Traffic Injury Prevention Techniques Using STATS19 Road Safety Data: A Model-comparison Approach

Meisam Siami Doudaran

One Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Civil Engineering

Eastern Mediterranean University
September 2020
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in Civil Engineering.

Prof. Dr. Umut Türker
Chair, Department of Civil Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Civil Engineering.

Prof. Dr. Ersun İşçioğlu
Supervisor

Examining Committee

1. Prof. Dr. Hilmi Berk Çelikoğlu
2. Prof. Dr. Özgür Eren
3. Prof. Dr. Ersun İşçioğlu
4. Prof. Dr. Zalihe Nalbantoğlu Sezai
5. Prof. Dr. Serhan Tanyel

ABSTRACT

In this thesis, five models, which are based on artificial neural network (ANN) and support vector machine (SVM), were proposed for prediction of personal injury severities. The models were examined by two case studies using STATS19 road safety data that occurred in the city of London and Cambridge. The main purpose of the first case study was to identify the group most in need of road safety intervention by predicting the severities sustained by all road users. Using Radial Basis Function (RBFNN), different factors and areas of concern contributing to direct actual influences in both case studies were identified and ranked. In more detail to the first case study, non–motorised road users were recognised to benefit from the interventions. Therefore, the second case study aims to predict cyclist injury severities. Furthermore, most of the key factors in both case studies were in connection with busy junctions and poor turn / manoeuvres, here, truly protected junctions might be the best answer to create space for everybody. On focus to the two-wheeled group, there were limited crossing facilities near to where they cycled. Importantly for Britain's everyday cycling capital, narrow bike lane defenders are needed to provide a fully segregated solution where road width is too limited.

In order to increase prediction accuracies, key factors were applied to multi–layer perceptron neural network (MLPNN) and SVM in both case studies. The models were selected as the benchmark due to their popularity in prediction modelling. Although the results of the predictions are encouraging, the models were not able to overcome incorrect predictions for 'fatal' and 'serious injury' severities due to limited data for those classes. In response to this, the two well-known models were combined as a hybrid MLPNN-SVM, and a learning vector quantization neural network (LVQNN) was improved for the first time ever to verify the best-fit model. Following this, different comparisons were made to evaluate the performance of the models in different classes. In addition, the models' fitting results were presented and discussed, suggesting that all proposed models have ability to achieve satisfactory predictions, nevertheless, the improved LVQNN model performed better than others and was properly able to solve the incorrect predictions. This thesis concludes by identifying evidence-based road safety intervention options to mitigate the identified concerns. The general conclusion that can be drawn from this study is that most of the factors directly blame some kind of human error with high injury concentration being linked to junction actions. Therefore, in to crack down on bad driving / cycling, besides the road engineering interventions, it is recommended to deliver innovative road safety education and broadcast promotional messages for the recognised groups.

Keywords: cyclist, driver, injury severity prediction, road safety intervention, STATS19.

ÖZ

Bu tezde, kişisel yaralanma şiddetini tahmin edebilmek için, yapay sinir ağına (YSA) ve destek vektör makinesine (DVM) dayalı beş farklı model kullanılmıştır. Önerilen yöntemler, Londra ve Cambridge şehirlerinin kentsel sokak ağlarında meydana gelen, STATS19 verilerinin kullanıldığı iki vaka çalışmasıyla gösterilmiştir. İlk vaka çalışmasının temel amacı, tüm yol kullanıcılarının maruz kaldığı kişisel yaralanma şiddetlerini tahmin ederek, karayolu güvenliği müdahalesine en çok ihtiyaç duyan grubu belirlemekti. Radyal temel fonksiyonu (RTF) kullanan ilk vaka çalışmasında, doğrudan gerçek etkilere katkıda bulunan faktörler ve endişe alanları tanımlanmış ve sıralanmıştır. Sonuç olarak, ilk vaka çalışmasında, motorsuz karayolu kullanıcılarının müdahalelerden faydalandığı belirlenmiştir. Bu nedenle, ikinci vaka çalışması, özellikle bisikletçi yaralanma şiddetlerinin tahminine odaklanmıştır. RTF'nin tahmin sonuçları, faktörlerin çoğunun yoğun kavşaklar ve zayıf dönüşler / manevralarla bağlantılı olduğunu göstermiştir. Burada gerçekten korunan kavşaklarda herkes için alan yaratmak en iyi çözüm olabilir. İki tekerlekli araç gruplarına odaklanıldığında, bisiklet sürüş alanlarının yakınında sınırlı sayıda geçiş tesislerinin var olduğu belirlenmiştir. Özellikle İngiltere'nin bisiklet başkenti için, yol genişliğinin çok sınırlı olduğu yerlerde tamamen ayrılmış bir çözüm sağlamak için dar bisiklet şeridi savunucularına ihtiyaç olduğu söylenebilir. Bu göstergelerin ardından, her iki vaka çalışmasında da tahmin doğruluğunu en üst düzeye çıkarmak için çok katmanlı algılayıcı sinir ağına (ÇKA) ve DMV'ye temel yaralanma şiddeti etki faktörleri uygulanmıştır. Elde edilen tahminlerin sonuçları her iki model için iyi olsa da, 'ölümcül' ve 'ciddi yaralanma' şiddetleri sınıflar için sınırlı sayıdaki veri

nedeniyle hatalı tahminlerin olmasını engelleyemedikleri görülmektedir. Yanlış tahminleri düzeltmek için, daha önce kullanılan iki modeli birleştirilerek karma bir model (hibrit ÇKA-DMV) yaratılmıştır. İlaveten öğrenen vektör niceleme (ÖVN) sinir ağı olan, verilerinin tahmini için en uygun modeli doğrulamak için geliştirilmiştir. Önerilen tüm modellerin tatmin edici tahminlere ulaşma yeteneğine sahip olduğu, bununla birlikte, geliştirilmiş ÖVN modelinin diğerlerinden daha iyi performans gösterdiği ve belirli sınıflar için verilerin sınırlamasını düzgün bir şekilde çözebildiği görülmüştür. Bu tez çalışması, ciddi boyutlu yaralanmaları azaltmak için kanıta dayalı yol güvenliği müdahale seçeneklerini belirleyerek sona ermektedir. İncelenen faktörlerin çoğu, yüksek yaralanmalı sonuçlarla, insan hatası ve kavşaklarda hareket öncelikleri ile bağlantılı olduğunu ortaya koymuştur. Dolayısı ile çalışma sonuçlarına göre, kötü sürüş / bisikletçilik davranışlarını değiştirmek için, yol mühendisliğine dayalı müdahalelerinin yanı sıra, belirli gruplara yenilikçi karayolu güvenliği eğitiminin verilmesi ve tanıtım mesajlarının yayınlanmasının yapılması önerilmektedir.

Anahtar Kelimeler: bisikletçi, sürücü, yaralanma şiddeti tahmini, yol güvenliği müdahalesi, STATS19.

ACKNOWLEDGMENT

Undertaking this PhD has been a truly challenging experience and long-winded process for me and it would not have been possible to carry out without the support and guidance that I received from a number of people. First of all, I would like to thank my supervisor, Prof. Dr. Ersun İşçioğlu, for his patient guidance and support during this research. Prof. Dr. İşçioğlu gave me freedom to make my own choice in terms of what I am very passionate about, but supported me when I needed it. I have been lucky to have such a supervisor because of his approachability and prompt responses.

My gratitude extends to all the Civil Engineering Departmental staff who helped me during these years of my study both academically and administratively. Notably, the head of the department Prof. Dr. Umut Türker, my course advisor Prof. Dr Zalihe Nalbantoğlu Sezai, and Prof. Dr. Özgür Eren who was the head of the department as well as the head of graduate studies during the majority part of my study in Eastern Mediterranean University (EMU).

Additionally, I would like to express gratitude to all jury members for their invaluable insights and suggestions which highly improved my thesis. Particularly, in the field of transport, my sincere appreciation is reserved for Prof. Dr. Serhan Tanyel and Prof. Dr. Hilmi Berk Çelikoğlu. Their contribution at different stages of this thesis was fundamental in improving this work.

It is also a great pleasure to pay tribute to the UK Department for Transport (DfT) and Cambridgeshire County Council for providing the road safety data used in this thesis.

Finally, a special thanks to my parents who have supported me throughout my travel overseas and my studies. I would like to dedicate this degree to them as an indication of their importance in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A	Number of Neurons Linked to Input Layer
α_i	Lagrange Coefficients
B	Number of Neurons Connected to Concealed Layer
b_j	Bias Value of j^{th} Concealed Layer Neuron
b_k	Bias Value of k^{th} Output Layer Neuron
C	Fine Cause
f	Activation Function in Concealed Layer
g	Activation function for output layer
j^{th}	Concealed Layer Neuron
K	Kernel Trick
k^{th}	Output Layer
L_p	Saddle Point
M	Distance between Support Hyperplanes
p_j	Value of j^{th} Concealed Layer Neuron
w	Neuron Weight
w_{ij}	Amount of Weight between i^{th} and j^{th}
w_{jk}	Amount of Weight between j^{th} and k^{th}
x	Input Layer Neuron
x_i	Value of Input Layer Neuron
δ	Distance of Classes
δ_i	Slack Variable
i^{th}	Model's Input

μ_i	Lagrange Coefficients
μ_k	Centre of Network
ϕ_k	Radial Function
ψ	Radial Function Predicted Result

LIST OF ABBREVIATIONS

AACN	Advanced Automatic Collision Notification
ACC	Accuracy
AIDS	Acquired Immune Deficiency Syndrome
AI	Artificial Intelligence
ART	Adaptive Resonance Theory
ANN	Artificial Neural Network
BBC	British Broadcasting Corporation
BSfE	Better Streets for Enfield
BNN	Bayesian Neural Network
BPNN	Back-Propagation Neural Network
CAR	Conditional Autoregressive
CART	Classification and Regression Tree
CHAID	Chi-squared Automatic Interaction Detection
CIHT	Chartered Institution of Highways and Transportation
COVID-19	Coronavirus Disease 2019
CPM	Crash Prediction Model
CYCLOPS	Cycle Optimised Protected Signals
DfT	Department for Transport
EMU	Eastern Mediterranean University
EV	Expected Value
FB	Full Bayesian
FP	False Positive
FN	False Negative

GA	Genetic Algorithm
GB	Great Britain
GCP	The Greater Cambridge Partnership
HIV	Human Immunodeficiency Viruses
ICE	Institution of Civil Engineers
IHE	Institute of Highway Engineers
IMD	Index of Multiple Deprivation
ISJ	International Security Journal
ISP	Injury Severity Prediction
KSI	Killed or Seriously Injured
LM–BP	Levenberg–Marquardt Backpropagation
LMIC	Low and Middle-Income
LVQNN	Learning Vector Quantization Neural Network
MATLAB	Matrix Laboratory
MLP	Multi-Layer Perceptron
MLPNN	Multi-Layer Perceptron Neural Network
MNL	Multinomial Logit
MSE	Mean Squared Error
MVC	Motor Vehicle Collision
NB	Negative Binomial
NHS	National Health Services
NNC	Nearest Neighbor Classification
OP	Ordered Probit
PKC	Perth & Kinross Council
PLN	Poisson Lognormal

PRE	Precision
PSV	Public Service Vehicle
R	Correlation Coefficient
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network
RF	Random Forests
RIPLN	Random Intercepts Poisson Lognormal
RLS	Recurrent Least Squares
RMSE	Root Mean Square Error
RoSPA	Royal Society for the Prevention of Accidents
RPPLN	Random Parameters Poisson Lognormal
SEN	Sensitivity
SPLN	Spatial Poisson lognormal
SWOV	Institute for Road Safety Research
SVM	Support Vector Machine
SOM	Self Organising Maps
TfGM	Transport for Greater Manchester
TfL	Transport for London
TN	True Negative
TRL	The Future of Transport (Transport Research Laboratory)
TP	True Positive
UK NARIC	United Kingdom National Recognition Information Centre
VRU	Vulnerable Road User
WHO	World Health Organization

Chapter 1

INTRODUCTION

1.1 Background

1.1.1 Road Traffic Injuries and Deaths – Global Concern

Road traffic injuries are always a major worldwide health concern as well as main barrier in the transport industry. Accident occurrence causes immense losses from the human, economic, and social sides, especially the injury and fatal crashes since current efforts to address road safety are minimal in comparison to this growing human suffering. As a result of increasing advance in technology and growing human population, road traffic injuries have become as a serious public health problem and the most important causes of unnatural losses in today’s world. According to statistics from World Health Organization (WHO) (2018a), every day more than 3700 people are killed, and thousands are seriously injured on the world's roads. People driving, riding, biking, walking to work or school, children playing in streets, will never return home. A lot of people every year spend long times in hospitals after severe accidents and many of them never are able to continue their normal life (WHO, 2018a). Kids, pedestrians, riders and elderly people are between the most vulnerable road users (VRUs). Unfortunately, road injuries are the eighth leading cause of fatalities worldwide for all ages. In addition, the leading cause of death among young people, aged 5–29 years.

1.1.2 Can Traffic Injuries Be Compared to AIDS, HIV, or COVID-19?

A report from WHO displays that due to the road traffic related injuries more people now are killed than from tuberculosis and the human immunodeficiency viruses (HIV) / acquired immune deficiency syndrome (AIDS). This organization also predicted that there will be around a 67 % by the year 2020 increase of the already existing 1.35 million deaths each year, unless there is new commitment to prevention or reduction of these injuries. Road traffic accidents are not contagious like Coronavirus disease 2019 (COVID-19) but more than 1.35 million people lose their lives every year on the roads plus an extra 50 million suffer non-fatal injuries, often resulting in long-term disabilities. (WHO, 2018b). However, nobody has suggested a lockdown that will for sure save many lives and prevent many more hospitalisations with the additional bonuses of cleaner air and fighting climate change.

1.2 UK's Road Safety Stats Raise Serious Concerns

The main barrier also affects safety of Britain's roads. Though Britain regularly has one of the greatest road safety records in the world but there are still numerous people die and many road users seriously injure day-to-day on the roads. The effects of the casualties are devastating, for public, for bereaved, and loved ones, and for victims who suffer the injuries, some of whom possibly will have lasting life-changing consequences (DfT, 2019a). Therefore, here is much more work to be done across the country in connection with traffic collision prevention and reduction techniques. It is vital to track the safety and reliability of Britain’s roads, in particular, where most of serious injuries and fatalities are concentrated, and which should be targeted (RoSPA, 2018).

1.2.1 Multi-million-pound Investment to Deliver Road Safety

The UK government's vision is multi-million-pound investment with respect to achieve a genuine avoidance or reduction of the appalling concerns resulting in mortalities and injuries. Hence, quickly mitigate the issues and make sure that the country lasts to retain as a strong global leader in road safety. Safer road infrastructures for all road users and safer road users will definitely save more lives, but this strategy will also help to decrease pressure on the national health services (NHS) and emergency services in the UK. Likewise, efficient movement of traffic and finally, keep the economy growing (DfT, 2018a).

1.2.1.1 Road Danger Reduction and Active Travel Plan

Road danger reduction and active travel plan for the central of London is part of this multi-million-pound investment. This plan seeks to set out important aims to sustain a safe environment for all road users as well as improve the quality of life in the City of London. It goals to work towards removing the annual number of people killed and seriously injured in road accidents to zero before 2041 (City of London Corporation, 2018).

1.2.2 Safety of Vulnerable Road Users

As a result of the dramatic increase in traffic related injuries, in recent years, road safety authorities have become more committed and have given full attention to solve any road safety concern. Therefore, it is necessary to encourage all road safety stakeholders to work better in partnership helping the government in building vision as a reality. Along with this goal, it is also important to prevent the fairly great danger of some clusters more quickly, in particular for VRUs such as; pedestrians, pedal riders and motorcyclists. More than half of all road traffic losses happen among VRUs, therefore, this group require effective road safety interventions due to the lack of personal protection. For example, cyclists are particularly more vulnerable in event of a motor vehicle collision (MVC) since they do not have the same additional protection which an enclosed vehicle provides (DfT, 2018b).

1.2.3 Road Danger as Biggest Barrier for UK Cycling

Cycle traffic over the past decade has highly increased in the UK as it is a brilliant approach to get about and delivers a wide range of health and environmental welfares (DfT, 2018b). Specially, COVID-19 has created more cyclists to ease demand on the public transport, thereon, local authorities encourage people to access goods, services and activities within their local area maximally by cycling, so as to retain a safe distance from others and keep active (PKC, 2020). According to rule 66 of the Highway Code, cyclists must be allowed to ride side by side (DfT, 2015a). Therefore, it's vital to create safer roads for cyclists as well as making them more visible to motorists. As a result, this should also make it easier for careful drivers to overtake as pedal riders are not spread out along the road. However, the vulnerability to serious injuries has grown more rapidly than traffic, to illustrate, DfT reports that there was a 48% rise of serious injuries in recent years (2018b). Each year in Britain about 18,500 pedal riders are killed or injured as stated by police reported collisions, plus nearby 3,500 who are killed or seriously injured (KSI). In addition, numerous cyclist casualties are not reported to the police despite some of them being critical enough to be hospitalised. (RoSPA 2017c). As a result, statistics of the DfT determines that percentage share of bicyclist fatality is higher than car occupants (DfT 2018b).

1.2.4 Cambridge as UK's Cycling Capital - It's on Decline

Cambridge city is UK's everyday cycling capital and the Cambridgeshire council aims to seriously rival with world's most bicycle friendly cities (CityLab, 2015).

With this ambition, at the moment it is not surprising that cycling in Cambridge became an appealing means of transport (The Guardian, 2011). The city demonstrates to be extremely perfect for pedal riders. You just must venture out on a weekday to understand why this city is considered such a great biking destination. Of course, the city hasn't got a lot of hills but there are many other reasons too. It boasts the wide-ranging cycle route networks of over 800 miles connecting the city with around towns and villages. Cambridge railway station bike park hosts space for over 3,000 bicycles, and soon after it being opened to riders in 2016 it gained the top honour in the National Cycle-Rail Awards. Another reason biking is so widely held in Cambridge is that the residents are very, very pro-environment. They believe pedal riding is superior for the planet as well as better for their health. (CambridgeshireLive, 2018a). A large population of students hop into the saddle to access to the venerable university campuses (DfT, 2016). Most of the routes across the campuses are entirely open to cycling, thus up to half of all journeys in the city centre are made by bicycle (GCP, 2017).

1.2.5 General Facts and Figures about British Road Users

In Cambridge, out of the commuting residents, 32.5% cycle to work which is by far the highest ratio in the UK, while the rate across England is only 3.1%. (The Cambridge News, 2017). In total, pedal riding made up just 1% of the mileage aggregated by all vehicular traffic across Great Britain throughout 2018. In comparison, cars and taxis accounted for approximately 77% in mentioned year. The percentages of road traffic (vehicle miles) for all the vehicle types in the country are shown in the chart below (Cycling UK, 2018; DfT, 2019b).

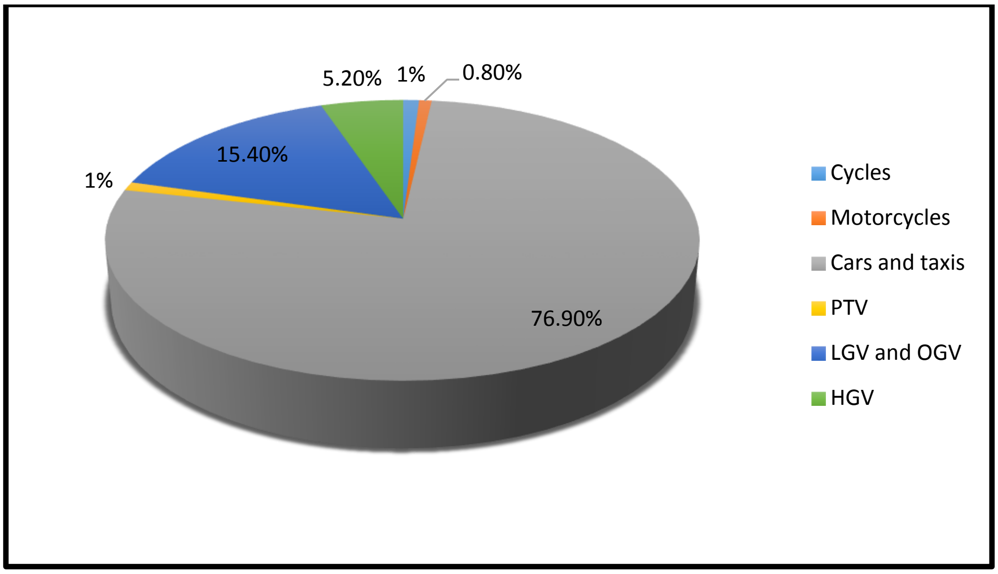

Figure 1: Proportion of Traffic by Vehicle Type in Great Britain (Cycling UK, 2018)

In particular, two-wheel use as a percentage of all vehicle miles for the same year are shown in the Table 1 for England, Scotland, Wales as well as through Great Britain. In the Table below, the amount for biking includes riding on cycle paths and public roads. It also does not contain biking movement elsewhere such as; on byways, bridleways or towpaths (DfT, 2019b). In the Table, bvm denotes billion vehicle miles (Cycling UK, 2018).

Table 1: Wheel use as a ratio of all vehicle miles

Region	Motor vehicle (bvm)	Cycle (bvm)	All bvm (motor + cycle)	**Cycled** (%)
Great Britain	328.1	3.3	331.4	1.00
England	280.1	3.0	283.1	1.10
Scotland	29.7	0.2	29.9	0.60
Wales	18.3	0.1	18.4	0.60

Figure 2 shows a proportion of miles cycled across Britain roads. The large percentage of British biking trips are referred to England. In line with England, two-wheel use increase has been greater in some urban zones. For instance, in London nearby 27,000 people biked among the central London cordon in 1977, in comparison to 162,000 in 2017. This amount is six times as many. Another example can be the tens of thousands of riders who use their bicycles in Cambridge each day. Cambridge leads the way in biking, and it is increasing day after day (Cycling UK, 2018).

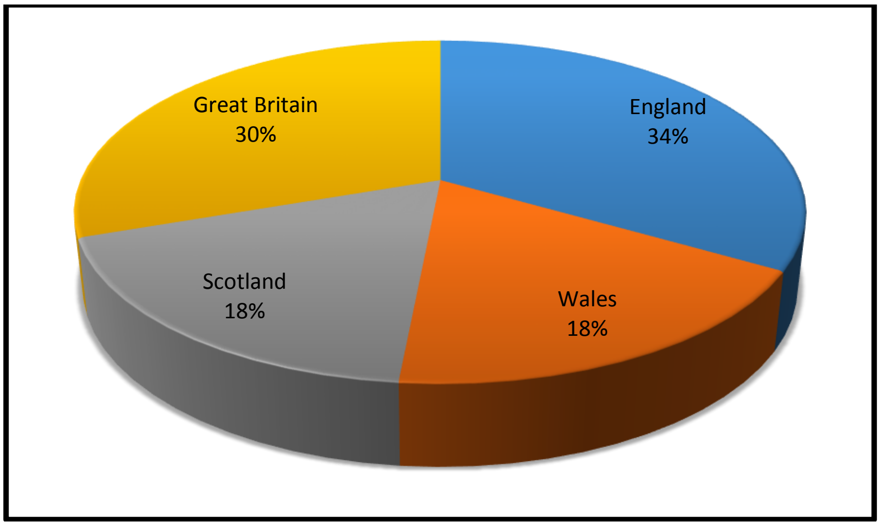

Figure 2: Proportion of British Biking Trips Across Britain (Cycling UK, 2018)

1.2.6 Automatic Cycle Counters across the UK

The UK government installed a wide number of cycle counters across the country. For example, Cambridge’s bike counter recorded that there are more than a million trips each year and around 3,000 journeys on a daily basis (GCP, 2017). The Figure 3 shows the counter which was installed in a bend of a cycle path in a corner of Parker's Piece.

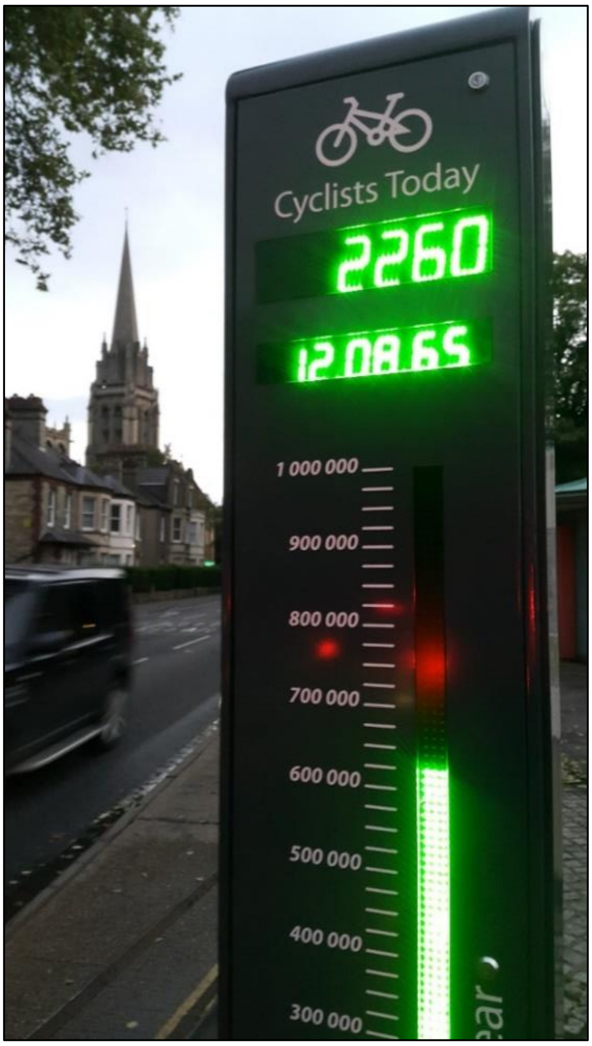

Figure 3: Cycle Counter in Cambridge

Another counter in Cambridge is located on Hills Bridge road adjacent to the station which is used by more than five thousand pedal riders in a day. Accordingly, as result of the large database with cycling statistics, first priority has been given to pedal riders in Cambridge (CambridgeshireLive, 2018a).

In order to visually understand why Cambridge leads a large percentage of population of pedal riders, we compared the number of people biking regularly in Cambridge with other cycling destinations in the UK. For example, cycles were the ‘major mode’ for 1% of all trips in in 2018 across Northern Ireland (Cycling UK, 2018).

A national survey by Cycling Scotland discovered the scots' attitudes to getting on a bike. The study found that the most of the people living in Scotland trust biking isn't for them and they have an 'entrenched reluctance' to getting on two wheels more often. Unfortunately, a particular number of those who were hesitant to biking either didn't know how to ride or didn't believe they were fit enough. The proportion of British biking trips shown in Figure 2 and the live cycle counter data in Figure 4 indicates that there is very small daily and annual number of people cycling in Scotland. The finding of the survey acknowledges that their absence of investment is turning a lot of people away from biking.

Due to the existing lack of cycling infrastructure, most of the Scots don't consider themselves as regular cyclist and don't support cycling as a means of transport. Compared to cycling destinations similar to Netherlands, where physically separated bike lanes are common, Scots share roads with motor vehicles. Scottish government sees motor vehicle traveling and flying as main priorities in their transport networks. However, Cycling UK reported that Scottish Government increases cycling investment in 'active and sustainable transport' (Inews, 2018). Therefore, number of trips undertaken by bike will increase in Scotland as they find it as cheap travel, affordable and convenient daily transport. Cycling UK hopes that this raise is just the start of a riding revolution in this country (Cycling Weekly, 2018).

Figure 4: Cyclist Counting Machine in Scotland, PKC

1.2.7 Cycling Numbers in UK Jumped during Covid-19 Pandemic

Fear of catching COVID-19 on public transport has helped lead to a boom in cycle-to-work schemes in the UK. During Covid-19 lockdown and post-lockdown cycling numbers have also jumped in Scotland. The cycle counters at many sites found in some locations the number of pedal riders more than doubled. Scots have rediscovered riding during the virus, for essential journey and exercise. It is hoped that people will continue to ride and carry on benefiting from the great helpful influence biking has on physical and mental health (BBC, 2020a; BBC, 2020b). In this regard, local authorities have invested hundreds of millions of pounds since the lockdown in cycling in response to the Coronavirus (PKC, 2020).

At the same time, on the other side of the UK in Cambridge biking has regularly continued as the leading means of transport. Many investments into pedal riding have sustained over the years (CambridgeshireLive, 2018a). The latest model of cycling network is now open to traffic as part of a multi-million pound' government schemes towards producing more cycle-friendly routes, which serve to improve road safety. Although, Cambridge is the top bike-friendly city in the UK but cycling collisions unsurprisingly are greater than any other vehicles (The Cambridge News, 2018). And unfortunately, the achievement of cycling capital is on the decline now. The reality is that this reward has all been achieved without any real and enough cycling infrastructure for such a big cycling city (CambridgeshireLive, 2018b). Therefore, it is vital to deliver a better and inclusive cycling infrastructure which could be used specifically by cyclists to reduce the accidents and support the riders feel safer on the street. Accordingly, encourage more people to bike as well as keep the city's prominence as an advanced biking destination in the world (DfT, 2018b). As a result, there are still extra improvements which are essential and need to be applied with the purpose of boosting Cambridge's prominence as a global leading innovative riding city (DfT, 2018c).

1.3 Personal Injury Severity Prediction

Although improving geometric design of roads, traffic management systems, better road safety educations and sometimes enforcement running by police forces are helpful to reduce number of injuries, however, these traditional mitigation methods are not sufficient enough and often are not feasible or prohibitively are very expensive. The reality is that the road traffic accidents are usually occur by multiple reasons resulting from complex interactions. In this circumstance, influencing factors which are involved in accident remain nonlinear and complicatedly have profound effects on outcome of injury (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). And therefore, the traditional methods along with analysing police report data, are perhaps not enough to describe connotation between the factors. In this event, personal injury severity prediction (ISP) model is a key tool as it unveils the relationship between the injury severities and various explanatory variables, particularly, to clearly understand this relationship in detecting the influences in a much wider area.

1.3.1 Benefits of Injury Severity Prediction

The benefits of ISP are numerous. This kind of study is vital since professionals in roadway design, freeway management, public health, enforcement, emergency and trauma, policy, and education and awareness could benefit from the results to reduce the occurrence of injury and fatality crashes from different aspects. Moreover, the contributory factors which are reported by police officers can be subjective based on the officer's opinion rather than fact at the scene. Likewise, reliability and results of the personal ISP models have a vital meaning for improvement of road safety (TRL, 2010; Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

1.4 Research Objectives

This thesis tests the performance of machine learning techniques in modelling and predicting personal injury severities in road traffic accidents. Using STATS19 road safety data, this study explores the relationship between the severity of injuries and the contributing factors under different circumstances. Indeed, this study is looking for evidence being led by the data to detect the group most in need of intervention. This approach would lead us to look more at VRUs such as; pedestrian and cyclist groups rather than other groups (e.g. drivers) which already have benefited from many interventions.

Previous studies have highlighted a considerable increase in published research on driving (Abdelwahab and Abdel-Aty, 2001; Abdelwahab and Abdel-Aty, 2002; Abdel-Aty and Abdelwahab, 2004; Delen et al., 2006; Kim et al., 2007; Xie et al., 2007; Li et al., 2008; Karlaftis and Vlahogianni, 2011; Li et al., 2012; Zeng and Huang, 2014; Chen et al., 2016b; Sharma et al., 2016; Yu et al, 2016; Li et al., 2016; Alkheder et al., 2017; Iranitalab and Khattakb, 2017; Li et al, 2018; Zhang et al., 2018; Hasheminejad et al., 2018; SiamiDouadarn and Iscioglu, 2019; Siamidoudaran et al. 2019a; Venkat et al., 2019; Amiri et al., 2020; Pradhan and Sameen, 2020), however, walking and cycling related predictions are severely limited (Kim et al., 2007; Siddiqui et al., 2012; Wei and Lovegrove, 2013; Vandenbulcke et al., 2014; Osama and Sayed, 2016; Prati et al., 2017a; Prati et al., 2017b; Lee and Abdel-Aty, 2018; Guo et al., 2018; Zhai et al., 2018; Siamidoudaran et al., 2019b). Encouraging and enabling people to walk or cycle needs action on many fronts especially in terms of safety.

Although, COVID-19 has been great for cycling (BBC, 2020a; BBC, 2020b; PKC, 2020), there is urgent need for improvements since every year in the UK thousands of cycling accidents occur on the roads (RoSPA, 2017c; DfT 2018b). For that reason, main aim of this research is to examine an active travel intervention with a cycling focus in a city wherein cycling is the main mode of transport and is not merely just a nice activity of having fun during weekend (The Guardian, 2011; CityLab, 2015; DfT 2018b).

1.4.1 Prediction of Personal Injury Severities

Crash prediction models (CPM) have been very popular in road safety analysis and particular attention has been on modelling of them in the reviewed literature. However, the prediction of injury severities is seldom (Abdelwahab and Abdel-Aty, 2001; Abdelwahab and Abdel-Aty, 2002; Abdel-Aty and Abdelwahab, 2004; Delen et al., 2006; Chang and Wang, 2006; Li et al., 2012; Chen at el., 2016), if ever, the injury severity related outcome wasn't a common significant focus to detect contributory factors for personal injury severities (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In addition, none of those predictions have further investigated the severity of injuries individually for a real-life case study to identify a set of road safety intervention options for specific groups and to suggest appropriate remedial measures.

Based on this framework, numerous road safety studies have been conducted over the years aiming at recognising factors that may affect both the frequency and the severity of road traffic accidents or injuries. However, as advised by Savolainen et al. (2011), one has to be aware that the variables contributing accident frequency and accident severity may vary from the ones influencing the severity of injuries; therefore, this thesis aims at examining prediction of injury severities individually.

In addition, all the UK leading accident software systems such as; KeyACCIDENT, AccsMap, MAST, CrashMap, iMAAP etc. are all based on analysis of reported factors by police. However, the factors may not have been based on a wide-range of research and probably are only attributed to the police's subjective judgement (TRL, 2010). Likewise, to cover all these gaps and further explore contributory factors, this thesis predicts personal injury severity levels of two case studies; the City of London and Cambridge. This thesis is conducting six predictions by applying STATS19 road safety data which has never been predicted before by previous researchers. This study primarily aims to identify the group most in need of intervention so that findings can be focused on in more detail in a second case study.

1.4.2 Determination of Evidence-based Road Safety Interventions

Some road safety projects do have negative influences on road safety, despite the greatest aims of the road safety practitioners who designed them. Following the findings in both case studies, this thesis attempts to suggest evidence-based road safety engineering interventions to mitigate the identified poor road designs. In this vein, STATS19 casualty data is the best type of evidence that can be used to determine whether an intervention is needed (RoSPA, 2017a; DfT, 2020). In addition, by proposing educational interventions, this study outlines behavioural change theories intending to change road users' behaviour (Road Safety Scotland, 2020). In this field, unfortunately, relatively little assessment carries out in road safety thus it is so difficult to find evidence-based intervention related research and evaluation (RoSPA, 2017a; DfT, 2020). As a consequence, this thesis can be really valuable to support future works to design different interventions and predict potential effects.

1.4.3 Predictive Modelling with Big Data to Improve Accuracy

Enhancing a model performance can be challenging at times. In this connection, presence of more data always permits the 'data to direct for itself', in place of trusting on assumptions and poor associations which will help to reduce pain of working on limited data sets. As a result of the solid relationship between injury severities and related factors, using more data leads to getting higher prediction accuracy, otherwise the result of the prediction could be less accurate (Ray, 2015; Li et al., 2018; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a;

Siamidoudaran et al., 2019b; Hébert et al., 2019). In response to this, a large mass of input data applies to this thesis as an attempt to better understand the connection of independent variables with target variable that will for sure decrease the prediction error and result in better and correct predictions (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

1.4.4 Rank Analysis of Personal Injury Severity Predictors (RBFNN)

For the rank analysis, radial basis function neural network (RBFNN) with varying levels of success in previous studies addresses as an identification method for sensitive traffic injury severity predictors (Yan and Guang-si, 2008; Yu and Liu, 2010; Huang et al., 2016; Pradhan and Sameen, 2020). Within this framework, the RBFNN applies to this study to predict personal injury severity sustained by all road users (including driver, motorcyclist, cyclist and pedestrian) into different classes to detect several crash related factors and areas of concern. There is no personal injury severity related studies which focus on all road users in a single prediction task in order to detect the specific group to be focused on, in a next case study.

1.4.5 ISP of Group Most in Need of Intervention (Cyclist Group)

This section refers to the main concern of this thesis which needed to be focused on in more detail in a different case study. The results of the first case study showed that the city of London’s cycling boom needs more road safety intervention. Therefore, this thesis suggests that an intervention to reduce bicycle injuries among the STATS19 data is extremely important. Cycling in the UK is on the rise since it is a brilliant style to get about and delivers a wide range of environmental and health profits. Notably, the COVID-19 pandemic has produced many more cyclists as a result of people’s desire to avoid crowded public transport as well as follow the government’s full guidance for physical distancing (PKC, 2020). However, this mode of transport also holds a certain amount of danger, and so it is essential to explore the factors that affect the cycling related injuries (RoSPA, 2017c). In response to the detected concern in the first case study connected to the cyclist group, this thesis separately predicts cycling related injury severities. What's more is that the second case study took place in Cambridge, known as the best cycling destination in the UK (The Guardian, 2011; CityLab, 2015; DfT 2018b). The road safety problem is a persistent barrier in the UK's cycling capital that does lead to more cycling casualties compared to other groups (The Cambridge News, 2018; DfT 2018b; Siamidoudaran et al., 2019). Importantly, in response to the severely limited number of previous studies on prediction models in particular for pedal riders related injury severities (Kim et al., 2007; Siddiqui et al., 2012; Wei and Lovegrove, 2013; Vandenbulcke et al., 2014; Osama and Sayed, 2016; Prati et al., 2017a; Prati et al., 2017b; Lee and Abdel-Aty, 2018; Guo et al., 2018; Zhai et al., 2018; Siamidoudaran et al., 2019b), this thesis presents a different prediction task that applies particularly to cyclist group.

1.4.6 Maximising Predictive Accuracy Using Key Factors (MLPNN and SVM)

There are a scarce number of literature that applied a rank analysis to increase performance of ISP model (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In relation to this limitation, additional prediction applies to each case study using the most important contributory factors identified in the first prediction task to increase prediction performances. To achieve this goal, multilayer perceptron neural network (MLPNN) was used for the City of London case study and Cambridge related data was predicted through support of vector machine (SVM) network.

1.4.7 Additional Trials to Scrutinise Data (Hybrid MLPNN–SVM, and LVQNN)

This thesis attempts to apply a hybrid ANN–SVM, and an improved type of ANN to overcome data limitation as well as to select the best fit model for prediction of STATS19 data. For the first time ever, this thesis aims at using learning vector quantization neural network (LVQNN) and hybrid MLPNN–SVM for accident injury related studies (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

1.4.8 Comparison of MLPNN, SVM, Hybrid MLPNN– SVM, and LVQNN

A comparison of SVM and ANN always provides fruitful outcomes; therefore, this study also aims to compare the performance of the proposed MLPNN, SVM, Hybrid MLPNN–SVM, and LVQNN predictive models by applying the most sensitive predictors. Accordingly, this study performs a comparison between actual and predicted target values of the injury severity for each class including different levels to evaluate prediction accuracy from the proposed models obtained through each class. To end, this study evaluates the performance of the predictive models in different classes according to confusion matrix, accuracy, error, and sensitivity.

The techniques of this thesis can be used in identify major contributing factors or understand relationship between severity of injuries and explanatory accident variables. The findings of this thesis can play an important role in helping road casualty reduction and prevention targets as well as handling several road safety problems. It is hoped that the findings can shed some light on the potential of remedial measures in order to mitigate such severity of personal injuries resulting from accidents.

Chapter 2

LITERATURE REVIEW AND BACKGROUND

2.1 Advantage of ANN and SVM Over Statistical Methods

In recent years, researchers from multiple disciplines have been carried out various studies on CPMs. Using this technology, some models were built to predict road traffic accidents. The related models mainly are grouped into following methods such as; statistics and machine learning techniques in which ANN is an example of a machine learning method. In this vein, statistical models are the most traditional ones and they can explicitly demonstrate affects of observed crash related factors. However, they struggle while dealing in circumstances of outliers, missing or mass of noisy nonlinear dataset. Also, weak performance in using many separable factors accompanied by big numbers of subdivision data is one more disadvantage (Principe et al., 2000; Abdelwahab and Abdel-Aty, 2002; Kim et al., 2007; Li et al., 2008; Karlaftis and Vlahogianni, 2011; Tabachnick et al., 2012).

Despite the wide range of studies about the statistical models, many researchers have developed a serious of ANN and SVM models in order to deal with weakness of the statistical approaches. The models have shown to be more beneficial in working with massive amounts of multidimensional accident data and achieving better prediction accuracy in comparison with the statistical techniques (Abdelwahab and Abdel-Aty, 2002; Xie et al., 2007; Li et al., 2008; Li et al., 2012; Zeng and Huang, 2014; Li et al, 2018; Zhang et al., 2018).

2.2 Most Popular Classification Algorithms

In general, the first classification algorithm was presented by Fisher (1936). In this algorithm, minimising the classification error of train data is evaluated as an optimization criterion. This method has been used in many classification algorithms, yet there are some problems encountered mainly the generalization properties of the classifiers, which are not directly involved in the error function. The function measures the deviation of an observable value from a prediction. Indeed, the purpose is to discover values of model predictors for which returned number is as great as possible. Among these methods, MLPNN model is definitely the most frequently used by researchers. Another example is RBFNN which is popular type of feedforward neural networks. However, the determination of number of neurons in the hidden layers of MLPNN or the number of Gaussian functions in RBFNN is one of the most important and time consuming task because there may be concern of underfitting and overfitting caused by number of hidden layer neurons. SVM model is also another commonly method for supervised machine learning which is reviewed in this literature (Cortes and Vapnik, 1995). SVM is an elegant and powerful algorithm which is used for both classification and regression problems though it is frequently used to solve classification tasks. However, the faster SVMs may struggle when the data set has more noise.

2.3 Driving Related Predictions

In several studies, SVM models are used for driver injury severity prediction (Sharma et al., 2016; Yu et al, 2016; Li et al., 2016; Alkheder et al., 2017; SiamiDouadarn and Iscioglu, 2019; Venkat et al., 2020). An example of SVM refers to a comparison between SVM with an ordered probit model (OP). The result of the comparison between SVM and the proposed statistical model displayed that, the

SVM attained a superior performance in terms of prediction accuracy. The study also examined the potential of applying the SVM method for assessing the influences of external predictors on the severity of injuries. The sensitivity analysis outcomes showed that the SVM model made capable outcomes concerning the key influences. For numerous crash related factors, the outcomes of the SVM model were more reasonable than those found from the OP method (Li et al., 2012). Another SVM related study has been carried out for predicting MVCs (Li et al., 2008). The findings of SVMs have been compared with different types of negative binomial (NB) regression using traffic accident data obtained on rural frontage roads in Texas. The comparison demonstrated that SVM methods have shown their superior ability to predict accident data than other statistical models. Furthermore, SVM models were faster to implement than back-propagation neural network (BPNN) models. Their study suggested that using SVM models are more appropriate if the objective of the research involves predicting motor vehicle accidents.

Yu and Abdel- Aty (2014) used SVM model, random parameter logit model, and fixed parameter for performing accident injury severity predictions. Along with the data collected from accident reports, weather and real-time traffic data were also applied into the models. Injury severity of accidents was classified into severe and non-severe collisions. Accordingly, random forest model initially predicted to select key influences linked with severe accidents. Temperature, speed standard deviation, snow season indicator, and steep grade indicator were selected by the random forest method as inputs for analysing the data. For the aim of finding genuine associations between severe accident incident and the crash related factors as well as enhancing goodness-of-fit for the model, several methods were developed such as; fixed parameter logit model, random parameter logit model, and hybrid SVM–RBF to identify non-linearity. The comparison outcomes presented that both random parameter and SVM established better prediction accuracy. Additionally, the results of their research showed that real-time traffic and weather factors have substantial impacts on accident injury severity.

In most recent comparisons associated to SVM, Chen et al. (2016a) used this model with polynomial and gaussian RBF kernels to evaluate the performance of the proposed model. In addition, a classification and regression tree (CART) which is one of the most widely used data mining was applied with intention of detecting contributory factors. Consequently, the model predicted the driver injury severity levels on rollover accident relaying on two-year road safety data collected in New Mexico. In several studies, artificial intelligence (AI) and non-parametric approaches have been used to dominate the disadvantage of the traditional models. In this matter, Chang and Wang (2006), applied a classification and CART to observe the link between injury severity classes and the accident related factors. CART model has been verified to be an effective technique, mostly for dealing with prediction. They used one-year traffic crash data for Taipei, Taiwan. As a result, vehicle type was discovered as the main predictor connected to accident injury severity. All VRUs were recognised to have greater dangers of being injured than car drivers. The impacts of several contributory factors were observed with reference to accident and environmental factors, driver characters, and vehicle related factors. The results of a comparison indicated that that the SVM models verified superior prediction accuracy. The polynomial kernel showed an outperformance compare to the Gaussian RBF kernel. Explanatory variables related to comfortable driving environment circumstances, seatbelt used, alcohol or drug involvement by driver, travel lanes, driver demographic characters, vehicle damages in accident, location of collision, and time band were discovered as the main contributory factors.

Iranitalab and Khattakb (2017) used several statistical and machine learning models including SVM, random forests (RF), multinomial logit (MNL), and nearest neighbour classification (NNC) for ISP in two-vehicle accidents. The data was obtained from Nebraska, US between 2012–2015. The results of the comparison between statistical methods and machine leanings displayed that the NNC had the greatest prediction performance in overall in more serious accidents. RF and SVM models had the next two satisfactory performances and MNL was the poorest technique. Hasheminejad et al. (2018) used a hybrid clustering and classification method to examine the accident injury severity of rural network in Tehran, Iran between 2011-2013. For this purpose, a novel rule-based genetic algorithm (GA) was considered to predict the injury severities, which was assessed by performance measures in comparison with ANNs. The result indicated that GA model outperformed other classification in terms of prediction accuracy.

Chen et al. (2016b) used ordinal logistic regression to create several predictive methods on a probabilistic basis. The main aim of their study was to evaluate factors for drivers both under and not under alcohol involvement in China. The results indicated that several contributory factors were identified to be significantly linked with the injury severity, including accident partner and type of junction. Age band was detected as the main key influence under alcohol involvement. In other hand, collision pattern, junction type, light condition, sex of driver, and time band were detected as the most important key predictors connected with serious driver injuries involving motorists not under the impact of alcohol.

Another comparison related to ANN and statistical models was done by Xie et al. (2007). Bayesian neural network (BNN) models were used for predicting MVCs. To achieve this aim, a series of models was examined using data captured from rural networks in Texas. The models have been presented to accomplish superior than BPNN methods while at the same time decreasing the difficulty connected with over-fitting the data. Accordingly, BPNN, BNN, and the NB regression models were compared. Although the BPNN approach was able to deliver superior prediction performance compared with the BNN method, in most circumstances its prediction performance was inferior than the BNN method. Moreover, the data fitting performance of the BPNN approach was reliably poorer than the BNN model, which suggested that the BNN method had higher generalization facilities than the BPNN method and was able efficiently alleviate the over-fitting issue without particularly compromising the nonlinear estimate capability.

Abdelwahab and Abdel-Aty (2001) used MLPNN model and fuzzy adaptive resonance theory (ART) along with ordered logit methods for prediction mission. Accordingly, the association between several crash related factors and injury severity of motorist was observed covering driver, vehicle, roadway and environment characteristics. The analysis focuses on two-vehicle crashes that happened at signalised junctions. The outcome verified that MLPNN predicted the injury severity classes higher than other methods. Also, the outcome of the prediction displayed that rural junctions were more unsafe than urban junctions. Moreover, sex of driver, speed ratio, vehicle type, seatbelt used, and point of impact were very likely to contributed to a serious injury. A comparison was made between the models, showing a greater accuracy for the ANN model.

Again, after three years, Abdel-Aty and Abdelwahab (2004) used the same models in another study and compared the outcomes with the previous research. The objective of their research was to examine the capability and potential profits of applying the ANN to injury severity prediction. The outcome demonstrated that a higher accurate prediction ability for ANN model over statistical models. Likewise, the findings of the predictions proved that point of impact, sex, speed, area type, seatbelt, and vehicle type contributed to probability of injuries.

Delen et al. (2006) used a series of binary MLPs to investigate the potentially non-linear associations among the injury severity classes and the related factors. The injury severity levels classified into five categories (damage only, possible injury, minor non-incapacitating injury, incapacitating and fatality). Sensitivity analysis was applied on the models to recognise and rank the accident related variables as they applied to different injury severity classes. Using appropriate parameter selection, 17 factors were selected that mostly affect the injury classes of drivers. However, applying more injury severity levels accompanied by the outcomes of the prediction did not determine any enhanced results than other earlier researches. Another example can be Alkheder et al. (2017) prediction task which was focused on developing ANN to predict injury severity in traffic crashes recorded over a six-year in Abu Dhabi. Compared with an ordered probit method with a forecast accuracy of 59.5%, outcomes from the ANN made a much more accurate prediction rate by 74.6% accuracy. However, their obtained accuracy outcomes were slightly lower than the results achieved by other ANN related studies. Using more sensitive predictors, this thesis seeks to achieve higher accuracies for predictions results.

Among all existing machine learning methods, SVM model has been increasingly applied in different areas of road safety studies for example; incident detection (Yuan and Cheu, 2003), VRU detection (Cheng et al., 2005), lane changing detection (Mandalia and Salvucci, 2005). It has been verified by the researches that SVM successfully achieved more accurate outcomes compared with other existing methods. ANN and SVM models have been used to examine injury severity impact factors (Delen et al., 2017). Using a few comparison metrics in their study (e.g. accuracy, sensitivity), the proposed SVM obtained the best outcomes compared with the other existing methods in their study. Using SVM in this thesis, additional comparison metrics are used in order to scrutinize the model performance which are discussed in results section of this thesis.

Liz et al. (2012) used SVM for injury severity prediction and the accuracy result received by SVM was higher than ordered probit. In addition, Yu and Abdel-Aty (2014) compared the performance of SVM with fixed parameters and random parameters binary logit models. Again, the outcomes showed that SVM method perform superior than the fixed parameters model. Mokhtarimousavi et al., (2019) used SVMs for work zone crash injury severity prediction and the contributing factors by applying a parametric method. They used the mixed logit modelling structure and a non-parametric machine learning method applying SVM. The mixed logit model was fitted to the level of random parameter models in which the impacts of flexible factors across several observations were recognised, that is, data heterogeneity was considered. The results showed that SVM offers greater prediction accuracy and outperforms the other model.

As indicated in the literature, the accuracy obtained from different existing methods are clearly less than the ANN and SVM accuracies which highly encouraged us to apply these methods for prediction tasks in this thesis.

2.4 STATS19 Related Predictions

Previous literature displays that ANN and SVM models are very useful tools in road safety, given their potential for detecting crash frequency occurrence, and class severity of accidents and injuries. However, a careful and thorough reviewing of the literature shows that very few researchers (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b) have examined the UK related data in prediction works. Using data analytics method, a recent study extracted some information for preventing possible accidents in rural and urban areas of the UK. The study applied several methods such as; SVM, data integration, correlation machines, and multinomial goodness. The research mainly offered a new framework model which could be trained and adapt itself to new data and create correct predictions. Using SVM, the study attempted to shed light on the use of SVMs to improve road safety. However, the study used very limited crash related factors and also failed to focus more on the outcomes of predictions which refers to contributory factors. In addition, the examined factors were insignificant (e.g. vehicle make and model) and were not common factors identified by DfT and other related previous studies. Using data for a particular site or a specific group (e.g. urban or rural, accidents at junctions, driver or VRU group, head-on collisions, animal-related accidents) could provide more accurate outcomes than focusing on all reported accidents in both rural and urban areas in a single task (TRL, 2009; TRL, 2010; DfT, 2014; DfT, 2018a; DfT, 2018b; DfT, 2018c; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Rural areas are the opposite of urban areas and therefore, some common factors that generally contribute to accidents could be totally different. For example, majority of wildlife-vehicle collisions occur on low-volume, high-speed rural roads, which are also likely to be areas with high animal populations (Hughes and Amis; 1996; Taylor et al., 2002; Wilkins et al., 2019). In addition, no technique was used to analyse ranked data in order to select actual influences for predictions (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). As a final point, using SVM the predictions were not really satisfactory, therefore, the study suggested that the use of ANN might project more correct outcomes and create more optimum effects to attain better road safety (Lokala et al., 2019).

To overcome the gap identified in the UK data related studies, for the first time ever, Siamidoudaran et al. (2019a) used LVQNN to predict injury severity in traffic accidents classified into fatality, serious injury, slight injury, and only damage to property. Their case study focused on particular road safety problems in the city of London. Once again, Siamidoudaran and Iscioglu (2019) compared a series of machine learning models for same case study including; MLPNN, SVM, and combination of the two models. As a result of the performance evaluation, their proposed MLPNN-SVM model demonstrated superiority in predicting the injury severities. Several predictors discovered as contributory factors and significant accident cluster sites. A specific finding revealed that a number of road safety interventions are required for the cycling group which have not been widely focused on by previous researchers. In the intervention group, road safety education was the key intervention measure suggested in the evaluation of the intervention outcome. As a result of this specific finding, in this thesis we attempted to pay particular attention for cycling group. To this end, in the next section of the literature review, cycling related predictions and their associated factors were also reviewed. However, published articles which focus on this topic are still very limited (Kim et al., 2007; Siddiqui et al., 2012; Wei and Lovegrove, 2013; Vandenbulcke et al., 2014; Osama and Sayed, 2016; Prati et al., 2017a; Prati et al., 2017b; Lee and Abdel-Aty, 2018; Guo et al., 2018; Zhai et al., 2018; Siamidoudaran et al., 2019b).

2.5 Cycling Related Predictions

Considerable research has been carried out in recent years to establish associations between accidents and environmental influences, traffic flow, and elements relating to geometric road design. Injury severity prediction models focused on cycling, however, have rarely been examined. In addition, most research has paid but little attention to the safety effects of variables which STATS19 data specifically focuses on such as; crossing facilities, dazzling sun, bicycle and vehicle location, bicycle and vehicle manoeuvre, junction detail, junction location of bicycle and vehicle etc. (TRL, 2010; DfT, 2011; SiamiDoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Kim et al. (2007) developed a multinomial logit model for cycling related predication activities. For this purpose, they used the accident data between bikes and MVCs in order to predict the severity of injuries. The findings displayed that several factors had higher affect in increasing fatal accidents. The factors were included; speed of motorised vehicle, goods, weather condition, rush hour, head-on crash, alcohol related accidents, elderly riders, and light condition. Along with these factors, the results of several studies showed that the pedal riders were very likely to be at fault than motorists in events of serious and fatal accidents (Kim et al. 2007; Wei and Lovegrove 2013; Prati et al. 2017b). In addition, likelihood of injury for group of

VRUs increases in urban roads in the vicinity of schools (Kim et al. 2007). Another cyclist related prediction was carried out by Vandenbulcke et al. (2014) using Bayesian statistics. A number of contributing factors were discovered for instance; tramway systems road, lack of crossing services for pedal riders on bridges, riding in wrong direction, riding in heavy traffic, type of intersections, type of roundabouts, and car parks at shopping centres.

Prati et al. (2017a) used logistic regression analysis to examine association among different crash outcomes and contributory factors. Consequently, behaviour of drivers and riders were recognised as the key influences on generation of cycle – MVC’s. Other variables were related to gender and age of pedal riders. Once more, in the same year, Prati et al. (2017b) applied the Chi-squared automatic interaction detection (CHAID) decision tree model to prioritise the cycle crash related predictors. The factors refer to road type, crash type, cyclist age, road signage, sex of pedal rider, month, oncoming traffic, and segment classification. Thus, they applied Bayesian network analysis into same accident data and the outcomes indicated that crash type, road type, and opponent vehicle had higher affect in risk of cycle accidents.

Several road safety studies used macro-level models in their cycle related predictions (Siddiqui et al., 2012; Osama and Sayed, 2016; Lee and Abdel-Aty, 2018; Guo et al., 2018; Zhai et al., 2018). For instance, Siddiqui et al. (2012) used macro-level crash prediction models for pedestrian and bicycle related accidents in Florida. The important indexes were included numerous highway characteristics, and several demographic and socio-economic variables. Another alternative approach to incorporate spatial dependency in count models in the past has been to use a conditional autoregressive (CAR) or a joint prior on a spatial random effect term that is introduced multiplicatively in exponential form in the parameterization of the expected value of the discrete distribution for the count variable. The result showed that there might be major spatial correlations in accident event across different sites. A separate distinct set of key influences were identified for the model related to cycle accident. In all circumstances, the Bayesian approaches with spatial correlation achieved superior performance compared to other the techniques that did not account for spatial correlation between traffic analysis zones. The outcomes indicated that spatial correlation must be examined while modelling pedal rider related collisions at the aggregate or macro-level. Osama and Sayed (2016) developed macro-level for predicting bicycle accidents. The researches applied generalised linear regression and full Bayesian (FB) approaches. They attempted to examine the models with and without spatial effects. They assessed the association among many cycling routs related indicators and the accidents outcomes. Following the predictions, the outcomes showed that there is a significant connection among cycle network infrastructure and the cycle crash related influences.

Another macro-level related study carried out by Lee and Abdel-Aty (2018) in order to identify key influences in cycle accidents. A multivariate Bayesian PLN conditional autoregressive (CAR) model was used to recognise the indexes which contributed to bicycle collisions. As a result, accident hotspots for the target areas were identified based on the modelling outcomes. Using their study can help to recognise contributing factors for cycle related accidents. Cyclist exposure is one of the most important factors that contribute to risk of accidents in bicycle–MVCs. Due to lack of this factor in a cycle accident related study, the researches focused on zonal characteristic and applied proxies in order to overcome the limitation of the data. The outcomes of a recent study indicated that there were strong associations between cycle accident related factors and cycle safety at a zonal level. The contributory factors refer to demographic, cycle routes and accessibility. Additionally, the results showed that, junction's actions had a higher affect in likelihood of cyclist injury severity classes. In respect of this outcome, recently, numerous approaches of zonal configurations were developed applying a FB model. Accordingly, cycle related macro-level models were carried our using Poisson lognormal (PLN), random parameters PLN (RPPLN), random intercepts PLN (RIPLN), and spatial PLN (SPLN). The outcomes showed that bike collisions were positively connected with several factors such as, signal density, traffic exposure, high density commercial zone, and households. On the other hand, negative association among the accidents and the factors related to cycle infrastructure indicators. As a final point, the outcome of the comparison between the several approaches displayed that goodness of fit test for the SPLN model performed far better than others (Guo et al. 2018). Alternatively, Zhai et al. (2018) only applied a few factors to evaluate the prediction performance of zonal configuration associated to macro-level model. The related task was carried out within a FB technique applying multivariate PLN approaches and multivariate conditional auto-regressive priors. The outcome of their research showed that there were significant variations between several zonal configurations. As a result of this outcome, zonal configuration with greater zone achieved superior overall fit.

2.6 Cycling Related Prediction Using STATS19 Data

Most of the researches had been conducted to display the effect of using CPMs or ISP models. Moreover, very limited studies were applied on prediction of cycling related studies. Furthermore, none of the bicycling related predictions have been conducted in the UK, apart from one recent study (Siamidoudaran et al., 2019b). The research predicted cyclist injury severity classes using LVQNN. Following this, relationship between injury severity levels and cycling related factors were discovered. As a result, their proposed prediction technique was able to find several predictors influencing the injury severity of the riders involved in MVCs. As a significant finding, T and staggered junctions where dedicated right-turn were noticed as the worst case scenario. The researchers applied the most important contributory factors into the model to maximise the prediction accuracy. Accordingly, they outlined nature of cycling accidents and casualties and made valuable recommendations for making a safer riding environment while sharing the road. Their findings can significantly help to reduce the possibility of injury and improve cycling uptake, and support people who need to cycle, but are discouraged from doing so because they believe it is dangerous. As a final point, they discovered how driver attitude and behaviour towards each other can be improved through education based on a number of factors.

2.7 Comparative Literature of Various Models

This section of literature review discusses several comparisons of existing methods to find appropriate algorithms for the prediction tasks in this thesis.

2.7.1 Advantages and Disadvantages of Statistical Methods

In summary, the reviewed literature exposes that statistical models have the ability to identify several crash related factors. However, the mass of noisy data makes it very difficult to better understand the relationship between the factors. Additionally, poor performance in using many subdivision variables (e.g. STATS19 data) is another weakness (Tabachnick and Fidel, 2012). In respect of this issue, a number of researchers have notified the linearity and some distributions of error terms by the statistical methods. Thus, the applied statistical models fail when dealing with complex and very nonlinear datasets (Principe, 2000; Abdelwahab and Abdel-Aty, 2002; Li et al, 2008; Karlaftis and Vlahogianni, 2011; Tabachnick and Fidel, 2012). Therefore, using statistical models has been ignored from this thesis.

2.7.2 ANN and SVM to Overcome Weakness of Statistical Methods

In recent years, to overcome the disadvantage of the statistical models, many researches have also used machine learnings techniques instead of the statistical models due to their advantage of dealing with mass of noisy data like STATS19 data as well as superior predictive ability. As a result, those models especially, ANN and SVM models have showed to be more useful tool and have achieved a good model's fit accompanied by prediction accuracy in comparison to other existing algorithms (Abdelwahab and Abdel-Aty, 2002; Zhang, 2006; Xie, 2007; Li et al., 2008; Li et al., 2012; Zeng and Huang, 2014; Li et al., 2018; Zhang et al., 2018).

2.7.3 Two Well-known Architectures of ANN

Two well-known architectures of ANN; MLPNN and RBFNN have achieved varying degrees of success in reviewed literature and are shown to be very useful in road safety research (Abdelwahab and Abdel-Aty, 2001; Abdelwahab and Abdel-Aty, 2002; Abdel-Aty and Abdelwahab, 2004; Delen et al., 2006; Yan and Guang-si, 2008; Yu and Liu, 2010; Huang et al., 2016; SiamiDoudaran and Iscioglu, 2019; Pradhan and Sameen, 2020).

In the authors' previous application of machine learning in a road safety research, the MLPNN performed a little better than SVM for predicting drivers' injury severity (SiamiDoudaran and Iscioglu, 2019). This thesis examines a potentially stronger technique, the RBFNN to detect most import contributory factors (Abdelwahab, Abdel-Aty, 2002; Yan and Guang-si, 2008; Yu and Liu, 2010; Zeng and Huang

2014; Huang et al., 2016; Pradhan and Sameen, 2020). In this regard, recursive least squares (RLS) learning is a powerful learning method for RBFNN (Park and Sandberg, 1991; Chen, 1995; Yu and Liu, 2010). For that reason, this thesis aims at using the RLS learning algorithm for improving performance of RBFNN (Wang and Zhu, 2000).

The commonly used MLPNN also verified to be very capable when categorising the severity into different classes (Abdelwahab HT, Abdel-Aty, 2001; Abdel-Aty MA, Abdelwahab; 2004) like outputs of STATS19 data (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Additionally, in the reviewed literature this network was very successful when was addressed as an identification technique of most important factors (Delen, 2006; SiamiDouadarn, 2019). In this thesis, we also aim at maximising the MLPNN performance through applying most important predictors of injury severities. Many studies used series of MLPNNs for road safety related prediction tasks (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab, 2004; Delen, 2006; Alkheder et al., 2017; Shamsashtiany and Ameri, 2018), however, applying more injury classes along with the results of their predictions did not suggest any better solutions than other earlier researches. It is hoped using MLPNN along with applying several effective methods in methodology such as rank analysis and using more data can shed some new light on cause of personal injury severities as well as areas and groups of intervention (SiamiDoudaran and Iscioglu, 2019). This thesis applies Levenberg–Marquardt backpropagation (LM–BP) algorithm to attain superior performance of MLPNN through sensitive injury related factors (SimiDoudaran and Iscioglu, 2019). The LM–BP algorithm seems to be the fastest technique for the training process of the network. In addition, the LM–BP algorithm operates efficiently in matrix laboratory

(MATLAB), because the solution of the matrix equation is a built-in function, so its properties become even more capable in a MATLAB environment (Hagan and Menhaj, 1994). This algorithm blends the gradient descent and gauss–newton algorithm in training process of the network (Suratgar et al., 2005; Costa et al., 2007; Kwak et al., 2011; Yu and Wilamowski, 2011; Wondimagegnehu and Alemu, 2017). Using this strategy, this thesis is seeking to achieve better result from MLPNN compare with other existing algorithms by importing a larger mass of input data, applying different types of road user, analysis of rank ordered data, and maximising predictive accuracy using most important injury severity impact factors (SimiDoudaran and Iscioglu, 2019).

2.7.4 SVM as a Powerful Tool for Prediction

Literature relating to the SVM showed that this model also like ANNs has ability to predict severity of accidents or injuries within acceptable satisfactory. (Sharma et al., 2016; Yu et al, 2016; Li et al., 2016; Alkheder et al., 2017; SiamiDouadarn and Iscioglu, 2019; Venkat et al., 2020). The comparison results with different models indicated that, SVM model produced results with a better prediction accuracy in recognising the significant predictors (Li et al., 2008; Li et al., 2012; Yu et al., 2014; Yu et al, 2016; Zhang et al., 2018). Furthermore, SVMs were very successful in resolving the limitation of statistical methods (Chang and Wangm, 2006) to determine the relationship among several classes of accidents and reasons influencing crash severity. SVMs also verified that they work quite well when there is clear margin of separation among classes. SVM is also showed to be more effective in high dimensional spaces, it is also effective in cases where number of dimensions is higher than the number of examples. Therefore, we believe SVM can be very suitable for prediction of STATS19 due to the nature of data (SiamiDoudaran and Iscioglu, 2019). In this connection, kernel trick sounds like a perfect plan that we can consider to use for this thesis. Kernel trick is considered due to being able to bridge linearity and non-linearity. Using kernel function is a method used to take data as input and transform into the required form of processing data (Theodoridis, 2008; Murty and Raghava, 2016).

2.8 Decided Models Relying on Comparison of Existing Methods

Based on the literature review, there is an urgent need for studying safety of VRUs of STATS19 data in order to cover the specified gaps identified in the reviewed literature. The major objective of this thesis is to detect contributory factors, site clusters and groups of intervention through RBFNN. In addition, to apply the identified factors to MLPNN and SVM, hybrid MLPNN–SVM, and LVQNN in order to maximise prediction accuracies and find out which model is the best fit for STATS19 data, finally, to suggest evidence-based intervention options to mitigate the concerns.

2.8.1 RBFNN (Both Cases), MLPNN (City of London) and SVM (Cambridge)

In recent years ANN and SVM based methods are becoming very popular due to their good predictive performance. Indeed, relying on the reviewed literature, the first three models were chosen for the thesis due to the empirical analysis and being notoriously good at detecting nonlinearities (Abdelwahab and Abdel-Aty, 2001; Abdelwahab and Abdel-Aty, 2002; Abdel-Aty and Abdelwahab, 2004; Delen et al., 2006; Yan and Guang-si, 2008; Yu and Liu, 2010; Huang et al., 2016; SiamiDoudaran and Iscioglu, 2019; Pradhan and Sameen, 2020).

2.8.2 Hybrid MLPNN–SVM (City of London)

As a result of MLPNN and SVM success in gaining the targets in this thesis, this study presents a hybrid MLPNN–SVM model for prediction of the injury serveries.

This hybrid architecture is proposed to significantly improve the MLPNN performance (Bellili et al., 2003). Section 3 in this thesis details the idea of applying SVM in a hybrid combination architecture to develop the overall performance of MLPNN.

2.8.3 LVQNN (Cambridge)

As a powerful predictor, LVQNN is considered to examine the merged classes (KSI) to overcome the limitation of the data for ‘fatal’ and ‘serious injury’ classes. This model of ANN is a precursor to self-organising maps (SOM) that can be applied where there is labelled input data. As the value of the STATS19 data used in this thesis is label, this method is more suitable for predicting the injury severities in comparison to other types of ANNs. Previous related studies showed that LVQNN is also an accurate application for road safety data analysis (Priyono et al., 2005; Shen and Chen, 2009) as well as it successfully being used for predicting data with subdivision values (Chen and Marques, 2009) such as STATS19 (Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Therefore, using LVQNN, this thesis is investigating a model that fits the data better than other commonly used models (Priyono et al., 2005; Chen and Marques, 2009; Shen and Chen, 2009; Al-Daoud, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

In addition, rank analysis, large number of subdivision data, different kernels, activation functions and additional algorithms are considered to achieve higher percentage of correct predictions in test data as well as the models’ performance (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Importantly, rank analysis leads to select correct input selection for models which is a necessary step to assure the successfulness of the model performance achieving accurate prediction accuracy for the model's output (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b, Alizamir et al., 2020).

Chapter 3

METHODOLOGY

This study uses a series of ANN and SVM models for prediction of personal injury severities. For this purpose, two case studies were considered by applying STATS19 road safety data. The aim of the first case study is to identify the group most in need of road safety intervention by predicting personal injury severities suffered by all road users. Therefore, to overcome the identified concern in the first case study, the second case study specifically focuses on prediction of cyclist injury severities. Safety concerns are the main barrier to more cycling in the UK and importantly for the second case study (The Cambridge News, 2018; DfT 2018b; Siamidoudaran et al., 2019). Furthermore, there are extremely limited number of previous prediction studies for cyclist injury severity which is the main contribution of this thesis to the literature (Kim et al., 2007; Siddiqui et al., 2012; Wei and Lovegrove, 2013; Vandenbulcke et al., 2014; Osama and Sayed, 2016; Prati et al., 2017a; Prati et al., 2017b; Lee and Abdel-Aty, 2018; Guo et al., 2018; Zhai et al., 2018; Siamidoudaran et al., 2019b).

Using MLPNN and SVM models, this thesis also aims to increase the prediction accuracies by applying the most important injury severity impact factors. In addition, using hybrid MLPNN-SVM and LVQNN for the first time ever, the models attempt to maximise the accuracies as well as to overcome limitations of the data in predictive analytics (Priyono et al., 2005; Al-Daoud, 2009; Shen and Chen, 2009;

Chen and Marques, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Also, this thesis attempts to carry out different comparisons through applying the most sensitive predictors to evaluate the performance of the models (MLPNN, SVM, hybrid MLPNN-SVM, and LVQNN) in different injury severity classes. Finally, the study ends by suggesting evidence-based road safety intervention options to help reduce the identified concerns. Accordingly, the structure of this thesis and the relationship between stages are shown in the below flowcharts (Figures 5, 6, and 7) for each case study.

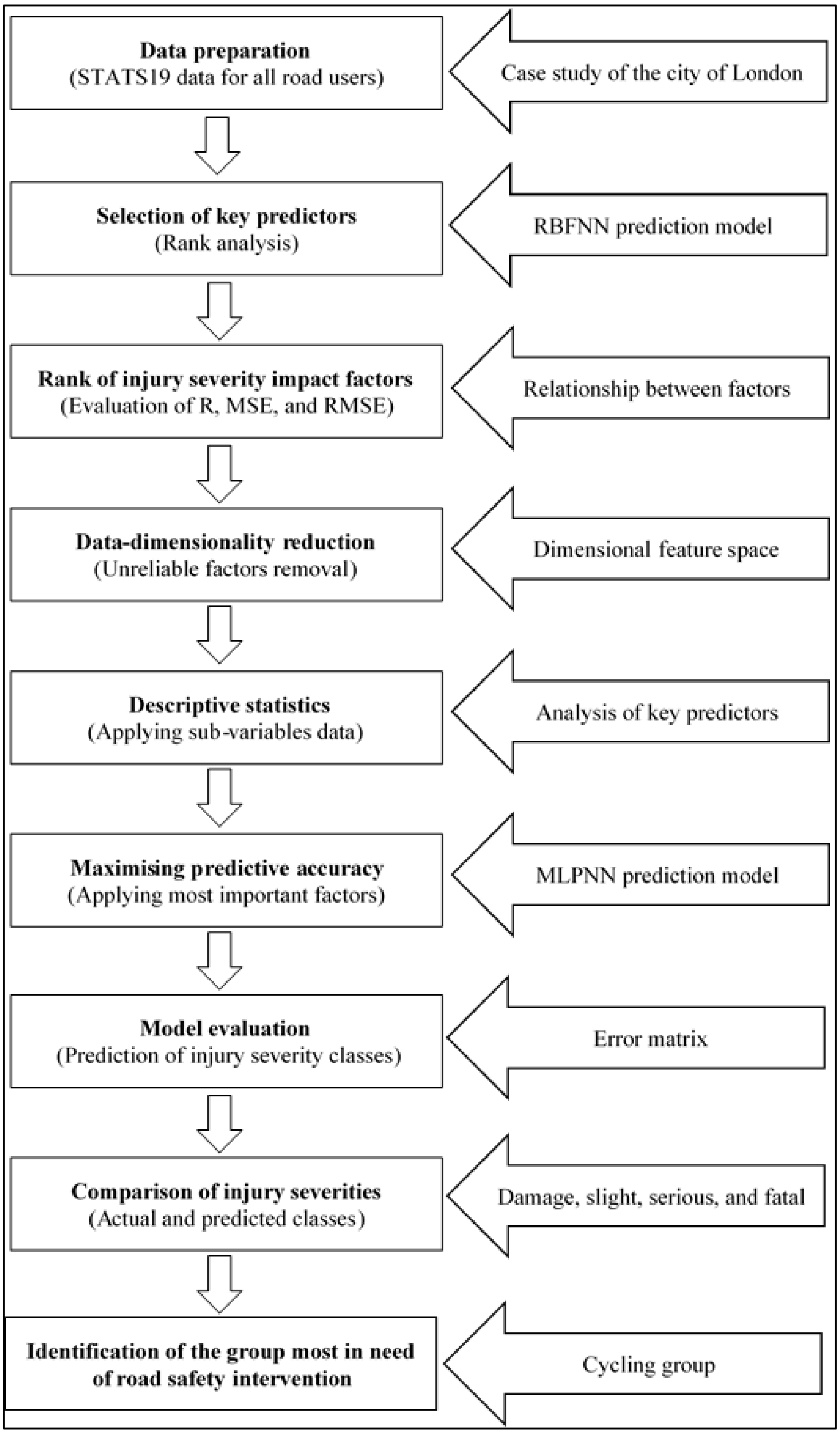

Figure 5: Flowchart Showing Structure of Thesis– First Case Study

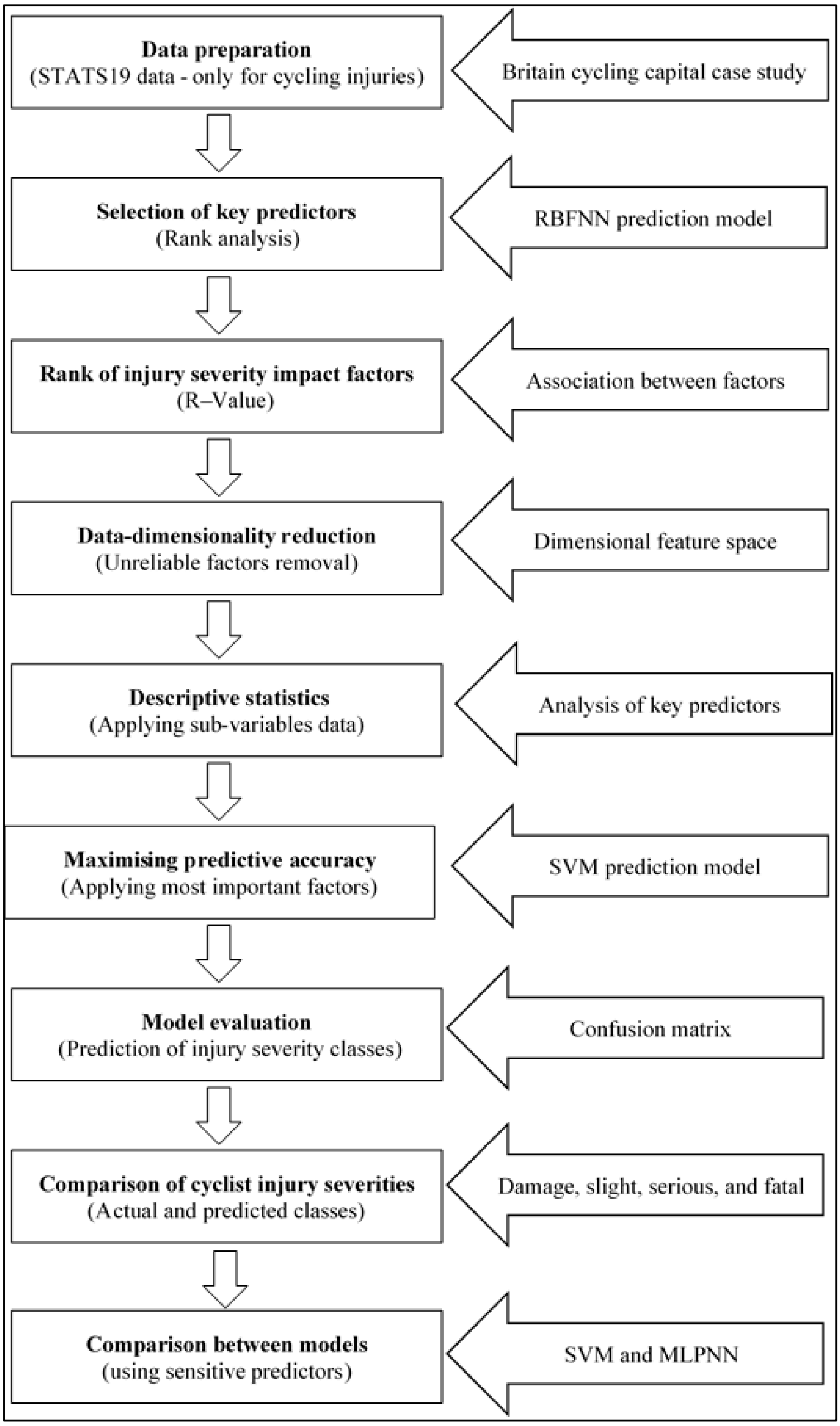

Figure 6: Flowchart Showing Structure of Thesis– Second Case Study

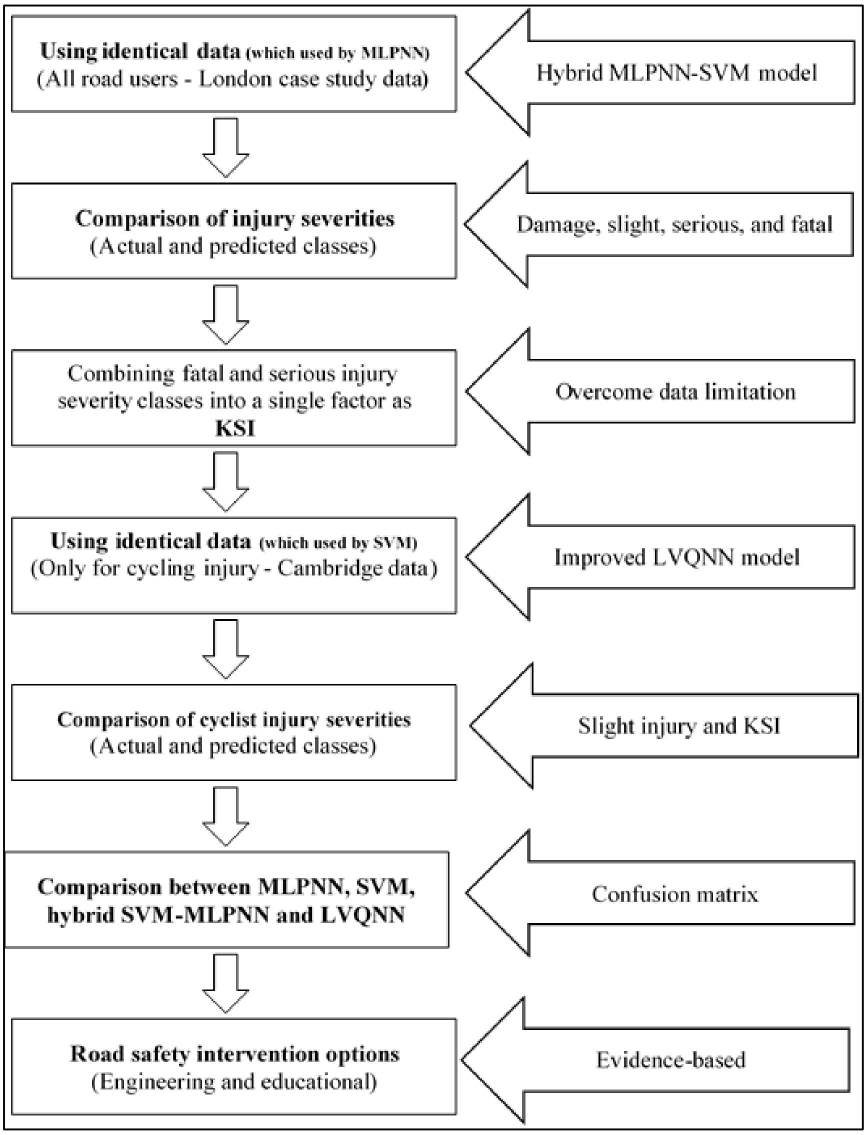

Figure 7: Flowchart Showing Structure of Thesis– Additional Trials on Data

3.1 Data Preparation

The focus of this thesis is to predict and analyse personal injury severities involving road casualties in Great Britain. This study conducts different predictions by applying STATS19 road safety data which has never been forecasted before by previous researchers. The name comes from a UK police form titled STATS19. The STATS19 database consisting of a set of all the collisions that caused in a personal injury, where, the accidents are informed to the police forces within 30 days of the crash. The data is obtained directly by the police at the roadside or when the crash is reported to a police station. The explanatory variables and data guide available in the dataset are defined by the DfT. Following this, the data is either sent to DfT or to the relevant local authorities. Some of the factors are sensitive (as is the breath test result variable) so the government doesn't make them publicly available and the related data was issued under an end user licence for this thesis. In addition, according to our agreement by DfT we were only able to publish the status as aggregations and not as specific accidents (or of a small set of accidents). This means that it was possible, for instance, publish the fact that "25% of the accidents on an M–class road had both 'turning manoeuvre' and 'inclement weather' as a contributory factor", but we were not able to point out "the accident that occurred on T junction 3 of the M6 on the 06/06/2012 had 'speeding' as a contributory factor". Therefore, it should be noted that identifying any specific circumstance or location was overlooked form this thesis.

3.1.1 Case Study Areas

The dataset along with the sensitive data used in this thesis was obtained officially from DFT and Cambridgeshire county council. The data includes different datasets of personal injury road accidents from 2007 to 2016 for the City of London and Cambridge city's total road networks.

3.1.1.1 The City of London Case Study

This case study focuses on developing personal injury severity prediction of all road users including driver, rider, pedal rider and pedestrian in the area situated in the City of London which is actually in the original London. The City is a local government district that covers the historic centre and the main central business district of London. It is now only a small part of the metropolis of London, though it remains a famous part of central London. Administratively, it forms one of the 33 local authority districts of London; however, the City of London is not a London borough.

The City of London is the smallest city in England with residential population of 8,000 people. However, over 500,000 people travel each day into the area for work and at the same time, more than 30 million tourists travel each year to the world's popular tourist destination (The City of London Corporation, 2020). Figure 8 details the locations where it is proposed to predict the injury severity levels (Google Maps, 2020).

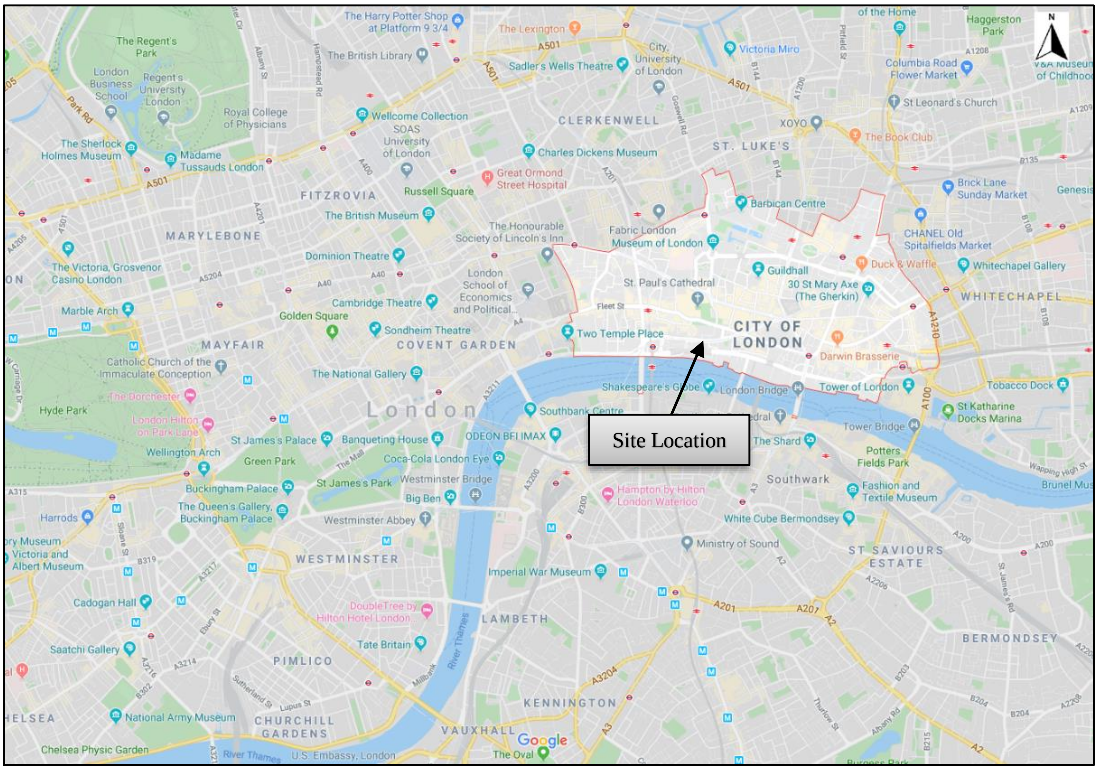

Figure 8: Site Location – City of London (Google Maps, 2020)

3.1.1.2 Cambridge Case Study

Cambridge case study focuses only on developing of pedal rider injury severity prediction. Cambridge is a university city that is located on the River Cam nearly 50 miles north of Greater London. Cambridge is the county town of Cambridgeshire council area which its population is around 129,000 as well as 25,000 students from all over the world. However, the population is expected to continue to increase but constant population growth with natural growth and development of the famous university of Cambridge. There is rather to Cambridge than a university, the city is the third most popular tourist destination in Britain for international visitors. There are many museums, extraordinary cultural sites, great places for adventure and family fun (World Population Review, 2020). The site location is enclosed by red colour in Figure 9 where it is proposed to predict the injury severity levels (Google Maps, 2020).

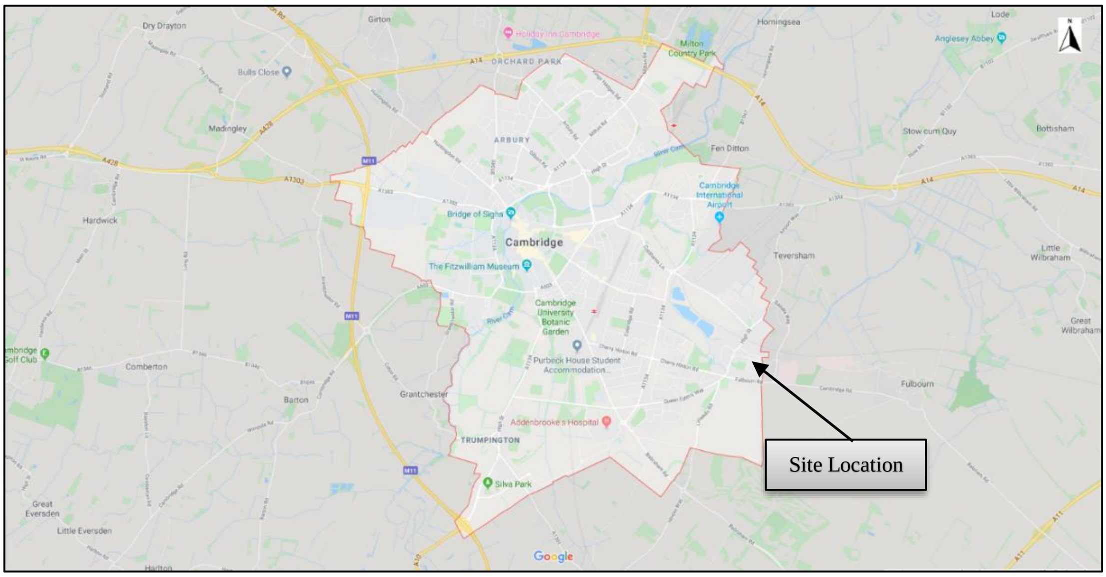

Figure 9: Site Location – Cambridge City (Google Maps, 2020)

3.1.2 Description of Data Variables

Three major variables in the dataset are included; collision circumstances, vehicle indexes, and casualty data. The variables detail all the information concerning crash types, locations, occurrence times, roadway geometric characteristics, weather conditions, vehicle characteristics, driver demographic, behaviour information etc. These variables were applied to the prediction models as input parameters and the injury severity classes were considered as output indexes. The information of each variable is given in Table 2.

Due to the large number of input factors within STATS19 data, merely the most important factors identified by the models are given in more detail in the results section of this thesis. In addition, the sub-variables (labels) of the input data are shown and analysed in more detail in Tables 5 and 6. Therefore, the detailed explanation of insignificant factors has not been provided in this thesis. However, there is another document called STATS20, which is a full detailed guidance in relation to STATS19 input data and parameters. The STATS20 document (DfT, 2011) aimed at providing instructions for the completion of traffic collisions reports and a detailed explanation of the information associated to the data used in this thesis (TRL, 2010; DfT, 2011).

Table 2: Implemented major input variables (DfT, 2011)

Major explanatory variables	
Collison circumstances (input)	**Vehicle related variables (input)**
Time	Skidding
Day	Vehicle type
Month	Overturning
Year	Engine capacity
Road type	Engine capacity
Speed limit	Junction location*
Police force	Sex of driver/rider
1st road class	Vehicle location
2nd road class*	Vehicle propulsion
1st Road Number	Vehicle manoeuvre
2st Road Number	Age of driver/rider
Junction detail	1st point of impact
Junction control*	Alcohol involvement
Lighting condition	Towing and articulation
Weather condition	Vehicle propulsion code
Urban/rural area	Journey purpose
Carriageway hazards	Hit object in carriageway
Number of vehicles	Age of vehicle (manufacture)
Numbers of casualties	Vehicle leaving carriageway
Road surface condition	Hit object off carriageway

Special conditions at site Pedestrian crossing– human control Pedestrian crossing – physical**	Was vehicle left hand drive Driver/rider home area type Driver IMD
Casualty class (output)	Personal injury severity class

* This index is employed when the collision is at junction.
** This parameter is applied for monitoring bicyclist and pedestrian's movement.

3.1.3 Data Presentation of Incident Severity Classes

The figures in this section cover official DfT statistics about reported injury severity levels in traffic accidents resulting from 250 accidents as an illustration. The incident outcomes extracted from CrashMap licensed to Agilysis (2020) along with Google Map (2020) for the background map. The related incident severity classes are simply viewed in different casually types in both case studies. The data only indicates to personal injury collisions on public roads, using the STATS19 collision reporting form. Data on damage-only crashes, with no human injuries or crash on unadopted roads or car parks are not involved.

Figure 10 refers to the 250 accidents involving young drivers resulting in 250 casualties of all road users as a sample in a small part of the city of London.

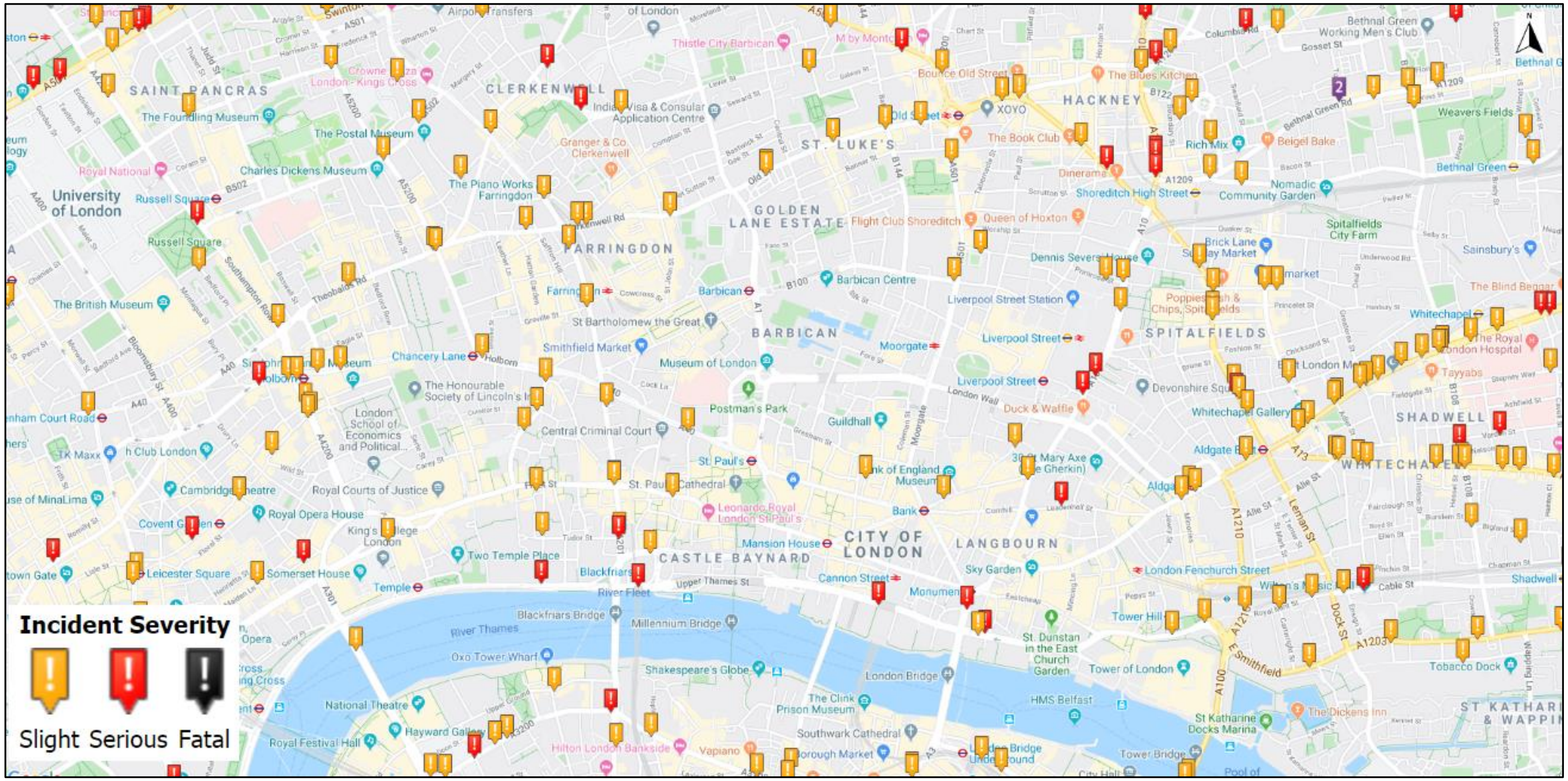

Figure 10:Incident Severity Levels by All Casualty Types – City of London (Google Maps, 2010; Agilysis, 2020)

Figure 11 refers to the 250 accidents involving young drivers resulting in 95 bicycle casualties for the same specified area. The result shows that cyclists made up around 40% of the injuries between the 250 accidents.

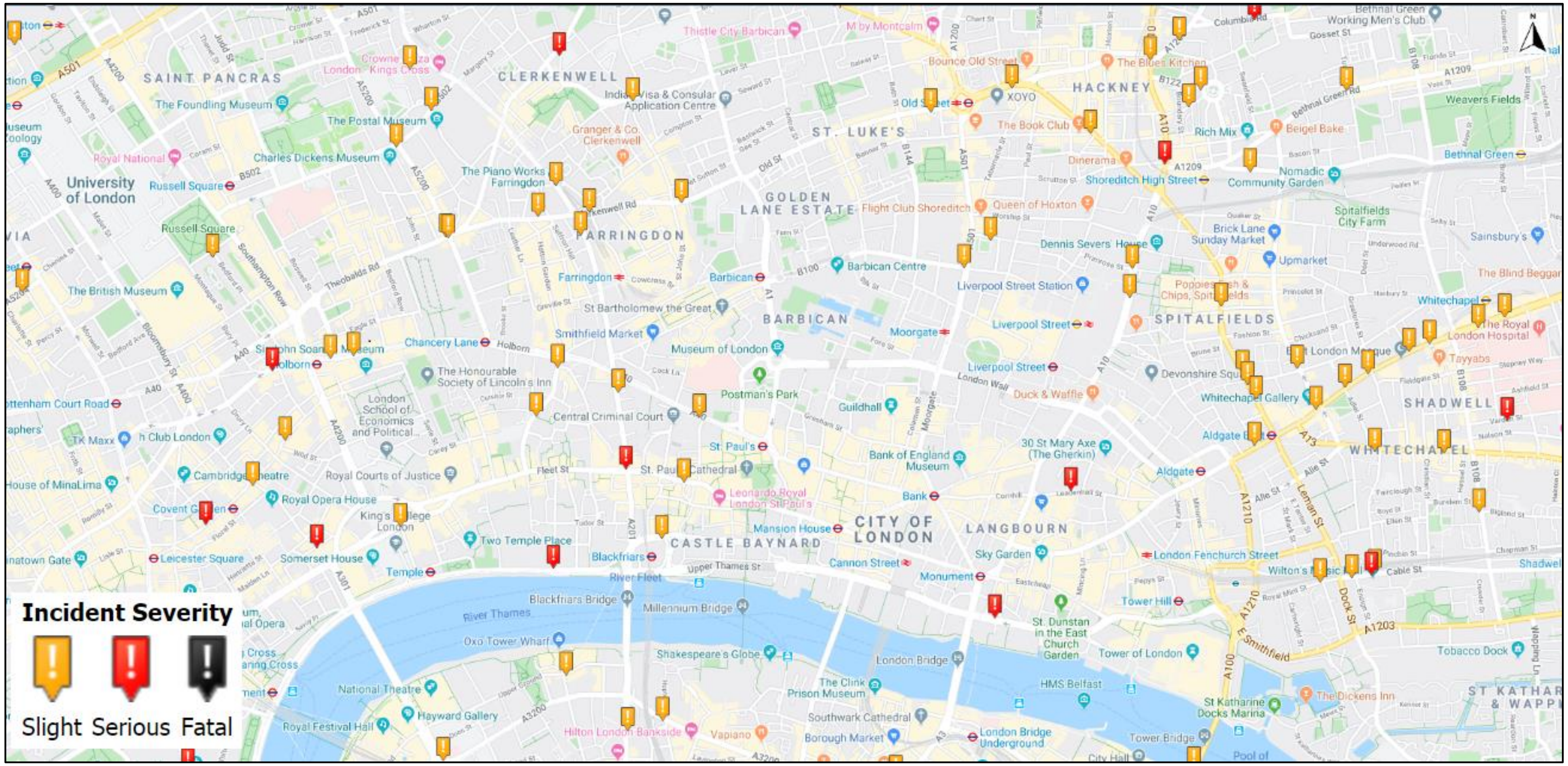

Figure 11: Incident Severity Levels by Cyclist Casualty Type – City of London (Google Maps, 2010; Agilysis, 2020)

Figure 12 refers to the same number involving young drivers resulting in 42 pedestrian casualties for the same site. The proportion of the pedestrian injuries is approximately 20% compere to the other road users which is considerable losses to individuals.

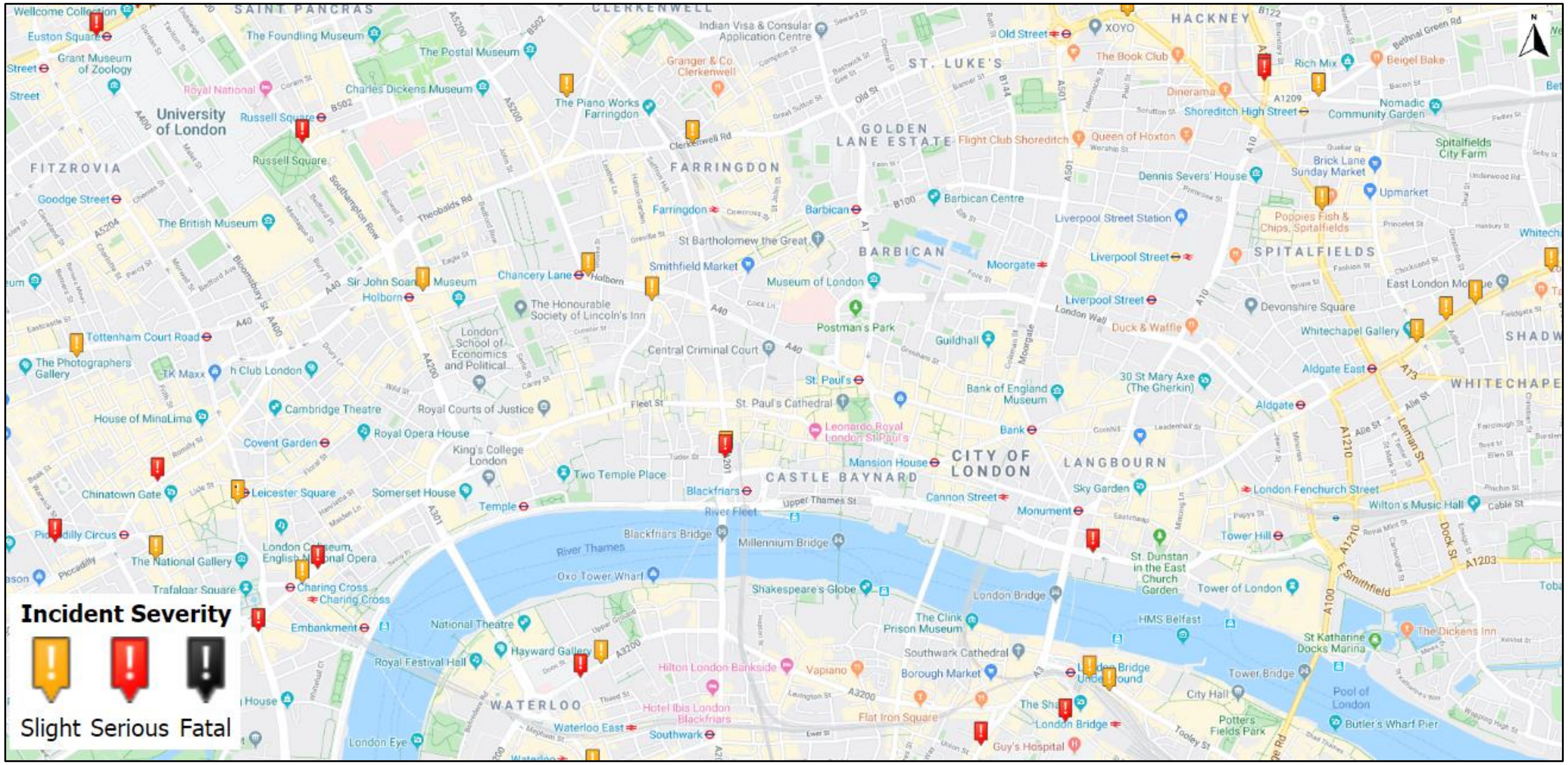

Figure 12: Incident Severity Levels by Pedestrian Casualty Type – City of London (Google Maps, 2010; Agilysis, 2020)

Figure 13 is 250 crashes involving all vehicle types resulting in 250 casualties in a small part of Cambridge as an illustration.

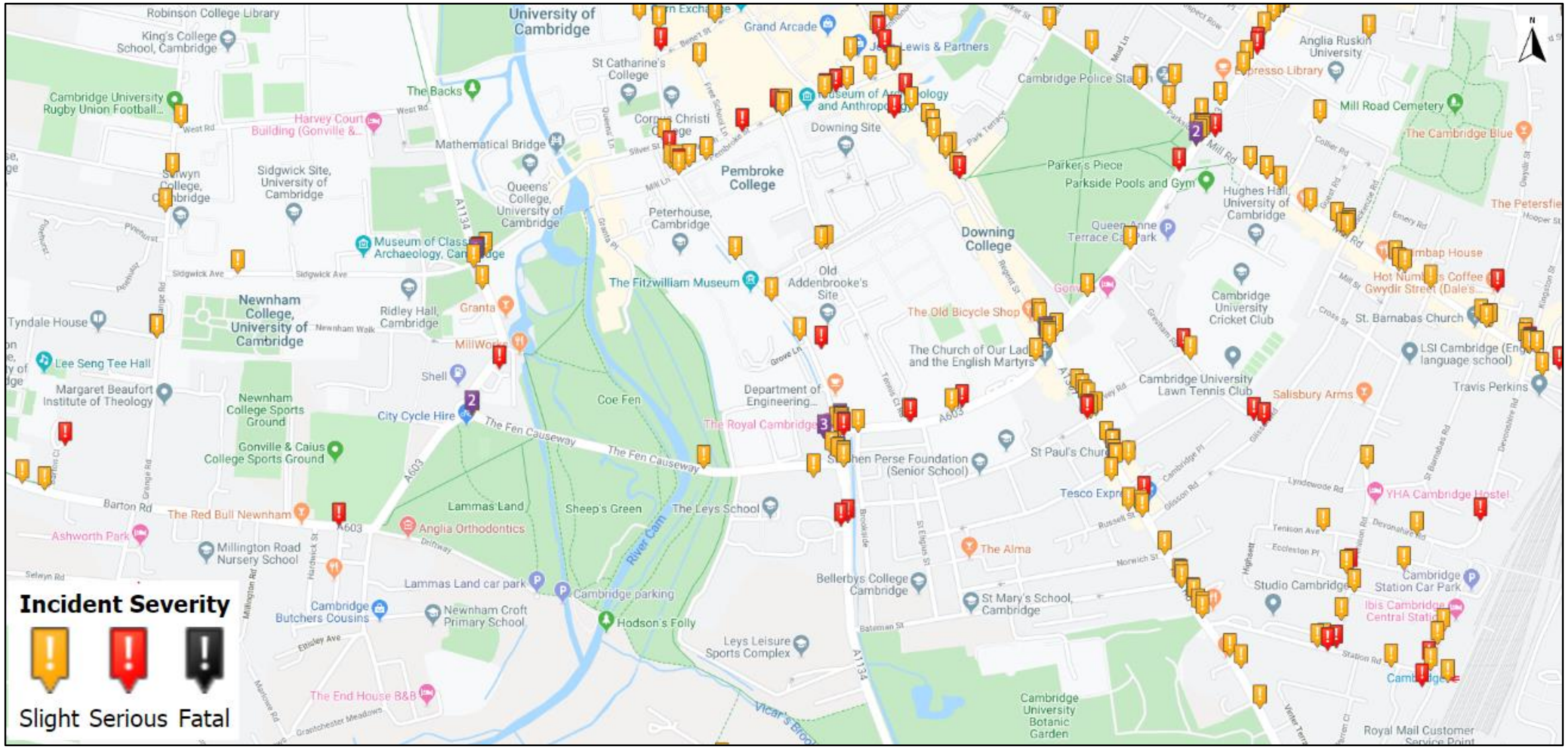

Figure 13: Incident Severity Levels by All Casualty Types – Cambridge (Google Maps, 2010; Agilysis, 2020)

Figure 14 refers to the 250 crashes involving all vehicle types resulting in 180 pedal rider casualties (more than 70%) in the same area. The majority of the injuries were caused by cyclists which is much more compared to the cycling injuries in the city of London.

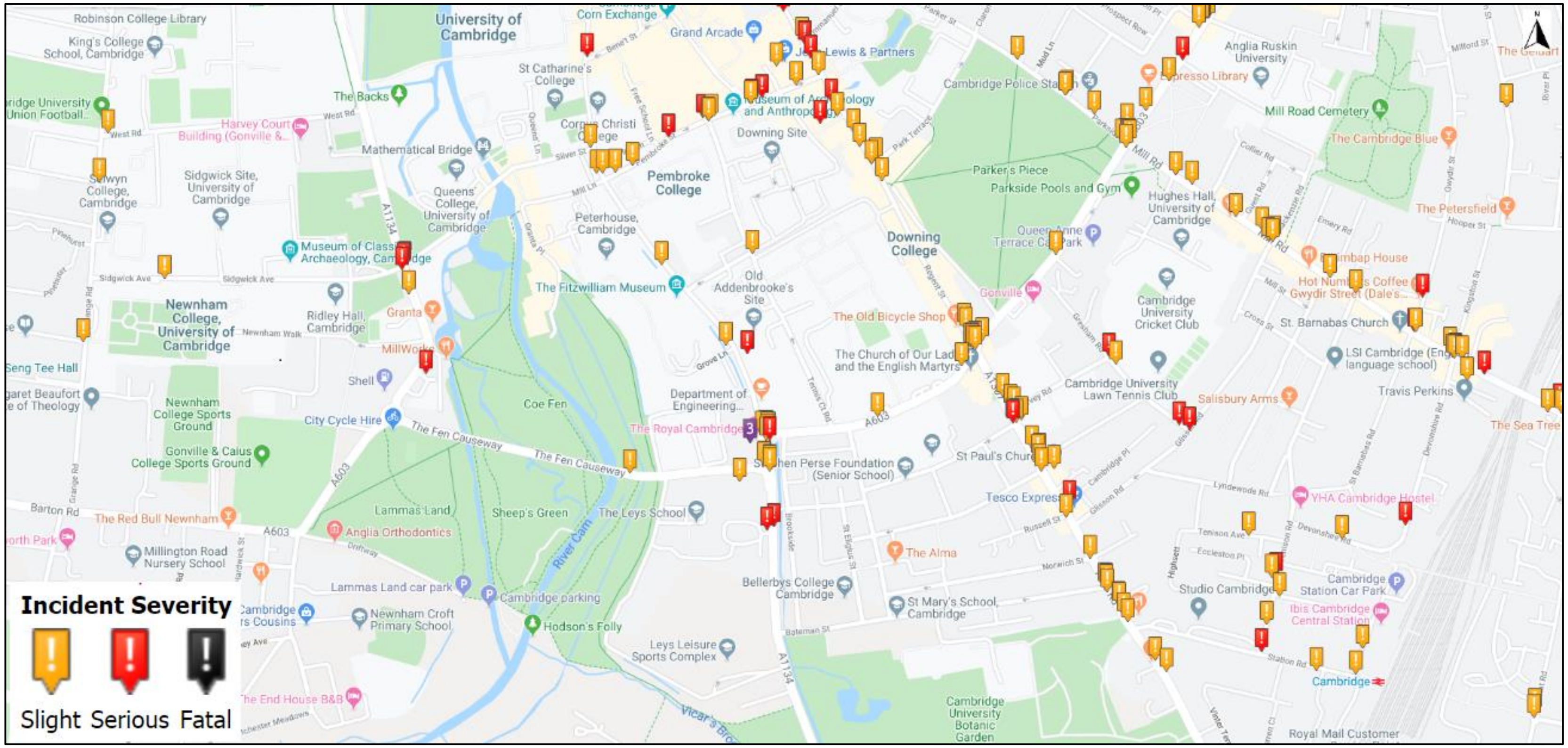

Figure 14: Incident Severity Levels by Cyclist Casualty Type – Cambridge (Google Maps, 2010; Agilysis, 2020)

Typically, the indication of the above figures simply gives an idea about the results of injuries caused by different road users. According to the data analysis as well the comparison between the incident severity levels, it is verified that Cambridge's pedal riders have more collisions than other vehicle types. Therefore, we only focus to pedal cycle casualties in Cambridge as they need more road safety intervention.

3.2 Why Proposed ISP Models Were Selected

Like CPM, ISP model is a mathematical model that defines associations between risk of personal injury severity and various road accident influences such as; environment variables, vehicle related factors, roadway geometric features, human's behaviours etc. Being one of the major steps of road safety, ISP model can provide crucial information to evaluate the severity level of injuries, estimate the potential impacts, and implement efficient accident management procedures. There is no accurate statistical model that can describe the association since the influencing factors are nonlinear. However, ISP model can mathematically describe the relationship between a set of independent variables and a dependent variable. In this regard, many road safety problems involving complex interrelationships can be efficiently solved using machine learning algorithms. There are numerous types of machine learning models that we can use for prediction tasks. For example, ANNs or SVM models can be very useful to determine general solutions for irrelevant data which then causes an extracting pattern for types of road safety problems. These types of models have many benefits but one of the most important of them is the fact that it can essentially learn from assessing data sets. The models apply as a random function approximation technique and support the prediction through the most efficient approaches for gaining more accurate results while describing distributions and computing functions. These models indicate a more accurate prediction capability over other traditional methods. The predictions models are measured as nonlinear data modelling techniques where the complex associations between inputs and outputs are used. (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

This thesis examines the viability and potential benefits of using a series of ANNs and an SVMNN model in predicting personal injury severities for number of road users. RBFNN, MLPNN, SVMNN were selected as the benchmark simply because of their popularity in injury severity modelling (Abdelwahab and Abdel-Aty, 2001; Abdelwahab and Abdel-Aty, 2002; Abdel-Aty and Abdelwahab, 2004; Delen et al., 2006; Yan and Guang-si, 2008; Yu and Liu, 2010; Huang et al., 2016; SiamiDoudaran and Iscioglu, 2019; Pradhan and Sameen, 2020).

In relation to Hybrid MLPNN-SVM, combining two powerful methods in a single model is a great idea to achieve better accuracy. Notably, it has never been used by other researchers in an accident/injury prediction (SiamiDoudaran and Iscioglu, 2019) and there are also few previous articles in different fields using this model for prediction tasks (Bishop, 1995; Bellili et al., 2003; Tifani et al., 2017).

Lastly, LVQNN has been identified as a more powerful model for prediction to overcome the limitation of data in connection to 'fatal' and 'serious' injury classes. Importantly, it is a very suitable model where there is labelled and subdivision data such as STATS19 data (Priyono et al., 2005; Shen and Chen, 2009; Chen and Marques, 2009; Al-Daoud, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In addition, it has never been used by other researchers in an accident/injury prediction (Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

3.2.1 Introduction to Artificial Neural Networks

ANNs are the most leading and powerful algorithms used in machine learning. Various fields of sciences such as engineering, medical science, mathematics etc. use ANNs for linear and non-linear regression, function approximation, classification and other technical and scientific applications. ANNs are mathematical models which is inspired by the human brain and learning rules to enhance existing data analysis tools. As the 'neural' part of their title proposes, they are brain-inspired methods which are aimed to reproduce the approach that we humans find out. However, much is still unclear regarding how the human brain trains itself to process a lot of information. Therefore, there are many concepts against this background. In the brain, a neuron gathers signals from others by the use of fine structures known as 'dendrites'. The neuron dispatch spikes of electrical action via a long, thin stand called an 'axon'. Likewise, axon ruptures into thousands of branches and after the completion of each branch, structure of 'synapse' transforms the movement from the axon into electrical influences that excite or inhibit task from the axon into electrical influences that excite or inhibit task in the associated neurons. When a neuron obtains stimulator input that is appropriately great in comparison to its inhibitory input, it directs a spike of electrical action down its axon. In this case, learning task happens through varying the validity of the synapses. as a result, the effect of one neuron on alternative neuron modifies (Hinton, 1992).

ANNs seem to be a modern development. However, it was created before the advent of computers, and has sustained at least one main setback and numerous eras. Several significant developments have been boosted through the use of low-cost computer simulations. After a preliminary period of eagerness, it continued a period of disappointment and dishonour. At the time, when professional support was insignificant, great progresses were done by quite a small number of scientists. These innovators were able to improve convincing approach which surpassed the limitations recognised by Minsky and Papert (1969). The researches defined the limitations of the neural network and summed up a wide-ranging feeling of obstruction against these models. Their finding was accepted by maximum devoid of additional examination. However, at the present time, the neural network technique enjoys a resurgence of interest. The initial artificial neuron was created in 1943 by American neurophysiologist Warren McCulloch and the logician Walter Pits who worked in the area of computational neuroscience. However, the lack of the technology at that time did not let them to perform too much works related to this field (McCulloch and Pitts, 1943).

ANNs have three layers that consist of input layer, output layer along with a hidden layer in most circumstances. Data that flows by the network influences the structure of the model as a neural network changes or learns, in a particular way referring to the input and output layers. The layers comprise of a number of interconnected 'nodes' which hold an 'activation function'. Patterns are offered to the model via the input layer and those neurons communicate to the second layer. Accordingly, they refer the data on to the hidden layer where the actual processing is completed via a method of weighted 'connections'. The second layer covers neurons that transform the input into a position that the output layer can apply. The second layers then connect to an 'output layer' where the answer is output as displayed in the graphic below. Figure 15 describes by the way wherein circles (neurons) are linked together by lines (synapses).

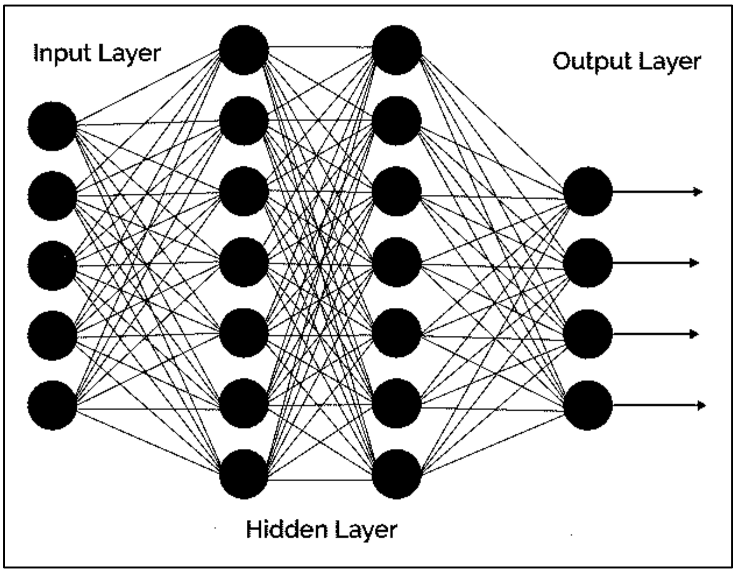

Figure 15: ANN – Neurones and Synapses (SiamiDoudaran and Iscioglu, 2019)

At first glance, ANNs may look like a black box. As seen in Figure 15 the input layer catches the data into the hidden layer and following a magic trick, the information is obtained by the output layer. Nonetheless, important phase of ANN application and optimization is to clearly understand the responsivity of the second layer. Learning rule of ANNs changes the weights of the network in relation to the input patterns that it is offered with. ANNs learn by instance as prepare their biological counterparts; a kid learns to recognise cats from examples of cats. While there are numerous different types of learning rules applied by AANs, this substantiation is interested only with the delta rule. This rule is frequently used by the most common level of ANNs titled BPNNs. Using backpropagation technique, ‘learning’ is a supervised process that arises with each epoch. Means that, in every turn, the network is offered with a new input form. This implementation is carried out via a forward activation flow connected with outputs, and the backwards error propagation for weight modifications. Accordingly, once an ANN is firstly denoted with a shape it creates a random ‘guess’ as to what it might be. It then gets how much further its response was from the actual one and creates a suiTable modification to its linking weights

(Schalkoff, 1997). Training an ANN involves choosing from allowed models for which there are several associated algorithms. They are usually nonparametric tools that are represented by connections between a very large number of simple computing processors or neurons, have been used for a variety of classification and regression problems. They also very useful techniques for discovering methods which are far too complex or busy for a human programmer to exploit and train the machine to identify.

There are many types of ANNs, but we shall concentrate briefly on our proposed RBFNN and MLPNN. This study also presents an SVM model in order to improve the average accuracy of the cycling related prediction. ANN and SVM are two typical classifiers and hold similar idea using linear learning method for recognition task. However, they are two different algorithms and the main difference is on how non-linear data is predicted. Mainly, SVM model uses non-linear mapping in order to build the data linear detachable, accordingly, the kernel function is the significant strategy. On the other hand, an ANN model works multi-layer association and numerous activation functions in the direction of dealing with nonlinear difficulties.

3.2.2 Proposed MLPNN Designed for ISP Model

MLPNN has applied several learning rules designed for training networks. This learning rule is one of the most widely used type of ANNs which has been proved as a universal predictor. An MLPNN is a class of feedforward used for function approximation tasks. It holds three layers comprising of input layer, hidden layer, and output layer. Associated data run from the first layer and pass through the second one to the third layer to build outputs. An MLPNN with one hidden layer is able to predict any finite nonlinear function with great accuracy (Schalkoff, 1997). The example below displays an MLPNN architecture for proposed traffic ISP.

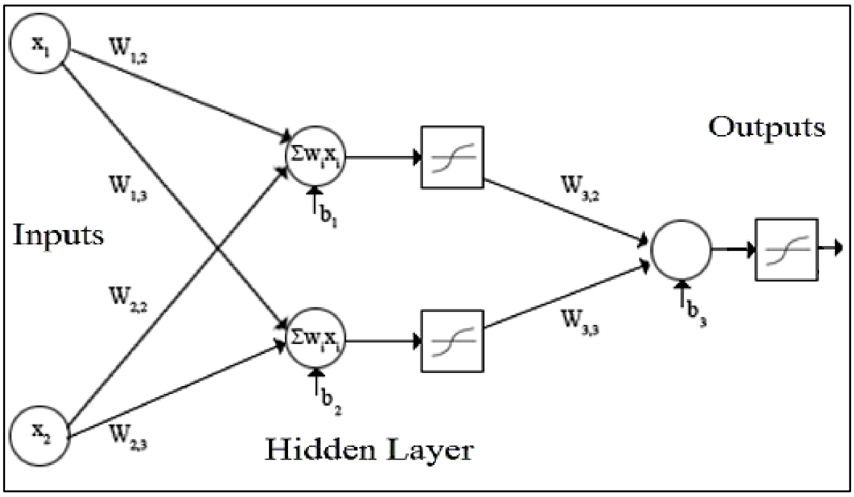

Figure 16: MLPNN Architecture (SiamiDoudaran and Iscioglu, 2019)

In Figure 16, each layer contains of neurons that are handling elements of network. Each neuron in any linked layer with the total of the next layer neurons via lines covered with coefficients named ‘weight coefficients’. Any variation in coefficients modifies the function of the model. Actually, the major aim of the network training is to establish the top weight coefficients to gain the preferred outcome. The outcomes recovered from the earlier layer are summarised with determined weights, specific for each neuron along with the bias term. The sum in a function named ‘activation function’. The function of a node describes the output of that node provided an input or group of inputs.

The proposed MLPNN used in this study applies three standard activation functions include; identity function, sigmoidal function, and Gaussian. The neurons of the MLPNN keep the equal functions, with the same free factors that are specified by user and are not altered through the training algorithms. Consequently, all the trained network acts when the feature vector is proceeded as input. In this case, the size of vector is the same as the input layer size. It then, passes values as input to the initial hidden layer. Finally, outputs of the hidden layer are completed applying the activation functions and the weights. Therefore, to figure the network, we must know all the weights. The weights are calculated by the training algorithm. The algorithm pulls a set of training, several input vectors with the corresponding output vectors, plus repetitively corrects the weights to allow the model to provide the chosen answer to the delivered input vectors. In some ways, to fix the weights and bias terms for learning set, a suitable algorithm is desirable, and it is straight depended on input data.

3.2.2.1 LM–BP Algorithm

In this thesis LM–BP algorithm was applied to achieve greatest performance for the prediction tasks. The LM–BP algorithm is a combination of the gradient descent and Gauss–Newton algorithm is used for training process of the network. The algorithm is examined on numerous function approximation problems; it is found that the algorithm is much more efficient than either of the other techniques when the network contains no more than a few hundred weights (Suratgar et al., 2005; Costa et al., 2007; Kwak et al., 2011; Yu and Wilamowski, 2011; Wondimagegnehu and Alemu, 2017; Siamidoudaran and Iscioglu, 2019). The LM–BP algorithm is more robust and trains ANNs at a rate of 10 to 100 times faster than the typical gradient descent backpropagation technique (Hagan and Menhaj, 1994). Training of the network automatically finishes when generalization ends improving, as showed by a rise in the mean square error (MSE). The algorithm is recognised as a technique of damped least-squares for reducing a function by using a numerical clarification. The process of the least squares is a system to fix the best fit line to data. In this thesis the BP algorithm was applied in the training that comprises of propagation and weight update. Hence, with the aim of performing this development, earlier layer data (front-propagation) was considered for calculation of the neuron's outputs for every single layer. Following this, in reference to the training outline target, the gradient of the weights was calculated via the difference between the target and the output. In conclusion, the weights of the layer are updated which is called 'weight update'. The value for every single neuron in the hidden layer is considered as below:

$$P_j = f\left(\sum_{i=1}^{A} x_i^T . w_{ij} + b_j\right) \quad (1)$$

In the above equation f is the activation function for hidden layer which is determined relaying on minimum test error. A denotes the number of neurons linked to input layer; x_i refers to the i^{th} model's input; interconnection among i^{th} neuron connected to input layer and j^{th} hidden layer neuron specified by w_{ij} . The bias term of the j^{th} is presented by b_j . Moreover, the value of every single neuron in the output layer is considered as following equation.

$$y_k = g\left(\sum_{j=1}^{B} p_j^T . w_{jk} + b_k\right) \quad (2)$$

In this equation activation function is shown by g for output layer which is called 'linear transfer function', B denotes the number the neuron connected to hidden layer, p_j is the amount of j^{th} the neuron associated to hidden layer. Moreover, Interconnection among j^{th} hidden layer neuron is presented through w_{jk} and k^{th} refers the neurons in output layer and b_k shows the bias term concerning k^{th} neurons in output layer (Suratgar et al., 2005; Costa et al., 2007; Kwak et al., 2011; Yu and Wilamowski, 2011; Wondimagegnehu and Alemu, 2017; Siamidoudaran and Iscioglu, 2019).

3.2.3 Proposed RBFNN Designed for ISP Model

In the field of mathematical modelling, a RBF network is an ANN that uses RBFs as activation functions. Likewise, the indication of RBFNN originates from the concept of function approximation. As previously explained in the MLPNN section, MLPNN with a hidden layer of sigmoidal units is able to learn in the direction of approximate functions. In this vein, RBFNN also acts as an insignificantly different method. An individual neuron in MLPNN model takes the weighted total of its input factors. Specifically, each factor is multiplied through a coefficient, and as a result, the outcomes are all summed together. A neuron refers to a simple linear predictor, but complex non-linear predictor can be made through merging the neurons accords a network. From this perspective, RBFNN seems to be an intuitive method compared to MLPNN model.

Figure 17 simply displays the typical structure of an RBFNN model developed to predict injury severity classes. As seen in the figure the model is a special form of the three-layer feedforward ANNs. The structure of RBFNN comprises of an input vector, the second layer refers to the model' neurons, and an output layer holding one node for each category of data. This model employs radial basis functions as activation functions and is very useful model for classification and prediction task.

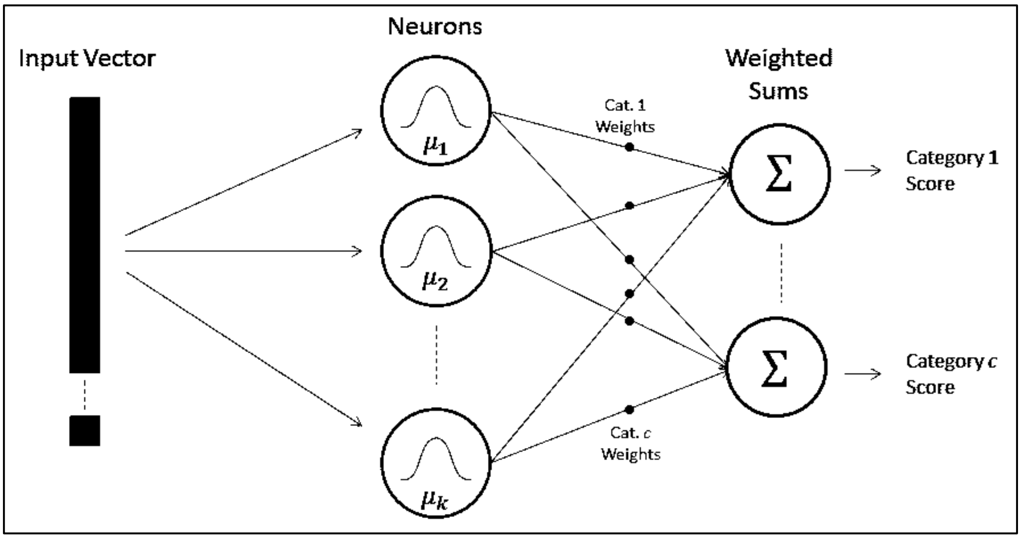

Figure 17: RBFNN Structure (McCormick, 2013)

In Figure 17, the first layer is the(n)–dimensional vector which is accounTable for classification task. The total vector is presented to each neuron in the above figure. In the figure μ_k is the prototype which denotes the neuron's centre. The model makes prediction through evaluating the input's similarity to samples from the training data. Each RBFNN neuron stocks a prototype which is one of the instances from the training data. When a new input factor is classified, each neuron figures the Euclidean space among the input factors and its archetype. Likewise, a comparison is applied by each neuron among the input vector and the prototype. As a result, output value refers among 0 and 1 that is kind of similarity measure. If the input is equivalent to the neuron's centre, then the output will be 1. On the other hand, if the space among the input and the centre raises, the answer drops off to 0. Likewise, if the input factor is more similar to the category 1 examples than the category 2 examples, it is predicted as level 1. The figure of the model neuron's reaction is a bell curve which is activation value. There are several potential sets of similarity functions. However, in this thesis the Gaussian was used which is the most common method.

The output comprises of a set of nodes which will result in prediction for each class. An output calculates a sort of mark for the connected class. Generally, a prediction result is completed by allocating the input to the class with the maximum mark. The mark is assessed by taking a weighted figure of the activation function from each neuron. By means of weighted sum, an output node links a weight assessment with each of the model neurons and grows the neuron's activation via the weight before adding it to the entire reaction. For the reason that an individual output node is figuring the mark for a different class, each output node takes its own set of weights. The output node normally provides a positive weight to the model neurons which is part of its class, and a negative weight in the direction of the rest (Park and Sandberg, 1991; Chen, 1995; Yu and Liu, 2010). The below network architecture typically explains RBFNN nomenclature.

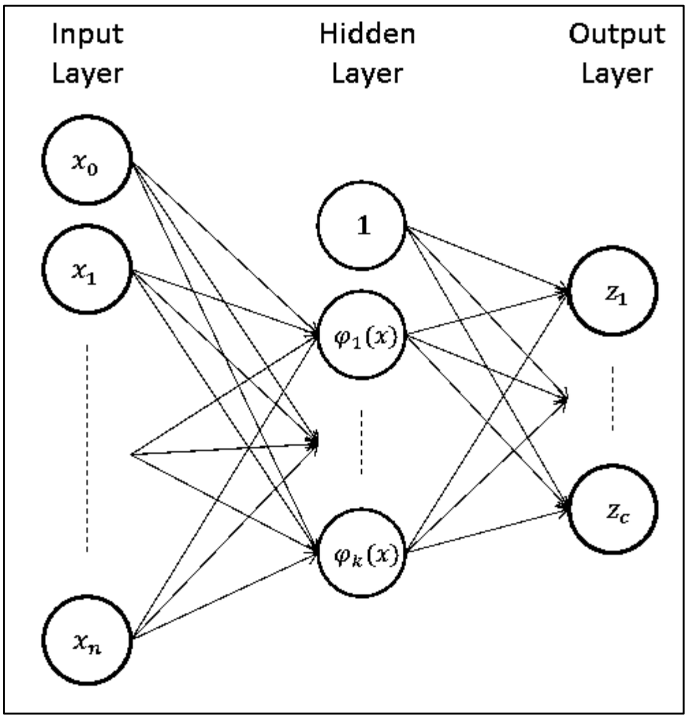

Figure 18: Structure of RBFNN Nomenclature (McCormick, 2013)

Figure 18 illustration simply displays that the model is intended to primary implement nonlinear mappings from are mapped from the input space into the hidden space. Nonlinear function is referred to as RBF. In the next circumstance, patterns mapped from the hidden space toward the output space through a linear function (Broomhead and Lowe, 1988a; Broomhead and Lowe, 1988b; Haykin, 2009). Below is the equation for the Gaussian function which is generally applied in structure of all the RBFNN for the hidden layer.

$$\phi_k(\mathrm{x}) = \exp = \left(-\frac{1}{2\sigma_1^2}\|\mathrm{x} - \mu_k\|^2\right), \quad k = 1, 2, 3, \ldots, K \tag{3}$$

In the above equation, x is input and μ_k denotes the centre of the network $\phi_k(\cdot)(k = 1, 2, \ldots, K)$. $\|\mathrm{x} - \mu_k\|$ is the Euclidean distance among x and μ_k. In this connection, Euclidian norm which refers for every single vector the length of its arrow and is recognised as the magnitude. Symbol σ_1 is the spread of function and manages the smooth function approximation. The basis refers via w_0 and the connection weight among the output node and the network defines by $w_k(k = 1, 2, \ldots, K)$. As a result, the predicted injury severity of driver or rider (ψ) is specified as below (Huang et al., 2016).

$$\psi = w_0 + \sum_{k=1}^{k} \left(w_k \exp \frac{1}{2\sigma^2} \|x - \mu_k\|^2 \right) \tag{4}$$

The training and leaning processes, RBFNN is too fast and the model is achieved wonderful performance at interpolation. The structure of training section is allocated into two phases; preliminary the weights from the first to second layer are established and the next stage refers the weights from the second to the last layer of neurons that creates given outputs for the database (Haykin 2009).

The learning and training technique of RBFNN is used to estimate the association among input factors and the injury severity classes. Initially, the centres of the prediction model are specified through a K–means clustering. It is a technique of vector quantization, initially from signal processing, that goals to partition n observations into kclusters in which each observation belongs to the cluster with the closest mean , allocating as a prototype of the cluster. Following this, an RLS algorithm is used for prediction of the basis and weights among the output node and RBFs (Chen, 1995; Wang and Zhu, 2000; Abdelwahab, Abdel-Aty, 2002; Zeng and Huang 2014).

3.2.4 Proposed SVM Designed for ISP Model

In machine learning, SVMs are supervised learning algorithms that analyse data used for both binary regression and classification. Like ANNs models, SVMs are also one of the most popular applications used for prediction of crash injury severities

(Siamidoudaran and Iscioglu, 2019). They are discriminative classifier and important characteristic of this technique is the power to mitigate the classification errors. This classification method is introduced by Cortes and Vapnik (1995). SVM reduces the operational hazard as an objective function in place of decreasing the classification fault and this advantage is the main difference of this model with other algorithms.

An SVM creates an optimal hyperplane as a decision surface such that the largest margin of separation among the two levels is improved. The model performs prediction task through discovering the hyperplane and the hyperplane maximizes the margin among the levels. The vectors that describe the hyperplane stand as the support vectors. The model makes subset of the training observations that are applied as support used for the optimal location of the decision surface. Actually, the model efforts to detect a separating hyperplane by decreasing the space of misclassified points to the decision boundary. In order to fit the data into the SVM model, a nonlinear mapping is applied to transfer the input factors from the primal space to the higher dimensional feature space. This action supports to discover the appropriate hyperplane and extended to solve multi-class issues.

Figure 19 typically shows the structure of an SVM developed to predict injury severity classes.

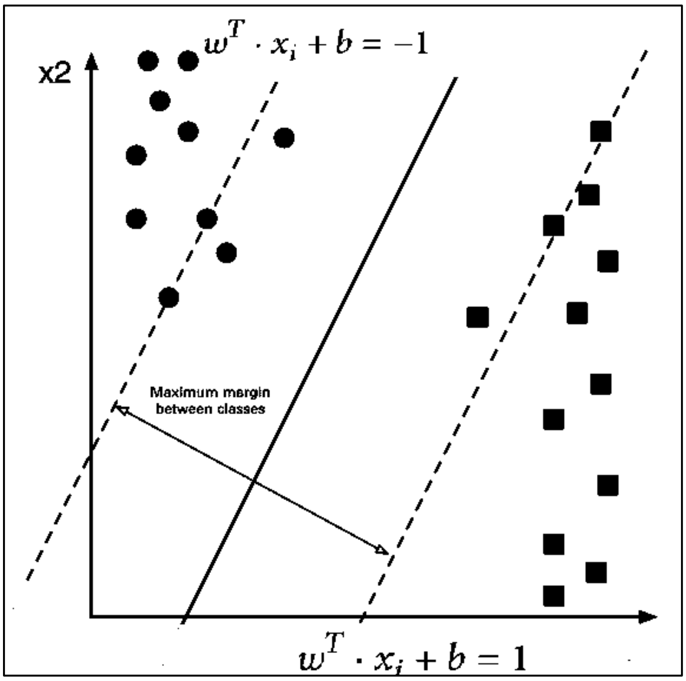

Figure 19: SVM Architecture (SiamiDoudaran and Iscioglu, 2019)

Using SVM model in predicting traffic injury severity classes is basically a binary classification and associated training data are given as $(x_1, y_1), \ldots, (x_i, y_i)$ where $xi \epsilon Rd$ and $yi \epsilon \{+1, -1\}$. The methodology is applied using a hyperplane to separate the data from one dimension to high dimensional space plus the support vectors which refer to the points lying on the boundaries. Accordingly, the two-dimensional space, a line typically divides the associated levels in middle of the margin. The middle of the margin is a discriminator and the margin of separation is maximised through this method. Margin of separation refers to the separation among the hyperplane and the nearest data point for bias and an assumed weight vector. On the other hand, in the multidimensional spaces circumstances which hold more than three dimensions, a hyperplane separates the levels as usual and the SVM classier function is assumed as below.

$$\begin{cases} w^T \cdot x_i + b \geq 1 \quad , & if y_i = 1 \quad i = 1, 2, \ldots, n \\ w^T \cdot x_i + b \leq -1 \, , & if y_i = -1 \quad i = 1, 2, \ldots, n \end{cases} \tag{5}$$

In this case, in order to attain the required discrimination, the margins among the levels must be the greatest. For that reason, by assessing the space among the support hyperplanes and the obtained margin is considered as $\frac{1}{2}\|w\|^2$. Nonetheless, in actual fact, the data in this thesis was not regularly separated and the data slightly took place in multiple datasets. As a result of this interaction, a hyperplane is achieved in relation to the minimum number of errors. Related indexes of each level are identified by examining the space from its own level which denotes by (δ) borderline. This technique is achieved using the strategy of soft margin formulation. This strategy allows the model to create a specific number of errors and retain margin as wide as possible ensure that other points are classified properly. Non-negative parameters (δ_i) is measured as inactive indexes appertaining to (s.t.) $\delta_i \geq 0$. Accordingly, the primal problem is calculated as following equations.

Minimize $\frac{1}{2}w^T \cdot w$

s.t.:

$$y_i(w^T \cdot x_i + b) \geq 1 - \delta_i \qquad i = 1, 2, \dots, n \tag{6}$$

$\delta_i \geq 0$

so,

$$L_p = \frac{1}{2}w^T \cdot w - \sum_{i=1}^{n}[y_i(w^T \cdot x_i + b) - 1 + \delta_i] \qquad i = 1, 2, \dots, n \tag{7}$$

The main requirement in this condition is to answer with a quadratic issue that desires some struggles to be fixed. In the meantime, input vector predictors are not the only indexes to be measured. Specifically, the problem reasonably requests to allow for a number of other indexes. Thus, the original formula modifies a dual form through applying Lagrange multipliers (α_i, μ_i). Hereon, the Lagrange coefficients (α_i, μ_i) are fixed to be non-negative real predictors, the equation is reformed in this way.

$$L_p = \frac{1}{2} w^T \cdot w + C \sum_{\mathrm{i}=1}^{\mathbf{n}} \delta_i - \sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}} \left[y_i (w^T \cdot x_i + b) - 1 + \delta_i \right] - \sum_{\mathrm{i}=1}^{n} \mu_i \cdot \delta_{\mathrm{i}} \qquad (8)$$

$\alpha_{\mathrm{i}}, \mu_{\mathrm{i}} \geq 0$

Where C is reflected to be the fine cause and L_p is a saddle factor, for that reason, the bottom volume must be used by using the factors w, b and δ. At this rate, the maximum amount wishes to be set by examining the Lagrange multipliers. With the aim of shifting the major problem to a maximized problem, the partial derivative of w, b and δ requests to be assessed. Following this induction, the dual problem is assumed as below.

$$L_D = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \cdot \alpha_j \cdot y_i \cdot y_j \cdot x_i^T \cdot x_j + \sum_{i=1}^{\mathrm{n}} \alpha_i. \qquad (9)$$

Best decision function used for hyperplane is achieved applying a developed formula.

$$y_i = \mathrm{Sign}(w^T \cdot x_i + b) \qquad (10)$$

It is extremely complex to pick the right hyperplane. This selection is in the direction of dividing the nonlinear data. To fix this sensitive issue, the Hilbert-Schmidt operator has been introduced to modify the D–dimensional feature vector of x to a high dimensional feature space. This action is carried out using the multidimensional support vector function and the relationship is considered as: ($\emptyset : \mathrm{R}^{\mathrm{d}} \rightarrow \mathrm{R}^{\mathrm{N}}$). Where in the link, the achieved decision function connected with the optimal hyperplane for an SVM is defined as following equation.

$$y = \mathrm{sign}\left(\sum_{i=1}^{n} \alpha_i \cdot y_i \cdot K(x_i, x) + b \right) \qquad (11)$$

Here, $K(x_i, x_j)$ refers to SVM kernel function.

3.2.4.1 The Kernel Trick

The task of the kernel function is inclusive of taking the data as input and then transforms the date into the essential formula. In this context, the functions can be in different format as the potential SVM algorithms are different. These functions can be different types. For instance, the following methods are different types of SVM kernel functions; sigmoid, RBF, polynomial, linear, nonlinear, kernel trick etc.

In this thesis, kernel trick was used to transform the prediction system from linear situation into nonlinear. From this perspective the Gaussian kernel function computed as below.

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2}\left\|x_i - x_j\right\|^2\right) \tag{12}$$

In the above, x_i , x_j refer to the input vectors and σ is the kernel function (Siamidoudaran and Iscioglu, 2019). This kernel is figured with a support vector which remains as an exponentially decaying function in the input. Following this, the determined value is reached at the support vector and fades equally in all ways nearby the vector. In circumstances where various levels get up, one level is typically compared with one more class that regularly creates the n classifier. Against this background, process of solving a special form of mathematical optimisation problem is presented involving a clarification achieved by the quadratic programming using n variables.

In this thesis, a constructive strategy was raised as 'one-versus-rest'. This method is involved training a single predictor per level with the examples of that level as positive example and all other examples as negatives (Cortes and Vapnik, 1995; Vapnik, 2013).

3.2.5 Proposed Hybrid MLPNN-SVM Designed for ISP Model

The combination of MLPNN and SVM networks is mainly aimed at connecting the output layer of an MLPNN classifier by means of optimal margin hyperplanes. The indication of using hybrid architecture of MLP-SVM is perceived to be the significantly improved model of MLPNN in terms of performance (Bellili et al., 2003; Tifani et al., 2017; SiamiDoudaran and Iscioglu, 2019). That is, it determines a distinctive solution to the last layer factors by using convex optimisation accompanied by a primal-dual understanding, as well as guaranteeing a higher bound concerning testing errors.

The hybrid model was applied with extra efficiency in comparison to the non-linear SVM which are trained in the input space. This was because of a nonlinear SVM requiring selection and turning a kernel to reach a respectable nonlinear mapping through the input space to a transformed feature space in which data was seemingly more linear separable. Concerning the hybrid model, this nonlinear mapping was discretely optimised all through the MLPNN training in the structure of the sigmoid kernel. In the phase of training in MLPNN, the function approximation application was used to map the input and output data in the first layer.

In an attempt to optimise the network, back-propagation was used to decrease the relative entropy among the output delivery and the true label delivery by constructing use of the optimised input aimed at hidden layer factors. Then, the input data dimension was reduced to one-dimension space fixing the operation for its next stage. For the next training period, the output data from MLPNN was reserved and was imported to SVM using the SVM scheme of ‘one-versus- the rest’ to forecast the personal injury severity levels (Bishop, 1995; Bellili et al., 2003; Tifani et al., 2017; SiamiDoudaran and Iscioglu, 2019).

Figure 20 simply shows the structure of the hybrid MLPNN-SVM used in this thesis in order to increase prediction performance.

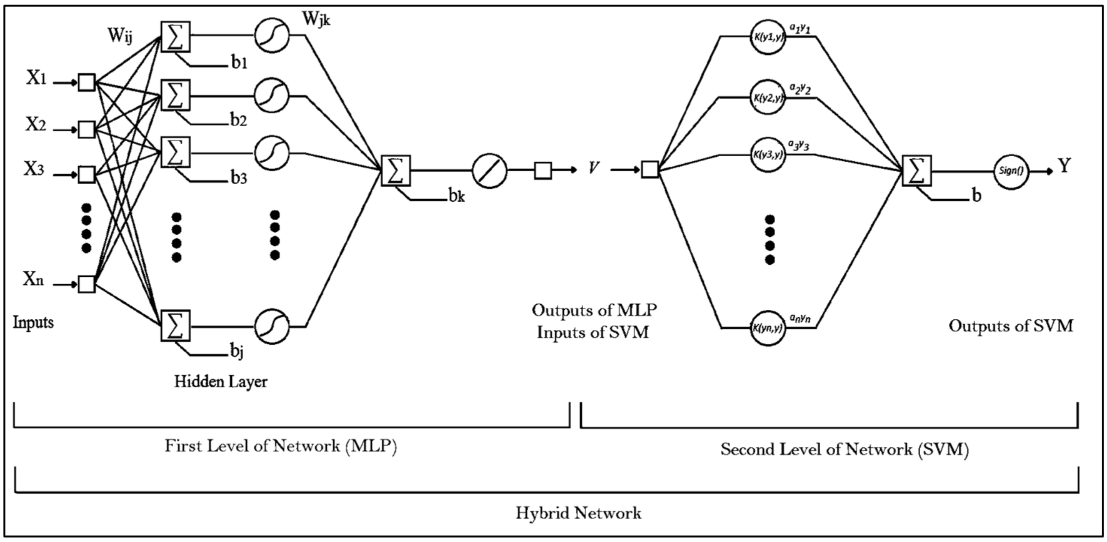

Figure 20: Hybrid SVM–MLPNN Architecture (Bellili et al., 2003)

3.2.6 Proposed LVQNN Designed for ISP Model

In this thesis LVQNN classification algorithm was used for personal injury severity prediction by applying sensitive predictors (Priyono et al., 2005; Al-Daoud, 2009; Shen and Chen, 2009; Chen and Marques, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). LVQNN is special example of a feed forward ANN assumed from supervised learning which are demonstrated as a competitive ANN. LVQNN is one of the most powerful methods for prediction (Villmann et al., 2017) and has reached best overall accuracy in comparison with other existing AANs (Al-Daoud, 2009; Thanasarn and Warisarn, 2013). Previous related studies revealed that LVQNN is also a proper application for road traffic data analysis (Priyono et al., 2005; Shen and Chen, 2009) as well as it effectively being used for classifying data with categorical values (Chen and Marques, 2009). Thus, by reason of using a large number of subdivisions for variables in this thesis, LVQNN was considered to scrutinise the prediction of personal injury severity classes (Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

This algorithm was devised by Teuvo Kohonen (2012). This model of ANN is a precursor to SOM which can be used where there is labelled input data. In this regard, the value of the STATS19 data is label, therefore, this learning method is more suitable for predicting the injury severities in comparison to other types of ANNs. Therefore, LVQNN is looking to maximise the prediction accuracies for injury severity classes. LVQNN uses the level data to relocate the Voronoi vectors slightly, in order to recover the quality of the classifier decision areas. It is a two stage procedure which contains a SOM trailed by LVQNN as presented in Figure 21.

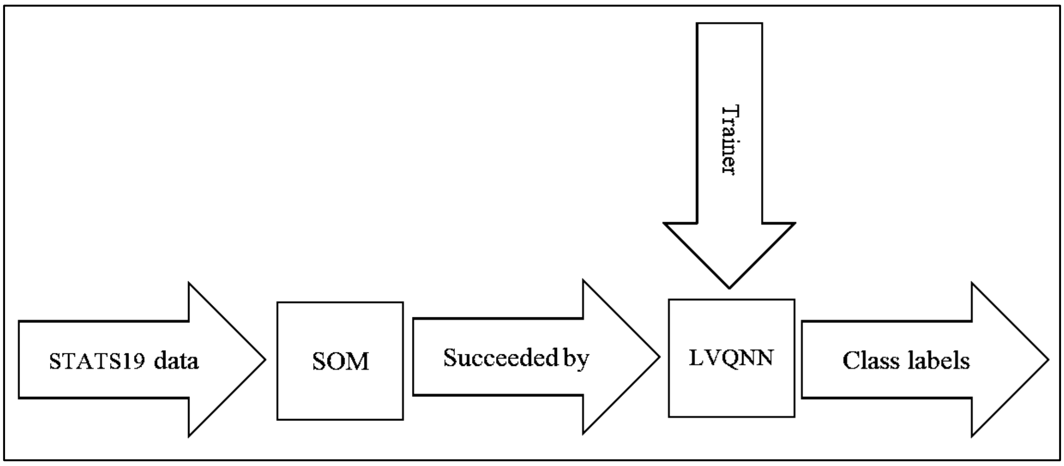

Figure 21: Two Phase Process Built from a SOM (Siamidoudaran et al., 2019a)

The model used in this thesis is an improved method of prediction and exactly appropriate for prediction concerns. The first phase is a collection of features that the unsupervised recognition of a reasonably minor set of specifications in which the main statistics content of the input data is focused. The second phase is the classification where the feature scopes are referred to individual classes. By using an encoder pattern for a big number of input vectors $x \in IR^n$, and transforming the input into an i-value which establish fewer major factors and attain a superior prediction to the unique input space. Below input vector x and suppose $x \in IR^n$, the LVQNN transforms the label input factors into an i-value with an encoder form which $i \in \{1,2,3, \dots, k\}$. Possibly the most efficient means to reflect the LVQNN is concerning about to the common encoders and decoders. After a decoder process has been applied to the above i, the vector $m \in IR^n$ is gained. Indeed, m is an approximation of x, where the quantum error value is attained from the Eq. 13. The Figure 22 simply shows that the architecture of the LVQNN includes two components as an encoder and a decoder.

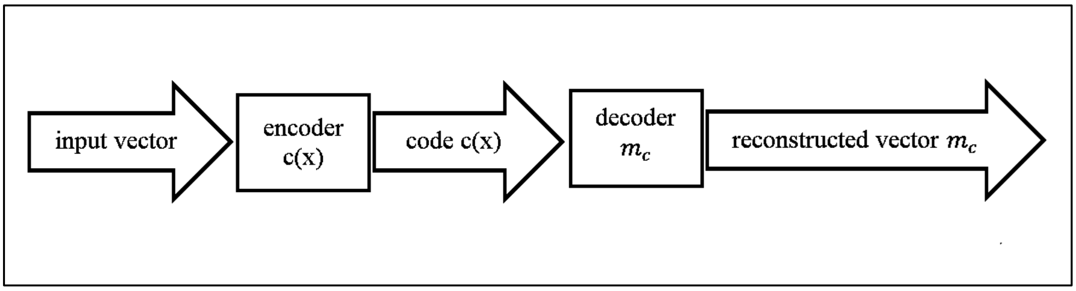

Figure 22: Encoder-Decoder Architecture in LVQNN (Siamidoudaran et al., 2019a)

Typically, x is selected at random in connection to some probability function $p(x)$. At that point the optimum encoding-decoding shape is recognised by adjusting the functions x and m_c to mitigate the expected distortion described by eq. 13.

$$E = \varepsilon\{\|x - m_c\|^2\} = \int \|x - m_c\|^2 p(x) d(x) \qquad (13)$$

Here, ε is the expected value (EV) and m_c is defined as centre of the winner. Once a decoder procedure is applied to i, the vector $m \in IR^n$, is gained and m remains an approximation of x, in the error of vector quantization approximation equation. The EV and the winning neuron are attained from the following equation in which C is the winner and obtained from eq. 14.

$$C = arg\ min_i \|x - m_i\|^2 \tag{14}$$

To identify the limit of each level, it is essential to display the midline of the line segment designed for m_1 , m_2 . Actually, the midline identifies a route that the space of all points on that route, is equal from the centres of m_1 and m_2 $(d_1 = d_2)$. In terms of three-dimensional space, the midline performs as a midplane, and typically it is offered as a hyperplane. The algorithm initiates through a trained SOM with input vector and uses weight/Voronoi diagram if the circumstance for a range of more centres is recognised (Kohonen, 2012; Nova and Estévez, 2014; Heris, 2016; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

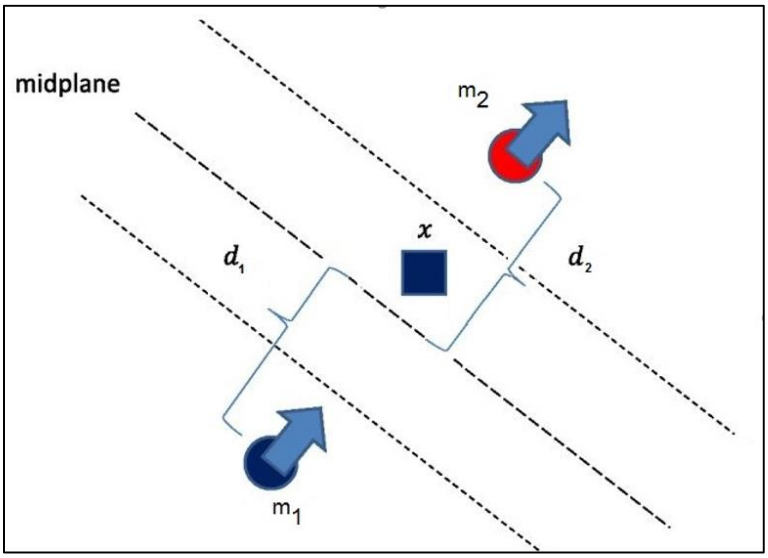

Figure 23: Model Updating Rule (Siamidoudaran, et al., 2019)

The classification labels of the inputs are used to notice the highest classification label for each Voronoi neuron. As the Voronoi neuron boundaries do not match the classification boundaries, the model is efforts to fix this problem through shifting the boundaries.

If $x(t)$ does not exist on the boundary ($d_1 \neq d_2$), the associated centre encourages the classified integer level ($m_i(t)$ to become closer to $x(t)$) and informs as revealed in the below equations (e.g. the weight of the winning output node has the same class label, which then moves them closer together by $\Delta m_i(t)$ as in the SOM algorithm).

$$m_i(t+1) = m_i(t) + \Delta m_i(t) \quad i = 1,2,3,\ldots,k \tag{15}$$

$$\Delta m_i(t) = \delta_{ci}.\alpha(t).[x(t) - m_i(t)] \tag{16}$$

$$\delta_{ci} = \begin{cases} 1 & c = i \\ 0 & c \neq i \end{cases}, \; 0 < \alpha(t) < 1$$

Here $\alpha(t)$ is a learning rate which decreases with the number of iterations/epochs of the training process. In this method better classification results are achieved than by the SOM alone. In fact, the learning rate falls by the number of the iterations/epochs, and in each progress, this coefficient is dropped among 0 and 1. In this regard, Voronoi vectors/weights corresponding to other input regions are left unchanged with $\Delta m_i(t) = 0$ (Bullinaria, 2004) .

LVQNN1 is an improved form of LVQNN and is updated similarly in the adjacent centre. Nevertheless, if input $x(t)$ and associated Voronoi or weight such as, winning output node is properly classified and has the similar label of class, encouragement the of $m_i(t)$ to $x(t)$ moves closer together as in the SOM network. If they have the different level labels, at that point, it is penalised, and $m_i(t)$ moves apart from $x(t)$. Voronoi vectors or weights corresponding to other input areas are left unmoved with $\Delta m_i(t) = 0$. Consequently, the following equations are attained.

$$\Delta m_i(t) = \delta_{ci}.f_i(t).\alpha(t).[x(t) - m_i(t)] \quad , \quad i = 1,2,3,\ldots,k \tag{17}$$

$$m_i(t+1) = m_i(t) + \delta_{ci}.f_i(t).\alpha(t).[x(t) - m_i(t)] \tag{18}$$

$$f_i(t) = \begin{cases} +1 & if \;\; m_i(t)\,, x(t) \; have \; the \; same \; class \; label \\ -1 & if \;\; m_i(t)\,, x(t) \; have \; the \; different \; class \; label \end{cases}$$

Consequently, the following equations are considered.

$$m_i(t+1) = m_i(t) + s_i(t).\alpha(t).[x(t) - m_i(t)] \tag{19}$$

which $s_i(t) = \begin{cases} +1 & correct\ Classification \\ -1 & incorrect\ Classification \\ 0 & if\ not \end{cases}$

In the case of optimised LVQNN1, $\alpha(t)$ in place of being similar for all centres, it performs as an individual learning rate for each centre. Thereby, the superior classification is achieved through the SOM alone. This will be reached if the ranking of the input data does not oppose in relation to the timeframe in a manner that the effect of the initial data has no significant difference with the last input, and also, all the data have the equal class labels. Therefore, in such circumstances, the following is developed:

$$Weight\ of\ x(t) \rightarrow \alpha_i(t) \tag{20}$$

$$Weight\ of\ x(t-1) \rightarrow [1 - s_i(t).\alpha_i(t)].\alpha_i(t-1) \tag{21}$$

As a result of equating the above relationships, the following is gained:

$$\alpha_i(t) = [1 - s_i(t).\alpha_i(t)].\alpha_i(t-1)] \tag{22}$$

$$\alpha_i(t) = \frac{\alpha_i(t-1)}{[1+s_i(t).\alpha_i(t-1)} \qquad 0 < \alpha_i(t) < 1 \tag{23}$$

$$\begin{cases} 0 < \alpha_i(t) < 1 \\ \alpha_i(0) = 0.3 \sim 0.5 \end{cases}$$

LVQNN2 is the second developed type of the LVQNN that is used in this thesis and it moves closer in effect to Bayesian decision theory. LVQNN2 which is opposite of LVQNN1 is updated at the parallel nearer to the centre. The method uses the correct and incorrect classification update equations. So, in this case, the winners are two members as $m_i(t)$, $m_j(t)$. The input vector x gives the correct classification through the associated Voronoi vector ($m_i(t)$) and the other nearest centre is incorrectly classified ($m_j(t)$). Moreover, the input vector x is well near to the decision boundary and $x(t)$ is in a specified range (W) (Bullinaria, 2004). Consequently, the following equations are obtained as below.

$$d_i = \|x(t) - m_i(t)\| \tag{24}$$

$$d_j = \left\|x(t) - m_j(t)\right\| \tag{25}$$

Learning here is similar to that in the LVQNN2 weight learning function seeks two vectors of layer one that are closest to the input vector can be updated, provided that one belongs to the true level and one belongs to an incorrect level, and further produced that the input falls into a 'window' adjacent the midplane. If d is the Euclidean distance between x and m, the window is defined by the Eq. 26 (MathWorks, 2020).

$$min\left(\frac{d_i}{d_j}, \frac{d_j}{d_i}\right) > s \tag{26}$$

where

$$s = \frac{1-w}{1+w} \tag{27}$$

where w is the boundary width and normally is $0.2 \ll w \leq 0.3$ and that results to

$$\frac{7}{13} \leq s \leq \frac{2}{3}$$

$$0.5 < min\left(\frac{d_i}{d_j}, \frac{d_j}{d_i}\right) \leq 1$$

where $d_i < d_j$ then we will have the following equations:

$$\frac{d_i}{d_j} = \frac{\frac{d_i+d_j}{2} - \frac{d_j-d_i}{2}}{\frac{d_i+d_j}{2} + \frac{d_j-d_i}{2}} = \frac{1-\frac{d_j-d_i}{d_j+d_i}}{1+\frac{d_j-d_i}{d_j+d_i}} \tag{28}$$

$$w = \frac{d_j - d_i}{d_j + d_i} \quad (29)$$

$$\frac{d_i}{d_j} = \frac{1-w}{1+w} \quad (30)$$

where m_i is considered as correct classification and m_j as an incorrect classification, we will have the equations as below.

$$m_i(t+1) = m_i(t) + \alpha(t).\left[x(t) - m_i(t)\right] \quad (31)$$

$$m_j(t+1) = m_j(t) + \alpha(t).\left[x(t) - m_j(t)\right] \quad (32)$$

LVQNN acts as a differential mode and moves one centre nearer together while moving another node apart. Alternatively, the preliminary selection of nodes for the LVQNN2 is more complicated, and to work out this weakness, initially the runs were made using the LVQNN1 and then retrieved by the LVQNN2 below (Bullinaria, 2004; Kohonen, 2012; Nova and Estévez, 2014; Heris, 2016; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

LVQNN3 is another variation on this theme which is also used for producing additional superior classification systems. Following this, where m_i , m_j and x have the same class label, the equation is obtained as below (Kohonen, 2012; Nova and Estévez, 2014; Heris, 2016; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

$$m_k(t+1) = m_k(t) + \epsilon\, \alpha(t).\left[x(t) - m_k(t)\right] \quad , \qquad k = i, j \quad (33)$$

In the above equation ϵ is dependent on w and $0.1 \leq \epsilon \leq 0.5$.

LVQNN as a great pattern recognition performance in many more complex predictions, and as a different type of ANN, apart from the commonly reviewed ANNs is considered in this thesis on the modelling of injury severity prediction. In this context, the model is used by applying personal accident injury data for Cambridge case study (Priyono et al., 2005; Shen and Chen, 2009; Chen and Marques, 2009; Al-Daoud, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

3.3 Application of Rank Analysis

Selection of key input factors is always an arduous task which needs a sensible engineering judgment and a good understanding of building classification. Before implementation of the prediction models, it is important to identify what input factors are to be examined. which index is more sensitive and which is less sensitive (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In this regard, application of rank analysis is valuable technique for recognising significant model factors, testing the model understanding, and improving the model performance in terms of accuracy. This application is essential to classify the predictors that affect the injury severity classes as outputs. This tool is the examine of how the improbability in the output of a model can be distributed to different sources of doubt in its inputs. Accordingly, it helps to achieve better model efficiently and performance.

3.3.1 Using RBFNN for Rank Analysis

In this thesis, sensitivity test was applied using RBFNN to study of whether and how the output is influenced by different inputs. Thus, it offers an insight as to which factors are most important and which ones are not in the injuries. As we were dealing with hierarchical data along with large number of accident factors, rank analysis was essential for calibrating all the proposed models with purpose of focusing only on the sensitive inputs. And so, uncertainty reduction was applied for each prediction progression in order to assessment of strength of injury outcomes in the attendance of uncertainty. By way of encountering unexpected associations between crash related factors and injury severity classes, searching for errors in the model was carried out. As a result of data simplification, the factors that had no effect on the injury severity were identified. Accordingly, redundant factors of the model were removed, and the space of input factors dropped to minimum. Likewise, appropriate value settings on the injury severity related factors was improved with regard to predicting result (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

3.4 Predictive Models' Performance Evaluation

In this thesis, different accuracy measures were used to assess the participants' value. From this perspective, MSE delivers an accurate picture of prediction quality along with Root Mean Square Error (RMSE) which is also a right measure of accuracy. They are the most important standards for the goodness-of-fit if the major aim of the model is prediction. MSE is probability function, in proportion to the predicted assessment of the squared variation between the fitted values denoted by the predictive function and the assessment of the non-observable function. In other terms, it is an adjustment or curve fitting process in prediction model (Lehmann and Casella, 2006; Siamidoudaran and Iscioglu, 2019).

RMSE is a commonly used assessed value of the differences among values predicted by a model and shows the rate of how spread out prediction errors are. It implies the typical deviation of the residuals in a prediction model Likewise, RMSE verifies exactly how focussed the data is nearby the line of better fit (Hyndman et al., 2006). In this connection, normally the threshold related to good predictive model is lower values of MSE and RMSE. Correlation coefficient (R) is also successfully performed to measure the correlation between outputs and target and is also a measurement for the strength and direction of the relationship between outputs and targets (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

As a result of the rank analysis, unreliable factors were eliminated. The reduced data was designed at application of the final predictions to increase the accuracies. The confusion matrix is used to assess the performance of the prediction models.

3.4.1 Error Matrix

The error matrix was primarily presented to assess outcomes from binominal classification. In view of this, 'accuracy (ACC)', 'sensitivity (SEN)', and 'error' parameters are used to compute the number of correct and incorrect predictions of each level. Accordingly, one of the two classes as the class of interest is taken (e.g. the positive class). one value as the positive class is chosen arbitrarily in the target column. Following this, the other value is randomly measured as the negative class. Table 3 shows an error matrix displaying actual and predicted positive and negative levels in test set. The associated equations of the matrix are also defined in table 3.

Table 3: Sample of error matrix (Siamidoudaran and Iscioglu, 2019)

n		**Class1 (predicted)**	**Class2 (predicted)**	Total
		(+)	**(–)**	
Class 1 (actual)	**(+)**	True positives (TP)	False negatives (FN)	N
Class 2 (actual)	**(–)**	False positives (FP)	True negatives (TN)	P
Total		N	P	

In the above matrix, the data rows refer to positive class and properly classified as such. They are named ‘true positives (TP)’ and their room is located in the top left cell of the matrix. The data rows associate to the positive class and wrongly classified as negative. They are ‘false negatives (FN)’. The sum of false negatives is in the top right. The data rows related to the negative class and inaccurately classified as positive which is ‘false positives (FP)’, they are in the lower left. The last position is when the data rows in regard to the negative class and properly classified as well. They are named as ‘true negatives (TN)’ which are placed in the lower right cell of the error matrix. Consequently, accurate predictions are on the diagonal with a grey colour; those without any colour are false predictions. Using the four counts in the above matrix, the injury severity class assesses in terms of accuracy as well as the model performance using the formulas in the below. In the equations, how many positive and negative actions are predicted properly or wrongly are evaluated. (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{34}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{35}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{36}$$

Likewise, the average predicted class will be compared with actual class and the accuracy of the network will be calculated to obtain prediction performance (Hyndman et al., 2006; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Chapter 4

RESULTS AND DISCUSSION

This thesis used a series of machine learning algorithms to predict personal injury severity classes for different road users which are included; RBFNN, MLPNN, SVM, hybrid MLPNN–SVM, and LVQNN. In both case studies, RBFNN was used for rank analysis practice. MLPNN and hybrid MLPNN–SVM were applied for prediction of all road users (driver, rider, cyclist and pedestrian) injury severity levels in the city of London. SVM and LVQNN models predicted pedal rider injury severities in Cambridge. Pedal rider related prediction was to overcome the identified cycling safety concern in the first case study.

4.1 The City of London Case Study

In this section, the case study focuses on predicting of personal injury severity aim at all road users in the city of London.

4.1.1 Data Cleansing & Analysis Methodology – City of London

Data cleansing is the process of identifying and modifying or eliminating incorrect records or corrupt data. Accordingly, referring to classify inaccurate, incorrect, incomplete or irrelevant sections of the data and then replace, adjust, or remove the useless or coarse factors (Han and Kamber, 2001).

The data used in this thesis is delivered from the UK government, however, due to the large and complex data, the realisation of the factors is extremely complicated. In addition, the version of the database is mainly designed to be used by in-house software types which makes it very difficult to use for research purpose using other technical methodologies. For this reason, this study represents a perfect implementation to get stuck into a chunky data processing challenge. Accordingly, appropriate coding was carried out in order to clean and label the data in a usable format for predictions models.

Using MATLAB programming language as a powerful numerical computing environment, quality of the data was developed much more useful after it has been cleaned and labelled for the prediction models, compared to the national STATS19 database provided by the government. The analysis was run by RBFNN, and figures of the most important predictors were calculated. As a result, the relationship between accident related factors and the injury severity outcomes for the databases shown in Table 2 were examined. Table 5 breaks down the descriptive statistic of the sub-variables involved in the accidents and validates the distributions class for count data (e.g. the distribution among factors intended for 2nd Road class that are include the following labels; not at junction or within 20 metres, motorway, A(M) class road, A class road, B class road, C class road, and unclassified road) as well as percentage rate used for the injury severity outputs for the total frequency. It should be noted that 'unknown' code is only used in exceptional circumstances where no information on related factor is available, or where the other related labels codes are inappropriate (DfT, 2011; Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

4.1.2 Analysis of Rank Ordered Data by RBFNN – City of London

The question refers to this stage is to find the most effective inputs from a set of STAST19 data as pre specified inputs. Accordingly, RBFNN technique was used to predict the outcome of a decision given a certain range of crash related factors.

MATLAB software was used for each task in building a given set of the data. In this activity, the prediction model was able to identify how changes in one accident related variable affect the injury severity outcome. In fact, we simply specified the input factors through the model for feature selection, trained the model, and then systematically varies each of the inputs in a range relevant to its domain while keeping all other inputs fixed, and then measuring the change in the output. This supplied an estimate on the variability of the output conditional on the input, and hence the relative importance.

4.1.2.1 Implementation of Rank Analysis – City of London

Using engineering judgment 28 explanatory variables were selected from the data listed in the Table 2 as being involved in results of injury severities. Other factors mentioned in the Table hold very insignificant value or is impacted by missing data so they clearly were noticeable by visual perception. In addition, some factors were impacted by a large increase of unknown values, therefore, the related factors simply were eliminated from the data designed for rank analysis. The selected explanatory variables are shown in Table 4. Following this, initially the data was shuffled and normalised for each iteration to gain an equal series of value. It then, the data was applied individually as an input factor into the prediction model in an attempt to comparing the influence of each crash related factor on outcome of the injury severity. Therefore, the model predicted the injury severities for 28 times by changing an input factor in each period while retaining the other crash related factors fixed.

As a result of sensitivity analysis, more influential indexes on prediction of the injury severities were identified. Following this, the poor predictors were removed. This reduction was to recover the quality of the data and referred to deduct the input factors to less dimension space (Broomhead and Lowe, 1988a; Broomhead and Lowe, 1988b). This status strongly helped to boost the speed of the models, and subsequently to better understanding of the close associations between input and output indexes in the predictions (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.2.2 Rank Analysis Evaluation – City of London

Table 4 presents the amount of MSE, RMSE, and R–value versus absence of each index. On the basis of their MSE values, the most sensitive predictor variables were discovered and ranked as the most important contributory factors. R–value measures how accurately the model fits the dataset. If the R–value is close to 1 (good) then it displays that the RBFNN prediction exists very close to the actual dataset. Then again, if the value is zero (bad) then it displays that the model fully fails in building an accurate prediction (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Based on the results of the rank analysis, the effects of the predictors were identified and listed in the below Table. The sensitive predictors from 1 to 9 are considered as the main factors which contributed to the injury severity outcomes according to their MSE, RMSE, and R values. In different circumstances, the variables after rank of 9 refers lower sensitivity value which are considered as the poor predictors. By reason of using large number of subdivision variables, unreliable indexes after rank of 28 haven't been pointed out in the Table, as a result of their immaterial association. However, their major variables are itemised in Table 2. Furthermore, it can be referred to (DfT. 2011) for full specifics of the eliminated indexes in this stage.

Table 4:Sensitivity of RBFNN against the absence of each variable–First case study

Rank	**Contributory factors**	**MSE**	**RMSE**	**R–value**
A01	Pedestrian crossing facilities	0.009	0.197	0.925
A02	Junction location	0.012	0.212	0.895
A03	2^{nd} road class	0.055	0.226	0.885
A04	Vehicle type	0.065	0.251	0.872
A05	Road type	0.074	0.260	0.868
A06	Vehicle manoeuvre	0.081	0.265	0.856
A07	Junction detail	0.088	0.285	0.832
A08	1^{st} Point of impact	0.098	0.295	0.829
A09	Alcohol/drug involvement	0.101	0.301	0.821
A10	Time frame	0.106	0.303	0.798
A11	Condition of lighting	0.116	0.319	0.795
A12	Condition of weather	0.124	0.320	0.792
A13	1^{st} road class	0.137	0.322	0.788
A14	Age of driver	0.147	0.335	0.783
A15	Vehicle no. involved	0.142	0.340	0.779
A16	Road environment	0.150	0.342	0.772
A17	Vehicle propulsion	0.153	0.351	0.769
A18	Age of vehicle	0.155	0.358	0.765
A19	Engine capacity	0.156	0.361	0.758
A20	Casualty no. involved	0.157	0.369	0.757
A21	Urban or rural area	0.158	0.398	0.756
A22	Month band	0.167	0.409	0.739
A23	Journey purpose of driver/rider	0.171	0.413	0.739
A24	Sex of driver/rider	0.172	0.415	0.730
A25	Driver home type area	0.176	0.420	0.723
A26	Speed limit	0.177	0.421	0.721
A27	Driver IMD decile	0.184	0.430	0.712
A28	Carriageway hazards	0.204	0.452	0.669
	All STAS19 data	0.175	0.419	0.728

4.1.2.3 Elimination of Unreliable Variables – City of London

At this stage, as a result of the prediction designed for large number of the STATS19 subdivision variables, the poor factors after rank 20 were eliminated. This elimination is an effective way to generalise the dimensional feature space and to improve the limitation of poor quality data. Thus, the unreliable factors with minor R–value were dropped to 20 factors in ordered to apply to MLPNN model aimed at implementation of final prediction. All the sensitive predictors were normalised between 0 and 1 and run was completed using random division of 70 % and 30 % into training and testing datasets. Rank and affect for all the input parameters are listed in the Table 4. In addition, the specific result for total of the data is presented in Figure 25 (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3 Important Contributory Factors – City of London

In this thesis, the contributory factors were found to help provide some insight into why and how road injuries happened. The factors were detected to reveal the key actions and failures that led directly to the actual impact to help investigation of how severity of injuries might be reduced. Table 5 in this thesis provides detailed information about the subdivisions of the identified factors (labels). It is also strongly recommended that readers of this thesis refer to the STATS20 document (DfT, 2011) for more detailed information as this explains to police officers how they should complete the STAST19 reporting forms.

4.1.3.1 Crossing Facilities (A01) – Most Important Factor

The first identified factor is A1 which refers to pedestrian crossings including shared facilities with cyclists (physical facilities such as; footbridge, central refuge, zebra crossing, pelican, puffin, toucan or similar non-junction pedestrian light crossing) (TRL, 2010; DfT, 2011). This identified factor verified that the city of London is the area where pedal riders and pedestrians are most in danger of injuries. The findings along with the exploratory data analysis of the related labels (shown in Table 5) displayed that injuries involving the specific groups (cyclist and pedestrian) happened frequently nearby the crossings locations controlled by physical facilities. 'Pedestrian phase at traffic signal junction' is any pedestrian crossing at a junction controlled by traffic lights which has an indicator light for pedestrians and cyclists (as seen in the below example, Figure 24). This means that these facilities which are designed for pedestrians and cyclists were not necessarily good enough to avoid the severity of injuries in accidents (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Figure 24: Example of Pedestrian Phase at Traffic Signal Junction

As seen in Table 4, the influence of this factor (A01) is overhead compared to the other predictors. The outcome of A01 among with MSE, RMSE, and R–value verifies greater relationship between injury severity severities and their crash related factors, thus this index is considered as the main contributory factor. In this thesis generally, the R–value has been considered as the main measurement used for assessment of the most important contributory factors. As seen in Table 4, the R–

value for A01 is 0.925 which is close to 1 and this result displays that the prediction exists very close to the actual dataset.

In addition, Table 4 displays the results of the MSE and RMSE as 0.009 and 0.197, respectively. In this regard, both values are small and refer to a good prediction in comparison to other input factors. Furthermore, subdivided predictors of this variable (A01) are shown in the analysis of STATS19 data (Table 5) which are specified for pedestrians and pedal riders to cross a route.

The identified factor fits the author previous studies verifying that a greater part of the injuries took place for vulnerable road users in this area. In those researches, series of ANNs (MLPNN and LVQNN), SVM and hybrid SVM–MLPNN were applied to predict STATS19 data in different case studies. Nonetheless, all the results blamed the same factor which refers to pedestrian crossing facilities (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In this regard, Chang and Wang (2006) found that the pedestrians and cyclists were at a greater risk of injury in accidents. In addition, a Dutch crash statistics indicated that more than half of the KSI accidents which vulnerable road users are involved in occur while crossing the street (SWOV, 2010). This result also is in line with a previous research showing that the increase of crossing facilities at intersections have an extra positive safety outcome in reducing the number of the cycling injuries (Gårder et al., 1998).

4.1.3.2 Improving Safety at Pedestrian Crossings and on Cycle Tracks

To overcome the first detected factor, it is essential in some areas that crossings facilities for pedestrians and cyclist be improved and separated from the other traffic. In this connection 'protected junction' can definitely complement new crossing facilities as majority of the 'pedestrian phase at traffic signal junction' failed to prevent the severity of injuries (RoSPA, 2017a; RoSPA, 2017b; RoSPA, 2017c; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020; DfT, 2020).

4.1.3.3 Junction Location of Vehicle (A02)

The next sensitive predictor attributes to junction location of vehicle (A02) showing a higher impact and playing a key role in outcome of the injury severities (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). R–value for A02 is 0.895, and results of MSE and RMSE are 0.012 and 0.212 respectively. 'Junction location of vehicle' is an aspect covering the geographical location of junction where each accident happened (e.g. vehicles waiting to enter the roundabout/main road). The labels of this variable (analysed in Table 5) indicate the position of vehicles at or nearby busy junctions.

This finding fits the DFT's report that many British pedal riders are injured at junctions during recent years in build-out areas. Accordingly, large amount of bicyclist deaths happened at these locations (DfT, 2018b). Furthermore, UK's Transport Research Laboratory reinforced the same outcome verifying that about 75% of vulnerable road users' crashes take place at or near junctions (Knowles et al, 2009).

4.1.3.4 Improving Junction Safety

To solve the second identified concern, again, it is recommended to make junctions more safer for VRUs using 'protected junction' designed for heavy traffic congestion in the city centre wherein people travelling on foot, by bike. Using the protected junctions' vehicles are all separated as they cross through the intersection. Unlike traditional intersections, which typically need right-turning pedal riders to wait in the centre of the intersection for an appropriate gap in the traffic, the protected junction donates a safer advantage (RoSPA, 2017a; RoSPA, 2017b; RoSPA, 2017c; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020; DfT, 2020).

4.1.3.5 Second Road Class (A03)

A03 refers to '2^{nd} road class' in the dataset which shows the class of the road upon which a collision took place (DfT, 2011). The letter labelling for each type of road classified road customs a prefix to its number of road. The '2^{nd} road class' is used when the accident is on a junction. The '1^{st} road class' is the road that the vehicle was on at the time of the collision; the 2^{nd} road class is the road that the junction was with (if relevant). For collisions at intersections which cannot clearly be allocated to one certain road, the class of the major road is considered. The major road is specified as the road which has priority. For signalised intersections and roundabouts, the major road is the one with the greatest class of all the roads entering the intersection. Accordingly, as shown in the results, A02 predictor refers to the contributory factor which linked with junction actions. The R–value related to this factor is 0.885, and results of the MSE and RMSE are 0.055 and 0.226 respectively. As a result of the prediction in this stage, an alarming number of accidents in or around junctions are caused by drivers, cyclists and riders (Knowles et al, 2009; DFT, 2018b; Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

4.1.3.6 Vehicle Type (A04)

In relation to A04, influence of the vehicle type was well recognised and demonstrated as a significant role in result of injury severities. The amounts of R–value, MSE, and RMSE are as 0.872, 0.065 and 0.251 respectively. ‘Type of vehicle’ identifies the nature of each vehicle involved in an accident (DfT, 2011). Vehicle type category is incorporated into two major ranges; motor vehicles and non-motor vehicles, and are shown in Table 5 (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab, 2004; Chang and Wang, 2006; DfT, 2014; Li et al., 2018; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.7 Road Type (A05)

A05 ‘Road type’ directs the type of road upon which an accident took place and refers to physical nature of the road, for example number of carriageways completely is different with road class (DfT, 2011). Roads are adapted to a large variety of infrastructures and types with the purpose of attaining a common objective of transport under a huge and extensive series of circumstances. As a result, factor A05 contributed the higher impact of road types in outcome of the injuries. The results are as 0.868 for R–value, 0.074 for MSE, and 0.260 for RMSE. The different range of roads are described in more detail in Table 5 which contributed to the road injuries (DfT, 2014; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.8 Vehicle Manoeuvre (A06)

'Vehicle manoeuvre' (A06) implies to vehicle manoeuvre which directs the junction actions taken by a motor vehicle/non-motor vehicle instantly earlier it became involved in an accident (DfT, 2011). The results for R–value, MSE, RMSE are as follow respectively; 0.856, 0.081, 0.265. Based on RBFNN prediction model as well as the data analysis shown in Table 5, 'going ahead bend' and 'turning manoeuvres' were identified as direct contributory factors to the outcome of the injuries. In addition, parked vehicles are contained public service vehicle (PSV) such as; coaches and buses stationary at a bus stop which is very common in the central London. Collisions involving stationary queues of the traffic in resulting the associated pedestrian casualties running out from in front of the bus (DfT, 2014; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.9 Junction Detail (A07)

Next identified factor contributing to the injury severities refers to junction detail (A07). 'Junction' is defined as a location where two or more streets meet whatever the angle of the axes of the roads. 'Junction detail' specifies the general layout of the intersection where an accident happened (e.g. mini roundabout, slip road, T or staggered junction) (DfT, 2011). Only collisions which took place at or within 20 metres of an intersection are implied by this index. The results of the prediction are as following: R–value = 0.832, MSE = 0.088, and RMSE = 0.285. If two or more intersections are available within 20 metres, the intersection nearest to the to the collision location is applied. Table 5 simply defines the specific varieties of junction (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab, 2004; Chang and

Wang, 2006; DfT, 2014; Li et al., 2018; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.10 First Point of Impact (A08)

1st Point of impact' (A08) contributed to the injuries severities where initial point to arise into contact with another motor vehicle, bicycle, pedestrian, or other object. If a vehicle breaks suddenly to avoid contact with another vehicle, bicycle, pedestrian or object in a street, but there is no impact, then 'did not impact' label is recorded by the police officer (shown in Table 5). R–value, MSE, and RMSE obtained the following results; 0.829, 0.098, and 0.295 (DfT, 2014; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.11 Alcohol/Drug Involvement (A09)

A09 refers to the pedestrian(s) impaired by alcohol/drug and behaved in a way which contributed to traffic as well as the severity of injuries. The R–value, MSE, RMSE are 0.821, 0.101 and 0.301, respectively. The contribution of the alcohol involvement (whether or not completely drunk) among the road users are shown in the Table 5 (DfT, 2014; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.1.3.12 Overall Prediction by RBFNN– City of London

As a final point, Figure 25 shows an overall prediction of the STATS19 data obtained from the total crash related variables affecting the injury severity classes (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). R–value for the total data is 0.728, and results of the MSE and RMSE are 0.175 and 0.419, respectively that are presented in Figure 25 (small MSE and RMSE values, and R–value close to 1). The figure displays that the proposed ANN has an acceptable performance, however, the learning process is relatively difficult. The error diagram is exposed where the values of MSE and RMSE also are marked, in which both of the values are small. Indeed, the idea of MSE isn't equal to zero, since then we would have a model that completely forecasts the training data, which in this case would not be capable of totally estimating any other data. In this subfigure the error value is mostly between −0.1 and 0.1, and the error instances and its relation to the typical distribution; the value of μ and σ is shown which is indicative of the good performance of the ANN. In Figure 25, the prediction chart is shown where output and the target are drawing closer together; the R–value is close to 1, this means that most of the data is truly fitted. However, due to lack of data for first class and limited data for second class, the model totally failed to predict fatal class and had some incorrect predictions for serious injury severity class. Likewise, they could not be precisely fitted due to inadequate data for these classes. On the other hand, the predictions of class 3 and class 4 were satisfactory which referred to achieve acceptable R–value due to sufficient data for these classes (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

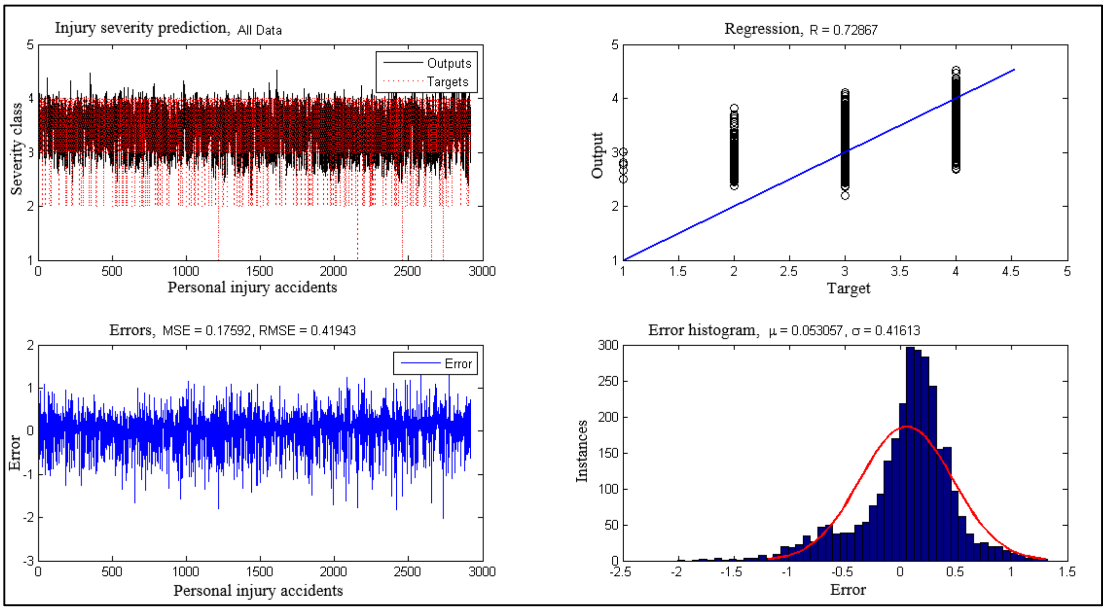

Figure 25: Sensitivity by Overall Data – City of London Case Study

Consuming professional judgment, threshold of the measured values is considered as percentage of 25 which is 75% of the range. Likewise, predictors over the threshold of 25% had a larger impression in performance of the prediction in terms of accuracy. As a result, the factors from A1 to A20 proved greater association to outcome of the injury severities. These factors have been listed in specific order (Table 5) according to their influence as a result of their R–value; time frame, lighting condition, weather condition, 1st road class, age of driver, number of vehicle involved in accident, road surface, vehicle propulsion, age of vehicle, engine capacity, casualty number involved in accident. Examples for coding the input variables are shown in table 5 in more details. In various situations, variables after A21 indicated minor R–value which are considered as unreliable indexes. Because of their poor affect, they were eliminated from the impending implementations. This removal was essential aimed at extra performance improvement of the models. The factors that are included; urban or rural area, month, journey purpose, gender, driver home type area (urban, small town or rural), speed limit, driver IMD decile (an IMD decile is a dimension which places the deprivation scores of individual zones), carriageway hazards (this factor is used for all personal injury accidents and includes various types of hazards such as; animal or any object). For more information about the factors, the STATS20 document (DfT, 2011) details exactly what data were required to be collected by police officers as part of the STATS19 system.

4.1.3.13 Summary of RBFNN Prediction – City of London

In summary, the influence of the pedestrian physical crossing facilities is overhead compared to the other factors. This clearly indicates that, pedestrians and pedal riders' groups were most VRUs and suffered the majority of the injuries (Gårder et al., 1998; Chang and Wang, 2006; SWOV, 2010; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Other discovered key factors are related to junction actions include; junction location, 2nd road class, vehicle manoeuvre, and junction detail (DfT, 2014, DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Vehicle and road types also played massive roles in result of the injuries and, were discovered as the important contributory factors/ area of concerns (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab, 2004; Chang and Wang, 2006; DfT, 2014; Li et al., 2018; DfT, 2018d; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). The details of the contributory factors/ area of concerns are shown in the Table 5.

4.1.4 Maximising Predictive Accuracy by MLPNN– City of London

After elimination of the insignificant crash related factors, the indexes from A1 to A20 were detected as the most important sensitive predictors. At this section, the 20 key predictors were applied to MLPNN model in the input layer which included; pedestrian crossing facilities, junction location, 2nd road class, vehicle type, road type, vehicle manoeuvre, junction detail, 1st point of impact, alcohol/drug involvement, time frame, condition of lighting, condition of weather, 1st road class, age of driver, number of vehicle involved, road environment, vehicle propulsion, age of vehicle, engine capacity, and casualty number.

4.1.5 Analysis of Key Predictors – City of London

Analysis of key factors contributing to the injury severities are described in Table 5. The Table displays the percentage distribution of the data across levels of each variable (e.g. the distribution among vehicle types; goods, PSV, car/taxi/private hire car, motorcycle, pedal cycle, and other vehicle) as well as the injury severity outputs (damager only, slight, seriously injured, and fatal) for the whole occurrence (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Table 5: Description of key contributing factors – The city of London case study

Input	Subdivision	Variable type	Code	Label	Total (%)
1	6	Pedestrian crossing facilities	1 2 3 4 5 6	Pelican, puffin, toucan or similar light crossing No physical crossing facilities within 50 metres Footbridge or subway Zebra Central refuge Pedestrian phase at traffic signal junction	02.77 46.14 00.41 02.12 00.44 48.12
2	9	Junction location	1 2 3 4 5 6 7 8 9	Not at or within 20 metres of junction Approaching / waiting / parked at junction approach Leaving from roundabout Leaving from main road Entering from slip road Entering roundabout Entering main road Mid Junction Cleared junction / waiting / parked at junction exit	10.93 43.67 00.03 02.16 00.01 00.00 01.11 30.98 11.11
3	6	2nd Road Class	1 2 3 4 5 6	U C B A A(M) Motorway	17.45 52.01 02.89 27.65 00.00 00.00
4	6	Vehicle type	1 2 3 4 5 6	Goods PSV Car/Taxi/Private hire car Motorcycle Pedal cycle Other vehicle	15.01 11.01 31.23 13.02 29. 25 00.39
5	6	Road type	1 2 3 4 5 6	One-way street Single carriageway Dual carriageway Roundabout Slip road Unknown	04.50 83.34 10.60 01.69 00.00 00.00
6	9	Vehicle manoeuvre	1 2 3 4 5 6 7 8 9	Overtaking Moving off Waiting to go - held up Reversing Slowing or stopping Parked Going ahead bend Turning / Waiting to turn Changing lane to left or right	09.12 04.09 05.15 01.05 03.12 07.31 46.89 20..41 02.86

7	9	Junction detail	1	T or staggered junction	61.69
			2	More than 4 arms (not roundabout)	03.65
			3	Crossroads	18.79
			4	Not at junction or within 20 metres	12.11
			5	Roundabout	01.88
			6	Private drive or entrance	00.43
			7	Slip road	00.49
			8	Mini-roundabout	00.00
			9	Other junctions	00.96
8	5	1st Point of impact	1	Did not impact	04.06
			2	Front	36.22
			3	Nearside	23.74
			4	Back	12.65
			5	Offside	22.79
9	3	Alcohol/drug involvement	1	Pedestrian	68.23
			2	Rider	19.15
			3	Driver	12.62
10	6	Time	1	00:00 - 03:59	06.25
			2	04:00 - 07:59	11.36
			3	08:00 - 11:59	26.87
			4	12:00 - 15:59	18.28
			5	16:00 - 19:59	27.53
			6	20:00 - 11:59	09.71
11	5	Condition of lighting	1	Darkness - no lighting	00.00
			2	Darkness - lights unlit	00.40
			3	Darkness - lights lit	26.01
			4	Daylight	73.33
			5	Darkness - lighting unknown	00.26
12	8	Condition of weather	1	Fog or mist	00.06
			2	Snowing + high winds	00.00
			3	Snowing no high winds	00.26
			4	Raining + high winds	00.20
			5	Raining no high winds	06.91
			6	Fine no high winds	90.12
			7	Fine + high winds	00.40
			8	Other	02.06
13	6	1st Road class	1	U	01.17
			2	C	33.04
			3	B	01.77
			4	A	64.02
			5	A(M)	00.00
			6	Motorway	00.00
14	9	Age of driver	1	11 - 15	00.17
			2	16 - 20	01.67
			3	21 - 25	09.24
			4	26 - 35	31.04
			5	36 - 45	25.99
			6	46 - 55	20.49
			7	56 - 65	08.95
			8	66 - 75	02.10
			9	over75	00.36
15	6	Vehicle no. involved	1	One vehicle	23.02
			2	Two vehicles	70.82
			3	Three vehicles	05.28
			4	Four vehicles	00.57
			5	Five vehicles	00.14
			6	Over five vehicles	00.17

16	6	Road environment	1 2 3 4 5 6 7	Mud Wet or damp Flood over 3cm. deep Oil or diesel Dry Snow Frost or ice	00.00 11.54 00.40 00.00 87.98 00.06 00.40
17	10	Vehicle propulsion	1 2 3 4 5 6 7 8 9 10	New fuel technology Petrol Gas Gas/Bi-fuel Petrol/Gas (LPG) Electric Hybrid electric Heavy oil Steam Fuel cells	00.00 38.05 00.06 00.27 00.05 00.27 03.26 57.99 00.00 00.00
18	6	Age of vehicle	1 2 3 4 5 6	00 - 05 06 - 10 11 - 15 16 - 20 21 - 25 26 - 30	53.19 32.09 12.49 10.33 00.74 00.16
19	5	Engine capacity (EC)	1 2 3 4 5	0 - 1000 1001 - 2000 2001 -3000 3001-5000 Over 5000	21.22 34.54 29.08 04.22 10.93
20	5	Casualty no. involved	1 2 3 4 5	One casualty Two casualties Three casualties Four casualties Over four casualties	88.95 09.11 01.31 00.57 00.06
Output	4	**Personal injury severity class**	B1 B2 B3 B4	**Fatal injury** **Serious injury** **Slight injury** **Damage only**	00.39 12.25 42.77 44.59

4.1.6 Implementation of MLPNN Applying Key Predictors – City of London

At this stage, typically, random division was used for 70% of the data in in training process, and 30% for testing section. For hidden layer, the tangent sigmoid function was used for finding the relation between the inputs and outputs indices. The layer also made of twenty-five (optimum value) neurons which each neuron relates to each input parameters and this relation is controlled by weight factors and bias terms. In training process, the LM–BP algorithm using MATLAB was applied. Furthermore, the MSE was used as a performance point of the network. The neural network structure and training process are shown in Figure 26. In the training process of MLPNN, the network stopped after 13 iterations with MSE and gradient equal to 0.0399 and 0.0407, respectively. The performance of the model trained with LM–BP algorithm using MATLAB programming language, is presented in the figure (Siamidoudaran and Iscioglu, 2019).

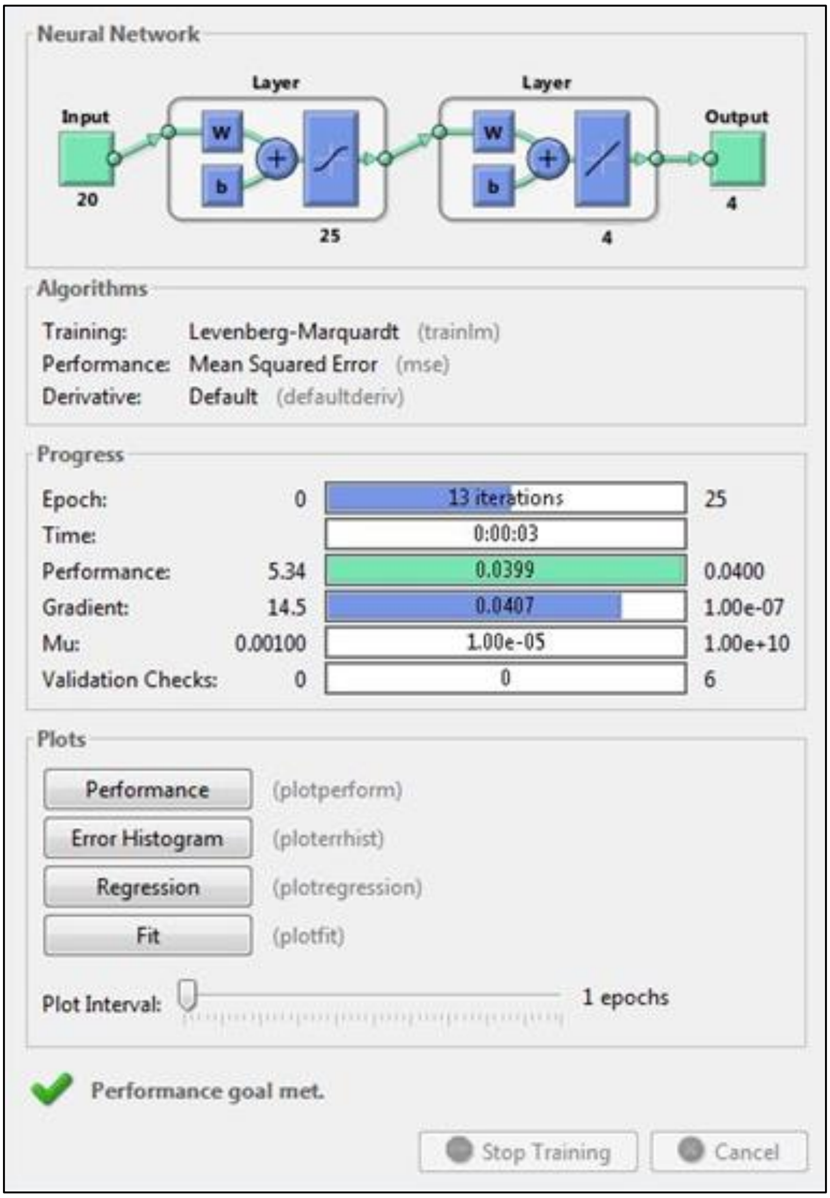

Figure 26: Final Structure and Train Process of MLPNN

4.1.7 Train Regression by MLPNN – City of London

The Regression plot defines connotation between the outputs and the targets. Figure 27 shows train regression for prediction of the data. R–value for this phase was calculated to measure the relationship between outputs and targets as 0.905 for the training response. An R–value close to 1 means a close association, and close to 0 is a random correlation (Siamidoudaran and Iscioglu, 2019).

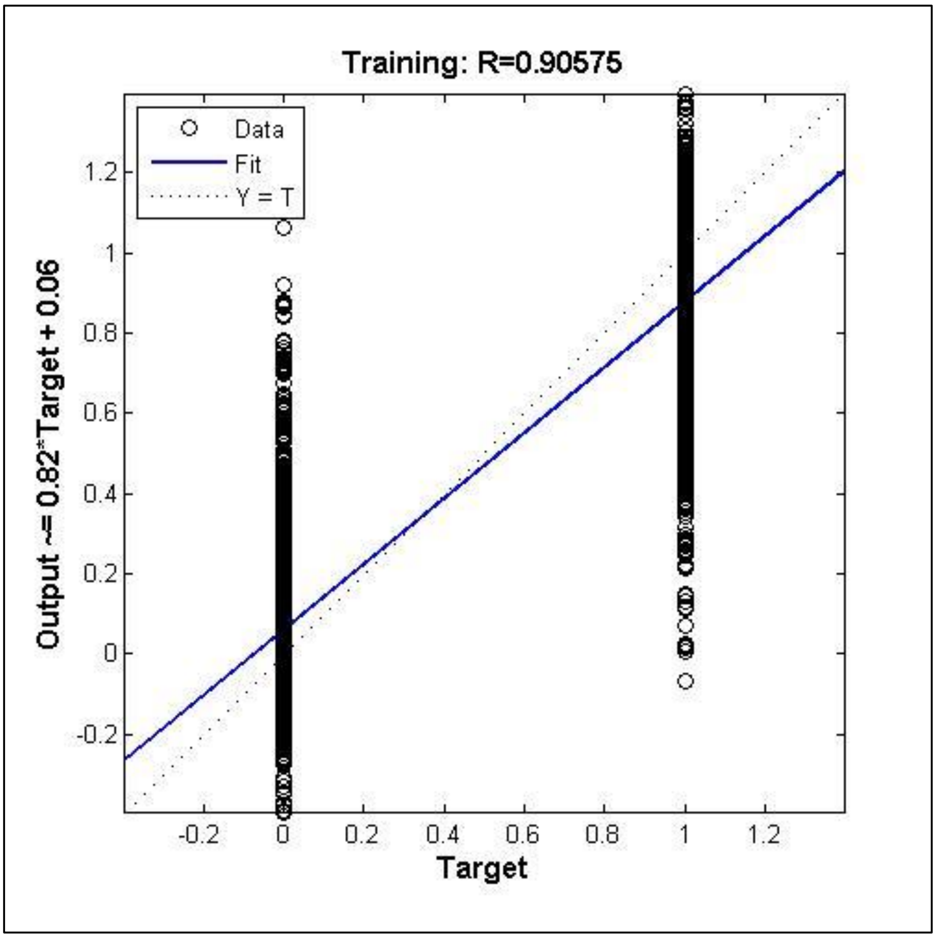

Figure 27: Train Regression Result by MLPNN – City of London

4.1.8 Error Histogram by MLPNN – City of London

Error histogram displays how accurately the trained MLPNN model fits the STATS19 dataset. In this point, the error values were computed as the difference between target values and predicted values. The graph in Figure 28 displays an error histogram with 20 bins for all the data. The plot indicates how to picture errors among target values and estimated values after training a feedforward neural network. The concentration of error bins around the zero line shows that the network was able to predict the injury severity classes with high accuracy (Siamidoudaran and Iscioglu, 2019).

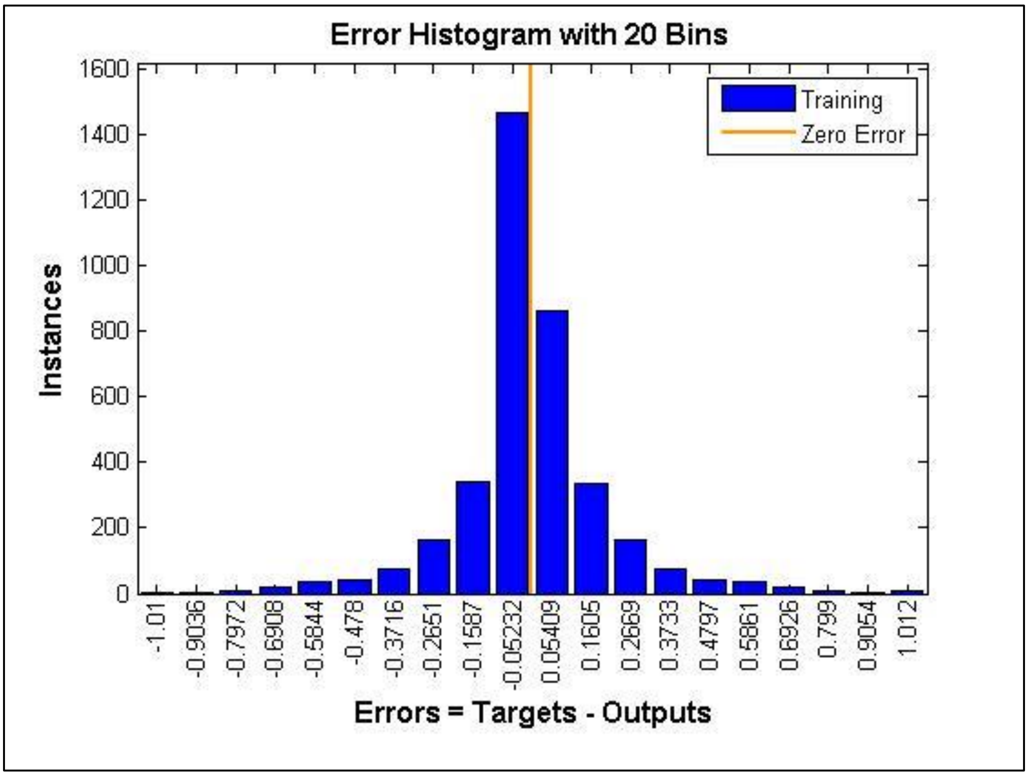

Figure 28: Error Histogram Result of MLPNN – City of London

4.1.9 Best Training Performance by MLPNN – City of London

Figure 29 plots the training errors and, in this regard, the amount of MSE was considered. The amount of MSE for each train steps is shown in the figure. From the figure it is observed that the best performance was obtained at epoch 13 with MSE equal to 0.0399. As seen in Figure 29, the MSE of the network started from the top value and decreased to the minor value and means that the MLPNN learning was improving (Siamidoudaran and Iscioglu, 2019).

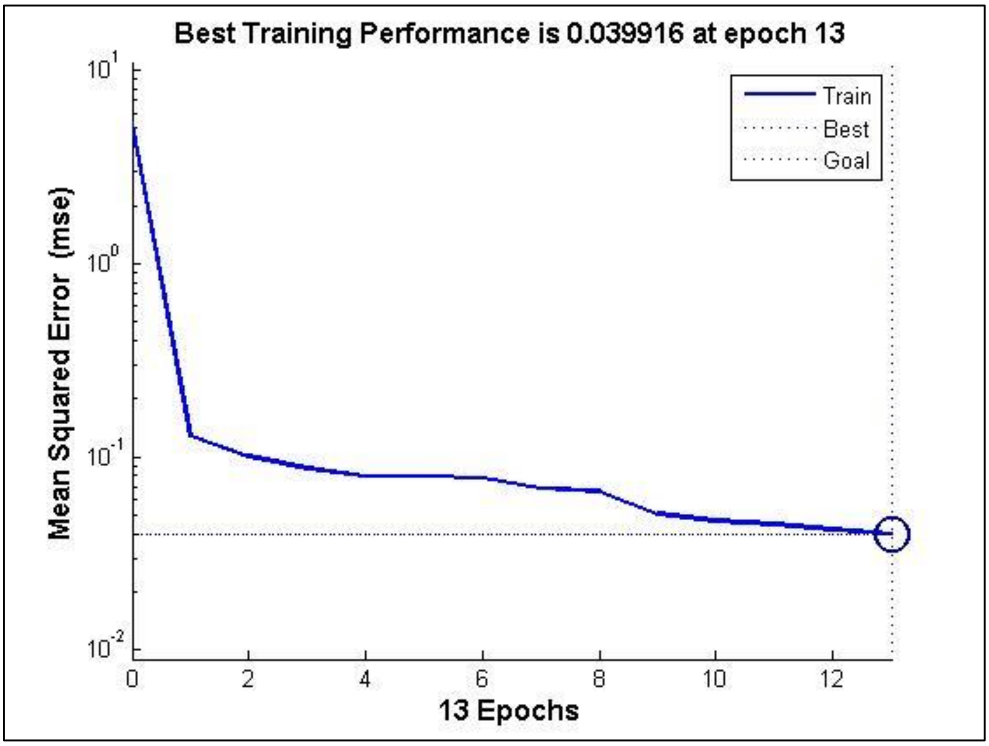

Figure 29: Best Training Performance MLPNN – City of London

4.1.10 Error Matrix by MLPNN – City of London

Confusion matrix is always used in machine learning field specifically for classification problem. Using this matrix, performance of the prediction model is well-defined on the set of the test data for which the true values are recognised (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). The error matrix resulting from the train and test phases of MLPNN are shown as the specific Table layouts in the below in that allows understand the model performance. The rows of the matrix show the instances in the predicted classes of the injury severity. In other hand, the columns signify the instances in actual classes (or vice versa) (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

$$\text{Confusion matrix for the train data} = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 6 & 121 & 0 \\ 0 & 19 & 789 & 136 \\ 0 & 0 & 104 & 1289 \end{bmatrix}$$

Total accuracy = 93.75 %, MSE = 0.0399, RMSE = 0.1998 ;

$$\text{Confusion matrix for the test data } = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 49 & 0 \\ 0 & 11 & 317 & 71 \\ 0 & 0 & 37 & 571 \end{bmatrix}$$

Total accuracy = 75.38 %, MSE = 0.1979, RMSE = 0.4448 .

Table 6 simply shows the summary of MSE, RMSE, total accuracy (ACC) along with confusion matrix for the train data and test phase of MLPNN.

Table 6: Summary of MSE, RMSE, total ACC, and error matrix by MLPNN

Prediction results in the training data set				
MSE	RMSE	ACC (%)	Error matrix	Class
0.0399	0.1998	93.75	$\begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 6 & 121 & 0 \\ 0 & 19 & 789 & 136 \\ 0 & 0 & 104 & 1289 \end{bmatrix}$	B1 B2 B3 B4
Prediction results in the testing data set				
MSE	RMSE	ACC (%)	Error matrix	Class
0.1979	0.4448	75.38	$\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 49 & 0 \\ 0 & 11 & 317 & 71 \\ 0 & 0 & 37 & 571 \end{bmatrix}$	B1 B2 B3 B4

4.1.11 Sensitivity, Precision, Accuracy and Error – City of London

Based on extracted confusion matrix from MLPNN, the amounts of SEN (sensitively), PRE (precision), ACC and error for each class and each set of data are presented in Table 7. The amount of error for train and test data were obtained equal to 06.25% and 24.62%, respectively. The results show that the best classes are numbers 3 and 4 (slight injury and damage) because of the sufficient number of the data for these classes in training phase. The detailed results of the performance for other classes are defined in the Table below.

Table 7: The SEN, PRE, ACC and error for each class by MLPNN

Values (%)	**Train Data**				**Test Data**			
	Fatality	Serious	Slight	Damage	Fatality	Serious	Slight	Damage
SEN	NaN	65.08	91.29	92.49	NaN	16.09	82.14	87.07
PRE	18.57	63.74	96.42	98.42	0	20.01	80.22	84.21
ACC	93.75				75.38			
Error	06.25				24.62			

4.1.12 Comparison of Actual and Predicted Classes of MLPNN

The predicted results by the training and testing phases of MLPNN for the city of London case study are interpreted in the next Figures in which allow visualisation of the model performance in different injury severity classes as well as comparison between actual and predicted classes. The obtained findings are breakdown by different colours in each level.

4.1.12.1 Interpretation of the Analysis

The blue symbols imply the actual levels of injury severity of the data and the pink symbols show the predicted classes resulting from MLPNN. The interpretation of the analysis specifies that, if the pink symbols are integrated with the blue marks, the MLPNN was successful to predict the severity of injury outcomes with highest accuracy. Against this background, if there is no integration, the situation specifies that the model was failed to predict the outputs of train or test data, or the model made less accurate predictions. Figures 30 and 31 also display comparison of the actual and the predicted target values of the injury severities in the train data and test data. The results show that best performance obtained through classes number B3 (slight injury) and B4 (damage only) as the number of the data in these classes was more than enough for tanning process of network. At that point, the prediction accuracy was satisfactory for B2 which implies the serious injury severity class. On the other hand, due to the insufficient number of the data for B1, the network was not able to attain a good performance in predicting for fatal class (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al.,2019b).

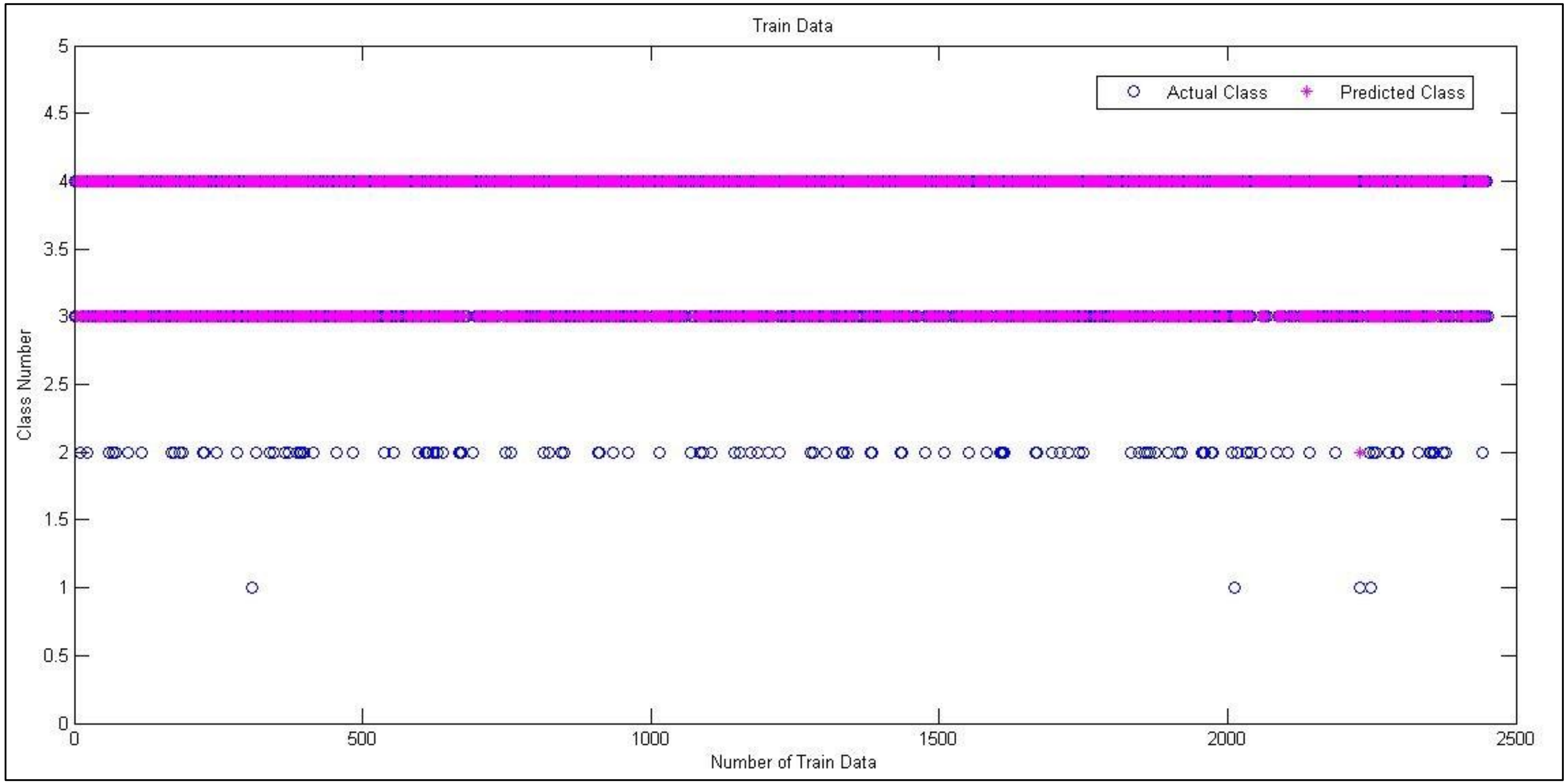

Figure 30: Actual and Predicted Classes of Injury Severity by MLPNN – Train Data

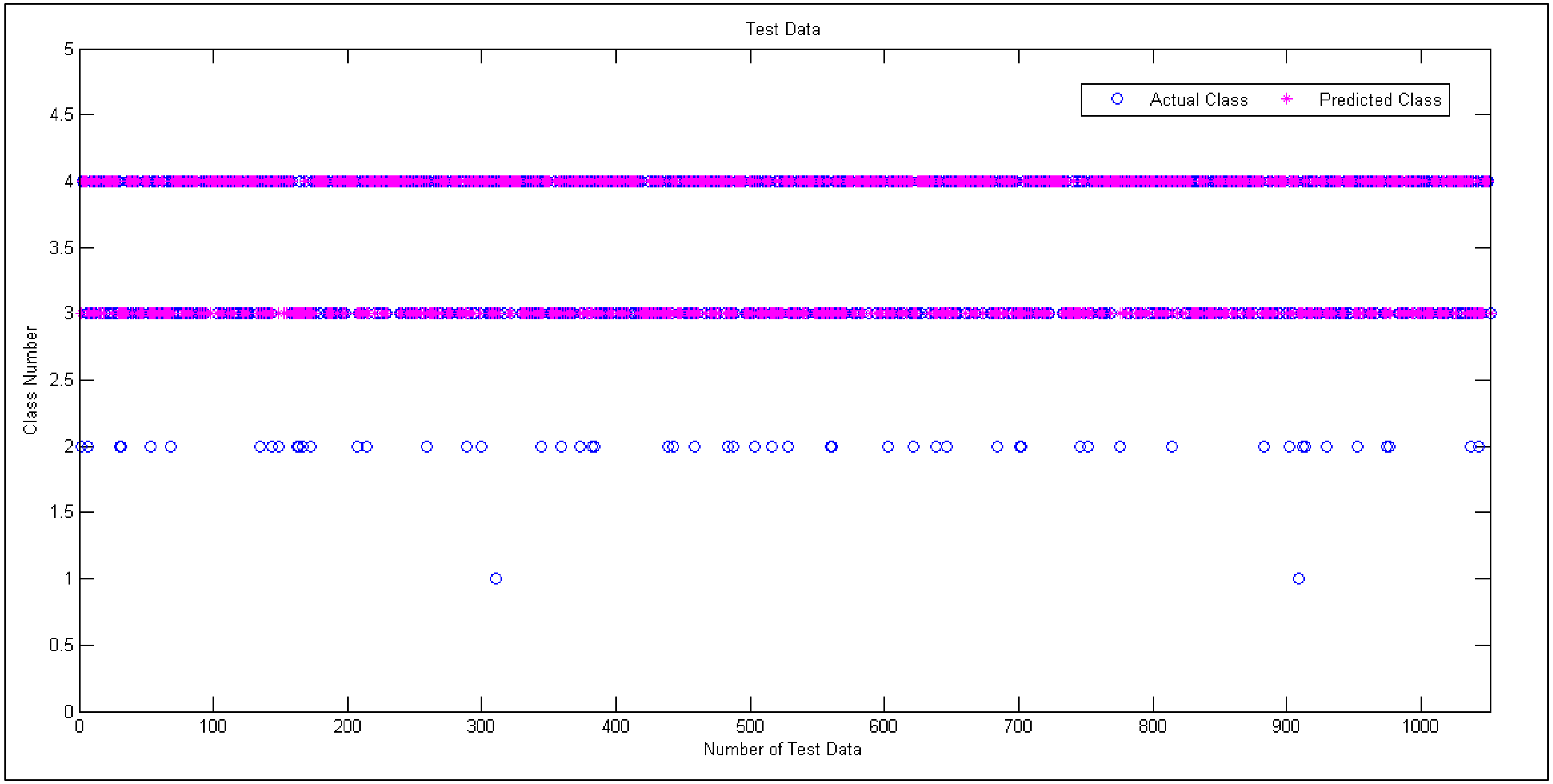

Figure 31: Actual and Predicted Classes of Injury Severity by MLPNN – Test Data

4.2 Cambridge City Case Study

As a result of higher accident injury severity involvement in the previous STATS19 prediction and potential of using data analysis in relation to bicycle–motor vehicle collisions, this section of thesis was observed for the two-wheels. Therefore, this case study only focuses on developing of cyclist injury severity prediction in Cambridge city**.**

4.2.1 Data Cleansing & Analysis Methodology

Data cleaning was also carried out by MATLAB programing language for Cambridge data due to the large and complex factors to adjust coarse factors as well as to detect preliminary unreliable data (Han and Kamber, 2001). As a result of this stage, quality of the data was improved and formatted in more appropriate figure in order to use for prediction tasks. Data on road safety in the UK are typically based on traffic crashes reported to the police via the STATS19 format. This allows police forces to report all personal-injury crashes to several departments. Table 5 splits down the data into related important labels as well as percentage rate used for the outputs. In this section, the output was also classified into following factors; fatal, serious injury, slight injury, and damage only collision. (DfT, 2011; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.2.2 Analysis of Rank Ordered Data by RBFNN – Cambridge Case Study

To determine the sensitive predictors as well as to rank the most important factors, error of each input factor resulting from RBFNN prediction model was measured. In reality, the sensitivity of RBFNN against the absence of each predictor for the presence of all factors was compared. As a result of applying the most important factors to the final prediction, the accuracy of prediction designed for the cyclist injury severities will be increased.

4.2.2.1 Implementation of Rank Analysis – Cambridge Case Study

In order to carry out the rank analysis, the factors listed in the Table 2 were applied to the model. All the factors were normalised between 0 and 1 and run was done by random division of 70 % and 30 % into training and testing datasets. As a result of the rank analysis, the most operative inputs were identified using RBFNN prediction model. The model was able to find how changes in one injury incident factor contribute the injury severity classes. Accordingly, each input factor was changed in a range related to its domain while retaining all other labels fixed, and then evaluating the change in the outcome. Likewise, the influence of each crash related factors on the injuries were compared and significant factors on the prediction were recognised. At the same time, error of each input factor was evaluated, and unreliable factors were eliminated in order to increase the final prediction accuracy.

4.2.2.2 Rank Analysis Evaluation – Cambridge Case Study

Similar to the previous case study, MSE, RMSE, and R–value measured the strength and the direction of the association between injury severity class of cyclist and the injury severity impact factors (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). Figure 32 interprets the graphical representation for the output in relation to overall training response. In this phase, the results of the strength measurements are as; MSE = 0.156, RMSE = 0.395, and R–value is around 0.742 for the total response, the values of μ and σ also is also determined. More information about the below figures (Figures 32, 33, and 34) are given in the explanation of Figure 25.

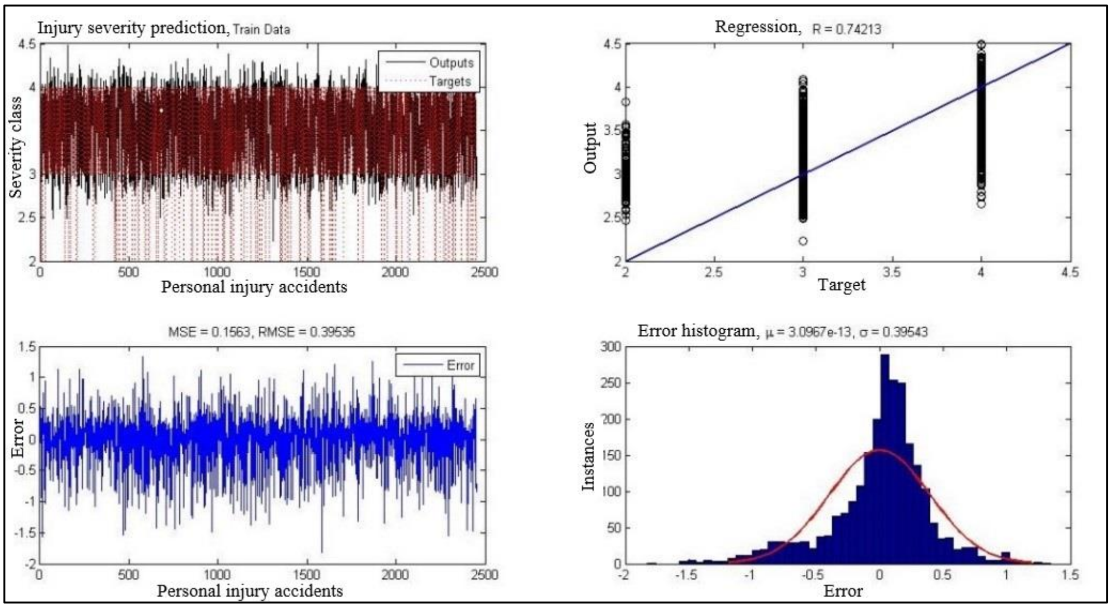

Figure 32: Sensitivity by Train Data Phase – Cambridge Case Study

Figure 33 depicts the sensitivity plots with respect to test data. In this stage, effects of iteration on MSE = 0.207, RMSE = 0.455, and R–value is 0.692.

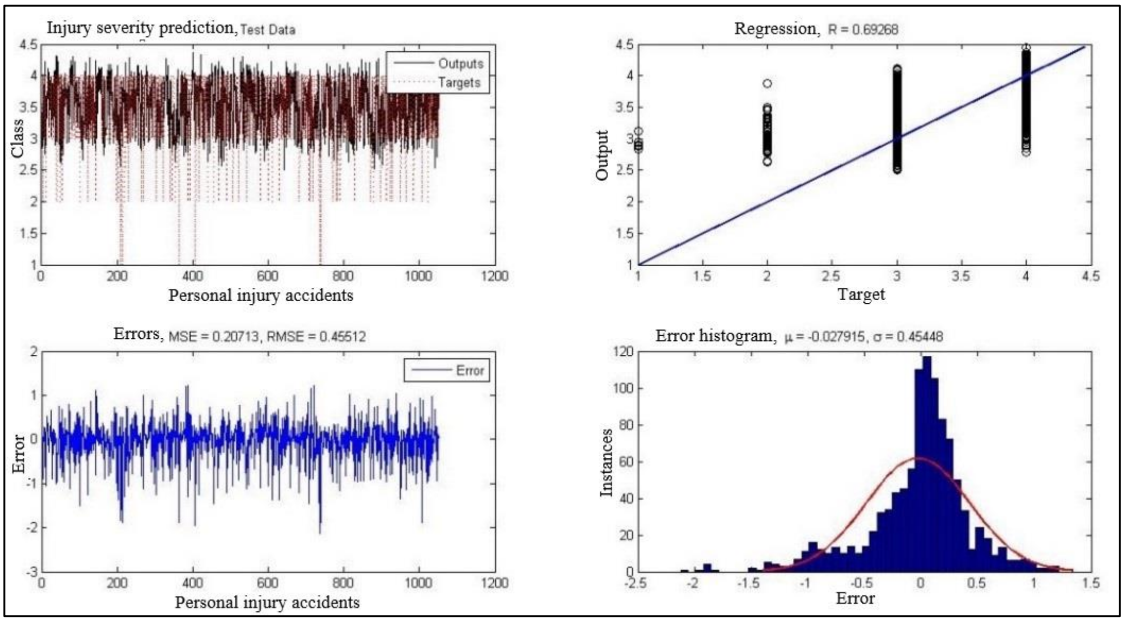

Figure 33: Sensitivity by Test Data Phase – Cambridge Case Study

Figure 34 shows the rank analysis plots for the output for total data in which MSE = 0.171, RMSE = 0.414, and R–value is 0.726.

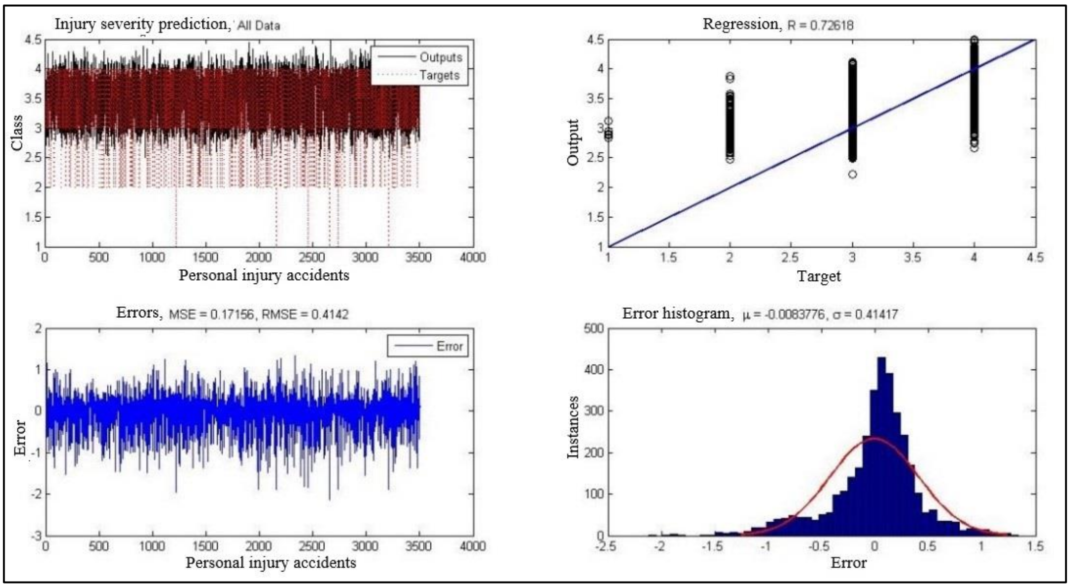

Figure 34: Sensitivity by Total Response – Cambridge Case Study

In general, the above figures (Figure 32, Figure 33, and Figure 34) display that the model achieved a satisfactory performance as the values of MSE and RMSE are small and the R–value is close to one. This means that most of the data has been accurately fitted. However, due to lack of the data for 'fatal' class and limited data for 'serious injury' class, the model totally failed to predict the first class and had many incorrect predictions for the second class. In fact, they could not be accurately fitted due to the insufficient data for these classes. This shortfall was also recognised in the first case study which must be focused on in a different task to resolve the incorrect predictions. On the other hand, the predictions of 'slight' class and 'damage only' class were satisfactory which resulted in achieving acceptable R–value due to sufficient data for these classes (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

4.2.3 Results of the Rank Analysis – Cambridge Case Study

In general, R–value is considered to measure the correlation between the predictors for each injury related factor (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). Based on the outcomes of the RBFNN against the absence of each variable, affect of each factor was recognised and listed in Table 8. The results of the factors after A30 are not mentioned on the Table due to their insignificant effect. The factors from A01 until A30 are included; bicycle manoeuvre, time, junction location of bicycle, vehicle manoeuvre, physical crossing facilities, 2nd class, junction detail, junction location of vehicle, bicycle location, day, 1st point of impact–bicycle, junction control, vehicle type, weather condition, light condition, 1st point of impact–vehicle, road surface, speed limit, road type, journey purpose, driver sex, cyclist sex, 1st road class, age of driver, number of vehicles, month, skidding, carriageway hazards, cyclist age, and IDM.

Table 8: Rank analysis results by RBFNN – Cambridge data

Rank	Predictors	**R–value** (%)		
		Train data	**Test data**	**All data**
A01	Bicycle location	84.22	81.14	84.45
A02	Bicycle manoeuvre	82.04	77.02	81.62
A03	Junction location of bicycle	80.65	78.35	80.95
A04	Vehicle manoeuvre	80.16	75.24	78.97
A05	Pedestrian crossing–physical facilities	70.17	64.11	68.14
A06	2nd Road class	63.52	54.95	57.14
A07	Junction detail	61.01	52.01	51.34
A08	Junction location of vehicle	60.27	51.67	49.34
A09	Time	54.41	52.95	43.17
A10	Day	29.15	26.19	27.91
A11	1st Point of Impact – bicycle	26.34	25.99	26.51
A12	Junction control	24.42	22.15	23.19
A13	Vehicle type	25.52	24.17	25.01
A14	Weather condition	25.34	23.25	24.95
A15	Light condition	23.84	21.95	22.90
A16	1st Point of impact – vehicle	17.09	16.04	17.34
A17	Road surface	15.94	14.14	15.24
A18	Speed limit	11.87	10.45	10.84
A19	Road type	10.77	08.54	08.99
A20	Journey purpose	09.61	8.51	08.67
A21	Driver sex	07.55	07.6	07.2
A22	Cyclist sex	07.45	07.4	07.1
A23	1st road class	07.34	07.1	07.0
A24	Age of driver	06.94	06.6	06.2
A25	Number of vehicles	05.42	05.64	05.88

A26	Month	04.06	03.44	04.24
A27	Skidding	04.03	03.14	03.55
A28	Carriageway hazards	03.88	03.48	03.58
A29	Cyclist age	03.21	02.45	03.19
A30	Driver IMD	03.08	02.06	02.59

4.2.4 Most Important Contributory Factors – Cambridge

As the result of the prediction in this stage, the most important factors/areas of concern which contributed to the severity of cyclist injuries were discovered. At this stage, R-value measures the strength and direction of the relationship between variables. It ranges from -1.0 to +1.0. If the value is close to 1, it means there is a good relationship but if it is close to 0, it means there is no relationship between the variables. As seen in Table 8, R–values for A01, A02, A03, A04, A05, A06, A07, A08, and A09 are 84.45%, 81.62%, 80.95%, 78.97%, 68.14%, 57.14%, 51.34%, 49.34%, 43.17%, respectively. These values are almost between 0.84 and 0.43 which is more closer to 1 and are almost triple/double compared to other factors (A10 to A19) and so they were considered as the important contributory factors for the Cambridge case study. On the other hand, A20 to A30 have been identified as very poor predictors and played very insignificant roles in the outcome of injuries. Table 9 in this thesis offers detailed information about the subdivisions of the contributory factors from A01 to A19. In addition, STATS20 is the designation of a document entitled ‘instructions for the completion of road accident reports’ published by DfT (2011). This document defines the scope and meaning of each input factor used in this study in more detail.

4.2.4.1 Bicycle Location (A01) – Most Important Cycling Factor

A01 refers to bicycle location at the time of accidents (restricted lane/away from main road). For example, cycle lane wherein lane (advisory or mandatory) marked off within main roadway for use by bicycle only. Or cycleway which is forms part of the road but is not part of the main carriageway (DfT, 2011). Cycling across a certain lane (A01) had a massive influence in increasing the risk of injury severities (TRL, 2009; RoSPA, 2017c; DFT, 2018b; Siamidoudaran et al., 2019b). Although cycle facilities are used on Cambridge roads, but it appears there are still limited modern protected bike tracks associated with these locations or the existing sequences are not standardised. It seems that the painted bike lanes are not enough to protect riders but instead cause drivers to pass closer (e.g. Figure 49). The routs that are not physically separated from other traffic significantly decrease the space drivers provide because motorists do not feel the need to give cyclist space when they have their own way. Therefore, where the rider is in the same lane as the driver, the driver must make an overtaking. This circumstance is a commonly identified factor which contributes to the risk of cyclist injury wherein riding across the main carriageway and not in restricted lanes had a huge effect in increasing the numbers of injuries (TRL, 2009; RoSPA, 2017c; DfT, 2018b). On the other hand, where there are bicycle-specific infrastructure alongside other road users, the driver has a clear lane ahead to pass (Siamidoudaran et al., 2019b).

4.2.4.2 Making Space for Cycling

To improve safety of cyclists, narrow cycle lane defenders can deliver cycle lane segregation and are safer for riders because they offer continuous or light segregation by excluding other traffic from the bike lane (Siamidoudaran et al., 2019b; Rosehill Highways, 2020).

4.2.4.3 Cyclist Manoeuvre (A02)

A02 applies to any manoeuvre made by the cyclist which happened, or contributed to, the crash. Examples include changing lanes, overtaking reversing, U-turn, and turning right or left. Table 9 in this thesis lists more possible cyclist manoeuvres (DfT, 2011). This identified factor is in line with many studies where cyclists are far more likely to suffer when a motorist is waiting to turn in an incorrect place or when they made an unsafe left or right 'hook' turn suddenly across the path of a cyclist who is riding straight towards a junction on a bend ahead (TRL, 2009; DfT, 2015b; Siamidoudaran et al., 2019b).

4.2.4.4 Junction Location of Bicycle (A03)

A03 also played a large role in the injuries around and in junctions with collisions happening more frequently here than in any other zones in the city (SiamiDoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). ‘Junction location of bicycle’ refers to a position on the road where each crash occurred (e.g. cyclists waiting to enter the roundabout/main road). Cyclist absolute geographic location is defined by two coordinates, longitude and latitude. The labels of this variable (analysed in Table 9) indicate the position of pedal riders at or nearby busy junctions (DfT, 2011). In this vein, the UK government reported that many cyclists are injured around and at intersections, in particular, more than half of bicycle-MVCs happen at intersections in the UK, both the rider and the driver may be at fault (Knowles et al, 2009; DfT, 2018b).

4.2.4.5 Vehicle Manoeuvre (A04)

Vehicle manoeuvre (A04) contributed to the injuries which deal with action immediately prior to an accident A04 was caused due to poor driving manoeuvre nearby busy junctions (Siamidoudaran et al., 2019b). The most important configurations between bicycle–MVCs is the vehicle turning right or left while the rider is cycling straight ahead (DfT, 2015b). A similar outcome was also found by RoSPA (2017) which identified 'poor turn/manoeuvre' by drivers as a common contributory factor. Another finding by DfT (2018b) displays that between 2011 and 2016, 46% of the riders' KSI casualties that happened at T or staggered junctions occurred by the rider 'cycling ahead' and the driver turning right or turning left. In addition, UK police forces arbitrated stationary or parked vehicle as a contributory factor in accidents involving bicycle and motor vehicles for the above period (DfT, 2018b). This situation can reduce visibility especially at a junction onto a main road. Most likely, drivers or passengers in stationary or parked vehicles incautiously open doors without looking and slams it against a passing cyclist (TRL, 2009; DfT, 2015b; DfT, 2018b; Siamidoudaran et al., 2019b).

4.2.4.6 Crossing Facilities (A05)

A05 refers to pedestrian crossing facilities including shared facilities with bicyclists. This factor indicates that the local authority services were not able to make existing crossings safer for the road users (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). Injuries involving cyclists occurred frequently at facilities designed for cyclists (e.g. pedestrian crossings, cycle lanes, and cycle tracks). This means that these facilities were not necessarily good enough to avoid accidents in Cambridge. This factor was also identified as the main contributory factor for the previous case study. In this regard, Chang and Wang (2006) suggested that cyclists are at a superior danger of severity in injuries around crossing facilities. Furthermore, a road safety research showed that more than half of the KSI crashes which cyclists are also involved in happen while crossing streets (SWOV, 2010). This outcome is also in line with a previous research displaying that the increase of the crossing facilities at crossings have an extra positive safety outcome in dropping the number of the cycling injuries (Gårder et al., 1998).

4.2.4.7 Factors Related to Junction Actions (A06-A08)

Again, factors from A06 to A08 refer to some kind of junction actions including '2nd road class' (A06), 'junction detail' (A07), and 'junction location of vehicle' (A08) (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). These factors have also been identified as the most important contributory factors in the city of London case study. The explanation of the factors is given in the results section of the previous study. According to many studies, pedal riders are more often involved in accidents while trying to cross a multilane road. This finding fits many studies, while official government reports also indicate that many cyclists are injured at or nearby intersections during recent years (TRL, 2009; SWOV, 2010; RoSPA, 2017; DFT, 2018b; Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b). For that reason, large amount of bicyclist deaths occurred at these locations (Knowles et al, 2009; DfT, 2018b).

4.2.4.8 Improving Junction Safety

Like previous case study, truly protected junctions can offer specific cycle-signals which aid those on bikes to pass straight ahead or to carefully turn right, within the segregated bike lane (RoSPA, 2017a; RoSPA, 2017b; RoSPA, 2017c; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020; DfT, 2020).

4.2.4.9 Time (A09) and Day (A10)

Next factors (A09 and A10) show that the riders were far more expected to suffer during specific time bands and days (Siamidoudaran et al., 2019b). According to the data analysis in Table 9, the weekday peak time for cyclist injuries is from 7am to 9am and from 4pm to 6pm. This predicament is an injury severity result which is consistent with the UK government's official report which found that almost majority of cyclist injuries take place during rush hours on weekdays (DfT, 2015b). This finding is highly accurate since one in four Cambridge residents cycle to work during the identified hours. This amount is the greatest level of cycling to work in the UK. In addition, a very high proportion of the students hop into the saddle to get to their universities and colleges in Cambridge (CambridgeshireLive, 2018a; CambridgeshireLive, 2018b).

4.2.4.10 Other Cycling Predictors

Factors from A11 to A19 associated with increased risk of injuries including; 1st point of impact by bicycle, junction control, vehicle type, weather condition, light condition, 1st point of impact as a result of vehicle, road surface, speed limit, and road type, however, as seen in Table 9 their influences are not as great as A01 to A10 were (TRL, 2009; SWOV, 2010; RoSPA, 2017; DfT, 2018b; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.2.4.11 Unreliable Cycling Predictors

The variables from A20 to A30 are located at the bottom of Table 9 and have been shown to play insignificant roles on the outcome of injury severities while keeping their affects nearly equal (Knowles et al, 2009; DfT, 2018b; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). These factors comprise of journey purpose, driver sex, cyclist sex, 1st road class, age of driver, number of vehicles, month, skidding, carriageway hazards, cyclist age, and driver IMD (Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b).

4.2.5 Maximising Predictive Accuracy by SVM – Cambridge

In order to generalise the dimensional feature space before processing the data into the SVM network through MATLAB, poor predictors were eliminated. As a result of the previous prediction by RBFNN intended for large number of subdivision variables, the labels after A19 were removed due to their minor significance. Also related predictors from A1 to A19 were discovered as the most significant contributory factors. Following this, the sensitive predictors were applied to SVM as input factors in order to get a higher prediction accuracy (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

These factors are included; bicycle location, bicycle manoeuvre, junction location of bicycle, vehicle manoeuvre, pedestrian crossing facilities, 2nd road class, junction detail, junction location of vehicle, time, day, 1st point of impact by bicycle, junction control, vehicle type, weather condition, light condition, 1st point of impact by vehicle, road surface, speed limit, and road type.

4.2.6 Analysis of Key Predictors – Cambridge

Predictors from A1 to A19 were discovered as the most significant predictors. The descriptive statistics of these predictors are broken down in the Table 9. The Table shows the count and percentage distribution across levels of each cluster (e.g. the distribution among 20, 30, 40, 50, and 60 mph speed limits) and the injury severity classes (fatal injury, serious injury, slight injury, and damage only) for the total frequency (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Table 9: Description of contributing factors – Cambridge case study

Input	Subdivision	Variable	Code	Sub-variables	Total (%)
1	6	Bicycle location	1 2 3 4 5 6	Footway Bus lane Not in restricted lane (main Rd) Busway (including guided busway) Cycle lane (on main Rd) Cycleway or shared use F/W (not part of main Rd)	03.23 00.82 84.01 00.19 09.69 02.06
2	9	Bicycle manoeuvre	1 2 3 4 5 6 7 8 9	Parked Reversing Waiting to go - held up Slowing or stopping Moving off Turning / Waiting to turn Changing lane to left/right Overtaking vehicle Going ahead bend	00.05 00.00 01.88 02.95 03.45 11.02 01.82 11.09 67.74
3	5	Junction location of bicycle	1 2 3 4 5	Approaching junction Mid Junction / on roundabout Leaving / Entering roundabout Leaving / Entering main road Not at or within 20 metres of junction	17.22 46.12 04.02 04.06 28.62
4	9	Vehicle manoeuvre	1 2 3 4 5 6 7 8 9	Parked Reversing Waiting to go Moving off Slowing Changing lane Overtaking vehicle Turning / Waiting to turn Going ahead bend	09.17 01.04 02.77 11.02 03.10 01.03 07.98 41.03 22.86
5	6	Physical crossing facilities	1 2 3 4 5 6	Pelican, puffin, toucan or similar non-junction light No physical crossing facilities within 50 metres Pedestrian phase at traffic signal junction Footbridge or subway Central refuge Zebra	18.25 67.25 08.25 00.05 02.36 03.84
6	6	2st Rd class	1 2 3 4 5 6	Not at junction or within 20 metres Motorway (including A(M) Rd) A B C U	00.00 00.01 09.23 00.42 17.23 73.11
7	7	Junction detail	1 2 3 4 5 6 7	Slip road Roundabout Crossroads T or staggered junction Private drive or entrance Other junctions Not at junction or within 20 metres	00.15 19.14 07.53 41.25 07.25 01.66 23.02
8	5	Junction location of vehicle	1 2 3	Mid Junction - on roundabout Leaving/Entering roundabout Leaving/Entering main road	20.59 13.04 27.26

			4 5	Approaching junction or waiting in a queue Not at or within 20 metres of junction	15.20 23.91
9	4	Time	1 2 3 4	07:00–09:00 and 16:00–18:00 09:01–15:59 18:00–22:00 22:01–06:59	40.36 32.89 21.45 05.03
10	2	Day	1 2	Weekday Weekend	85.65 14.35
11	5	1st Point of Impact – bicycle	1 2 3 4 5	Front Back Offside Nearside Did not impact	42.12 13.32 24.55 16.99 03.02
12	5	Junction control	1 2 3 4 5	Stop sign Authorised person Auto traffic signal Give way or uncontrolled Not at junction or within 20 metres	00.43 00.35 10.79 88.43 00.00
13	3	Vehicle type	1 2 3	Car Motorcycle Buses, lorries and goods	79.23 01.03 19.74
14	2	Weather condition	1 2	Inclement Good	19.95 80.05
15	2	Light condition	1 2	Daylight Darkness	69.22 30.78
16	5	1st Point of impact – vehicle	1 2 3 4 5	Front Back Offside Nearside Did not impact	43.25 05.36 16.87 29.28 05.24
17	2	Rd surface	1 2	Wet Dry	23.19 76.81
18	5	Speed limit	1 2 3 4 5	20 30 40 50 60	09.86 89.23 00.41 00.23 00.04
19	4	Rd Type	1 2 3 4	Single carriageway Dual carriageway One-way street Roundabout	73.17 01.98 02.24 22.61
Output	4	**Injury severity class of cycles**	B1 B2 B3 B4	**Fatal injury** **Serious injury** **Slight injury** **Damage only**	01.86 14.78 39.80 43.56

4.2.7 Implementation of SVM Applying Key Predictors – Cambridge

In this section of the thesis, SVM prediction model was fit on training, validation, and testing subsets. For this prediction task, random division was used for 70% of the data in training. For validation phase 15% of data was used with intention of delivering an unbiased assessment. Lastly, the remaining 15% was considered in the test phase.

As seen in Figure 35, the second layer produced thirty neurons to discover the relationship between the inputs and outputs. Each neuron is related to each input factor and the weights associated with each predictor gives statistics about its weight for the discrimination of the levels. Accordingly, the hyperplane is implicitly defined in a higher dimensional space through the kernel trick. Again, like MLPNN model, the LM–BP algorithm was applied in the training phase. MSE was considered for evaluating the model performance (Siamidoudaran and Iscioglu, 2019). The structure of the network and training process are shown in Figure 35.

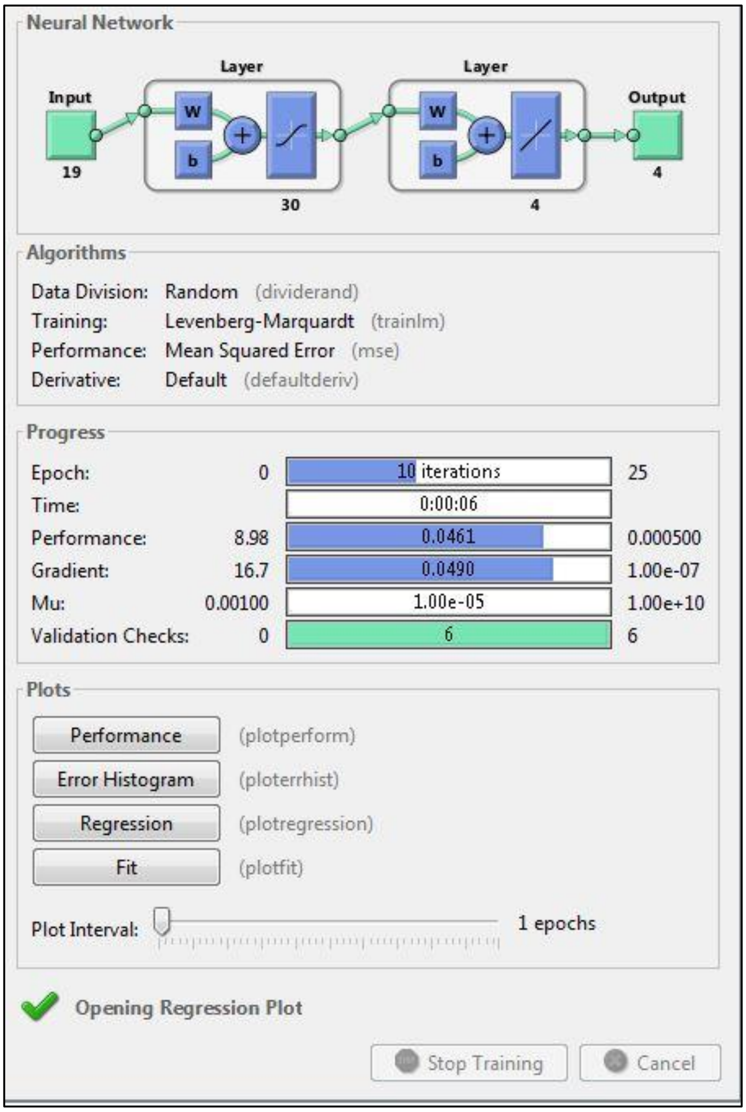

Figure 35: Final Structure and Train Process of the SVM

In the training phase, the network was discontinued after 10 iterations with MSE result of 0.0461 and gradient equal to 0.0490 as shown in Figure 35. As a result, the network was able to balance among overfit (very small MSE value for training) and underfit (great value of MSE for test or validation data). Therefore, the model should be suitable to predict the test data that we haven't yet seen. In fact, the indication of MSE isn't equal to zero, since then we would have a model that entirely predicts the training data, but which is impossible to completely forecast any other data (Siamidoudaran and Iscioglu, 2019).

4.2.8 Error Histogram by SVM – Cambridge

The below error histogram with 20 bins were achieved using SVM. The error histogram shows how to represent errors between target values and estimated values. In line with Figure 36, it is evident that the histogram has a peak about 0.04 that displays an accurate prediction for the model (Siamidoudaran and Iscioglu, 2019).

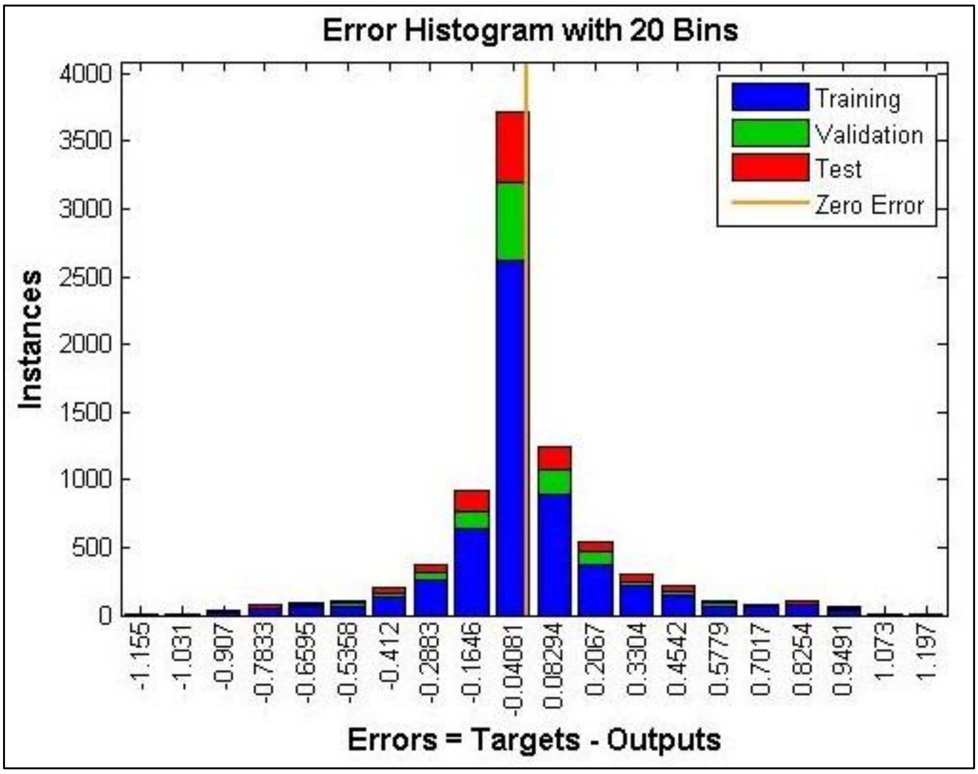

Figure 36: Error Histogram by SVM– Cambridge

4.2.9 Best Validation Performance by SVM – Cambridge

Figure 37 shows the training, validation, and test errors to identify the validation error in the training window. SVM structure generated the best model for predicting cyclist injury severity classes, with the best validation performance of 0.07878 at iteration of 4 as shown in Figure 37. The diagram also indicate that the validation and test curves are very close. If the outcome of test curve had particularly raised before the validation curve raised, then there would have likely been some error. As a result, it is evident that the model was able to predict the minimise error (Siamidoudaran and Iscioglu, 2019).

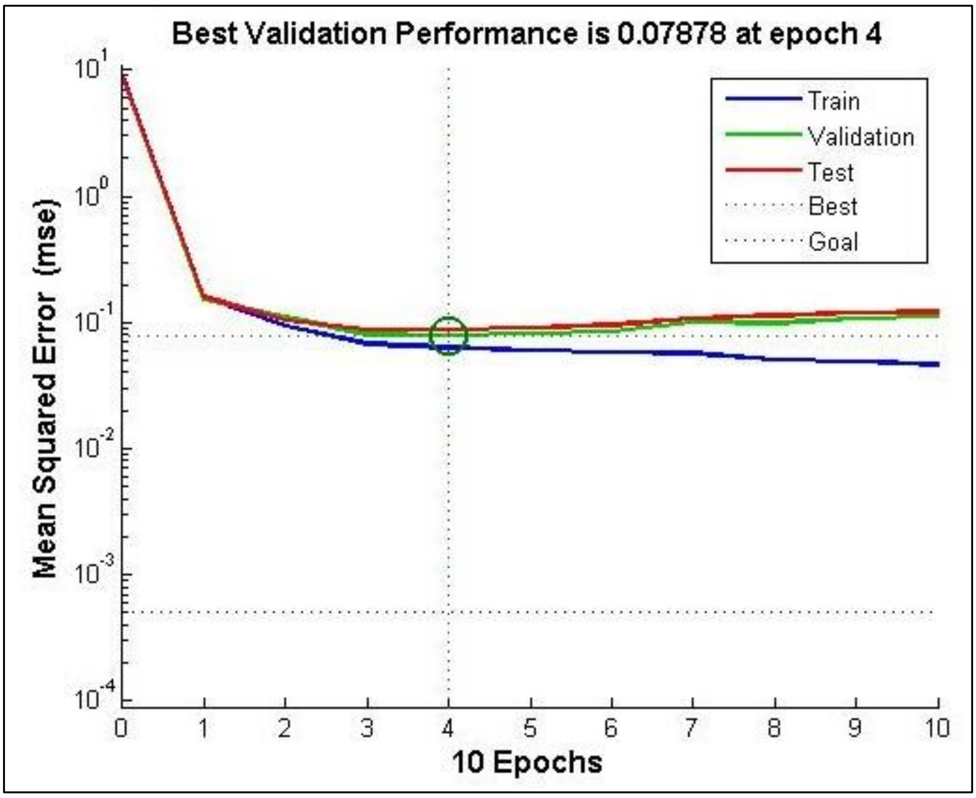

Figure 37: Best Validation Performance by SVM– Cambridge

4.2.10 Regression Results by SVM – Cambridge

Figure 38 plots the regression results by SVM for the output with in connection with training, validation, and test data. R–value was considered to understand the association between outputs and targets. An R–value close to 1 means a close association, and close to 0 is a random correlation. The performance of the SVM is presented in Figure 38 in which the regression results for training (0.815), validation (0.765), test (0.733) phases and total response of the data (0.795) (Siamidoudaran and Iscioglu, 2019).

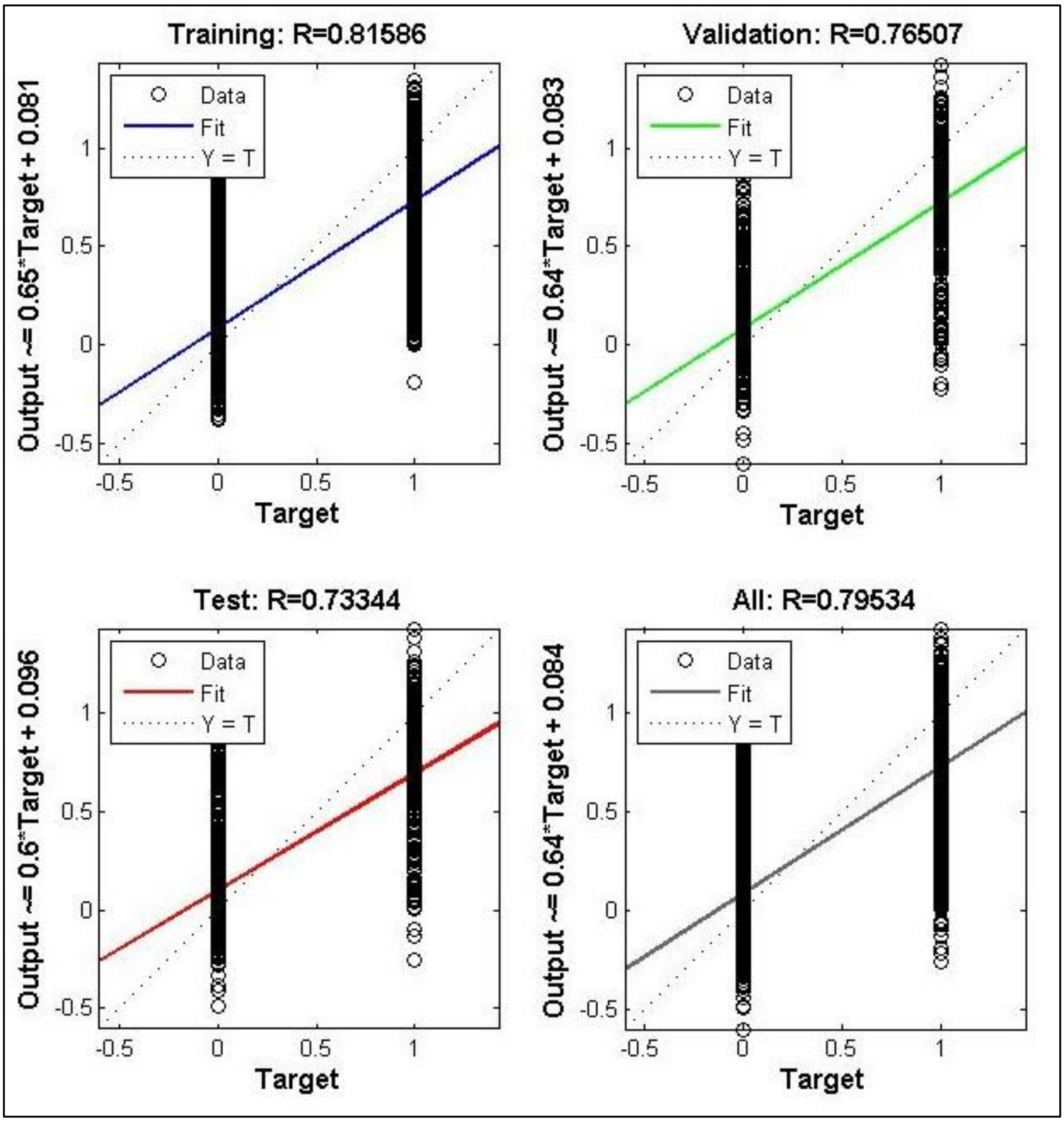

Figure 38: Regression Results by SVM – Cambridge

4.2.11 Error Matrix by SVM – Cambridge

This matrix is used as a valuable technique to summarise the performance of a classification algorithm (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). The matrix shows the various types of errors that happened for the prediction task by SVM. Visualization of the model performance is specified in the below matrix showing the true values for set of the accident data aimed at Cambridge case study. The rows display the actual classes (or output level that is gained from the dataset) and the columns indicate the predicted class (or target level that is predicted by the model after being trained). The Error matrix for the train data and test process are specified as below.

$$\text{Error matrix for the train data} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 23 & 116 & 7 \\ 4 & 56 & 931 & 41 \\ 0 & 0 & 31 & 1289 \end{bmatrix}$$

Total accuracy = 92.15 %, MSE = 0.0461, and RMSE = 0.1955

$$\text{Confusion matrix for the test data} = \begin{bmatrix} 0 & 0 & 2 & 2 \\ 5 & 37 & 61 & 2 \\ 0 & 52 & 353 & 31 \\ 0 & 0 & 31 & 524 \end{bmatrix}$$

Total accuracy = 72.85 %, MSE = 0.1911, RMSE = 0.4422 .

Table 10: Summary of MSE, RMSE, total ACC, and error matrix

Prediction results in the training data set				
MSE	RMSE	ACC (%)	Confusion matrix	class
0.0461	0.1955	92.15	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 23 & 116 & 7 \\ 4 & 56 & 931 & 41 \\ 0 & 0 & 31 & 1289 \end{bmatrix}$	B1 B2 B3 B4
Prediction results in the testing data set				
MSE	RMSE	ACC (%)	Confusion matrix	class
0.1911	0.4422	72.85	$\begin{bmatrix} 0 & 0 & 2 & 2 \\ 5 & 37 & 61 & 2 \\ 0 & 52 & 353 & 31 \\ 0 & 0 & 31 & 524 \end{bmatrix}$	B1 B2 B3 B4

4.2.12 Sensitivity, Precision, Accuracy and Error – Cambridge

In reference to the above results from SVM, the amounts of SEN, PRE, ACC and error for each level as well as for each set of the data were specified. The sum of error for train and test data were found as 07.85% and 27.15%. The outcomes display that the greatest levels refer to B4 damage only and B3 slight injury. The prediction accuracy for B2 serious injury is acceptable, however, lack of the data for B1 fatality class is still remained as the number of the data for this class wasn't sufficient for training of the network.

Table 11: The SEN, PRE, ACC and error matrix for each class

Values (%)	**Train Data**				**Test Data**			
	Fatality	Serious	Slight	Damage	Fatality	Serious	Slight	Damage
SEN	NaN	64.25	91.59	93.59	NaN	15.12	81.19	85.11
PRE	17.85	62.22	97.42	98.14	0	19.23	81.23	83.55
ACC	92.15				72.85			
Error	07.85				27.15			

4.2.13 Comparison of Actual and Predicted Classes of SVM

Moreover, the predicted outcomes of the cyclist injury severities in view of the train and test dataset are interpreted in Figure 39. The figure is able to visually compare the prediction classes for the injury severity wherein the training and testing phases are presented into each class.

4.2.13.1 Interpretation of the Analysis

The blue marks show the actual level of the data and the pink colours specify the predicted class of the injury severities. As seen in the graphical representation, the pink symbols are integrated with the blue results in class B4 (damage only) and class B3 (slight injury), it indicates higher accuracy in predicting cyclist injury severities for these classes. On the other hand, there is no much integration for class of B2 which refers to serious injury, this situation specifies that the SVM model was predicted with less accuracy. In addition, as a result of the limited data for B1 (fatality), the model failed in making correct predictions for this class. The comparison between the actual classes and the predicted classes for the train data and test data display that the greatest performances were achieved for B4 and B3. The number of the data in these levels was more than enough for tanning process of network (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

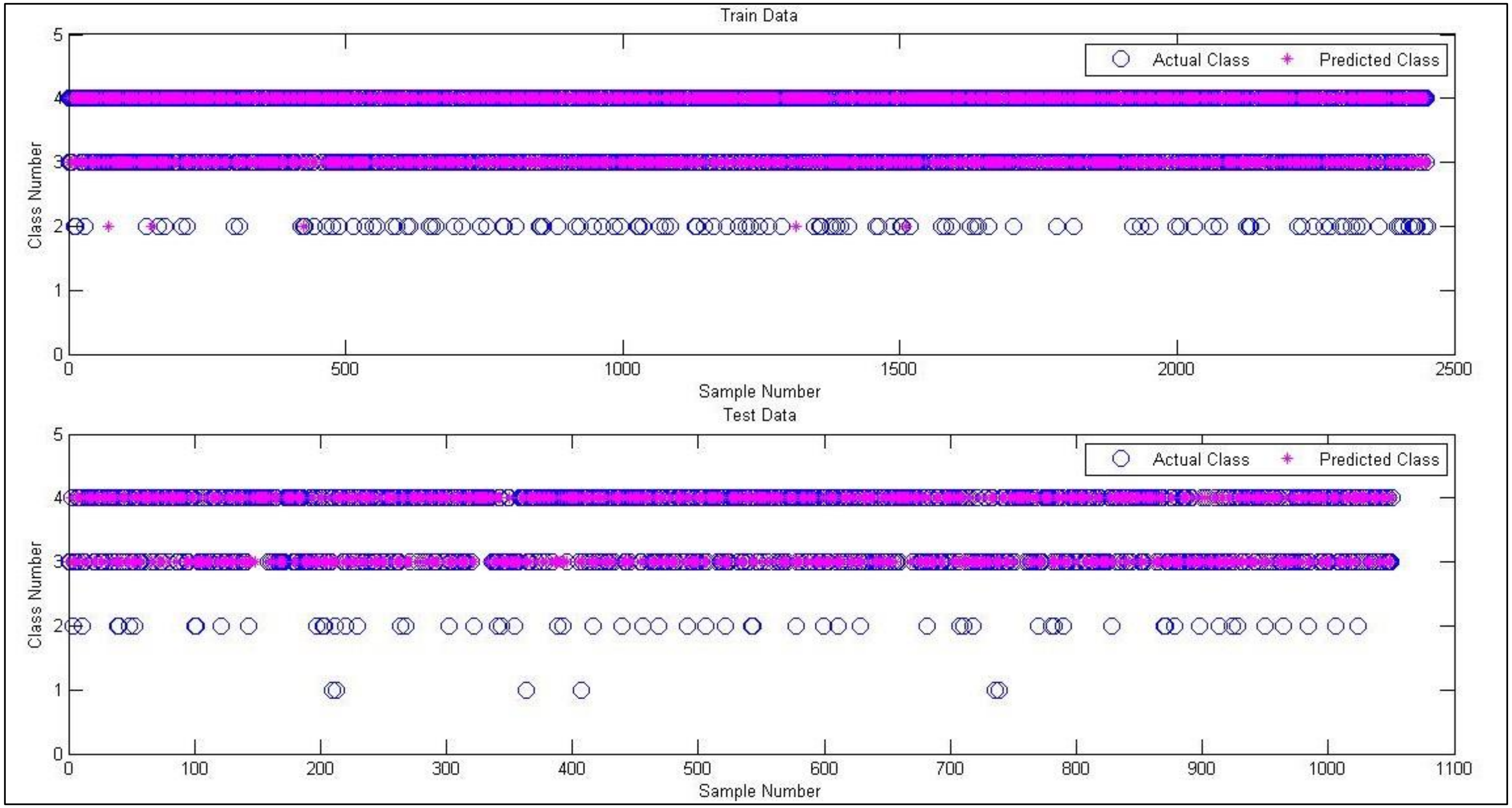

Figure 39: Actual and Predicted Classes of Cyclist Injury Severity by SVM – Train and Test Data

4.3 Comparison between MLPNN and SVM

Using STATS19 key predictors we aim at comparing the predictive performance, including prediction accuracy and error between the proposed ANN and SVM for personal injury severity classes. In this thesis, in the spotlight is the technique from the ANN family, most notably, MLPNN was considered for the comparison task due to the use of the most important contributory factors.

As two different algorithms, ANN and SVM share the similar theory using linear learning approach for pattern recognition. The main difference is mostly on how non-linear data is predicted. In general, SVM model uses nonlinear mapping to create the data linear separable, therefore, the kernel function is the significant tactic. On the other hand, ANN works multi-layer association and several activation functions to deal with nonlinear issues. Actually, single layer ANN can merely produce linear boundary, and the second layer can join the linear boundary together; while at least three layers are needed to create boundary of arbitrary forms (Ren, 2012).

At this stage, we compare the performance of MLP and SVM when applied to STATS19 data and explore the sensitivity of the models tuning of a successful MLPNN and SVM architectures. According to the obtained results from the confusion matrix obtained from MLPNN and SVM, the comparison of the actual and the predicted classes of models was applied and explained in the Table 12.

Table 12: Comparison between MLPNN and SVM using most important contributory factors of STATS19 data

Model	**Training process of ISP models**					**Testing process of ISP models**				
	ACC (%)	Error (%)	Error matrix	Severity class	SEN (%)	ACC (%)	Error (%)	Error matrix	Severity class	SEN (%)
MLPNN	93.75	06.25	$\begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 6 & 121 & 0 \\ 0 & 19 & 789 & 136 \\ 0 & 0 & 104 & 1289 \end{bmatrix}$	Fatal Serious Slight Damage	00.00 65.08 91.29 92.49	75.38	24.62	$\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 49 & 0 \\ 0 & 11 & 317 & 71 \\ 0 & 0 & 37 & 571 \end{bmatrix}$	Fatal Serious Slight Damage	00.00 16.09 82.14 87.07
SVM	92.15	07.85	$\begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 110 & 16 \\ 0 & 0 & 738 & 181 \\ 0 & 0 & 156 & 1246 \end{bmatrix}$	Fatal Serious Slight Damage	00.00 64.25 91.59 93.59	72.85	27.15	$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 53 & 1 \\ 0 & 0 & 325 & 69 \\ 0 & 0 & 76 & 525 \end{bmatrix}$	Fatal Serious Slight Damage	00.00 15.12 81.19 85.11

4.3.1 Results of Comparison between MLPNN and SVM

In the MLPNN model, the accuracy measure for ISP model in the training and testing phase were found 93.75% and 75.38%, respectively. However, the number of the fatalities was not particularly high, thus, due to the lack of data; the network wasn't able to evaluate accurate prediction for this class in the training process. Therefore, in the test phase, the amount of sensitivity for this class was equal to zero. For serious injury class, the network was unable to perform a good prediction. In fact, in the training process, the network accommodated the input parameters with slight injury class instead of serious injury class. For slight injury class, the classification was almost desirable and the amount of sensitivity for the training and test phase was obtained as 91.29 and 82.14%, respectively. Moreover, in this class, the classification was 'occasionally tended to damage only' class. Lastly, among the classes, the prediction of the injury severity for 'damage only class' was performed better than other classes (92.49% for train phase, 87.07% for test phase and 87.07% sensitivity).

In the SVM model, the amount of accuracy obtained was slightly less than MLP network (92.15% for the training phase and 72.85% for test phase). However, for fatal class, both in the training and testing process, like MLPNN model, the SVM totally failed in predicting the severity class (sensitivity equalled to zero). As for the prediction of serious injury, the SVM was unsuccessful and more data was incorrectly classified. As for 'serious injury' and 'damage only' class, compared to the two previous classes, the performance of SVM had improved and the sensitivity value for the training and testing process was obtained around 80% and 90%, respectively.

4.3.2 Performance Comparison of MLPNN and SVM

In conclusion, the performance from the two classifiers becomes very capable for predicting STATS19 data, but, it seems that MLPNN slightly outperforms SVM and the MLPNN has been able to improve the accuracy rate for predicting the injury severity by 5%. The finding fits many studies that indicated the ANN and SVM models are capable networks for understanding the nonlinear relationship between independent variables and dependent variables. In most cases, MLPNN is effectively capable of predicting injury severity classes, accident severity and accident frequency for high nonlinear data and gives better performance (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab; 2004; Delen, 2006; Alkheder et al., 2017; Shamsashtiany and Ameri, 2018; SiamiDouadarn and Iscioglu, 2019). Additionally, SVM fits data well so it is also a capable tool for prediction of injury severity data (Chang and Wangm, 2006; Li et al., 2008; Li et al., 2012; Yu et al., 2014; Li et al., 2016; Sharma et al., 2016; Yu et al., 2016; Alkheder et al., 2017; Zhang et al., 2018; SiamiDouadarn and Iscioglu, 2019; Venkat et al., 2020).

4.3.3 Using MLPNN-SVM to Achieve Better Accuracy – City of London

As a result of comparison between proposed MLPNN and SVM, the results indicated that both models are very capable in prediction of personal injury severities. However, as a result of poor prediction accuracies for serious injury and fatal classes, we combined the both algorithms in a single prediction model in order to predominantly link the output layer of an MLPNN classifier by means of optimal margin hyperplanes. Therefore, combining two powerful methods in a single model should be a great idea, importantly it has never been used by other researchers in an accident/injury prediction. There are also a few previous articles using this model for any prediction tasks. (Bellili et al., 2003; Tifani et al., 2017; SiamiDoudaran and Iscioglu, 2019).

In this section, the key predictors of STATS19 data for the city of London resulting from the rank analysis were applied into the hybrid MLPNN-SVM. MATLAB programming language was again used in the training, testing and structure algorithms of MLPNN-SVM. All the factors were normalised between zero and one. The delivered dataset was randomly separated to form two subsets; training data of 70 % and testing data of 30 % which lead to implementation of the structure optimization algorithm accompanied by the performance comparison between the injury classes. Following this, run was made using the random division. In the process of the hybrid MLPNN-SVM, at the first level, the MLPNN had reduced the dimension space of input data, and it led to facilitate the process of prediction for the second layer (SVM). The obtained results from the hybrid model are shown in the below Table using confusion matrix. For evaluation of the model, in addition to ACC, error parameters, and sensitivity SEN were used to show the rate of correct positive prediction of each injury severity classes (SiamiDoudaran and Iscioglu, 2019).

Table 13: Confusion matrix for hybrid model (SiamiDoudaran and Iscioglu, 2019)

Training results of hybrid MLPNN-SVM				
ACC	Error	Confusion matrix	Severity class	SEN
91.3%	08.6%	$\begin{bmatrix} 3 & 1 & 0 & 0 \\ 2 & 14 & 110 & 0 \\ 2 & 52 & 827 & 38 \\ 0 & 0 & 7 & 1395 \end{bmatrix}$	Fatality	7.5%
			Serious Injury	11.1%
			Slight injury	89.9%
			Damage only	99.5%

Testing results of hybrid MLPNN-SVM				
ACC	Error	Confusion matrix	Severity class	SEN
90.5%	09.4%	$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 4 & 6 & 44 & 0 \\ 0 & 30 & 348 & 16 \\ 0 & 0 & 4 & 597 \end{bmatrix}$	Fatality	5.0%
			Serious Injury	11.1%
			Slight injury	88.3%
			Damage only	99.3%

4.3.4 Comparison between MLPNN, SVM, and Hybrid MLPNN-SVM

According to the outcomes from the confusion matrix shown in the above Table, the comparison of the actual and the predicted classes of the injury severity for training and testing data was applied. As seen in Table 13 (results of hybrid MLPNN-SVM) and in Table 12 (comparison between SVM and MLPNN), hybrid model provided a superior fitting of model and increased the accuracy rate by 20% in comparison to MLPNN and SVM. This amount is greater than the accuracy rates which were previously received by similar hybrid ANN–SVM models (Bellili et al., 2003; Tifani et al., 2017). However, incorrect classifications for ‘serious injury’ and ‘fatal’ classes still remain. Therefore, using STATS19 dataset, it is suggested that the combination of MLPNN and SVM achieved better prediction accuracy.

A similar comparison outcome was attained by Bellili et al. (2003) in a classification task. The aim of their research was to reduce the recognition of error rate applying the hybrid MLP-SVM recogniser. Accordingly, their outcome proved that the hybrid MLP-SVM model significantly increased the performance. However, they used limited data in their study so application of a rank analysis was missing from the methodology. In a more recent comparison between SVM and the hybrid model, Tifani et al., (2017) proposed a similar hybrid ANN-SVM for an estimation task in which the hybrid model had more accuracy rate, by 5% compared to SVM.

4.3.5 Additional Trail Using LVQNN to Overcome Incorrect Predictions

As a result of the comparisons between SVM, MLPNN, and hybrid MLPNN-SVM different insights were generated which needed to be more focused on, by using a different model. Such a prediction will aim to overcome the limited data for 'fatal' and 'serious injury' classes and to make predictions for better outcomes (Priyono et al., 2005; Al-Daoud, 2009; Chen and Marques, 2009; Shen and Chen, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.3.6 Killed and Seriously Injured Casualties as a Single Injury Class (KSI)

The analysis obtained from the error matrix showed that the limited data for 'fatal class' led to poor classification. Moreover, for 'serious injury' class, quantitative effects of each input factor on the injury severity could not predict very well and tended to work with a different class. Therefore, in order to achieve better results for all classes, 'fatal class' and 'serious injury' classes were merged together in a single factor as KSI (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

4.3.7 Prediction of Cyclist Injury Severity Levels Using LVQNN – Cambridge

Due to the necessary cycling intervention, Cambridge data was considered in order to be examined by LVQNN. At this stage only two injury severity classes were used (slight injury and KSI) in order to overcome the incorrect predictions for 'fatal' and 'serious injury'. In addition to KSI class, 'slight injury' severity class was used to focus more on personal injury related classes. Again, MATLAB application was used to apply LVQNN into the STATS19 data. The most important contributory factors were directly applied to improve the data quality and transfer the data into a space of fewer dimensions which can aid to boost the network speed, optimise the efficiency and, as a result, it will impact the accuracy of the LVQNN. In addition, combining the ‘fatal’ and ‘serious injury’ severity classes, we aim to maximise the prediction accuracies in comparison to the other existing models.

Consequently, to achieve this aim, the data was applied separately as an input to the LVQNN to compare the influence for each injury related factor in the output (slight or KSI). In this section, the model was predicted with the more sensitive factors which had been found as the contributory factors. Typically, all data was firstly shuffled and also normalised for each iteration in order to have an equal series of feature values and to ensure that training, validation and testing sets are representative of the total distribution of the data. For this section, the LVQNN was fit on training, validation and testing subsets. 70% of the total data was randomly separated for training data with the purpose of recognising apparent association that doesn’t hold in common. 15% of data was used for validation, in order to provide an unbiased assessment, fit on the training dataset through tuning the model's hyperparameters. As a final point the remaining 15% of data specified in the test dataset which is independent of the training dataset but examines the similar probability distribution. Again, for evaluation of the LVQNN, the confusion matrix was used to summarise the performance of prediction results, which the classes comprised of KSI (no.1) and slight injury (no.2). According the LVQNN, the number of correct and incorrect predictions have been summarised as an error matrix and the value of each injury severity class broken down in Figure 40 (Siamidoudaran et al., 2019b).

Figure 40: Error Matrix – LVQNN (Siamidoudaran et al., 2019b)

The outcomes of accrual and predicted injury severity classes display the insights in which the LVQNN was confused or predicted properly during the prediction. The number and value of correct predictions are recognised by first two cells diagonally. For example, in the all error matrix, 145 injury severities were correctly predicted as KSI. This refers to 11.5% of all 1264 injury severities in the STATS19 data. Analogously, 1071 severities were properly predicted as slight injury. This refers to 84.7% of all injury severities. 9 numbers of the slight injury severities were wrongly predicted as KIS and this refers to 0.7% of all 1264 injury severities. In the same vein, 39 events of the KSI casualties were wrongly predicted as slight injury and this equals to 3.1% of all the data. Out of 154 KSI predictions, 94.2% were correct and 5.8% remained incorrect. Out of 1110 slight injury predictions, 96.5% were correct and 3.5% were incorrect. Out of 184 KSI casualties, 78.8% were correctly classified as KIS and 21.2% were classified as slight. Out of 1080 injuries, 99.2% were properly predicted as malignant and 0.8% were predicted as KSI. In addition, the four grey squares (starting the top right, clockwise) denote the net present value, precision, sensitivity and specificity. The blue cell in the bottom right displays the overall percent of properly predicted injury severities (in green) and the total percent of misclassified events (in red). As shown in the blue square, the overall accurate prediction rate is 96.2% for overall data and 3.8% were incorrect which was very accepTable. These amounts are highly improved compared to the previous models (MLPNN, SVM, and hybrid MLPNN-SVM) used in this thesis (Siamidoudaran et al., 2019b).

4.3.8 Comparison of Actual and Predicted Classes of LVQNN

Comparison between the actual and the predicted injury severity classes of the training, validation, testing, and total data are interpreted in Figures 41, 42, 43, and 44. The blue symbols denote the actual levels and the pink symbols display the predicted classes resulting from LVQNN. As seen in the below figures, the acquired results of the predictions for each injury severity class were very satisfactory as there is great integrations between pink and blue marks. Therefore, the LVQNN model succeeded to predict each cyclist injury severity class with an advanced accuracy. However, the class 2 (Slight injury) provided a greater fitting of the model plus an improved prediction accuracy in comparison with class 1 (KSI) in each stage as a result of having more data for ‘slight injury’ casualties (Siamidoudaran et al., 2019b).

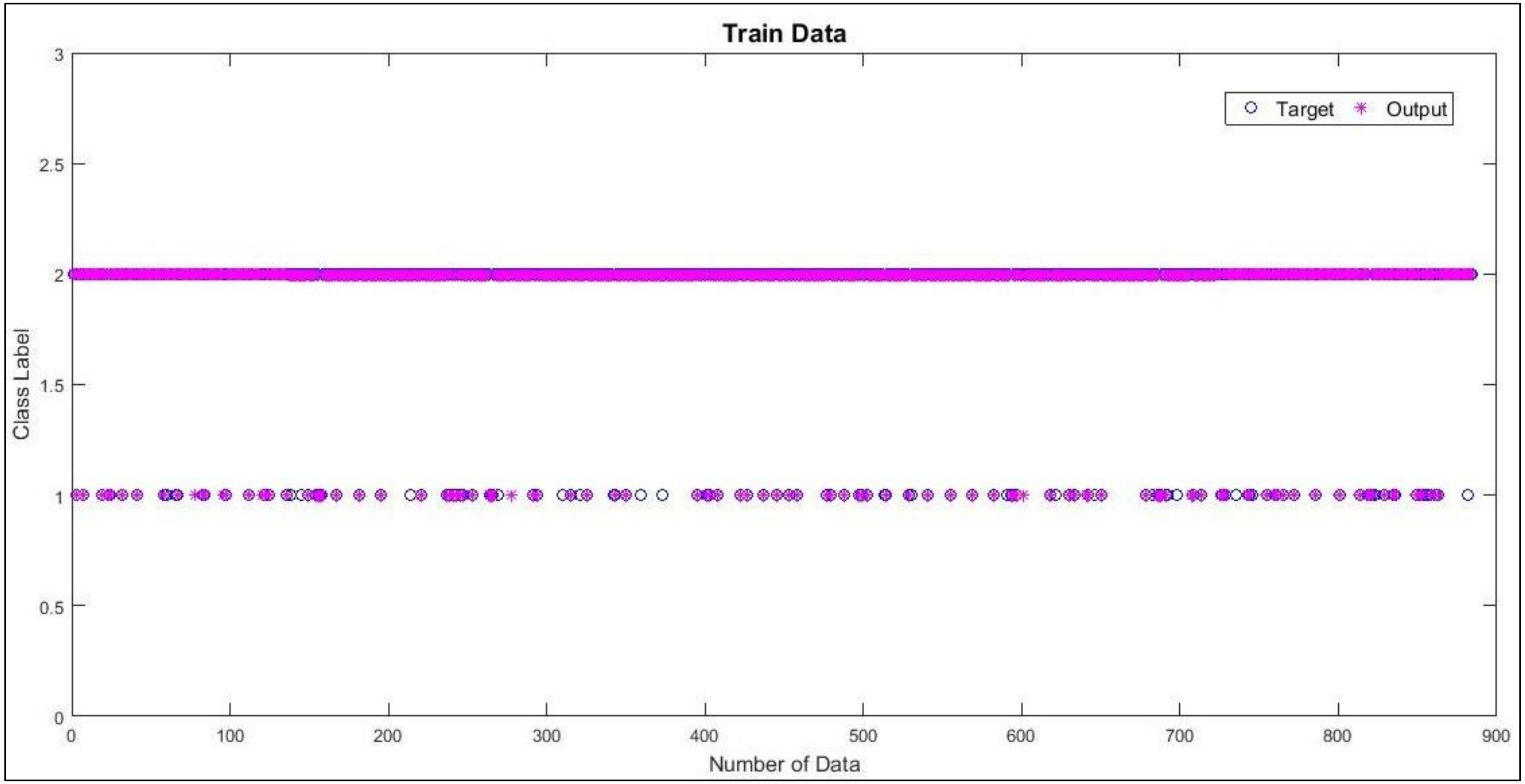

Figure 41: Actual and Predicted Classes of Cyclist Injury Severity by LVQNN – Train Data (Siamidoudaran et al., 2019b)

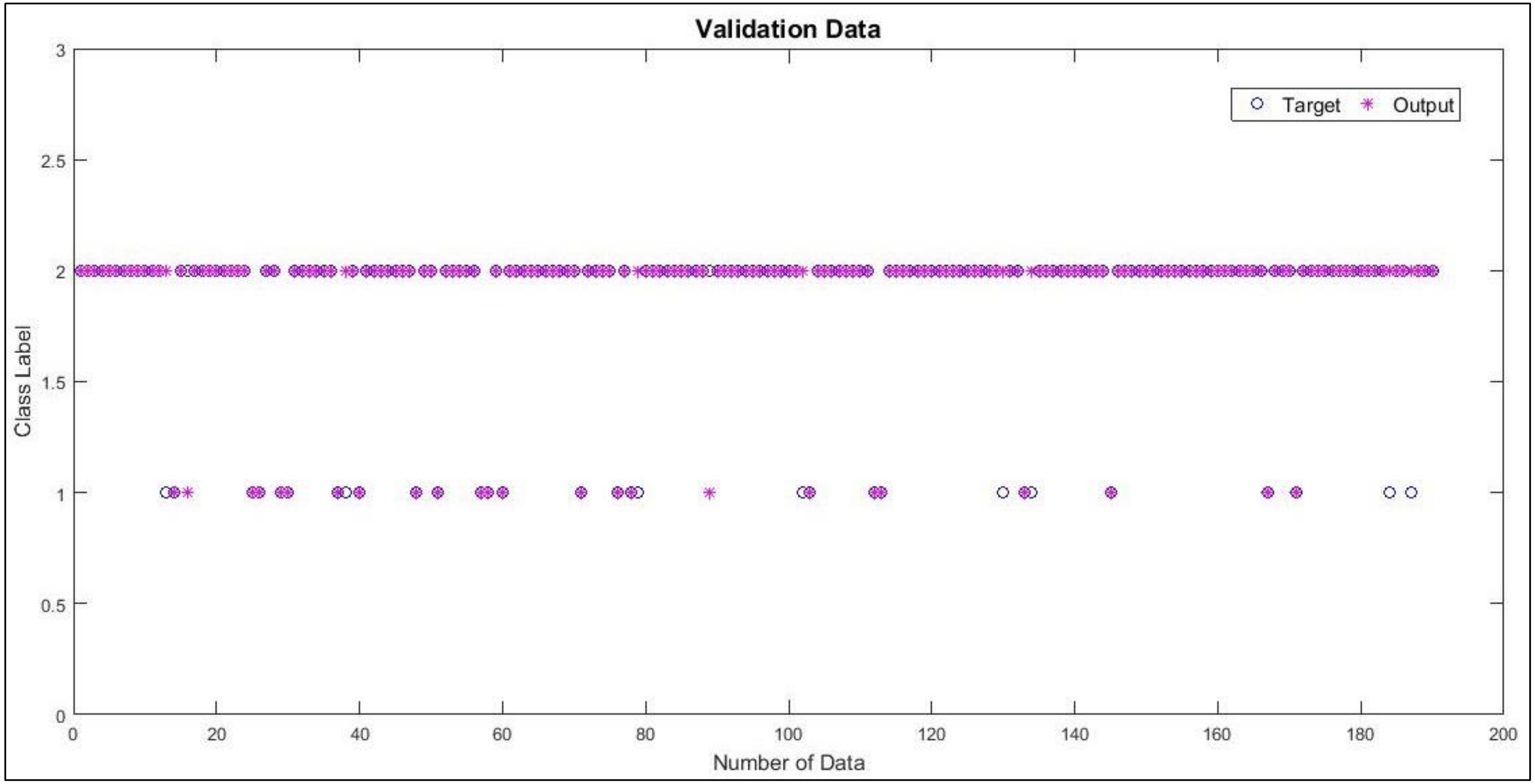

Figure 42: Actual and Predicted Classes of Cyclist Injury Severity by LVQNN – Validation Data (Siamidoudaran et al., 2019b)

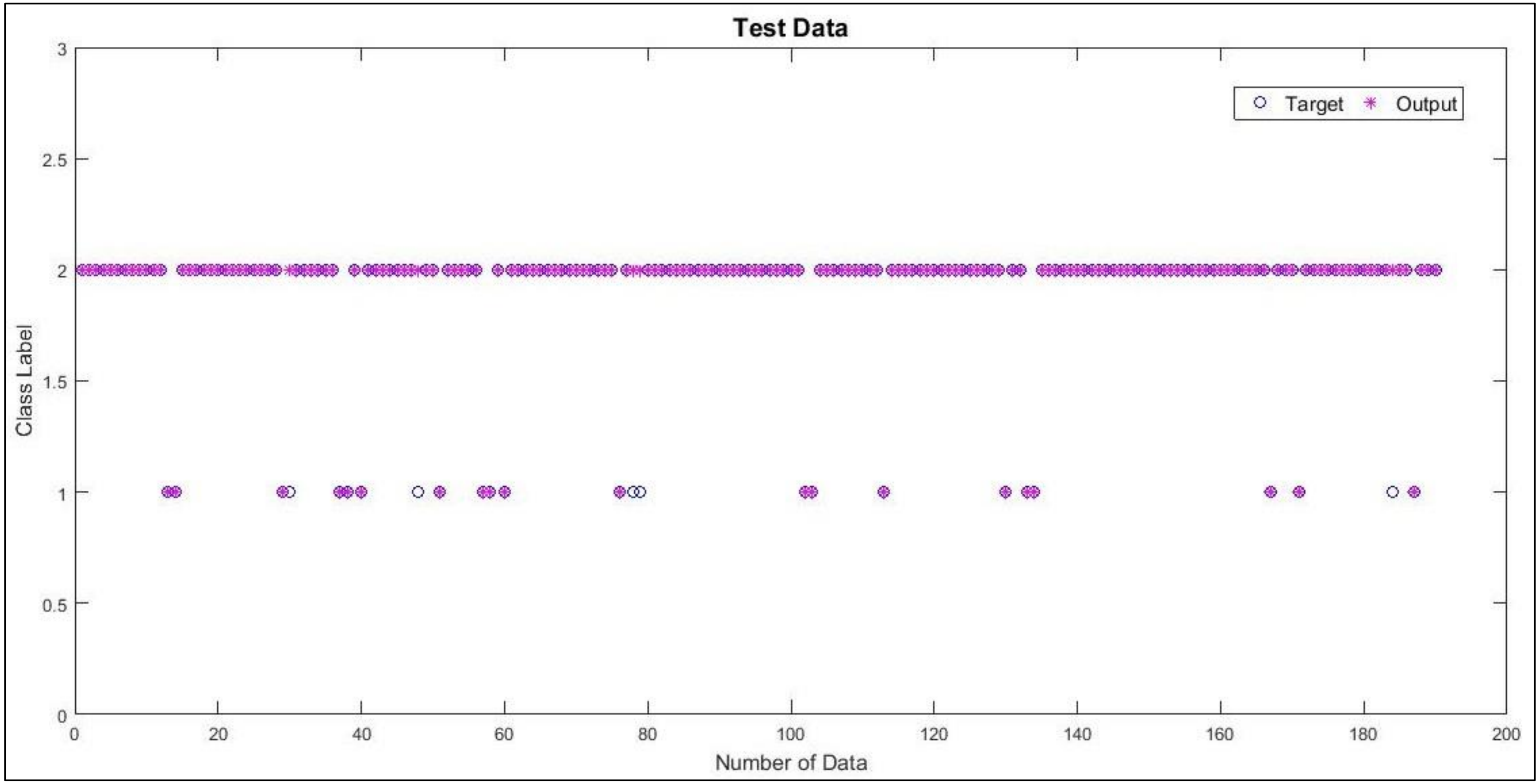

Figure 43: Actual and Predicted Classes of Cyclist Injury Severity by LVQNN – Test Data (Siamidoudaran et al., 2019b)

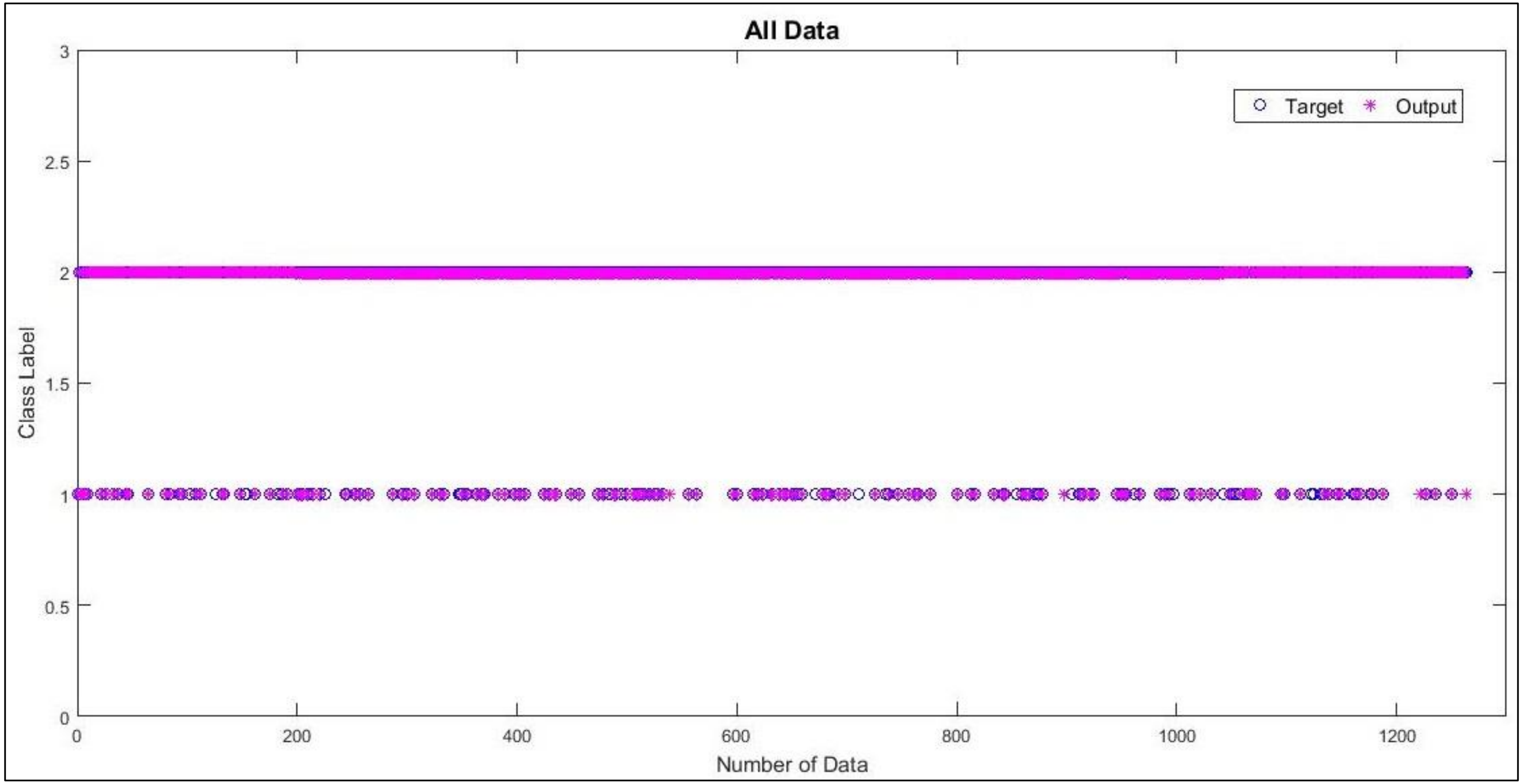

Figure 44: Actual and Predicted Classes of Cyclist Injury Severity by LVQNN – All Data (Siamidoudaran et al., 2019b)

4.3.9 Comparison between MLPNN, SVM, MLPNN-SVM, and LVQNN

At this stage, the prediction accuracies of MLPNN, SVM, hybrid MLPNN-SVM, and LVQNN models were also evaluated and compared. As the comparison results of MLPNN, SVM, hybrid MLPNN-SVM have already been discussed in the previous sections, we mostly focus on the testing results of the models along with overall prediction of LVQNN which are summarised and ranked in Table 14.

Table 14: Comparison results of MLPNN, SVM, MLPNN-SVM, and LVQNN

SVM – Case study of Cambridge (4^{th} rank)				
ACC	Error	Confusion matrix	Severity class	SEN
72.8%	27.1%	$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 53 & 1 \\ 0 & 0 & 325 & 69 \\ 0 & 0 & 76 & 525 \end{bmatrix}$	Fatality	00.00
			Serious Injury	15.1%
			Slight injury	81.1%
			Damage only	85.1%
MLPNN - Case study of the city of London (3^{rd} rank)				
ACC	Error	Confusion matrix	Severity class	SEN
75.3%	24.6%	$\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 2 & 49 & 0 \\ 0 & 11 & 317 & 71 \\ 0 & 0 & 37 & 571 \end{bmatrix}$	Fatality	00.00
			Serious Injury	16.0%
			Slight injury	82.1%
			Damage only	87.0%
Hybrid MLP-SVM - Case study of the city of London (2^{nd} rank)				
ACC	Error	Confusion matrix	Severity class	SEN
90.5%	09.4%	$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 4 & 6 & 44 & 0 \\ 0 & 30 & 348 & 16 \\ 0 & 0 & 4 & 597 \end{bmatrix}$	Fatality	5.0%
			Serious Injury	11.1%
			Slight injury	88.3%
			Damage only	99.3%
LVQNN - Case study of Cambridge (1^{st} rank)				

ACC	Error	Confusion matrix	Severity class	SEN
96.2%	3.8%	$\begin{bmatrix}145 & 09\\ 39 & 1071\end{bmatrix}$	KSI	78.8%
			Slight injury	99.2%

The outcome shows that the fitting and prediction performance of all models are satisfactory but the hybrid model performed better than SVM and MLPNN (Siamidoudaran and Iscioglu, 2019). This thesis suggests that the MLPNN, SVM, and the hybrid model fit the data well so they are promising tools for future accident injury severity studies (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab; 2004; Chang and Wangm, 2006; Delen, 2006; Li et al., 2008; Li et al., 2012; Yu et al., 2014; Yu et al., 2016; Sharma et al., 2016; Li et al., 2016; Alkheder et al., 2017; Alkheder et al., 2017; Zhang et al., 2018; Shamsashtiany and Ameri, 2018; Siamidoudaran et al., 2019a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019b; Venkat et al., 2020).

Using the identical dataset of the city of London for MLPNN and the hybrid model (Siamidoudaran and Iscioglu, 2019), this thesis suggests that the hybrid model was better able to decrease the recognition of error rate. Although the results of the comparisons are positive, the incorrect predictions remain for 'fatal' and 'serious injury' classes by all the models. Therefore, by using LVQNN as a powerful method for prediction, it was intended to verify that there might be other existing models that fit the data better than the proposed models (Priyono et al., 2005; Al-Daoud, 2009; Chen and Marques, 2009; Shen and Chen, 2009; Kohonen, 2012; Thanasarn and Warisarn, 2013; Nova and Estévez, 2014; Villmann et al., 2017; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). As a result of comparison shown in Table 14, it is suggested that LVQNN is the best method for prediction of STATS19 data (by applying the identical data which was used for the Cambridge case study). The model showed an improved overall accuracy (96.02%) and also maximised the prediction rates for each injury severity class (78.8% for KSI and 99.2% for slight injury class). Therefore, using LVQNN along with combination of 'fatal' and 'serious injury' classes within a single class (KSI), this model was able to overcome the limitation of data (Siamidoudaran et al., 2019a; Siamidoudaran et al, 2019b).

4.4 Discussion of Results

4.4.1 First-stage Prediction– Rank Analysis (RBFNN)

In this thesis we provide evidence of how prediction models of injury severities can truly help to understand the relationship between crash related factors. Using this technology can decrease severity of injuries directly and indirectly. Indeed, the main aim of this stage was to select key input factors for the next stage of prediction in order to maximise model performance. The RBFFNN model was applied to link the likelihood of injury occurrences at different severity levels with various traffic related factors.

4.4.1.1 First Case Study– Identification of Group Most in Need of Intervention

In the first stage, RBFNN was used to examine the influence of a number of factors on the injury severity faced by all road users involved in road accidents for the case study of London. The prediction results for the RBFNN model for city of London are shown in Table 4 for the city of London. In that stage, different variables were found to be significantly correlated with the likelihood of injury severities. The results displayed that different contributory factors affect different road users differentially. More specifically, the prediction model suggests that the severity of injury likelihood tends to be high where the crossing facilities were not available within 50 metres.

Here, the most common injuries were suffered by the VRUs, especially the cyclist group. In this vein, a previous research verified that the VRUs were considered to affect the high likelihood of being involved in accidents (Chang and Wang, 2006; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a). Moreover, a Dutch traffic collision study displayed that more than half of the KSI crashes which VRUs were involved in happened while crossing the road (SWOV, 2010).

4.4.1.2 Second Case Study – Response to First Case Study Concern

In response to the first case study concern regarding to the two-wheel group, the RBFNN model was applied to the Cambridge case study to investigate the influence of a number of crash related factors on the injury severity faced only by pedal riders. The prediction model indicated that a variety of cycling accident related factors can highly affect the severity of injuries. Most of the factors were in connection with busy intersections and poor turn/manoeuvres. Importantly, biking across the main carriageway and not in a restricted lane had a massive effect in increasing the severity of injuries. This scene is well-known and always contributes to risk of cyclist injury (Knowles et al., 2009; RoSPA, 2017c; DFT, 2018b; SiamiDoudaran et al., 2019b).

4.4.2 Second-Stage Prediction–Maximise Performance (MLPNN and SVM)

An accurate ISP requires a good insight into the factors that are believed to be related directly to severity of road accident injuries. Therefore, based on the selected contributory factors, a second stage of prediction was carried out for each case study to maximise the prediction accuracy. We applied MLPNN and SVM to STATS19 data to find out the relationship between injury severities and sensitive predictors which were identified in the previous stage of predictions (listed in Table 5 and Table 9). Using most important contributory factors at this stage we aim to improve performance of both models (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al.,2019b).

4.4.3. Extra Trials on STATS19 Data (Hybrid MLPNN-SVM and LVQNN)

ANN and SVM models were chosen as the benchmark in this thesis, particularly as a result of their popularity and success in injury severity modelling. However, using sensitive predictors, the MLPNN and SVM were not capable to overcome the incorrect predictions for 'fatal' and 'serious injury' severity classes. Therefore, the comparison between MLPNN and SVM couldn't verify that these models are the best tools for STATS19 data. Therefore, using hybrid MLPNN-SVM, and LVQNN were also necessary to examine the same datasets in order to outline a comprehensive comparison result.

4.4.3.1 KSI to Overcome Limitation of Data Using LVQNN

Using an improved LVQNN in this thesis, 'fatal class' and 'serious injury' classes were merged together in a single factor as KSI to achieve better results for all classes (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). As a result, applying same identical data which was earlier used in Cambridge case study, LVQNN was able to fit the same data better than SVM for cycling injury severities.

4.4.4 Model Comparisons (MLPNN, SVM, MLPNN-SVM, and LVQNN)

Initially, the prediction accuracies of the MLPNN and SVM models were evaluated and compared. The outcome shows that the fitting and prediction performance of both models is satisfactory but MLPNN performed a little better than SVM (Siamidoudaran and Iscioglu, 2019). This thesis suggests that the ANN and SVM models fit the data well so they are promising tools for future accident injury severity studies. Accordingly, if the main purpose of a research is to predict injury severity classes, the ANN and SVM models can be a very good choice (Abdelwahab and Abdel-Aty, 2001; Abdel-Aty and Abdelwahab; 2004; Chang and Wangm, 2006; Delen, 2006; Li et al., 2008; Li et al., 2012; Yu et al., 2014; Yu et al., 2016; Sharma et al., 2016; Li et al., 2016; Alkheder et al., 2017; Zhang et al., 2018; Shamsashtiany and Ameri, 2018; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b; Venkat et al., 2020).

An interesting direction was to discover some new methods and compare them with the proposed machine learning models. Therefore, additional trials using hybrid MLPNN-SVM (first case study dataset), and LVQNN (second case study dataset) on the same datasets were applied to draw a comparison. Hybrid MLPNN-SVM, and LVQNN methods were used for the first time in an injury severity related prediction by the author (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b).

Although the hybrid model was able to highly improve the overall prediction accuracy, LVQNN overperformed all the models used in this thesis plus correctly predicted the data for all the classes. Therefore, this comparison can suggest that the improved LVQNN is the best tool for prediction of STATS19 compared to other existing methods used in this thesis. In addition, in particular to LVQNN, further details can be referred to in the author' previous case studies for the city of London and Cambridge (Siamidoudaran el al., 2019a; Siamidoudaran el al., 2019b).

4.4.5 Requirement of Road Safety Interventions

The identification of effective road safety interventions is essential. In this regard, not only should the requirement for an intervention be based on evidence, the intervention chosen should also be based on evidence. A needs assessment involves using STATS19 data as evidence to recognise whether an intervention is required to address a particular road safety problem. This is a vital stage. If the intervention is not desired, there could be a pointless waste of resources. STAS19 data can be used to detect: the what, where, when and who in connection to the road safety problem. It can aid to identify specific roads, areas or road user groups such as pedestrian, cyclist, young drivers, children that might benefit from an intervention. Accordingly, their findings should have a main role in road safety design, policy and education (Ameratunga et al., 2006; Perel et al., 2007; RoSPA, 2017a; DfT, 2020).

Typically, to achieve this, the government collected statistics can be analysed using different types of in-house software (RoSPA, 2017a). However, all the types of software are based on the reported contributory factors which are subjective and are all based on the police's opinion at the scene, and perhaps are not based on a wide long-term investigation so may well not be absolutely reliable (TRL, 2010; DfT, 2014). In addition, some of the factors are less likely to be recorded since evidence may not be available after the accident (DfT, 2014). Therefore, we believe, the findings of this thesis are more meaningful compared to the contributory factors reported by police. Moreover, there is very limited researches carry out on road safety interventions thus it is so challenging to discover evidence-based intervention related evaluations (RoSPA, 2017a; DfT, 2020).

In this connection, the thesis ends by suggesting evidence-based road safety intervention options relying on the identified key injury severity impact factors in order to mitigate poor road designs and bad behaviour of road users. The suggestions in this thesis greatly matches the author's previous published articles (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In addition, we have cited different published articles and official government reports to support our ideas (RoSPA, 2017a; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020; DfT, 2020).

4.4.5.1 Evidence-based Suggestions

Evidence-based plans were created in the field of medicine, but have been translated into a lot of policy areas, including road safety (RoSPA, 2017a; DfT, 2020). From this perspective, many road safety intervention options are suggested to reduce severity of injury accidents. However, the suggestions often do not match up to reality plus some projects could make roads more dangerous (Luria et al., 2000; McKenna, 2010), and it is therefore vital that interventions are designed based on evidence (Connexions, 2001; Hauer, 2007; McKenna, 2010; RoSPA, 2017a; DfT, 2020).

This involves looking at collision, casualty and especially examining historical collision data to be sure that the safety concerns are addressed, and research and evaluations to investigate whether the intervention type being considered is likely to be helpful. Even if a previous intervention was not evidence based in order to start with, it is a great idea to go back and look at the evidence (RoSPA, 2017a; DfT, 2020). Although, this might be a long process, it can aid to improve the intervention based on the greatest practice of other road safety practitioners. Moreover, it will help to understand whether the intervention is really needed for the related site. For instance, although there might be a lot of public concern about a road safety in a specific area, examining the accident data might suggest that it is not a priority concern (RoSPA, 2017a).

Before planning a road safety intervention, it was essential to pin point the road safety issues that need to be tackled and then identify the most appropriate ways of dealing with them. In this regard, there are a number of types of evidence that can be used.; mainly, investigating STATS19 casualty data, as this will help to determine whether an intervention is really needed (RoSPA, 2017a; Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b; DfT, 2020). Unfortunately, relatively little assessment is carried out on road safety, and therefore it might be difficult to find research and evaluation reports that are related to a particular intervention (RoSPA, 2017a; DfT, 2020). In response to this shortage, this thesis attempted to suggest different evidence based road safety interventions.

The data was applied into the ANN model to detect specific groups and areas of concerns that require more intervention of road safety. The outcomes from the above predictions (shown in Tables 4 and 8) along with the data analysis for the most important contributory factors (shown in Tables 5 and 9) display that most of the factors were in connection with busy junctions and poor turn/manoeuvres. To overcome this concern, installing truly protected junctions in various site clusters is the key solution. In more detail to the first case study, non–motorised road users (pedestrians and cyclists) were recognised to benefit from road safety interventions.

On focus to the cycling, there were limited crossing services near to where they cycled. Importantly, for Cambridge, such a big cycling city, narrow bike lane defenders are needed to provide a full segregation where road width is too limited (RoSPA, 2017a; RoSPA, 2017b; RoSPA, 2017c; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020; DfT, 2020). The injuries resulting from car and non-motorised road users played a large part in the injuries. Following this, the most common injuries suffered by cyclists happened where the crossing facilities were not presented within 50 metres for the controlling of pedal riders crossing. Furthermore, junctions were definitely the accident hotspots, particularity when they attempted to approach or were located at the mid of the junction leading into another road. In a word, most of the factors identified in this study were in connection with busy junctions and poor turning manoeuvres. In addition, the majority of these actions were caused by human error or misjudgement. To solve the human error, intervention options concerning the specific groups should be applied through professional Road Safety Education, Training and Publicity (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). The followings examples of interventions are suggested by this thesis.

4.4.5.2 Infrastructure of Engineering Intervention

As a result of the factors identified in this thesis, it seems that the design of the roads may be flawed. In this regard, majority of the injuries caused by poor road design which tended to be extremely dangerous for road users, especially cyclist group often resulted in serious injuries. The suggestions of the protected junctions and the narrow road defenders in this section greatly matches the author’s previous publish articles (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In addition, the suggestions fit the local governments’ new cycling and walking infrastructure plan which are detailed below (RoSPA, 2017b; TfGM, 2019; Reid, 2019; Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020; Glasgow City Council, 2020). The concern resulting from junction actions taken by the two-wheel group was identified and the idea of the solution was suggested in general by the former UK Ministry of Transport about one century ago but surprisingly, has not been taken seriously enough until now (Reid, 2019).

4.4.5.2.1 Protected Junctions

It’s relatively not difficult to protect pedal riders on busy roads using curbs to carve out a cycleway. However, it is not at all easy to protect cyclist group at intersections. At this point, truly protected junctions can be a good solution for both case studies. Junctions like this, giving priority to pedal riders and pedestrians are common practice in Netherlands. Also, they are extensively and successfully used across Europe, however protected junctions are not common in the UK. Protecting pedal riders at intersections has been a known concern since the 1930s. “The benefit of the cycle-track is lost at the intersection (just where traffic segregation is most needed).” detailed Architectural Review magazine in year 1937. (Reid, 2019).

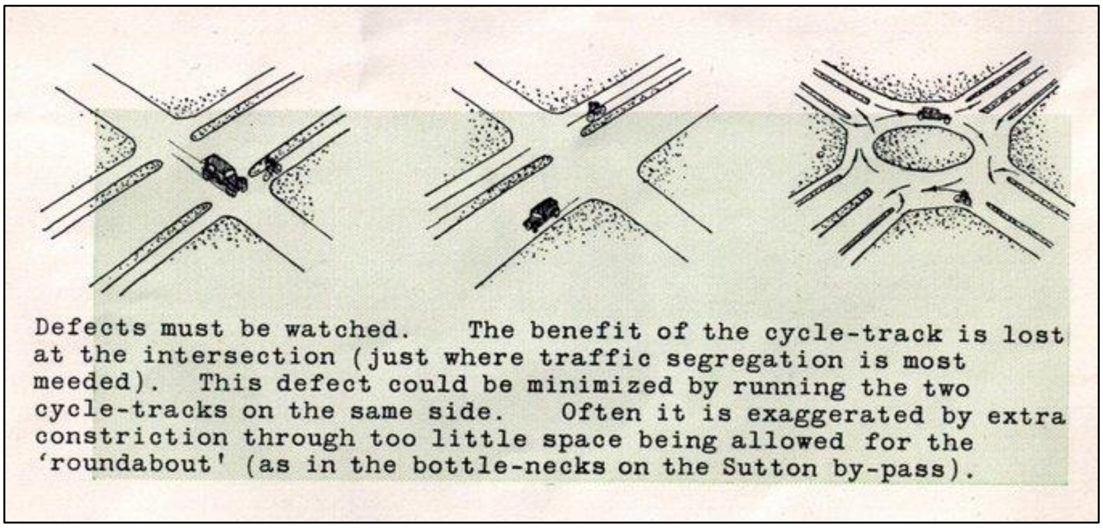
Defects must be watched. The benefit of the cycle-track is lost at the intersection (just where traffic segregation is most meeded). This defect could be minimized by running the two cycle-tracks on the same side. Often it is exaggerated by extra constriction through too little space being allowed for the 'roundabout' (as in the bottle-necks on the Sutton by-pass).

Figure 45: Importance of Protected Junctions in An Old Magazine, 1937 (Reid, 2019)

As it is shown in the above magazine, it is clear that protected junctions are really needed for cycling network in the UK. At the time, a junior engineer from the Ministry of Transport in the 1930s, complained that the UK’s putative cycle network system modeled in the Dutch style provided protection where the carriageway was safe but discharged the cyclists into the whirlpool of main traffic where the network was most dangerous. Again the ministry reported on 1946 showing that how cycle routes should be implemented nearby roundabouts, offering protection everywhere, but no instances were proposed at the time (Reid, 2019). Almost a century has gone by but only a few UK cities, including Cambridge, Glasgow, Manchester and Aberdeen, working on (very limited) protected junctions. The below types of protected junctions are some examples which are matched with findings of this thesis.

4.4.5.2.2 Glasgow Style Protected Junction

As part of the South City Way, the below junction shown in Figure is being trialled in Glasgow city, which is the first of this kind of junction in the UK outside of London (Glasgow city Council, 2020).

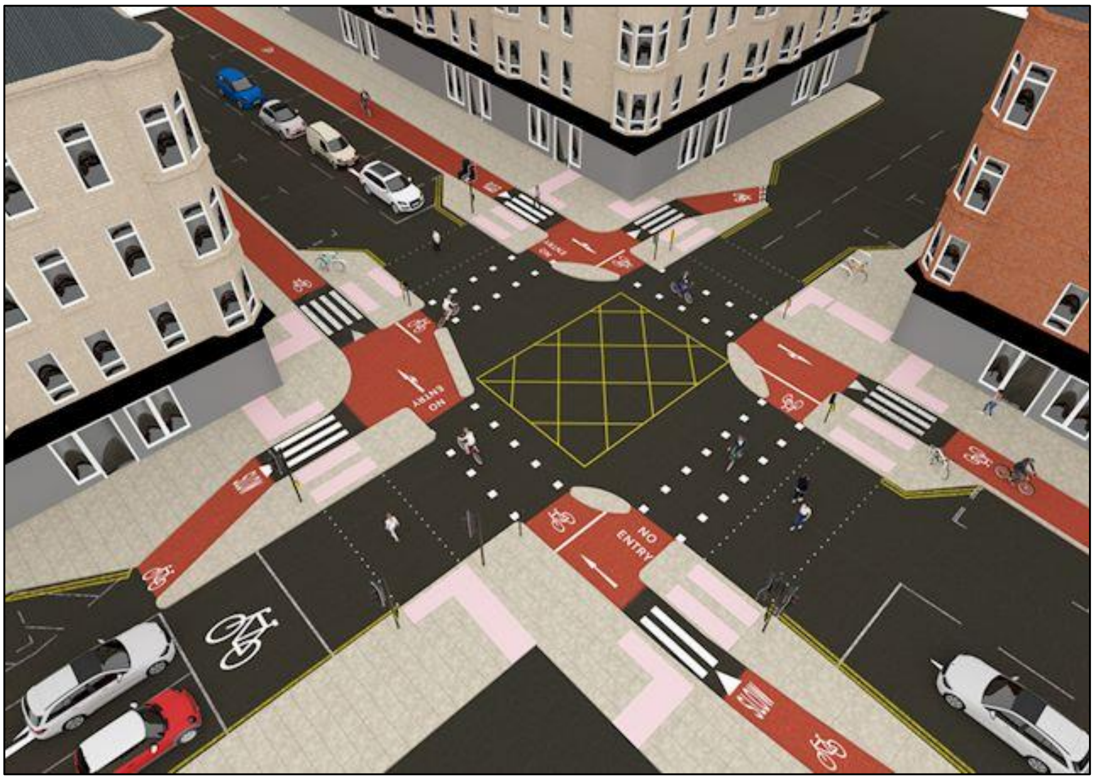

Figure 46: Proposed New Style of Junction, Glasgow (Glasgow city Council, 2020)

This kind of junction has been verified to be much safer because it reduces conflict between road users, and particularly for those on cycles. It is a road junction designed so that road users travelling on foot, by cycle, and in vehicles are all separated as they pass through the junction. Unlike traditional junctions, which usually requires right-turning bicyclists to wait in the mid of the junction for an appropriate gap in the traffic, this type of junction offers a safer alternative. Furthermore, protected cycle tracks like those on Figure 46 make it easier for people to have everyday travels by cycle instead of by motor vehicle. The protected junction design produces space for these segregated tracks to flow through and nearby the junction, making a continuous and safe road (Glasgow City Council, 2020).

4.4.5.2.3 Dutch-style Roundabout

The evidence based intervention identified in this thesis fits the installation of the UK's first Dutch-style junction which was officially opened during Covid-19 post–lockdown in Cambridge and is the first of its kind in the UK (Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020). However, in 2013 Transport for London (TfL) worked on a design with TRL the Future of Transport (Transport Research Laboratory) but Cambridge claimed it installed the initial truly Dutch style in Britain after building a semi Dutch example junction in 2013 by TfL (Reid, 2019).

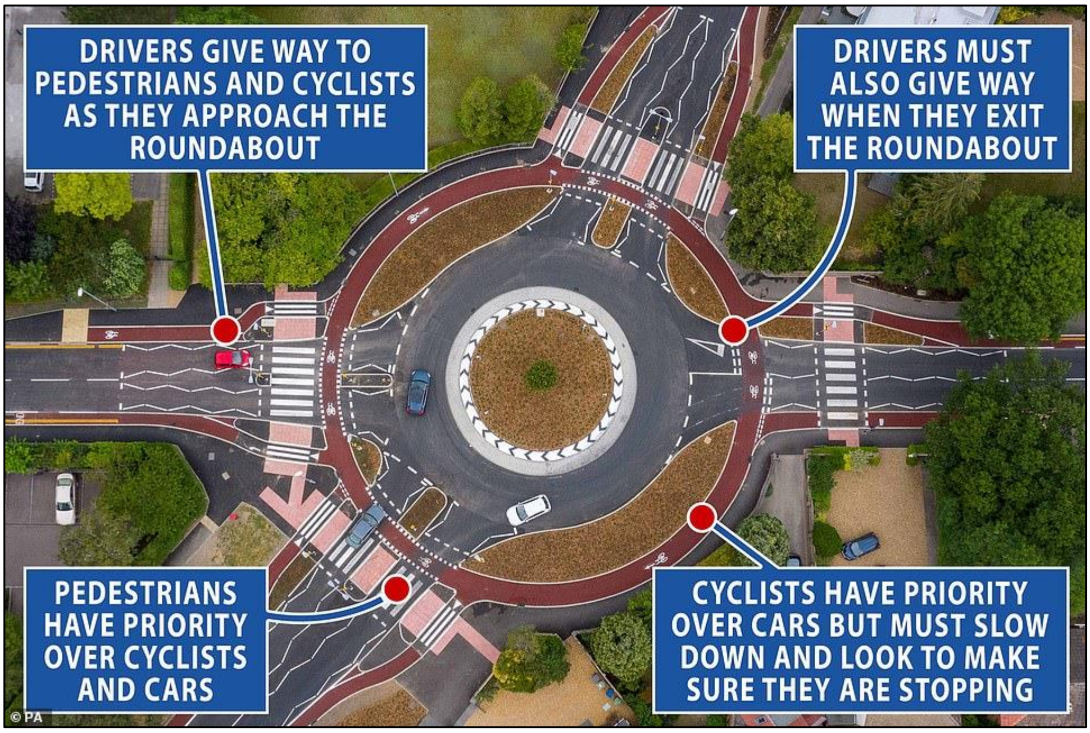

Figure 47: UK's First Dutch-style Junction, Cambridge (Daily Mail, 2020)

UK’s cycling capital celebrates the arrival of this style of junction as it was opened in a cyclist killing zone at a cost of £2.4m. The protected junction holds a red bike lane round to give cyclists and pedestrians priority (Cambridge Cycling Campaign, 2020). It has been called Dutch roundabout due to being developed in the Netherlands.

A Dutch-style roundabout has parallel cycle and pedestrian zebra crossings on each arm which allows the VRUs to have priority over drivers. The entry and exit arms are vertical, rather than tangential to the roundabout and have minimal flare. Moreover, by decreasing the width of the arms and circulatory carriageway, all vehicle speeds decrease. With speeds reduced, any collisions that do happen are probably to be of much lesser severity. A central over-run area permits big vehicles to manoeuvre round the roundabout (Cambridgeshire County Council, 2020; Cambridge Cycling Campaign, 2020; Daily Mail, 2020). This kind of scheme can encourage cycling amongst both adults and children, which in turn can bring health welfares from better physical activity of new riders (Cambridgeshire County Council, 2020).

4.4.5.2.4 CYCLOPS Junction

Another option regarding protected junctions can be Cycle Optimised Protected Signals (CYCLOPS) type which an orbital cycle route separates riders from motor vehicles, decreasing the likelihood of accidents or conflicts. The UK's first CYCLOPS cycling junction has been proceeded in South Manchester, planned to separate VRUs from traffic. Pedal riders approach the intersection from four 'arms', converging onto a cycle route which fully encircles the intersection, allowing bicycles to create a right turn while being protected from traffic flow, and to perform the manoeuvre in one movement (subject to signal timings) (TfGM, 2019).

Figure 48: Proposed UK's First CYCLOPS, Greater Manchester (TfGM, 2019)

The key innovation of the design in CYCLOPS is the cycle track being on the outside of the pedestrian crossings, which offers more space and means for all kinds of junction activities which can be incorporated within the external orbital bile route. It is estimated that the amount of people cycling and walking will rise in the future, therefore CYCLOPS junctions can simply accommodate this modal shift as cycle and pedestrian stages run in parallel simultaneously profiting from green time reallocated from traffic phases (TfGM, 2019).

4.4.5.2.5 Narrow Cycle Lane Defenders

Within framework of the second case study, the first key injury severity impact factor refers to location of bicycle. The reality is that the title ‘cycling capital’ has been achieved without any real and enough cycling infrastructure for such a big cycling city (CambridgeshireLive, 2018b). The majority of network for central routes in Cambridge, cyclists rely on narrow lanes which consist of merely a few inches of white paint (as example below) which give the riders a false feeling of safety on busy mandatory cycle tracks (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). The below photo shows an example related to the narrow bike lanes.

Figure 49: Narrow Bike Lanes Design Flaws Make A Bike Lane Incredibly Unsafe

Nonetheless, based on the first factor detected in this thesis, narrow protected cycle defenders are needed to segregate bikes from other traffic where road width is too narrow (ISJ, 2020). In addition, flexible, passively safe, highly visible bollards/wands should also be combined with defenders to clearly direct where a cycle lane is to other road users. Moreover, continuous segregation should be considered where passible and for over longer distances where vehicle speeds may be higher. The below photo is an example of narrow bike defender which are suggested be more effective for narrow roads (ISJ, 2020; Rosehill Highways, 2020). The photo shows Wand Orca defenders combining vertical cones with reflective markings, together with horizontal rubber modules (Rediweld, 2020).

Figure 50: Narrow Cycle Defender Incorporating Wands (Rediweld, 2020)

The defenders must provide cycle lane segregation and safety for cyclists as well as provide continuity by excluding other traffic from the cycle lane (where possible). Importantly, they must integrate reflectivity for maximum visibility.

Fully protected bike lanes are an effective and increasingly popular tool to elevate cycling mode shares and road safety in most cycling destinations in the world such as; Netherlands. However, most of the current defenders in the UK (e.g. Orcas shown in Figure 51) provide light segregation which are not really suitable for bike friendly cities where carriageway width is too limited. These kinds of defenders are a compromise and are definitely not very suitable for some sites in Cambridge as well as the City of London, the big cycling cities. Although, this type of separator discourages drivers from drifting or parking, there are reports of people not seeing the defenders as they cross the road, tripping over and hurting themselves (BSfE, 2018).

Figure 51: Example of Cycle Lane Soft Segregation

Perhaps a more visible form of separator, like the Wand-Orca is more noticeable to pedestrians. Ultimately it is hoped that the light segregations will lead to improve and be fully segregated cycle lanes in the future like the paths we see in the Netherlands. In this regard, where carriageway width is limited, the narrow cycle lane defenders are the perfect solution at just 235mm wide for continuous segregation which comply with Cycling England and Sustrans guidelines (Rosehill Highways, 2020). The finding of this study fits the UK government advice against the use of public transport due to physical distancing during Covid-19 post lockdown distancing and while promoting ‘active travel’ to work, local authorities across the UK are strongly seeking to rapidly expand their protected cycle lane networks (ISJ, 2020).

4.4.5.3 UK Cycling Is Booming During COVID–19

There is a 200% increase in bicycle orders from people working for some service area (BBC, 2020b). Restrictions and fear of catching the Covid-19 virus on public transport has inevitably helped lead to a boom in cycle-to-work schemes as well as number of people cycling to work in their area; and physical distancing guidelines mean they require more space in which to do so. In the meantime, it's vital that pedal riders feel safe when travelling to work (BBC, 2020a; BBC, 2020b; PKC, 2020).

4.4.5.4 Road Safety Education, Training and Publicity

As the majority of the identified predictors directly blame some kind of human error, what is clear from this important outcome is that it seems there is still lack of an effective education in road safety in both case studies and this privation poses significant dangers for VRUs, specifically for the cyclist group. Therefore, this strongly suggests that poor road safety skills of road users need to be improved by applying various and truly effective education as well as training and publicity programmes to further ensure VRUs are safer on streets, particularly around and in junctions. To overcome this, road safety intervention options concerning the specific groups should be applied (Luria et al., 2000).

4.4.5.5 Importance of Educational Interventions in Road Safety Engineering

Road engineering has traditionally been thought of as being an extrinsic method of improving road safety. However, road safety education, information, training and publicity can also be a vital part of the safer systems. Educational interventions have been a popular method to cope with road safety concerns as they satisfy a number of aims. Education interventions allow governments to address a problem of public concern, they are apparently reasonable and they are politically uncontroversial. Nevertheless, evidence offers that a lot of these interventions are ineffective

(McKenna, 2010) because they are frequently designed in the lack of theory or evidence such as examining STATS19 data. (RoSPA, 2017a). Lopez et al (2009) recommends that designing an educational intervention without any guiding model is similar to planning a medical intervention without an understanding of physiology. Just as it cannot be expected that aspirin decreases the risks of a heart disease without evidence, intuition and what seems plausible cannot give information on how much of delivering road safety education or publicity programme on seatbelt wear will change drivers and occupants' behaviour (Hauer, 2007). In addition, addressing road safety education and training are important requirements of road safety design in most of developed countries (Vardaki et al., 2018). Therefore, it is vital for all engineers and technicians to have solid knowledge of road safety education, training and publicity programmes.

To better understand the requirement, this thesis provides an example to show the importance of education, training and publicity in laying these foundations in the design process. For example, if a design plan (e.g. traffic calming measures) is to install bollards opposite a primary school, across a path which is approaching a crossing point in order to discourage child pedestrians from proceeding straight across road, it is great idea to use pencil shaped bollards and match the colours of the bollards with the pupils' uniform. Because this kind of bollard is a more suitable option for school environments compared to the traditional pedestrian guardrail. Its distinctive pencil shape provides a clear and effective form of pedestrian and vehicle demarcation for a wide range of landscapes (Marshalls, 2020). This is an effective idea which can emphasise the traffic calming features and can highly help encourage drivers to voluntarily slow down. In this regard, many previous studies have shown that traffic calming can reduce collision levels by up to 40%, and have an important impact on reducing the severity of injuries (Harvey, 1992; Elliott et al., 2003). The below photo shows installed Ferrocast PiPencil Bollards in a school environment.

Figure 52: Example of PiPencil Bollards

As it is clear in the example, the designers can fail to use an effective idea in the design process, if they not completely aware of importance of educational and publicity interventions in road safety engineering (Vardaki et al., 2018). Indeed, road safety education, training and publicity plans are complementary to road safety engineering and can provide more effective models for engineering roles. Institutions of Highways and Transportations in UK such as; The Institute of Highway Engineers (IHE), Road Safety GB Academy, and Chartered Institution of Highways and Transportation (CIHT) recognised this shortfall and already have included Education, Information, Training and Publicity in the content of their road safety engineering course as a specific subject (IHE, 2020 and Road Safety GB Academy, 2020). Unfortunately, educational and publicity interventions are ignored by most of the low and middle-income (LMIC) countries. Majority of the current road safety intervention researches (including SATAS19 data and interventions in this thesis) originate from high-income countries, despite LMIC countries bearing the more injury burden (Ameratunga et al., 2006; Perel et al., 2007).

4.4.5.6 Improving Effectiveness of Road Safety Campaigns through Humour

Online activities can help inspire the road users' interest in road safety learning. For example, game, video or ads with a series of humorous scenarios can be used to start learning travel, leading to discovering vital themes through fun, play and active experience. More specific to young people, this thesis believes that humour can be a key part of road safety to encourage safer behaviour and it is an effective way to engage the audience, and mainly a younger audience.

For example, a new campaign by Road Safety Scotland urges young people to drive like their grandmother is in the car with them. The 'Drive like Gran's in the car' humorous ads targeted at young male drivers in particular, shows a series of videos related to Grandmother characters, who unexpectedly appear in the car while their grandsons are driving (as shown in Figure 53) (Road Safety Scotland, 2020).

Figure 53: Drive Like Gran's in the Car (Road Safety Scotland, 2020)

Each video shows a different scenario, with the grandmothers addressing unsafe and bad driving behaviours including some of the key contributing factors that lead to casualties such as mobile phone use, alcohol and drug involvement, passenger distraction and driving too fast. It comes after research found that young lads change their driving behaviour depending on who they have in the car with them. It’s a fact that young people drive better with their grans in the car. According to an official survey, they drive more carefully when they are carrying "valuable cargo" like their grandmother in the car. Therefore, the ideology is that, next time during bad driving behavior, they should imagine Grandmother is in the car with them (Road Safety Scotland, 2020).

In addition, the concerns and suggestions made for decreasing the severity of injuries in traffic accidents are summarised in Table 15.

Table 15: The summary of concerns and suggestions made to improve road safety

Evidence-based Road Safety Interventions					
No	**Type of Intervention**	**Theme (s)**	**Location**	**Concern**	**Purpose of intervention**
1	Protected junctions (e.g. Glasgow style, Dutch-style roundabout, and CYCLOPS)	Travelling safely (All group specifically people travelling on foot and by cycle)	Intersections	Crossing busy junctions on foot or by cycle is a scary and difficult experience	Provide cyclists physical separation through an entire junction and eliminate stressful junctions with motor vehicles
2	Narrow cycle lane defenders	Cycling safety	Where carriageway width is too limited	Narrow bike lanes design flaws make a bike lane incredibly unsafe	Provide protected segregation for bikes from other traffic
3	Cycle lane soft segregation	Cycling safety	Where there is conjunction with mandatory cycle lane markings	Risk of cycling injury while sharing the road with motor vehicles	Provide light segregation by excluding other traffic from bike lane (not highly recommended due to being a compromise)
4	Continuous cycle lane segregation / Fully continuous	Cycling safety	Where passible and for over longer distances where	Vehicles entering a cycle lane must be deterred	Provide continuous segregation by excluding other traffic from

	cycle lane segregation		vehicle speeds may be higher		bike lane
5	Flexible, passively safe, highly visible bollards/wands	Cycling safety	Combined with defenders	Driver can crash into cycle lane bollards	Clearly direct where a cycle lane is to other road users/integrate reflectivity for maximum visibility
6	Trials of lower speed limit (e.g. 20mph)	Travelling safely (people travelling on foot and by cycle)	Busy and narrow roads in city centre	People biking and commuters encounter increasing traffic (especially when physically distancing due to Covid-19)	Provide safer conditions for VRUs by introducing a lower speed limit
7	Safety effectiveness of crossing enhancements (e.g. PiPencil bollard)	Crossing safety (VRUs)	At locations where designers wish to encourage VRUs to cross	Risk of injury to a cyclist/pedestrian while crossing a carriageway	Improving the visibility of pedestrian crossing points and providing more adequate facilities for crossing carriageways
8	Proper road safety education (e.g. humour ads)	Addressing the key contributing factors that led to casualties	Humorous scenarios addressing bad driving habits	Unsafe driving behaviours including distractions	Protecting specific road users and reducing devastating casualties by teaching life-saving messages

4.4.5.7 Future Study of Road Safety Education

In field of road safety education, publicity and training, further research needs to discover innovative teaching and learning approaches as well as platforms in direction of playing a dynamic role in laying these foundations. The evidence-based road safety intervention options of this thesis offer peace of mind for VRUs. The findings of this study can play a vital role in helping casualty reduction and prevention targets along with handling numerous road safety issues.

4.4.6 Advanced Automatic Collision Notification

The results can be also connected to vehicle safety for instance; advanced automatic collision notification (AACN). AACN captures contributing factors in real-time and announces the information to emergency responders, warning responders of the location and nature of the collision so they can be responded to more quickly. AACN is able to improve results among severely injured crash patients by predicting the probability of serious injury among vehicle occupants (Stitzel et al., 2016; Yoshida et al., 2016).

4.4.7 Standards and Guidelines on Road Safety Engineering

As Civil Engineering degrees from EMU (recognised globally as an international university) are accepted by UK National Recognition Information Centre (UK NARIC) and the Institution of Civil Engineers (ICE), this study aimed to shed some light on the potential of remedial measures and techniques while trying to meet road safety engineering requirements of the UK DfT, ICE, IHE, CIHT, Road Safety GB Academy, Road Safety Scotland, RoSPA, TRL, and local authorities.

Chapter 5

CONCLUSION

5.1 Thesis Statement

Although much of STATS19 data proves that UK has some of the safest roads in the world, there is still much more work to be done in order to improve the safety (Siamidoudaran and Iscioglu, 2019; Siamidoudaran et al., 2019a; Siamidoudaran et al., 2019b). In particular, collision history for South East England, London, and Cambridge alarm that more traffic injury prevention interventions should be done to protect specific groups. The UK government has a vision to avoid all road fatalities and mainly mitigate injury severities and subsequent costs and social influences from traffic collisions (DfT, 2018d). Therefore, there is an urgent need to detect the factors that significantly affect severity of the injuries caused by accidents.

5.2 Key Points and Contributions

In this thesis, RBFNN, MLPNN, SVM, Hybrid MLPNN-SVM, and LVQNN were applied to understand the relationship between injury severity levels as well as the factors that contribute to their generation. To achieve this goal, two case studies were considered on the basis of STATS19 road safety data.

5.2.1 Rank Analysis to Find Key Factors (RBFNN)

The objective of the first case study which refers to the city of London, was to predict personal injury severity levels through RBFNN for all road users. Using this technology that includes the interaction of input and output factors, the model predicted likelihood of the injury severities while classifying them into different levels. The prediction model was addressed as an identification system for key injury severity impact factors. The sensitive predictors were the key actions and failures that led directly to the actual influences.

5.2.2 Cycling as a Common Concern between Two Case Studies

As a result of the first case study, specific groups were recognised as needing more road safety interventions. In this regard, most vulnerable were detected as non–motorised road users, therefore, to solve the first case study's concern within an innovative way, the pedal rider group was considered for the UK's cycling capital case study. Although, Cambridge has clung on to this title, new statistics display that road safety concerns are the top barriers to pedal riding. Thus, RBFNN also shed some light on the potential of remedial techniques to predict injury severities sustained by the two–wheel group. In both case studies, the associated explanatory factors have been discovered and ranked as most import contributory factors or accident cluster sites.

5.2.3 Key Findings of the First Case Study – City of London

The first case study results warn that absence of required crossing facilities was responsible for the majority of injuries involving collisions with vehicles and VRUs wherein interventions should be a high priority. The probability of high injury severities increases as a result of poor manoeuvres at and around busy junctions; therefore, certain manoeuvres must be banned on some of the specific roads. Specific types of vehicles were correspondingly responsible for majority of the injuries. Certain point of the 'initial contact' in crashes was another contributory factor. Drug and alcohol consumption also had a negative effect on pedestrians' safety. Foot-travelers were injured as a result of being intoxicated by alcohol or illicit drugs in the city of London.

5.2.4 Key Findings of a Second Case study – UK's Every Day Cycling Capital

Within cyclist related injuries, the first factor refers to location of bicycle, in which cyclists relied on narrow lanes comprised of only a few inches of white paint to give riders on bikes a feeling of comfort and safety on busy mandatory cycle lanes. Modern protected cycle routes are suggested to fully separate bikes from other traffic where road width is too narrow. Like the first case study, T or staggered junctions on an unclassified bend was discovered to be particularly dangerous for the UK's cycling capital and was identified as a collision hotspot. Lack of a necessary number of crossing facilities for the riders were identified to have a higher affect on probability of the injuries. Following this position, possibly they just went across the street without a signal or they didn't completely appreciate how the facilities work leading to conflict or confusion. Installing protected junction in busy sites can definitely give priority to cyclists and pedestrians, which unfortunately is not a common practice in the UK. A higher accident involvement, with regard to their traffic volumes, was observed for specific times and days such as; rush hour during weekday or weekends.

5.2.5 Maximise Accuracy by Sensitive Predictors (MLPNN and SVM)

The models used with the intention of building an improved performance model through applying the key injury severity impact factors. To achieve this, rank analysis was done to find the key predictors, then, MLPNN was applied to the first case study and SVM was used for cycling injuries in Cambridge. MSE, RMSE and R–value was used to measure the relationship between actual crash related factors and predicted classes of the injury severity.

5.2.6 Comparison of MLPNN and SVM along with Severity Classes

As a result of a comparison among the classes, the predicted classes of 'slight injury' and 'damage only' achieved higher prediction accuracy in comparison to serious injury levels. In these classes, the models' predictions were very close to the actual dataset. Due to the lack of data for 'fatal class', the MLPNN and SVM totally failed to predict this class in both case studies. A comparison of model fitting results of the proposed ANN and SVM models show that the models are effectively capable in prediction of STATS19 data but MLPNN performs a little better than SVM. Howbeit, this thesis suggests SVM is also another viable modelling alternative for injury severity prediction.

5.2.7 Overcoming Limitation of Data (Hybrid MLPNN-SVM and LVQNN)

In response to this, a hybrid MLPNN-SVM and an improved LVQNN were considered to improve the prediction accuracies. In this regard, hybrid MLPNN–SVM was applied to the same dataset which was used for the city of London by MLPNN. Although the hybrid model was able to demonstrate an advanced development (approx. 20%) on predictions, the incorrect predictions for 'fatal' and 'serious injury' still persisted. Therefore, 'fatal' and 'serious injury' classes were combined within a single class as KSI to improve the quality of the data. In this vein, an improved LVQNN was successfully applied for building a better performance model by using the identical dataset which was used for Cambridge by SVM. As a result, the LVQNN prediction model achieved a higher overall accuracy compared to MLPNN, SVM, and hybrid MLPNN-SVM. The model was also able to obtain a supreme result for each injury severity class in terms of correct predictions.

5.2.8 Contribution of Study

The contribution of this thesis is multiple. First, the RBFNN model was used for traffic accident injury severity prediction. In general, a large series of ANN and SVM models have been used for traffic crash injury severity studies. However, after a thorough literature review, we identified a gap in the published studies on the methodology in traffic injury severity research through rank analysis. There are limited prediction applications using rank analysis to find key injury severity impact factors, which believes that applying most important contributory factors is the key approach for maximising prediction accuracy. This assumption may not be valid in predictions which use limited crash related factors.

There have been great efforts to develop prediction models for numerous road safety concerns, however, as far as the author is aware, prediction models have still not been developed in the UK for determining traffic accident or injury key predictors. Thus, in this thesis, prediction models for two case studies were set up on the basis of STATS 19 data.

STATS19 information is of great value to road safety practitioners, however, most research has paid only little attention to the contributory factors which STATS19 data specifically focuses on. Having more data is always a great way to improve prediction accuracy, therefore, the second major contribution of this thesis is applying mass of subdivision data to examine personal injury severity classes. We apply RBFNN to SATAS19 data to find out the relationship between injury severities and related contributing factors. Approximately 50 potential explanatory variables were examined initially using the RBFNN model. Each variable holds many new subdivision data (label) which are surely an innovation in the field of road safety. As a result of rank analysis, around 20 of them were found to be contributory in increasing the severity of injuries. Nevertheless, some commonly used variables such as weather condition, road surface condition; gender and age of road users were found to have no important impact on the injury severities. This outcome could have some key implications in road safety and warrant extra investigation which can be done in the future.

In addition, we discovered several groups and site clusters that needed to benefit from road safety interventions. In the reviewed literature, we also found a major gap on injury-related results of the predictions which was not a common key focus with the aim of discussing contributory factors and intervention options. To bridge this gap, we suggested several evidence-based intervention options and demonstrated how they could be applied to mitigate the concerns which have largely been ignored by researchers.

This thesis examined different models and compared them to find the best fit model for prediction of STATS19 data. In order to maximise model performance and to respond to the limitation of data, for the first time ever, a hybrid MLPNN-SVM and a LVQNN have been used in an accident injury severity related study by the author. Using the identical data for both case studies, performance of the predictive models along with prediction accuracies for each class were also compared and discussed. The result suggests that the models have the capability to handle sub-categorical data of STATS19 so they are promising tools for future crash injury severity studies with similar datasets. Nevertheless, LVQNN model was more accurate in predicting all the injury severity classes and fitted the data better than MLPNN, SVM, and the hybrid model.

Prediction models specifically focused on cycling injuries or accidents have rarely been investigated. Therefore, in response to the lack literature on prediction models for cyclist related injury severities, this study separately predicts cyclist injury severities for one of the most incredible cycling destinations in the world which is believed to be the most important contribution of this thesis. In this vein, bicycles were involved in the highest number of casualties on roads in Cambridge; one of the most bike-friendly cities in the world wherein cycling is considered as a significant means of transport.

5.3 The Significance of Research

The result of this thesis can be used as a function to assess the safety performance of the overall road networks. Having this methodology as a casualty reduction technique, local authorities and communities can point out the main safety concerns as well as their own priorities in their area to use to for engineering, education, and enforcement. We believe, the findings are more accurate compared to the reported factors which the attending police officer thought had contributed to the accident. Those factors are largely subjective, reflecting the opinion of the reporting police officer, and are not necessarily the outcome of wide-ranging investigation so may well not be completely reliable.

5.4 General Conclusion

The general conclusion that can be drawn from this thesis is that all the ANN, SVM, and the hybrid ANN-SVM have the ability to predict within accepTable limits. In view of that, the models are capable tools for predicting severity of personal injuries. However, combination of the fatal and serious injury classes is important to achieve accurate predictions.

Most commonly identified contributory factors occurring in both case studies were a result of human error with high accident concentration being attributed to busy junction actions and poor turning manoeuvres. To address this concern, along with the road engineering interventions (fully protected junctions and visible narrow cycle defenders) intervention options concerning the specific groups should be applied through professional road safety education, training and publicity.

REFERENCES

Abdel-Aty, M. A., & Abdelwahab, H. T. (2004). Predicting injury severity levels in traffic crashes: a modeling comparison. *Journal of transportation engineering*, *130*(2), 204-210.

Abdelwahab, H. T., & Abdel-Aty, M. A. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record*, *1746*(1), 6-13.

Abdelwahab, H. T., & Abdel-Aty, M. A. (2002). Artificial neural networks and logit models for traffic safety analysis of toll plazas. *Transportation Research Record*, 1784(1), 115-125.

Agilysis. (2020). Transport Behaviour & Safety Consultants. (2020, May 13). Retrieved from https://agilysis.co.uk

Al-Daoud, E. (2009). A Comparison Between Three Neural Network Models for Classification Problems. *Journal of Artificial Intelligence*, *2*(2), 56-64.

Alizamir, M., Kisi, O., Ahmed, A. N., Mert, C., Fai, C. M., Kim, S., ... & El-Shafie, A. (2020). Advanced machine learning model for better prediction accuracy of soil temperature at different depths. *Plos one*, *15*(4), e0231055.

Alkheder, S., Taamneh, M., & Taamneh, S. (2017). Severity prediction of traffic accident using an artificial neural network. *Journal of Forecasting*, *36*(1), 100-108.

Ameratunga, S., Hijar, M., & Norton, R. (2006). Road-traffic injuries: confronting disparities to address a global-health problem. *The Lancet*, *367*(9521), 1533-1540.

Amiri, A. M., Sadri, A., Nadimi, N., & Shams, M. (2020). A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers. *Accident Analysis & Prevention*, *138*, 105468.

BBC. (2020a). Coronavirus: Cycling numbers in Scotland jump during lockdown. (2020, June 18). Retrieved from https://www.bbc.co.uk/news/uk-scotland-52269964

BBC. (2020b). Coronavirus: Boom time for bikes as virus changes lifestyles. (2020, September 18). Retrieved from https://www.bbc.co.uk/news/business-52564351

Bellili, A., Gilloux, M., & Gallinari, P. (2003). An MLP-SVM combination architecture for offline handwritten digit recognition. *Document Analysis and Recognition*, 5(4), 244-252.

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford university press.

Broomhead, D. S., & Lowe, D. (1988a). *Radial basis functions, multi-variable functional interpolation and adaptive networks* (No. RSRE-MEMO-4148). Royal Signals and Radar Establishment Malvern (United Kingdom).

Broomhead, D. S., & Lowe, D. (1988b). Multivariable functional interpolation and adaptive networks, complex systems, vol. 2.

BSfE. (2018). Dangerous orcas and bus stop boarders? Better Streets for Enfield. (2020, October 01). Retrieved from http://betterstreets.co.uk/orcas-and-bus-stop-boarders-are-they-dangerous/

Bullinaria, J. A. (2004). Learning Vector Quantization (LVQ). *University of Birmingham*, UK.

Cambridgeshire County Council. (2020). Fendon Road and Queen Edith's Way Roundabout. (2018, July 13). Retrieved from https://www.cambridgeshire.gov.uk/residents/travel-roads-and-parking/transport-projects/cycling-pedestrian-improvements/fendon-road-and-queen-edith-s-way-roundabout

Cambridge Cycling Campaign. (2020). Protected junctions. (2020, October 02). Retrieved from https://www.camcycle.org.uk/blog/category/protected-junction

Cambridge News. (2017). Cambridge's position as UK's top cycling city challenged - by Cardiff. Cambridge: The Cambridge News. (2020, March 23) Retrieved from https://www.cambridge-news.co.uk/news/cambridge-news/cambridges-position-uks-top-cycling-12492349

Cambridge News. (2018). Five charts that show the shocking rise in carnage on Cambridgeshire's roads." Cambridge: The Cambridge News. (2019, March 21) Retrieved from https://www.cambridge-news.co.uk/news/local-news/road-accidents-cambridgeshire-death-injuries-14409910

CambridgeshireLive. (2018a). Why is Cambridge such a big cycling city?. (2018, July 13). Retrieved from https://www.cambridge-news.co.uk/news/cambridge-news/why-is-cambridge-cycling-city-14475326

CambridgeshireLive. (2018b). Cambridge is the UK's cycling capital - but it's on the decline. (2020, July 13). Retrieved from https://www.cambridge-news.co.uk/news/cambridge-news/cycling-cambridge-greenways-cam-cycle-15095519

CityLab. (2015). How Cambridge Became the U.K.'s model cycling city. Boston: Atlantic Monthly Group. (2017, July 18) Retrieved from https://www.citylab.com/transportation/2015/08/how-cambridge-became-the-uks-model-cycling-city/400550/

City of London Corporation. (2018). Road Danger Reduction and Active Travel Plan 2018 – 2023. Department of the Built Environment City Transportation.

(2018, July 15) Retrieved from https://www.cityoflondon.gov.uk/services/transport-and-streets/road-safety/Documents/road-danger-reduction-plan-2018.pdf

City of London Corporation. (2020). About the City Corporation. (2020, April 20) Retrieved from https://www.cityoflondon.gov.uk/about-the-city/about-us/Pages/default.aspx

Chang, L. Y., & Wang, H. W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis & Prevention*, *38*(5), 1019-1027.

Chen, S. (1995). Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning. *Electronics letters*, *31*(2), 117-118.

Chen, N., & Marques, N. C. (2009). Extending learning vector quantization for classifying data with categorical values. In *International Conference on Agents and Artificial Intelligence* (pp. 124-136). Springer, Berlin, Heidelberg.

Chen, C., Zhang, G., Qian, Z., Tarefder, R. A., & Tian, Z. (2016). Investigating driver injury severity patterns in rollover crashes using support vector machine models. *Accident Analysis & Prevention*, *90*, 128-139.

Chen, H., Chen, Q., Chen, L., & Zhang, G. (2016). Analysis of risk factors affecting driver injury and crash injury with drivers under the influence of alcohol (DUI) and non-DUI. *Traffic injury prevention*, *17*(8), 796-802.

Cheng, H., Zheng, N., & Qin, J. (2005). Pedestrian detection using sparse Gabor filter and support vector machine. *In IEEE Proceedings. Intelligent Vehicles Symposium*, 2005. (pp. 583-587). IEEE.

Connexions. (2001). A Little Book of Evaluation. Connexions. (2020, September 13). Retrieved from http://www.roadsafetyevaluation.com/docs/little-book-of-evaluation.pdf

Cortes, C. & Vapnik N. V. (1995). Support-vector networks. *Journal of Machine Learning*, 20(3), 273-297.

Costa, M. A., de Pádua Braga, A., & de Menezes, B. R. (2007). Improving generalization of MLPs with sliding mode control and the Levenberg–Marquardt algorithm. *Neurocomputing*, *70*(7-9), 1342-1347.

Cycling Weekly. (2018). Increase in Scotland's bike use is 'start of a cycling revolution. (2020, January 28) Retrieved from https://www.cyclingweekly.com/news/latest-news/increase-scotlands-bike-use-start-cycling-revolution-371156#8ZclHJ5f25SpyQCI.99

Cycling UK. (2018). Cycling UK's Cycling Statistics. (2019, January 28) Retrieved from https://www.cyclinguk.org/statistics

Daily Mail. (2020). 'A cyclist killing zone': Britain gets its first Dutch-style roundabout that gives priority to bicycles - but motorists condemn 'confusing' £2.3million project they fear will cause 'more accidents'. (2020, September 13). Retrieved from https://www.dailymail.co.uk/news/article-8599907/Britain-gets-Dutch-style-roundabout-gives-priority-cyclists.html

Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, *38*(3), 434-444.

Delen, D., Tomak, L., Topuz, K., & Eryarsoy, E. (2017). Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods. Journal of Transport & Health, 4, 118-131.

DfT. (2011). Instructions for the Completion of Road Accident Reports from non-CRASH Sources. Department for Transport. (2012, January 25) Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/230596/stats20-2011.pdf

DfT. (2014). Contributory factors to reported road accidents 2014. Department for Transport. (2020, September 25) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/463043/rrcgb2014-02.pdf

DfT. (2015a). The Highway Code. Department of Transport. (2018, January 25) Retrieved from https://www.gov.uk/guidance/the-highway-code/rules-for-cyclists-59-to-82

DfT. (2015b) Facts on Pedal Cyclists. London: Department for Transport. (2020, September 25) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/447674/pedal-cyclists-2013-data.pdf

DfT. (2016). Local area walking and cycling statistics: England, 2014/15, Department for Transport. (2017, January 24) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/536822/local-area-walking-and-cycling-in-england-2015.pdf

DfT. (2018a). Road Safety Statement Progress Report (Moving Britain Ahead). Department for Transport. (2019, March 25) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719259/road-safety-statement-progress-report.pdf

DFT. (2018b). Pedal cycling road safety factsheet. London: Department for Transport. (2019, April 25) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/686969/pedal-cycle-factsheet-2017.pdf

DfT. (2018c). Multimillion pound government funding boost for cycle safety, Department for Transport. (2020, January 20) Retrieved from https://www.gov.uk/government/news/multimillion-pound-government-funding-boost-for-cycle-safety

DfT. (2018d). Reported road casualties in Great Britain: 2017 annual report, Department for Transport. (2020, January 14) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/744077/reported-road-casualties-annual-report-2017.pdf

DfT. (2019a). The Road Safety Statement 2019 A Lifetime of Road Safety, Department for Transport. (2020, February 25) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/817695/road-safety-statement-2019.pdf

DfT. (2019b). Road traffic statistics, Department for Transport. (2020, January 25) Retrieved from https://www.gov.uk/government/collections/road-traffic-statistics

DfT. (2020). Cycling and Walking Safety: a rapid evidence assessment for the Department for Transport, NatCen Social Research. Department for Transport. (2020, September 19) Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/846488/Cycling_and_walking_safety_rapid_evidence_assessment.pdf

Fisher, A. R. (1936). The use of multiple measures in taxonomic problems. *Annals of Eugenics*, 7, 179-188.

Elliott, M. A., McColl, V. A., & Kennedy, J. V. (2003). Road design measures to reduce drivers' speed via'psychological'processes: a literature review. TRL.

Gårder, P., Leden, L., & Pulkkinen, U. (1998). Measuring the safety effect of raised bicycle crossings using a new research methodology. *Transportation Research Record*, 1636(1), 64-70.

GCP. (2017). Greater Cambridge–an innovative cycling destination, The Greater Cambridge Partnership. (2018, June 15) Retrieved from https://www.greatercambridge.org.uk/news/greater-cambridge-an-innovative-cycling-destination/

Google Maps. (2020). Cambridge, UK. Google. (2020, June 25) Retrieved from https://www.google.com/maps/place/Cambridge/@52.2012641,0.127004,13.04z/data=!4m5!3m4!1s0x47d8709be1a78bb3:0x40e1ca8bbc1e330!8m2!3d52.201776!4d0.1324319

Google Maps. (2020). The city of London, London, UK. Google. (2020, June 20) Retrieved from https://www.google.com/maps/place/City+of+London,+London/@51.5096348,-0.1148715,14z/data=!4m5!3m4!1s0x487603554edf855f:0xa1185c8d19184c0!8m2!3d51.5123443!4d-0.0909852

Guardian. (2011). What makes Cambridge a model cycling city. London: the Guardian. (2013, March 20) Retrieved from https://www.theguardian.com/environment/bike-blog/2011/aug/17/cambridge-model-cycling-city

Guo, Y., Osama, A., & Sayed, T. (2018). A cross-comparison of different techniques for modeling macro-level cyclist crashes. *Accident Analysis & Prevention*, *113*, 38-46.

Han, J., & Kamber, M. (2001). Data Mining Concepts and Techniques, Morgan Kaufmann Publishers. *San Francisco, CA*, 335-391.

Harvey, T. (1992). A review of current traffic calming techniques. Leeds, *UK: Institute for Transport Studies, University of Leeds.*

Hasheminejad, S. H. A., Zahedi, M., & Hasheminejad, S. M. H. (2018). A hybrid clustering and classification approach for predicting crash injury severity on rural roads. *International journal of injury control and safety promotion*, *25*(1), 85-101.

Hauer, E. (2007). A case for evidence-based road-safety delivery. *Improving Traffic Safety Culture in the United States*, 329.

Haykin, S. S. (2009). Neural Networks and Learning Machines. 3rd ed. New York: Prentice Hall.

Hébert, A., Guédon, T., Glatard, T., & Jaumard, B. (2019). High-Resolution Road Vehicle Collision Prediction for the City of Montreal. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 1804-1813). IEEE.

Heris, SMK., (2016). Pattern classification and pattern recognition training with LVQ neural networks in MATLAB. *FaraDars*. ID 8-12452-049624. (2018, October 01) Retrieved from https://faradars.org/courses/fvrml110-vector-quantization-neural-networks-matlab

Hinton, G. E. (1992). How neural networks learn from experience. *Scientific American*, *267*(3), 144-151.

Huang, H., Zeng, Q., Pei, X., Wong, S. C., & Xu, P. (2016). Predicting crash frequency using an optimised radial basis function neural network model. *Transportmetrica A: transport science*, *12*(4), 330-345.

Hughes, W., & Amis, G. (1996). Accidents on Rural Roads: Single Carriageway 'A' Class Roads. *AA Foundation for Road Safety Research*, Basingstoke, Hants.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, *22*(4), 679-688.

IHE. (2020). Institute of Highway Engineers. Road Safety Engineering course. (2020, October 01) Retrieved from https://www.theihe.org/courses/road-safety-engineering/

Inews. (2018). Most Scots think cycling is 'not for them', national survey finds. (2020, June 20) Retrieved from https://inews.co.uk/news/scotland/most-scots-think-cycling-is-not-for-them-national-survey-finds-239238

Iranitalab, A., & Khattak, A. (2017). Comparison of four statistical and machine learning methods for crash severity prediction. *Accident Analysis & Prevention*, *108*, 27-36.

ISJ. (2020). Rosehill protects cyclists and pedestrians with lane Segregation Products. International Security Journal. (2020, October 01) Retrieved from https://internationalsecurityjournal.com/rosehill-protects-cyclists-and-pedestrians-with-lane-segregation-products/#:~:text=Where%20carriageway%20width%20is%20limited,for%20continuous%20and%20light%20segregation

Jo, Y. B., Na, W. S., Eom, S. J., & Jeong, Y. J. (2014). Traffic Sign Recognition using SVM and Decision Tree for Poor Driving Environment. *Journal of IKEEE*, *18*(4), 485-494.

Karlaftis, M. G., & Vlahogianni, E. I. (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, *19*(3), 387-399.

Kim, J. K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle–motor vehicle accidents. *Accident Analysis & Prevention*, *39*(2), 238-251.

Kohonen, T. (2012). *Self-organizing maps* (Vol. 30). Springer Science & Business Media.

Knowles, J., Adams, S., Cuerden, R., Savill, T., Reid, S., & Tight, M. (2009). Collisions involving cyclists on Britain’s roads: establishing the causes. *London: TRL Limited*.

Kwak, Y. T., Hwang, J. W., & Yoo, C. J. (2011). A new damping strategy of Levenberg-Marquardt algorithm for multilayer perceptrons. *Neural Network World*, *21*(4), 327.

Lee, J., & Abdel-Aty, M. (2018). Macro-level analysis of bicycle safety: Focusing on the characteristics of both crash location and residence. *International journal of sustainable transportation*, 12(8), 553-560.

Lehmann, E. L., & Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

Li, T., Yang, Y., Wang, Y., Chen, C. and Yao, J., 2016. Traffic fatalities prediction based on support vector machine. *Archives of Transport*, *39*.

Li, X., Lord, D., Zhang, Y., & Xie, Y. (2008). Predicting motor vehicle crashes using support vector machine models. *Accident Analysis & Prevention*, *40*(4), 1611-1618.

Li, Y., Ma, D., Zhu, M., Zeng, Z., & Wang, Y. (2018). Identification of significant factors in fatal-injury highway crashes using genetic algorithm and neural network. *Accident Analysis & Prevention*, *111*, 354-363.

Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. *Accident Analysis & Prevention*, *45*, 478-486.

Lokala, U., Nowduri, S., & Sharma, P. K. (2017)., Road accidents big datamining and visualization using Support Vector Machines (SVM), *World Academy of Science, Engineering and Technology International Journal of Computer and Systems Engineering*, vol. 10(8), 2016.

Lopez, L. M., Tolley, E. E., Grimes, D. A., & Chen-Mok, M. (2009). Theory-based interventions for contraception. *Cochrane Database of Systematic Reviews*, (1).

Luria, J. W., Smith, G. A., & Chapman, J. I. (2000). An evaluation of a safety education program for kindergarten and elementary school children. *Archives of pediatrics & adolescent medicine*, *154*(3), 227-231.

Mandalia, H. M., & Salvucci, M. D. D. (2005). Using support vector machines for lane-change detection. *In Proceedings of the human factors and ergonomics* *society annual meeting* (Vol. 49, No. 22, pp. 1965-1969). Sage CA: Los Angeles, CA: SAGE Publications.

Marshals. (2020). PIPENCIL BOLLARD. (2019, June 25) Retrieved from https://www.marshalls.co.uk/commercial/product/pipencil-bollard#:~:text=The%20Ferrocast%20PiPencil%20Bollard%20is,a%20wide%20range%20of%20landscapes.&text=Bollards%20from%20Marshalls%20are%20root,can%20be%20remov

MathWorks. (2020). Learning Vector Quantization (LVQ) Neural Networks. (2020, December 01) Retrieved from https://www.mathworks.com/help/deeplearning/ug/learning-vector-quantization-lvq-neural-networks-1.htmlable%2.

McCormick C. (2013). Radial basis function network (rbfn). (2019, June 25) Retrieved from https://mccormickml.com/2013/08/15/radial-basis-function-network-rbfn-tutorial/

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), 115-133.

McKenna, F. (2010). Education in Road Safety: Are we getting it right? London: RAC Foundation. 4. (2020, September 10) Retrieved from https://www.racfoundation.org/wp-content/uploads/2017/11/education-in-road-safety-mckenna-080910-report.pdf

Minsky, M. & Papert, S., (1969). Perceptron: an introduction to computational geometry. *The MIT Press*, Cambridge, expanded edition, 19(88), p.2.

Mokhtarimousavi, S., Anderson, J. C., Azizinamini, A., & Hadi, M. (2019). Improved support vector machine models for work zone crash injury severity prediction and analysis. *Transportation research record*, 2673(11), 680-692.

Murty, M. N., & Raghava, R. (2016). Kernel-based SVM. In *Support vector machines and perceptrons* (pp. 57-67). Springer, Cham.

Ngan, H. Y., Yung, N. H., & Yeh, A. G. (2015). A comparative study of outlier detection for large-scale traffic data by one-class SVM and kernel density estimation. In *Image Processing: Machine Vision Applications VIII* (Vol. 9405, p. 94050I). International Society for Optics and Photonics.

Nova, D., & Estévez, P. A. (2014). A review of learning vector quantization classifiers. *Neural Computing and Applications*, 25(3-4), 511-524.

Osama, A., & Sayed, T. (2016). Evaluating the impact of bike network indicators on cyclist safety using macro-level collision prediction models. *Accident Analysis & Prevention*, *97*, 28-37.

Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. *Neural computation*, *3*(2), 246-257.

Perel, P., Ker, K., Ivers, R., & Blackhall, K. (2007). Road safety in low-and middle-income countries: a neglected research area. *Injury Prevention*, 13(4), 227-227.

PKC. (2020). Spaces for People bid plans for increased pedestrian and cyclist safety in Perth and Kinross. Perth & Kinross Council. (2020, July 15) Retrieved from https://www.pkc.gov.uk/article/21717/Spaces-for-People-bid-plans-for-increased-pedestrian-and-cyclist-safety-in-Perth-and-Kinross

Pradhan, B., & Sameen, M. I. (2020). Review of traffic accident predictions with neural networks. In *Laser scanning systems in highway and safety assessment* (pp. 97-109). Springer, Cham.

Prati, G., De Angelis, M., Marín Puchades, V., Fraboni, F., & Pietrantoni, L. (2017a) Characteristics of cyclist crashes in Italy using latent class analysis and association rule mining. PLOS ONE 12(2): 1–28.

Prati, G., Pietrantoni, L., & Fraboni, F. (2017b). Using data mining techniques to predict the severity of bicycle crashes. *Accident Analysis & Prevention*, *101*, 44-54.

Principe, J. C., Euliano, N. R., & Lefebvre, W. C. (2000). *Neural and adaptive systems: fundamentals through simulations* (Vol. 672). New York: Wiley.

Priyono, A., Ridwan, M., Alias, A. J., Rahmat, R. A. O., Hassan, A., & Ali, M. A. M. (2005). Application of LVQ Neural Network in Real–Time Adaptive Traffic Signal Control. *Jurnal Teknologi*, *42*(1), 29-44.

Ray, S. (2015). 8 Proven Ways for improving the "Accuracy" of a Machine Learning Model. Analytics Vidhya.. (2020, September 15) Retrieved from https://www.analyticsvidhya.com/blog/2015/12/improve-machine-learning-results/

Rediweld. (2020). Rediweld Rubber and Plastics Limited 2020 Greenwich Wand Orca (2020, October 02) Retrieved from http://www.rediweldtraffic.co.uk/products/cycle-lane-products/greenwich-wand-orca/

Ren, J. (2012). ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging. *Knowledge-Based Systems*, *26*, 144-153.

Road Safety GB 2020. Academy Introduction to Road Safety Delivery. Road Safety GB Academy. (2020, September 17) Retrieved from http://rsgbacademy.org.uk/training-courses/introduction-to-road-safety-delivery-online-course/

Road Safety Scotland. (2020). Drive like Gran's in the car. (2020, September 29) Retrieved from https://roadsafety.scot/topics/young-drivers/

Rosehill Highways. (2020). Cycle Lane Segregation. (2020, September 17) Retrieved from https://www.rosehillhighways.com/products/cycle-lane-defenders/

RoSPA (2017a). Guide, A. P. S. Designing Evidence Based Road Safety Interventions. (2020, September 25) Retrieved from https://www.rospa.com/rospaweb/docs/advice-services/road safety/practitioners/evidence-based-intervention-guide.pdf

RoSPA. (2017b). RoSPA Policy Paper: Cycling. Royal Society for the Prevention of Accidents. (2019, June 25) Retrieved from https://www.rospa.com/rospaweb/docs/advice-services/road-safety/cyclists/cycling-policy-paper.pdf

RoSPA. (2017c). Road safety factsheet - cycling accidents. Royal Society for the Prevention of Accidents. (2020, September 27) Retrieved from https://www.rospa.com/rospaweb/docs/advice-services/road-safety/cyclists/cycling-accidents-factsheet.pdf

RoSPA. (2018). Road Safety Advice, Information & Resource. Royal Society for the Prevention of Accidents. (2019, March 20) Retrieved from https://www.rospa.com/road-safety/

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accident Analysis & Prevention*, *43*(5), 1666-1676.

Schalkoff, R. J. (1997). *Artificial neural networks*. McGraw-Hill Higher Education. New York, PP.146-188.

Shen, X., & Chen, J. (2009, April). Study on prediction of traffic congestion based on LVQ neural network. *In 2009 International conference on measuring technology and mechatronics automation* (Vol. 3, pp. 318-321). IEEE.

Shamsashtiany, R., & Ameri, M. (2018). Road Accidents Prediction with Multilayer Perceptron MLP modelling Case Study: Roads of Qazvin, Zanjan and Hamadan. *Journal of civil Engineering and Materials Application*, *2*(4), 181-192.

Sharma, B., Katiyar, V. K., & Kumar, K. (2016). Traffic accident prediction model using support vector machines with Gaussian kernel. In *Proceedings of fifth international conference on soft computing for problem solving* (pp. 1-10). Springer, Singapore.

Siamidoudaran, M., & İşçioğlu, E. (2019). Injury severity prediction of traffic collision by applying a series of neural networks: the City of London Case Study. *Promet-Traffic&Transportation*, *31*(6), 643-654.

Siamidoudaran, M., Iscioglu, E., & Siamidodaran, M. (2019a). Traffic injury severity prediction along with identification of contributory factors using learning vector quantization: A case study of the city of London. *SN Applied Sciences*, *1*(10), 1268.

Siamidoudaran, M., Siamidodaran, M., & Iscioglu, E. (2019b). Predictive modelling of injury severity in bicycle–motor vehicle collisions utilizing learning vector quantization: a case study of Britain's cycling capital. *International Journal of Crashworthiness*, 1-11.

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis & Prevention*, *45*, 382-391.

Stitzel, J. D., Weaver, A. A., Talton, J. W., Barnard, R. T., Schoell, S. L., Doud, A. N., ... & Meredith, J. W. (2016). An injury severity-, time sensitivity-, and predictability-based advanced automatic crash notification algorithm improves motor vehicle crash occupant triage. *Journal of the American College of Surgeons*, *222*(6), 1211-1219.

Suratgar, A. A., Tavakoli, M. B., & Hoseinabadi, A. (2005). Modified Levenberg-Marquardt method for neural networks training. *World Acad Sci Eng Technol*, *6*(1), 46-48.

SWOV. (2010). SWOV Fact sheet: crossing facilities for cyclists and pedestrians. Leidschendam: Institute for Road Safety Research. (2020, September 27) Retrieved from https ://www.swov.nl/sites /defau lt/files /publi catie s/gearc hiveerde-facts heet/uk/fs_cross ing_facil ities archi ved.pdf

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2012). Using multivariate statistics. 6th edition. Boston, MA: Pearson

Taylor, M. C., Baruya, A., & Kennedy, J. V. (2002). *The relationship between speed and accidents on rural single-carriageway roads* (Vol. 511). Crowthorne: TRL.

TfGM. (2019). Transport for Greater Manchester. Greater Manchester's cycling and walking network CYCLOPS – Creating Protected Junctions. (2020, September 27) Retrieved from https://assets.ctfassets.net/nv7y93idf4jq/20Kq0JNhFmtlp5vm9glsQn/eaa4b299c69522e526b03180caaaa0b2/19-1369_CYCLOPS_technical_guide_A4_v3_Hi-Res.pdf

Thanasarn, T., & Warisarn, C. (2013, August). Comparative analysis between BP and LVQ neural networks for the classification of fly height failure patterns in HDD manufacturing process. *In Proceeding conference paper ICEAST* (Vol. 4)

Theodoridis, S. (2008). Pattern Recognition: Elsevier Academic Press. p. 203.

Tifani, Y., Lee, B., & Jeong, E. (2017). A patient's indoor positioning algorithm using Artificial Neural Network and SVM. *Journal of Theoretical & Applied Information Technology*, 95(16).

TRL (2010). TRL. Guidance for the analysis of STATS19 Contributory Factors. The Future of Transport. (2020, September 20) Retrieved from http://trlcms.demo.varistha.co.uk/uploads/trl/documents/PPR488.pdf

Vandenbulcke, G., Thomas, I., & Panis, L. I. (2014). Predicting cycling accident risk in Brussels: a spatial case–control approach. *Accident Analysis & Prevention*, *62*, 341-357.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Vardaki, S., Bekiaris, E., & Kanellaidis, G. (2018). Addressing education and training needs of road safety auditors. *Cogent Engineering*, *5*(1), 1525814.

Venkat, A., KP, G. V., & Thomas, I. S. (2020). Machine Learning Based Analysis for Road Accident Prediction. *Machine Learning Based Analysis for Road Accident* Prediction (March 7, 2020). IJETIE, 6(2).

Villmann, T., Bohnsack, A., & Kaden, M. (2017). Can learning vector quantization be an alternative to SVM and deep learning?-recent trends and advanced variants of learning vector quantization for classification learning. *Journal of Artificial Intelligence and Soft Computing Research*, 7(1), 65-81.

Wahab, L., & Jiang, H. (2019). A comparative study on machine learning based algorithms for prediction of motorcycle crash severity. *PLoS one*, *14*(4), e0214966.

Wang, Z. O., & Zhu, T. (2000). An efficient learning algorithm for improving generalization performance of radial basis function neural networks. *Neural Networks*, *13*(4-5), 545-553.

Wei, F., & Lovegrove, G. (2013). An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. *Accident Analysis & Prevention*, *61*, 129-137.

WHO. (2018a). Global status report on road safety 2018. World Health Organization. (2019, March 11) Retrieved from https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/

WHO. (2018b). Road traffic injuries. World Health Organization. (2020, March 17) Retrieved from http://www.who.int/mediacentre/factsheets/fs358/en/

Wilkins, D. C., Kockelman, K. M., & Jiang, N. (2019). Animal-vehicle collisions in Texas: how to protect travelers and animals on roadways. *Accident Analysis & Prevention*, 131, 157-170.

Wondimagegnehu, M.B. and Alemu, G., 2017, September. Short time traffic flow identification and prediction at road junctions using Recurrent Multilayer Perceptron. In *2017 IEEE AFRICON* (pp. 15-20). IEEE.

World Population Review. (2020). Cambridge Population 2020. (2020, June 17) Retrieved from http://worldpopulationreview.com/world-cities/cambridge-population/

Yan, B. A. I., & Guang-si, M. A. (2008). Traffic Prediction and Analysis Based on the Grey Radial Basis Function Neural Network Model [J]. *Computer Engineering & Science*, *10*.

Ye, F., & Lord, D. (2014). Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic methods in accident research*, *1*, 72-85.

Yoshida, S., Hasegawa, T., Tominaga, S., & Nishimoto, T. (2016). Development of injury prediction models for advanced automatic collision notification based on Japanese accident data. *International journal of crashworthiness*, *21*(2), 112-119

Yuan, F., & Cheu, R. L. (2003). Incident detection using support vector machines. *Transportation Research Part C: Emerging Technologies*, 11(3-4), 309-328.

Yu, R., & Abdel-Aty, M. (2014). Analysing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety science*, *63*, 50-56.

Yu, B., Wang, Y. T., Yao, J. B., & Wang, J. Y. (2016). A comparison of the performance of ANN and SVM for the prediction of traffic accident duration. *Neural Network World*, *26*(3), 271.

Yu, R., & Liu, X. (2010, December). Study on traffic accidents prediction model based on RBF neural network. In *2010 2nd International Conference on Information Engineering and Computer Science* (pp. 1-4). IEEE.

Yu, H., & Wilamowski, B. M. (2011). Industrial electronics handbook. *Levenberg-Marquadt Training*. *5*(12), p.1.

Xie, Y., Lord, D., & Zhang, Y. (2007). Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. *Accident Analysis & Prevention*, *39*(5), 922-933.

Xie, Y., Zhao, K., & Huynh, N. (2012). Analysis of driver injury severity in rural single-vehicle crashes. *Accident Analysis & Prevention*, *47*, 36-44.

Zeng, Q., & Huang, H. (2014). A sTable and optimized neural network model for crash injury severity prediction. Accident Analysis & Prevention, 73, 351-358.

Zhai, X., Huang, H., Xu, P., & Sze, N. N. (2019). The influence of zonal configurations on macro-level crash modeling. *Transportmetrica A: transport science*, *15*(2), 417-434.

Zhang, J., Li, Z., Pu, Z., & Xu, C. (2018). Comparing prediction performance for crash injury severity among various machine learning and statistical methods. *IEEE Access*, *6*, 60079-60087.