Pattern recognition and classification of fatal traffic accidents in Israel: a neural network approach

Carlo Giacomo PRATO *

Department of Transport

Technical University of Denmark

Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark

Victoria GITELMAN

Ran Naor Road Safety Research Center

Technion - Israel Institute of Technology

Technion Campus, Haifa 32000, Israel

Shlomo BEKHOR

Faculty of Civil and Environmental Engineering

Technion - Israel Institute of Technology

Technion Campus, Haifa 32000, Israel

* corresponding author

Telephone: +45.45256595, Fax: +45.45936593

E-mail: cgp@transport.dtu.dk

Abstract

This paper provides a broad picture of fatal traffic accidents in Israel in order to answer an increasing need of addressing compelling problems, designing preventive measures, and targeting specific population groups with the objective of reducing the number of traffic fatalities. The analysis focuses on 1,793 fatal traffic accidents occurred during the period between 2003 and 2006, and applies Kohonen and feed-forward back-propagation neural networks with the objective of extracting from the data typical patterns and relevant factors.

Kohonen neural networks reveal five compelling accident patterns: (i) single-vehicle accidents of young drivers; (ii) multiple-vehicle accidents between young drivers; (iii) accidents involving motorcyclists or cyclists; (iv) accidents where elderly pedestrians crossed in urban areas; (v) accidents where children and teenagers cross major roads in small urban areas. Feed-forward back-propagation neural networks indicate that socio-demographic characteristics of drivers and victims, accident location and period of the day are extremely relevant factors. Accident patterns suggest that countermeasures are necessary for identified problems concerning mainly vulnerable road users such as pedestrians, cyclists, motorcyclists and young drivers. A “safe-system” integrating a system approach for the design of countermeasures and a monitoring process of performance indicators might address the priorities highlighted by the neural networks.

Keywords: Accident Patterns; Accident Factors; Cluster Analysis; Kohonen Networks; Feed-Forward Back-Propagation Neural Networks.

1. Introduction

Traffic accidents are one of the major causes of trauma death among the Israeli population, as every year over 400 people are killed and thousands more are injured with foreseeable costs to society in terms of human lives, property damages and road users' delays. The recognition of accident patterns and the individuation of major accident factors constitute an increasing need in order to address compelling problems, to design countermeasures and to target specific population groups with the ultimate objective of reducing the annual number of traffic fatalities and accidents. This study intends to provide a broad picture of Israeli traffic accidents by solving a multi-faceted challenge in the data to be considered, the methodology to be applied and the results to be obtained. This study extends a preliminary study focusing on the specific problem of pedestrian accidents (Prato et al., 2011) and evaluates whether this latter problem is compelling in the global picture of road accidents in the country.

From the data perspective, this study focuses on all fatal accidents occurred in Israel during the four-year period between 2003 and 2006. The focus on fatal accidents has a twofold motivation. On the one hand, political commitment toward saving human lives has been traditionally measured in terms of reduction of the number of fatalities in Israel and all over the world (OECD 2008). On the other hand, considering severe injury accidents would introduce a bias in the analysis because of extensively documented problems in reporting both severe and light injuries in Israel (see, e.g., Peleg and Aharonson-Daniel 2004; Avitzour 2007). The focus on the four-year period has a twofold motivation as well. On the one hand, considering a long enough period allows limiting random fluctuations in the accident counts. On the other hand, analyzing a short enough period enables limiting the effect of changes in road and traffic conditions.

From the methodology perspective, this study applies Kohonen neural networks with the objective of extracting from the data relevant information about accident patterns without a priori assumptions about the expected outcome of the study. Complementarily, this study applies feed-forward back-propagation neural networks with the objective of underlying the most relevant factors behind crash occurrence. Existing methods for the individuation of accident patterns include a-priori definition of accident types (e.g., Preusser et al. 1995; Retting et al. 1995; Fleury and Brenac 2001; Laapotti and Keskinen 2004; Retting et al. 2000), frequency analysis combined with odd ratio calculation (e.g., Zegeer et al. 1993; Harruff et al. 1998; Ballesteros et al. 2004; Sivak and Schoettle 2010), factorial analysis of correspondence combined with hierarchical ascendant classification (e.g., Fontaine and Gourlet 1997; Berg et al. 2004; Hasselberg et al. 2005; Wang et al. 2008; Skyving et al. 2009), data mining techniques for accident clustering (e.g., Chang and Chen 2005; Geurts et al. 2003; Geurts et al. 2005; Tseng et al. 2005; Yan and Radwan 2006), and econometric models (e.g., Eluru et al. 2008; Milton et al. 2008). The advantage of implementing data mining techniques consists in the ability of unraveling accident patterns from a large number of accident characteristics without the need of assuming accident types prior to the data analysis as in a-priori accident type definition, ignoring the multidimensionality of the accident analysis problem as in frequency analyses, restricting significantly the number of examined variables as in factorial analysis of correspondence, or introducing restrictive assumptions about the relationship between dependent and independent variables as in econometric models.

From the results perspective, this study intends to consider the multidimensionality of the problem that is often neglected in previous studies about accident patterns. Some studies consider only the modality of the accident as the characteristics discerning the type of accident (e.g., Preusser et al. 1995; Retting et al. 1995; Retting et al. 2000). Other studies analyze separately characteristics of accidents in relation to the crash occurrence without examining their contextual effect (e.g., Zegeer et al. 1993; Harruff et al. 1998; Ballesteros et al. 2004; Laapotti and Keskinen 2004; Sivak and Schoettle 2010). Recent studies acknowledge the multidimensionality of the problem, although with limitations imposed by the applied methodology (e.g., Fontaine and Gourlet 1997; Geurts et al. 2003; Berg et al. 2004; Chang and Chen 2005; Geurts et al. 2005; Hasselberg et al. 2005; Tseng et al. 2005; Yan and Radwan 2006 ;Wang et al. 2008; Skyving et al. 2009). Also from the results perspective, this study intends to obtain patterns that are less predictable than obvious ones such as "accidents with rainy weather, wet surface and slippery conditions" (see Geurts et al. 2003). Data mining techniques are often criticized because of their inability to obtain non-trivial results in terms of safety recommendations and stable solutions in terms of accident patterns. Accordingly, this study partitions the fatal accident dataset with respect to different main accident characteristics and compares the results in order to uncover consistent accident patterns that represent compelling safety issues that need to be prioritized. Moreover, this study applies one-dimensional Kohonen networks (i.e., 1×J maps where J is the number of clusters) for pattern recognition, in order to find stable solutions converging to a global optimum (e.g., Ritter and Schulten 1988; Erwin et al. 1992; Cottrell and Rousset 1997) as demonstrated by their application to operations research for the Traveling Salesman Problem (e.g., Burke 1996; Aras et al. 1999) and to oncology research for anticancer drug analysis (e.g., Augen 2004).

The remainder of this paper is structured as follows. The next section introduces data and methods applied in the current study. The following section presents results from the implementation of Kohonen and feed-forward back-propagation neural networks. The last section discusses the findings of the study in terms of policy implications.

2. Methods

2.1 Accident dataset

The fatal accident dataset is prepared on the basis of data gathered by the Israeli Police and organized by the Israeli Central Bureau of Statistics in three different files, namely the accident file, the vehicle and driver file, the injured person file.

The accident file reports general information about the accident such as severity level (i.e., fatal, severe, light), date and time, and geographical location (i.e., urban vs. rural and section vs. intersection). Then, the accident file contains information about the infrastructure such as allowed speed, road width, presence and condition of median barrier, traffic lights, and road signals in general, conditions of the surface, and weather at the moment the accident happened. Last, the accident file comprises information concerning pedestrians and collided objects for specific types of accidents. The vehicle and driver file includes information about each vehicle and driver involved in the accident, and lists vehicle features (e.g., type, age, motor, weight and direction of travel of the vehicles) and driver characteristics (e.g., gender, age, licensing year and birth location). The injured person file contains information regarding each injured person including pedestrians and registers gender, age, birth location, population group, place of residence and type of injury sustained.

Accidents with fatal outcome are extracted from the accident file because of the extensively documented problems regarding the collection of reliable information on both severe and light injuries in Israel (see, e.g., Peleg and Aharonson-Daniel 2004; Avitzour 2007). After extracting the accidents with fatal outcome, the fatal accident dataset results from merging the three data files. As the accident file contains one accident for each record and the other two files contain multiple records corresponding to drivers, vehicles and persons involved in each accident, a unique file is composed by extracting fatal accidents from the accident file, and adding progressively vehicles, drivers and persons involved to the extracted records. The consistency of the coding across the four-year period allows avoiding potential problems related to inconsistencies in the data definition. The absence of missing values and the police attention in reporting fatal accidents allows avoiding potential problems related to the trustworthiness of the information collected (Peleg and Aharonson-Daniel 2004).

The fatal accident dataset consists of 1,793 records and the categorical variables considered for data analysis are summarized in table 1.

[Insert Table 1]

2.2 Data mining techniques

A conceptual description of Kohonen neural networks for accident pattern recognition and feed-forward back-propagation neural networks for relevant factor identification are provided in the following sub-sections.

2.2.1 Kohonen networks

Kohonen networks are unsupervised learning methods that allow treating large samples, obtaining comprehensible patterns without any predefined assumption and reaching stable optimal solutions under the one-dimensional configuration (see, e.g., Ritter and Schulten 1988; Erwin et al. 1992; Cottrell and Rousset 1997; Augen 2004).

A Kohonen neural network (Kohonen 1982; Kohonen 2001) is a self-organizing map that is trained in an unsupervised mode. The input layer is given by neurons that correspond to each fatal accident with its characteristics expressed in terms of categorical values of the considered variables, while the output layer corresponds to fatal accident patterns. The input neurons cause a reaction from the output neurons by being presented during a learning process that involves stepping through several cycles until the error of the network is below an acceptable level. Since the Kohonen network is trained in an unsupervised mode, the traditional definition of error as the difference between predicted and observed outputs does not apply. However, since the purpose of the network is to classify the input into clusters, the error for the Kohonen network is a measure of the ability of the network in classifying the records and computing the relevance of the input neurons to the process.

Initially, the network presents records of the input layer to neurons of the output layer and assigns random weights to the connections between the layers. A competitive learning algorithm compares the characteristics of each input neuron with those of all output neurons, and the output neuron with the most similar characteristics to the input record is the “winner”. Thus, the weight of the “winner” output neuron is adjusted to be more similar to that of the record just acquired in order to enhance the likelihood of similar records to be captured by the same node. Then, the network adjusts the weights of the neighboring neurons as well. The process is repeated until the weights converge, and the result consists of a map containing clusters corresponding to different patterns in the data, where similar clusters are close and dissimilar clusters are at opposite sides of the map.

The algorithm works in two phases that are defined through different learning parameters and different numbers of neighbors to be modified when the “winner” neuron is identified. High learning parameters cause fast training progress, but also generate the risk of non-convergent network because the oscillations of the weight vectors would be too great for the classification patterns to ever emerge. Accordingly, the learning parameter is usually programmed at 0.4 or 0.5

for learning faster in the first phase, and at 0.1 or 0.2 for fine-tuning of the map in the second phase. Similarly, the number of neighbors to be modified is 2 or 3 in the first phase, when weights are modified for several clusters to capture similarities among adjacent output neurons, and 1 in the last phase, when weights are modified to fine-tune differences among adjacent output neurons.

2.2.2 Feed-forward back-propagation neural networks.

Feed-forward back-propagation neural networks are supervised learning methods (see, e.g., Reed and Marks 1999) that connect an input layer, which corresponds to each fatal accident with its characteristics expressed in terms of categorical values of the considered variables, to an output layer, which corresponds to the categories of the expected output variable, through a hidden layer.

The term "feed-forward" indicates that the neural network connects neurons only forward, thus each layer of the network contains only forward links (e.g., from the input to the hidden layer) without any backward link. The term "back-propagation" indicates that the neural network performs supervised training by comparing anticipated outputs against the actual outputs from the neural network. Using the anticipated outputs, the "back-propagation" training algorithm calculates the error and adjusts the weights of the various layers backwards from the output layer to the input layer. The algorithm also calculates the relevance of each input neuron as value that expresses the importance of each variable in making a prediction.

Initially, the network learning process assigns random weights and obtains random answers. Then, the network examines single input neurons, forecasts the output and corrects the weights each time a forecast is incorrect with respect to the anticipated outcome. Several parameters determine the development of the learning phase: the alpha parameter refers to the weight update and tends to move the weight changes in a constant direction to reduce the training time; the eta parameter refers to the learning rate and determines how much adjustment is feasible at each update and decreases according to a predetermined number of cycles; the persistence parameter defines the number of cycles for which the network trains without improvement before reaching the stopping point.

Typically, the structure of the neural network depends on both the number of input and output neurons, and the number of hidden layers and their neurons. Ultimately, three processes are available: the "forward selection method" selects a small number of hidden neurons to train and to test the neural network, and iteratively increases the number of hidden neurons as long as the overall results of the training and testing improve; the "backward selection method" uses a large number of hidden neurons to train and to test the neural network, and iteratively decreases the number of hidden neurons until the performance improvement is no longer significant; the "pruning method" evaluates the weighted connections between the layers and removes from the network hidden neurons that contain only zero weighted connections.

2.2.3 Implementation of data mining techniques

The implementation of the described neural networks with Clementine software (SPSS Inc. 2007) aims to uncover accident patterns and relevant factors from the fatal accident dataset. Even though neural networks do not make any assumption prior to the analysis of the data, literature about data mining applications to traffic accident analysis shows the risk of obtaining results that are too trivial to be relevant. Specifically, typical results illustrate that main characteristics of the accident (e.g., location, accident type, victim type) tend to guide the clustering process toward the division of the records according to predictable and not very insightful patterns such as "accidents with rainy weather, wet road surface and slippery conditions" that do not really require complex statistical methods to be uncovered.

Accordingly, this study applies Kohonen networks to partitions of the accident dataset according to three main characteristics of the accidents (i.e., type of accident, location of the accident and type of victim) in order to limit the possibility of defining only trivial clusters and to verify the consistency of the clusters obtained with different data starting points. Four data partitions are defined in accordance with the three accident characteristics selected: accident type (i.e., pedestrian, single vehicle and multiple vehicle), urban location (i.e., urban and rural), section location (i.e., sections and intersections) and victim type (i.e., pedestrians, car drivers, car passengers, two-wheelers). For each partition, the initial number of clusters in the Kohonen networks is equal to three and the increase of one cluster at the time stops when the additional cluster either does not contain a significant number of records or the error of the network does not diminish. At convergence, the comparison between clusters obtained from different partitions allows individuating compelling patterns for fatal accidents in Israel during the study period.

In addition, this study applies feed-forward back-propagation neural networks with exhaustive pruning to the entire dataset with respect to the same variables used for the construction of the partitions (i.e., accident type, urban location, section location, victim type), in order to obtain a classification of the relevance of accident factors. The number of input neurons corresponds to the number of categories of the input variables, the number of output variables for supervised learning equals the categories of the accident characteristics chosen to partition the data, and the optimal number of neurons in the hidden layer is determined through the exhaustive pruning process. It should be noted that the output neurons are three for the accident type (i.e., pedestrian, single vehicle and multiple vehicle), two for the urban location (i.e., urban and rural), two for the section location (i.e., sections and intersections), and four for the victim type (i.e., pedestrians, car drivers, car passengers, two-wheelers). Over-fitting is avoided by considering randomly 50% of the dataset for training and the remaining 50% for test before the validation. Accordingly, the neural network minimizes the error by comparing the prediction following the training of the model with 50% of randomly drawn records with the actual outcomes of the remaining 50% of records.

3. Results

This section illustrates accident patterns and relevant factors resulting from the implementation of Kohonen and feed-forward back-propagation neural networks to the illustrated partitions of the fatal accident dataset. The interpretation of the accident patterns depends on the frequency of the categories of the variables that the Kohonen neural network presents to discern the clusters. The interpretation of the relevant factors depends on the variables that the feed-forward back-propagation neural network indicates as the most relevant.

3.1 Partition according to accident type

Three types of fatal accidents are considered for partitioning the dataset: pedestrian (603 records), single-vehicle (416 records) and multiple-vehicle (774 records). Table 2 summarizes the clusters obtained for these three types of fatal accidents.

[Insert Table 2]

Variables differentiating the five clusters obtained for pedestrian fatal accidents are the accident location, the road width, the crossing modality, the period of the day, and the age and population group of the pedestrians. Exploring the Kohonen map, major differences are characterized in terms of age, as the first cluster contains accidents occurred to elderly population while the last cluster includes accidents occurred to children and teenagers, population group, as the first clusters concern Jewish pedestrians and the last one non-Jewish ones, and geographical location, as from major metropolitan areas for the first two clusters the map moves to small urban areas for the last one.

Variables discerning the four clusters found for single-vehicle fatal accidents are the seat-belt law compliance, the location in either urban or rural areas, the age of the vehicle and the age and population group of the drivers involved. Examining the clustering map, major differences are defined according to the population group of the drivers, as the first cluster involves non-Jewish drivers and the last one Jewish ones, and the geographical location, as the first cluster contains accidents in small villages and the last cluster includes accidents in large urban areas.

Variable characterizing the four clusters uncovered for multiple-vehicle fatal accidents are the level of experience of the drivers, the accident location in either a section or an intersection, the period of the day, and the age and population group of the drivers. Observing the clustering map, major differences are defined according to population groups, as the first cluster includes accidents involving only non-Jewish drivers and the last one only Jewish ones, while passing through a cluster where both are contemporarily involved.

Table 3 summarizes the relative importance assigned by the feed-forward back-propagation neural network to the input variables when the dataset is considered with respect to the three possible accident types. The age of the victim is the most relevant variable, and in fact children, teenagers and elderly are mostly victims of pedestrian accidents, young drivers up to 24 years old are mainly victims of single-vehicle accidents, and experienced drivers are generally victims of multiple-vehicle accidents. The population group of victims and drivers is also very relevant, and in fact almost clear-cut separation between Jewish and non-Jewish population characterizes most clusters. The road width, the location and the period of the day also contribute to the discernment of the accident type. These findings suggest that improving road safety should probably focus on solving issues related to specific population groups.

[Insert Table 3]

3.2 Partition according to location type

Two different types of classification of accident locations are considered for partitioning the dataset: accidents in intersections (451 records) versus accidents in sections (1342 records), and urban accidents (749 records) versus rural accidents (1044 records).

Table 4 illustrates clusters obtained for accidents in either intersections or road sections.

[Insert Table 4]

Variables telling apart the four clusters obtained for fatal accidents that occurred in intersections are the location in either urban or rural areas, the period of the day, the involvement of pedestrians, and the age of both drivers and victims. Inspecting the Kohonen map, major differences are observed with respect to the type of accident, as multiple-vehicle accidents are at opposite sides of the map with respect to pedestrian accidents, and the age and role of the persons involved, as young drivers in the initial clusters are opposite to elderly pedestrians in the last ones.

Variables differentiating the six clusters found for fatal accidents that took place in road sections are the location in either urban or rural areas, the period of the day and the related existence of artificial illumination on the road, the type of accident, the type of vehicles involved, and the age and population groups of drivers and victims. Examining the clustering map, major differences are observed in terms of population groups of the drivers, as clusters contain accidents in which either only non-Jewish or only Jewish drivers were involved, period of the day, as the first clusters focus on night accidents without artificial illumination and the last clusters concentrate mainly on accidents during day, and type of accident, as initially the clusters refer to multiple-vehicle and then to single-vehicle and pedestrian accidents.

Table 5 shows the relative importance assigned by the feed-forward back-propagation neural network to the input variables when the dataset is considered with respect to the location of the accident being in either a section or an intersection. The accident type is the most relevant factor, and in fact single-vehicle accidents generally occur in sections and front-to-side or pedestrian-vehicle accidents generally take place in intersections. The age of the victim is also very relevant, and in fact young drivers tend to have mostly accidents in sections while experienced drivers are often involved in accidents in intersections. The crossing modality of pedestrians, the zone of the accident, the road width and the use of seatbelts also help discerning accidents occurring in either sections or intersections. These findings suggest that improving road safety should probably focus on young drivers and vulnerable road users, and possibly on infrastructural enhancements.

[Insert Table 5]

Table 6 presents clusters obtained for accidents in either urban or rural areas.

[Insert Table 6]

Variables discerning the five clusters attained for fatal accidents that happened in urban areas are the location in either metropolitan areas or small villages, the period of the day, the road width, the seat-belt law compliance, and the age and population group of both drivers and victims. Reading the map of clusters, major differences are observed with respect to the age of the victims, as young drivers in the first cluster are opposite to elderly pedestrians in the last one, and the geographical location, with different clusters varying between small villages and large metropolitan areas.

Variable characterizing the five clusters uncovered for fatal accidents that occurred in rural areas are the geographical location, the period of the day and the related existence of artificial illumination, the road width, the modality of the accident, and the age and population group of both drivers and victims. Examining the clusters, major differences are drawn by the population group of the drivers, as non-Jewish drivers and Jewish pedestrians are at the opposite sides of the map, and the type of accident, as multiple-vehicle, single-vehicle and pedestrian accidents differentiate the output neurons.

Table 7 represents the relative importance assigned by the feed-forward back-propagation neural network to the input variables when the dataset is considered with respect to the location of the accident being in either urban or rural areas. Unlike for other dataset partitions, the most relevant variables have similar relevance: accident type, age of the victim, road width, zone of the accident and median have all importance varying between 9.5 and 8.9%, which indicates that the distinction between urban and rural accidents is related to a combination of factors. For example, typical rural accidents are single-vehicle accidents where young drivers lost control of the car in narrow roads without median separation in the north of the country. These findings suggest that improving road safety should probably consider a systemic approach rather than focus on single issues.

[Insert Table 7]

3.3 Partition according to victim type

Four types of victims have been considered for partitioning the dataset: pedestrians (611 records), car drivers (528 records), car passengers (433 records) and two-wheelers (221 records). Table 8 summarizes the clusters obtained for the four types of victims.

[Insert Table 8]

As expected, the definition of the five clusters obtained for fatal accidents involving the death of a pedestrian proposes the same solution of the determination of the clusters for pedestrian accidents. This finding confirms that one-dimensional Kohonen networks produce optimal solutions as suggested by the literature in operational research (e.g., Burke 1996) and oncology research (Augen 2004). It should be noted that the number of pedestrian victims is slightly larger than the number of pedestrian accident, and this difference simply indicates that some accidents had multiple pedestrian fatalities.

Variables telling apart the four clusters found for fatal accidents resulting in the death of at least one of the car drivers involved are the accident location in either urban or rural areas, the type of accident, the period of the accident, and the age and population group of the victims. Observing the Kohonen map, major differences are determined mainly by the population group of the drivers, as Jewish and non-Jewish drivers are at opposite sides of the map.

Variables differentiating the four clusters attained for fatal accidents resulting in the death of at least one of the passengers of the cars involved are the accident location, the type of accident, the period of the accident, and the age of the drivers and the victims. Examining the clusters, major differences are observed with respect to the period of the day, as night or morning or afternoon periods characterize the different clusters, and the geographical location of the accidents, as accidents occurred in the south or the center are at opposite sides of the map with respect to accidents occurred in the north.

Variables characterizing the three clusters uncovered for fatal accidents involving the death of at least one road user on either a motorcycle or a bicycle are the type of two-wheel vehicle, the accident location, the modality of the accident and the age of the riders. Inspecting the clustering map, major differences are related to the type of two-wheel vehicle, as motorcycle accidents are distinguished from bicycle ones.

Table 9 presents the relative importance assigned by the feed-forward back-propagation neural network to the input variables when the dataset is considered with respect to the four possible victim types. Four variables appear extremely relevant, and are all related to the age and population group of drivers and victims. In fact, pedestrian victims are teenagers or elderly, car drivers and car passengers are young or middle-aged, two-wheelers are usually in their 20’s or 30’s. Moreover, Jewish and non-Jewish drivers and victims are typically distinguished when describing which fatalities result from the accidents, and this underlines the extremely high importance of the population group. These findings reiterate that improving road safety should probably address issues related to specific population groups.

[Insert Table 9]

3.4 Compelling accident patterns

The large number of clusters obtained from analyzing the different partitions is examined in order to comprehend whether some clustering solutions are more compelling than others, namely whether clusters present similarities once they are obtained from different starting points in the analysis. As some similar patterns are individuated among the 49 clusters, five compelling accident patterns are recognizable regardless of the dataset partition considered.

The first cluster contains single-vehicle accidents of young drivers who tend not to wear seatbelts and to lose control of their vehicles while driving at night in rural sections without artificial illumination. The second cluster includes multiple-vehicle accidents involving young drivers behind the wheel of relatively new vehicles, who have front-to-front and front-to-side accidents in both urban and rural areas at night where artificial illumination is missing. The third cluster consists of accidents involving either motorcycles or bicycles in urban sections. The fourth cluster comprises accidents where elderly pedestrians crossed on crosswalks far from intersections in large urban areas, mainly in the Tel Aviv metropolitan area. The fifth cluster includes accidents where mostly young children and teenagers crossed suddenly large roads in small urban areas. Interestingly, the accident patterns underline that the preliminary study about pedestrian accidents (Prato et al. 2011) recognize only a part of the most compelling road safety problems in the country, and hence this study contributes in broadening the picture for necessary road safety policies.

4. Discussion and conclusions

This study focuses on the recognition of accident patterns and the individuation of relevant accident factors to answer an increasing need of designing countermeasures, addressing compelling problems and targeting specific population groups with the ultimate objective of reducing the annual number of traffic fatalities and accidents. The successful implementation of neural networks to a dataset of fatal accidents in Israel allows observing patterns that do not intend to predict that the combination of the factors obtained in the analysis will result in a fatal accident, but rather to classify fatal accidents and suggest which areas are more relevant to intervene on.

From the implementation of Kohonen neural networks, five compelling accident patterns are recognized among the large number of clusters resulting from the analysis of different data partitions. These five patterns suggest that preventive measures are necessary for identified problems concerning mainly vulnerable road users such as pedestrians, cyclists, motorcyclists, and young drivers. Interestingly, similarities are observed with France with respect to the problems regarding pedestrian accidents (see Fontaine and Gourlet 1997), and with Sweden with respect to issues pertaining drivers in training (Berg et al. 2004).

The problem of elderly pedestrians killed while crossing roads on crosswalks in metropolitan areas suggests the need for a critical review of possible infrastructure solutions, and for a conscious design of traffic calming measures able to meet necessities and restrictions of the elderly population (see Gitelman et al. 2010). The problem of children and teenagers killed in small urban areas crossed by major roads, especially in the north of the country, calls for educating children and teenagers in road hazard perception, raising parents' awareness about the need of monitoring children’s behavior, lowering allowed speeds in major roads when approaching urban areas, and creating physical separation between vehicles and pedestrians. The problem of two-wheelers demands intervention toward the physical separation between cyclists and motorized road users with dedicated lanes, the education for correcting dangerous behavior of powered two-wheelers, and the information to the public with the objective of raising awareness of this problem. The problem of young drivers, involved mainly in single-vehicle accidents at night in rural areas and in multiple-vehicle accidents at night in both urban and rural areas, enforces the need for the correction of the existing graduated driver licensing program toward an extension of the accompanied driving period and the introduction of a minimum requirement of supervised driving hours with the aim to reduce crash risks and tendency to risky behavior of young drivers (Prato et al. 2010).

From the implementation of Kohonen and feed-forward back-propagation neural networks, relevant factors of accidents are recognized. Unlike previous studies who focused on accident modality (e.g., Preusser et al. 1995; Retting et al. 1995; Fleury and Brenac, 2001; Laapotti and Keskinen 2004; Chang and Chen 2005; Geurts et al. 2005; Skyving et al. 2009), urban or rural location (e.g., Berg et al. 2004; Wang et al. 2008; Skyving et al. 2009), speeding (e.g., Berg et al. 2004; Laapotti and Keskinen 2004; Wang et al. 2008; Skyving et al. 2009), and weather conditions (e.g., Chang and Chen 2005; Geurts et al. 2005; Skyving et al. 2009), this study shows the relevance of socio-demographic characteristics of drivers and victims. For example, age is relevant with respect to accident type (children, teenagers and elderly are mostly victims of pedestrian accidents, young drivers are mainly victims of single-vehicle accidents, and experienced drivers are generally victims of multiple-vehicle accidents) and location (young drivers are mostly involved in accidents in road sections while experienced drivers are mostly involved in intersections). Moreover, this study emphasizes the relevance of the population group of both drivers and victims in the distinction of the clusters, as Jewish and non-Jewish population are generally positioned at the opposite sides of the Kohonen maps obtained at convergence of the unsupervised learning process. Interestingly, some accident patterns involve drivers belonging to two different groups, such as a Jewish and a non-Jewish driver or an experienced and a young driver. This finding confirms a theory about inter-groups differences as traffic accident determinants in Israel (Factor et al. 2008).

After recognizing the most compelling problems and the most relevant factors with the neural networks, a system approach may be necessary for the selection of preventive countermeasures. Advanced countries such as the Netherlands, Norway and Sweden adopt a system approach that brings together the principles of functionality, homogeneity, predictability, forgivingness and awareness to promote infrastructural and vehicle improvements, educational campaigns, law enforcement and dangerous behavioral habit corrections (see, e.g., Wegman and Aarts 2006). Pattern analysis might be beneficial to this system approach in indicating compelling issues that should receive prioritized attention and resource allocation. Moreover, a systematic monitoring process and an analysis of performance indicators might support this system approach within an integrated "safe-system" approach (OECD 2008).

Summarizing, cluster analysis contributes to the general knowledge about accidents by presenting insight into patterns and factors that call for effective intervention measures through a system approach. Further research could benefit from an extension of the available data to information regarding alcohol or drug intoxication, socio-economic status and education that would allow further enhancement of the profiling of both drivers and victims in light of the evident relevance of social and demographic aspects in the recognition of accident patterns.

Acknowledgements

The authors would like to acknowledge the helpful comments of Ayala Cohen and Tsippy Lotan and the financial support of the no-profit organization Or Yarok to the Ran Naor Road Safety Research Center where the study has been conducted. The authors would also like to acknowledge the helpful remarks of four anonymous reviewers who greatly contributed to the improvement of the manuscript.

References

Aras, N., Oommen, B.J., & Altmel, I.K. (1999). The Kohonen network incorporating explicit statistics and its application to the Traveling Salesman Problem. *Neural Networks*, *12*, 1273-1284.

Augen, J. (2004). *Bioinformatics in the Post-Genomic Era: Genome, Transcriptome, Proteome, and Information-Based Medicine*. Addison-Wesley, Calgary, Canada.

Avitzour, M. (2007). *Injured Pedestrians: Problem's Scope, Circumstances, Injury Profile and Outcomes* (in Hebrew). Israel Center for Trauma and Emergency Medicine Research, Gertner Institute, Tel Aviv, Israel.

Ballesteros, M.F., Dischinger, P.C., & Langenberg, P. (2004). Pedestrian injuries and vehicle type in Maryland 1995-1999. *Accident Analysis and Prevention, 36*(1), 73-81.

Berg, H.Y., Gregersen, N.P., & Laflamme, L. (2004). Typical patterns in road-traffic accidents during driver training: an explorative Swedish national study. *Accident Analysis and Prevention*, *36* (4), 603-608.

Burke, L.I. (1996). Neural methods for the Travelling Salesman Problem: insights from operations research. *Neural Networks*, *7*, 681-690.

Chang, L.Y., & Chen, W.C. (2005). Data mining of tree-based models to analyze freeway accident frequency. *Journal of Safety Research*, *36*(4), 365-375.

Cottrell, M., & Rousset, P. (1997). The Kohonen algorithm: a powerful tool for analyzing and representing quantitative and qualitative data. *Lecture Notes in Computer Science, 1240*, 861-871.

Eluru, N., Bhat, C.R., & Hensher, D.A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention*, *40*(3), 1033-1054.

Erwin, E., Obermayer, K., & Schulten, K. (1992). Self-organizing maps: ordering, convergence properties and energy functions. *Biological Cybernetics*, *67*, 47-55.

Factor, R., Mahalel, D., & Yair, G. (2008). Inter-group differences in road-traffic crash involvement. *Accident Analysis and Prevention*, *40*(6), 2000-2007.

Fleury, D., & Brenac, T. (2001). Accident prototypical scenarios, a tool for road safety research and diagnostic studies. *Accident Analysis and Prevention, 33*(2), 267-276.

Fontaine, H., & Gourlet, Y. (1997). Fatal pedestrian accidents in France: a typological analysis. *Accident Analysis and Prevention, 29*(3), 303-312.

Geurts, K., Wets, G., Brijs, T., & Vanhoof, K. (2003). Profiling of high-frequency accident locations by use of association rules. *Transportation Research Record*, *1840*, 123-130.

Geurts, K., Thomas, I., & Wets, G. (2005). Understanding spatial concentrations of road accidents using frequent item sets. *Accident Analysis and Prevention*, *37*(4), 787-799.

Gitelman, V., Balasha, D., Carmel, R., Hendel, L., & Pesahov, F. (2010). Characterization of pedestrian accidents and an examination of infrastructure measures to improve pedestrian safety in Israel. *Accident Analysis and Prevention*, doi:10.1016/j.aap.2010.11.017.

Harruff, R., Avery, A., & Alter-Pandya, A. (1998). Analysis of circumstances and injuries in 217 pedestrian traffic fatalities. *Accident Analysis and Prevention*, *30*(1), 11-20.

Hasselberg, M., Vaez, M., & Laflamme, L. (2005). Socioeconomic aspects of the circumstances and consequences of car crashes among young adults. *Social Science and Medicine*, *60*(2), 287-295.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59-69.

Kohonen, T. (2001). *Self-Organizing Maps (3rd extended edition)*. Springer & Verlag, Berlin Heldelberg, Germany.

Laapotti, S., & Keskinen, E. (2004). Has the difference in accident patterns between male and female drivers changed between 1984 and 2000? *Accident Analysis and Prevention*, *36*(4), 577-584.

Milton, J.C., Shankar, V.N., & Mannering, F.L. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis and Prevention*, *40*(1), 260-266.

OECD (2008). *Towards Zero: Ambitious Road Safety Targets and the Safe System Approach*. OECD Publishing, International Transport Forum of the Organization for Economic Co-operation and Development, Paris, France.

Peleg, K., & Aharonson-Daniel, L. (2004). Road traffic accidents - severe injuries. How missing data can impair decision making? (in Hebrew). *Harefuah*, *143*(2), 111-115.

Prato, C.G., Gitelman, V., & Bekhor, S. (2011). Mapping patterns of pedestrian fatal accidents in Israel. *Accident Analysis and Prevention, in press, doi:10.1016/j.aap.2010.12.022*.

Prato, C.G., Toledo, T., Lotan, T., & Taubman Ben-Ari, O. (2010). Modeling the behavior of novice young drivers during the first year after licensure. *Accident Analysis and Prevention*, *42*(2), 480-486.

Preusser, D.F., Williams, A.F., & Ulmer, R.G. (1995). Analysis of fatal motorcycle crashes: crash typing. *Accident Analysis and Prevention*, *27*(6), 845-851.

Reed, R.D., & Marks, R.J. (1999). *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*. M.I.T. Press, Cambridge, MA.

Retting, R.A., Williams, A.F., Preusser, D.F., & Weinstein, H.B. (1995). Classifying urban crashes for countermeasure development. *Accident Analysis and Prevention*, *27*(3), 283-294.

Retting, R.A., Williams, J., & Schwartz, S.I. (2000). Motor vehicle crashes on bridges and countermeasure opportunities. *Journal of Safety Research*, *31*(4), 203-210.

Ritter, H., & Schulten, K. (1988). Convergence properties of Kohonen's topology conserving maps: fluctuations, stability and dimension selection. *Biological Cybernetics*, *60*, 59-71.

Sivak, M., & Schoettle, B. (2010). Toward understanding the recent large reductions in U.S. road fatalities. *Traffic Injury Prevention*, *11*(6), 561-566.

Skyving, M., Berg, H.Y., & Laflamme, L. (2009). A pattern analysis of traffic crashes fatal to older drivers. *Accident Analysis and Prevention*, *41*(2), 253-258.

SPSS Inc. (2007). Clementine® 12.0 User's Guide. Integral Solutions Limited, Chicago, Illinois.

Tseng, W.S., Nguyen, H., Liebowitz, J., & Agresti, W. (2005). Distractions and motor vehicle accidents: data mining application on fatality analysis reporting system (FARS) data files. *Industrial Management and Data Systems*, *105*(9), 1188-1205.

Yan, X., & Radwan, E. (2006). Analyses of rear-end crashes based on classification tree models. *Traffic Injury Prevention*, *7*(3), 276-282.

Wang, Y., Hasselberg, M., Wu, Z., & Laflamme, L. (2008). Distribution and characteristics of road traffic crashes in the Chaoyang District of Beijing. *Accident Analysis and Prevention*, *40*(1), 334-340.

Wegman, F., & Aarts, L. (2006). *Advancing Sustainable Safety: National Road Safety Outlook for 2005-2020*. SWOW Institute for Safety Research, Leidschendam, The Netherlands.

Zegeer, C.V., Stutts, J.C., Huang, H., Zhou, M., & Rodgman, E. (1993). Analysis of elderly pedestrian accidents and recommended countermeasures. *Transportation Research Record*, *1405*, 56-63.

Table 1 Categorical variables for data analysis

Variables	Categories
year of the accident	2003 – 2004 – 2005 – 2006
season of the accident	spring – summer – autumn – winter
region of the accident	3 metropolitan areas (Jerusalem – Tel Aviv – Haifa) and 3 regions (north – center – south)
day of the week	Sunday – Monday – Tuesday –Wednesday – Thursday – Friday – Saturday
day / night	day – night
period of the day	morning peak – off-peak – afternoon peak – evening/night
type of day	normal – pre-festive – festive – inner-festive
type of accident	single-vehicle – front-to-front – front-to-side – front-to-rear – side-to-side – pedestrian
cause of the accident	driver behavior – passenger behavior – pedestrian behavior – motorcyclist behavior – cyclist behavior – vehicle malfunctioning – other
location of the accident	urban intersection – urban sections – rural intersection – rural sections
allowed speed	50 km/h – 60 km/h – 70 km/h – 80 km/h – 90 km/h – 100 km/h
weather conditions	clear – rainy – hot – foggy – not specified
road surface conditions	dry – wet from water – wet from slippery material – covered with mud – covered with sand – not specified
number of ways	one-way road – two-way road with separation – two-way without separation – not specified
median	with separation fence – without separation fence – not specified
shoulders of the road	good conditions – bad conditions – rough road
width of the road	up to 5 m. – 5 to 7 m. – 7 to 10 m. – 10 to 14 m. – over 14 m.
illumination on the road	daylight – night without illumination – night with illumination
traffic control	no control – working traffic light – failing traffic light – blinking yellow – stop sign – priority sign – not specified
crossing location	on crosswalk with traffic light – on crosswalk without traffic light – outside crosswalk next to intersection – outside crosswalk far from an intersection
crossing modality	suddenly – from hidden places – normally – not specified
type of victim	car driver – car passenger – motorcycle driver – motorcycle passenger – cyclist – pedestrian
age of the victims	less than 14 – 15 to 19 – 20 to 24 – 25 to 34 – 35 to 44 – 45 to 54 – 55 to 64 – more than 65
gender of the victims	male – female
population groups of the victims	Jewish – non Jewish –not specified or others
type of vehicles involved	car – light truck – heavy truck – public transport –motorcycle –bicycle – not specified
size of the motor of the vehicles	less than 1000 cc. – 1000 to 1300 cc. –1300 to 1600 cc. – 1600 to 2000 cc. – over 2000 cc.

Variables	Categories
age of the vehicles	up to 2 years – 2 to 5 – 6 to 10 – 11 to 15 – more than 15 years
safety on board	seatbelts – helmets – child-seat – not used
property of the vehicle	owned – rented – stolen
owner of the vehicles	private – Army – Police – other
age of the drivers	17 to 19 – 20 to 24 - 25 to 34 –35 to 44 –45 to 54 –55 to 64 – more than 65
licensing years for the drivers	up to 2 years – 2 to 5 – 6 to 10 – 11 to 15 – more than 15 years
gender of the drivers	male – female
population groups of the drivers	Jewish – non Jewish –not specified or others
type of city	less than 10,000 inhabitants – 10,000 to 30,000 – 30,000 to 100,000 – 100,000 to 200,000 – more than 200,000 inhabitants

Table 2 Accident patterns: partition according to accident type

Pattern	Characteristics
Pedestrian accidents	
Elderly pedestrians in urban areas (185 records)	Jewish pedestrians pedestrians more than 65 years old pedestrians crossing on crosswalks pedestrian crossing far from intersections accidents in the Tel Aviv metropolitan area
Pedestrians and two-wheel vehicles (90 records)	pedestrians crossing outside crosswalks pedestrians crossing suddenly accidents in large urban areas two-wheel vehicles involved
Pedestrians in rural areas at night (133 records)	pedestrians crossing outside crosswalks pedestrians crossing suddenly at night accidents in wide road sections accidents in rural areas
Young pedestrians at night (77 records)	pedestrians less than 35 years old pedestrians crossing outside crosswalks pedestrians crossing suddenly at night accidents in wide road sections
Children and teenage pedestrians in small villages (118 records)	non-Jewish pedestrians pedestrians less than 14 years old pedestrians crossing suddenly from hidden places accidents in wide road sections accidents in small urban areas
Single-vehicle accidents	
Non-Jewish population in single-vehicle accidents (116 records)	non-Jewish drivers low seat-belt compliance accidents in small urban areas
Experienced drivers in single-vehicle accidents (66 records)	experienced drivers accidents in road sections accidents in rural areas
Night in new single-vehicle accidents (55 records)	new vehicles low seat-belt compliance accidents at night
Young drivers in single-vehicle accidents (179 records)	young drivers less than 24 years old low seat-belt compliance accidents in road sections accidents in large urban areas accidents at night without illumination
Multiple-vehicle accidents	
Non-Jewish population in multiple-vehicle accidents (212 records)	non-Jewish drivers young drivers less than 24 years old accidents during afternoon off-peak accidents in road sections accidents in rural areas
Experienced drivers in multiple-vehicle accidents (133 records)	one Jewish and one non-Jewish driver experienced drivers accidents in narrow road sections accidents in rural areas
Young and experienced drivers in multiple-vehicle accidents (129 records)	one experienced and one young driver accidents in intersections accidents in rural areas
Young drivers in multiple-vehicle accidents (300 records)	Jewish drivers young drivers less than 24 years old accidents during morning or afternoon off-peak

	accidents in road sections accidents in rural areas

Table 3 Variable relevance for partition according accident type

Pedestrian vs. single-vehicle vs. multiple vehicle accidents	
Variable	Relevance (%)
Age of the victim	16.7
Population group of the victim	9.1
Width of the road	9.0
Location (large vs. small cities)	8.7
Location (urban vs. rural)	7.3
Period of the day	7.0
Population group of the driver	5.5

Table 4 Accident patterns: partition according to intersection vs. section location

Accidents in intersections	
Night in rural intersections *(153 records)*	front-to-side accidents accidents in intersections accidents in rural areas accidents at night
Young drivers in intersections *(94 records)*	young drivers less than 35 years old front-to-side accidents accidents in large cities accidents in the center of the country
Pedestrians in urban intersections *(Tel Aviv area)* *(58 records)*	pedestrians more than 50 years old young drivers new vehicles accidents in the Tel Aviv metropolitan area
Pedestrians in urban intersections *(outside Tel Aviv area)* *(146 records)*	pedestrians more than 50 years old accidents in narrow roads accidents in intersections without traffic lights
Accidents in sections	
Non-Jewish population in rural sections at night *(229 records)*	non-Jewish drivers young drivers less than 35 years old multiple-vehicle accidents accidents at night without illumination accidents in road sections accidents in rural areas
Jewish population in rural sections at night *(246 records)*	Jewish drivers young drivers less than 35 years old multiple-vehicle accidents accidents at night without illumination accidents in road sections accidents in rural areas
Young drivers in single-vehicle accidents *(337 records)*	young drivers less than 24 years old accidents in road sections accidents in urban areas accidents at night without illumination
Two-wheel vehicles in sections *(107 records)*	motorcycles or bicycles involved accidents in road sections accidents in the Tel Aviv metropolitan area
Elderly pedestrians in urban areas *(261 records)*	Jewish pedestrians pedestrian more than 65 years old pedestrians crossing on crosswalks pedestrian crossing far from intersections accidents in the Tel Aviv metropolitan area
Children and teenage pedestrians in small villages *(162 records)*	non-Jewish pedestrians pedestrians less than 14 years old pedestrians crossing suddenly from hidden places accidents in wide road sections accidents in small urban areas

Table 5 Variable relevance for partition according to intersection vs. section location

Accidents in sections vs. accidents in intersections	
Variable	Relevance (%)
Accident type	13.1
Age of the victim	9.9
Crossing modality of the pedestrian	7.1
Zone of the accident	6.8
Age of the vehicle	6.4
Width of the road	6.3
Use of seatbelts	5.7
Period of the day	5.6

Table 6 Accident patterns: partition according to urban vs. rural location

Accidents in urban areas	
Young drivers in urban areas (246 records)	young drivers less than 24 years old low seat-belt compliance accidents in road sections accidents in small urban areas accidents at night without illumination
New vehicles in urban areas (69 records)	new vehicles low seat-belt compliance accidents in small urban areas accidents at day
Children and teenage pedestrians in small villages (126 records)	pedestrians less than 14 years old pedestrians crossing suddenly from hidden places accidents in wide road sections accidents in small urban areas
Pedestrians in urban sections (110 records)	pedestrians more than 50 years old accidents in road sections accidents in large urban areas
Elderly pedestrians in urban areas (198 records)	Jewish pedestrians pedestrians more than 65 years old pedestrians crossing on crosswalks pedestrian crossing far from intersections accidents in the Tel Aviv metropolitan area
Accidents in rural areas	
Non-Jewish population in multiple-vehicle accidents (229 records)	non-Jewish drivers young drivers less than 35 years old multiple-vehicle accidents accidents in road sections accidents in small urban areas in the north accidents at day
Multiple-vehicle accidents in narrow rural roads (177 records)	new vehicles multiple-vehicle accidents accidents in narrow road sections accidents in rural areas
Single- and multiple-vehicle accidents in rural areas (286 records)	new vehicles single-vehicle or front-to-side accidents accidents in road sections accidents in rural areas
Young drivers in rural sections (101 records)	new vehicles young drivers less than 35 years old accidents in road sections accidents in rural areas
Pedestrian in rural sections (181 records)	Jewish pedestrians pedestrians crossing suddenly from hidden places accidents at night without illumination accidents in road sections accidents in rural areas

Table 7 Variable relevance for partition according to urban vs. rural location

Accidents in urban areas vs. accidents in rural areas	
Variable	Relevance (%)
Accident type	9.5
Age of the victim	9.2
Width of the road	9.1
Zone of the accident	9.0
Presence of median barrier	8.9
Crossing modality of the pedestrian	7.0
Licensure years of the driver	5.7
Age of the driver	5.5

Table 8 Accident patterns: partition according to victim type

Pedestrian victims	
Children and teenage pedestrians in small villages (142 records)	non-Jewish pedestrians pedestrians less than 14 years old pedestrians crossing suddenly from hidden places accidents in wide road sections accidents in small villages
Young pedestrians at night (76 records)	pedestrians less than 35 years old pedestrians crossing outside crosswalks pedestrians crossing suddenly at night accidents in wide sections
Pedestrians in rural areas at night (122 records)	pedestrians crossing outside crosswalks pedestrians crossing suddenly at night accidents in rural sections
Pedestrian and two-wheel vehicles (77 records)	pedestrians crossing outside crosswalks pedestrians crossing suddenly accidents in urban areas two-wheel vehicles involved
Elderly pedestrians in urban areas (194 records)	Jewish pedestrians pedestrians more than 65 years old pedestrians crossing on crosswalks pedestrian crossing far from intersections accidents in the Tel Aviv metropolitan area
Car driver victims	
New vehicles in single- and multiple-vehicle accidents (176 records)	Jewish drivers young drivers less than 24 years old new vehicles accidents in road sections accidents in urban areas accidents at night without illumination
Young drivers in multiple-vehicle accidents in rural areas (130 drivers)	young drivers less than 35 years old accidents at night accidents in road sections accidents in rural areas
Young and experienced drivers in rural areas at night (50 records)	one experienced and one young driver accidents in road sections accidents in rural areas accidents at night
Non-Jewish drivers in single- and multiple-vehicle accidents (172 records)	non-Jewish drivers young drivers less than 35 years old low seat-belt compliance single- and multiple-vehicle accidents
Car passenger victims	
New vehicles at night (142 records)	young drivers less than 24 years old young victims less than 35 years old new vehicles accidents at night accidents in the south
Young drivers in rural areas (87 records)	young drivers less than 35 years old young victims less than 35 years old accidents in road sections accidents in rural areas of the center or the south accidents during morning and afternoon off-peak

Experienced drivers in rural areas at night (78 records)	experienced drivers young victims less than 24 years old accidents in road sections accidents in rural areas accidents at night accidents during weekends
Non-Jewish population in rural sections (126 records)	non-Jewish victims young victims less than 35 years old low seat-belt compliance accidents in road sections accidents in small urban areas in the north

Two-wheeler victims

Young riders on motorcycles in urban areas (85 records)	young riders between 25 and 35 years old high helmet compliance accidents in Tel Aviv and in the center
Motorcycles on rural sections (66 records)	young riders (less than 35 years old) high helmet compliance accidents in road sections accidents in rural areas
Cyclists (70 records)	front-to-side accidents with vehicles front-to-rear accidents with vehicles accidents in road sections

Table 9 Variable relevance for partition according to victim type

Pedestrian victims vs. car driver victims vs. car passenger victims vs. two-wheeler victims	
Variable	Relevance (%)
Age of the victim	15.3
Age of the driver	14.4
Population group of the victim	13.2
Population group of the driver	10.1
Width of the road	5.7
Zone of the accident	5.3
Location (large vs. small cities)	5.1