

Student Name: Mohamed Rafhan Abdul Azeez Mohamed Rifan	Matriculation Number:
Supervisor: Alroy Mascrenge, PhD	Second Marker:
Project Title: Accident Severity Modelling of Road Accidents in the UK using Random Forest, Decision Tree and Multinomial Regression	
Word count: 18800	Start Date: 01.04.2022
	Submission Date: 05.05.2023 Resubmission Date: 27.05.2023

**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:	Date Signed:
--------------------	--------------

## **Abstract**

Accidents causes considerable cost in terms of human life and financial costs every day. There is various research taken place to control the severity of the accidents. The data of accidents are being collected by various authorities to analyse the factors and understand how the accidents are happening, what are the measures can be taken to control the occurrence of the accident as well as the severity. This study focuses on the accident severity using the key factors that could be reasons for an accident which are road factors, weather factors and speed limits. Under each factor, there are multiple secondary factors which could contribute the accidents severity. This research focuses on the secondary factors and tries to find out how these factors are interrelated and how the prediction of accident severity changes with the speed limit. Features were selected manually and using the filter methods. Identified features were used to develop fourteen feature combinations and each combination was further analysed using random forest, decision tree and multinomial regression algorithms in R Studio. The analysis was focused on finding out the performance of each model and how the factors are importance in predicting the accident severity. It was identified that the combination which comprises the all the manually selected features of road, weather along with the speed limit performs well compared to all the other thirteen combinations analysed in this study. This combination of features performed better compared to the features selected using the Chi-Square testing. It is also identified that junction control and junction details play a major role in accident severity. And the speed limit is third to these two factors. Further research is possible with combining other human and vehicle related factors to the model and understand how the complex scenarios plays when deciding the accident severity class.