents that will be occurring in the future. The python scikit-learn package has been used on top of the anaconda framework to discover hidden patterns of data with the help of decision tree, support vector machine and logistic regression powered by a one-vs-rest classifier. The two well-known ensembles; random forest and gradient boost classifiers have also performed in this dataset in order to enhance the accuracy. The performance of each classifier has compared critically based on their results. The results obtained by performing various experiments show that the ensemble's work well when compare to other classifiers.

Table of Content

List of Tables

List of Figures

List of Abbreviations

SLP - Sri Lanka Police

USA - United States of America

CART - Classification and Regression Tree

ID3 - Iterative Dichotomiser

UAE - United Arab Emirates

ANN - Artificial Neural Network

SVM - Support Vector Machine

LR - Logistic Regression

OVR - One-Vs-Rest

OVO - One-Vs-One

OVA - One-Vs-All

GPS - Global Positioning System

SMS - Short Message Service

IOS - iPhone Operating System

LHS - Left Hand Side

RHS - Right Hand Side

KNN - K-Nearest Neighbor

PCA - Principal Component Analysis

DT - Decision Tree

ROS - Random Over Sampler

GS - Grid Search

RS - Randomize Search

RF - Random Forest

GB - Gradient Boost

API - Application Program Interface

AWS - Amazon Web Services

UI - User Interface

UX - User Experience

List of Appendices