

Introduction

10 Problem statement

The deaths and injuries from traffic accidents are now a world phenomenon. Many countries all over the world are greatly concerned about the growth of mortality and injury rate on the road. According to WHO, every year the lives of approximately 1.35 million people are cut short as a result of a road traffic crash. Between 20 and 50 million more people suffer non-fatal injuries, with many incurring a disability as a result of their injury.

Road traffic injuries cause considerable economic losses to individuals, their families, and to nations as a whole. These losses arise from the cost of treatment as well as lost productivity for those killed or disabled by their injuries, and for family members who need to take time off work or school to care for the injured. Road traffic crashes cost most countries 3% of their gross domestic product.

1.1 Target audience

Motorists, people who use public transport, police and medical personnel are usually inconvenienced greatly when accidents occur and especially severe accidents. For, severe accidents it takes quite some time and resources for the police and medical personnel to clear the scene. This thus leads to many hours of wait. Students are likely to be late for their classes, people going for interviews get late, those expected in important meetings may not be able to attend, surgeons and doctors expected to save lives run late, just to name but a few. Although severe accidents are unexpected, it would be possible to build a machine learning algorithm to predict severe accidents using various factors that are determined to be related to a severe accident occurring. This would thus serve as an automated reference to help people make decisions where and when there is a high chance of a severe accident occurring and hence plan accordingly.

1.2 Scope and main aim

Factors influencing accident frequency may vary from the ones affecting the severity; hence, it is suggested that their analysis should be performed carefully. A set of explanatory variables, which could include: driver attributes (e.g., age and gender, whether under influence of alcohol), vehicle features (e.g., body type, vehicle age and number of vehicles

involved in the accident), road characteristics (e.g., number of lanes, road surface conditions, intersection control and types of road), weather conditions, day of week, time of day, speed limit and accident characteristics (e.g., accident's main cause) have been shown to be possible predictors of accident severity.

The **aim of this project** is therefore to develop a machine learning algorithm that would warn motorists and other people using public transport where and when there is a high chance that a severe accident would occur and hence help them plan accordingly to avoid delays and inconveniences.