

Artificial intelligence techniques for driving safety and vehicle crash prediction

Zahid Halim¹ · Rizwana Kalsoom¹ · Shariq Bashir² · Ghulam Abbas¹

Published online: 23 February 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Accident prediction is one of the most critical aspects of road safety, whereby an accident can be predicted before it actually occurs and precautionary measures taken to avoid it. For this purpose, accident prediction models are popular in road safety analysis. Artificial intelligence (AI) is used in many real world applications, especially where outcomes and data are not same all the time and are influenced by occurrence of random changes. This paper presents a study on the existing approaches for the detection of unsafe driving patterns of a vehicle used to predict accidents. The literature covered in this paper is from the past 10 years, from 2004 to 2014. AI techniques are surveyed for the detection of unsafe driving style and crash prediction. A number of statistical methods which are used to predict the accidents by using different vehicle and driving features are also covered in this paper. The approaches studied in this paper are compared in terms of datasets and prediction performance. We also provide a list of datasets and simulators available for the scientific community to conduct research in the subject domain. The paper also identifies some of the critical open questions that need to be addressed for road safety using AI techniques.

Keywords Road safety · Accident prediction · Artificial intelligence techniques · Traffic datasets and simulators

1 Introduction

It has been reported by the world health organization (WHO) that, the average annual human fatality rate globally is 18 per 100,000 persons due to accidents (Juárez and Gayet 2014). This figure varies across countries and excludes major and minor injuries. Not only does a road

✉ Zahid Halim
zahid.halim@giki.edu.pk

¹ Faculty of Computer Sciences and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Pakistan

² Department of Information Studies, College of Computer and Information Sciences, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

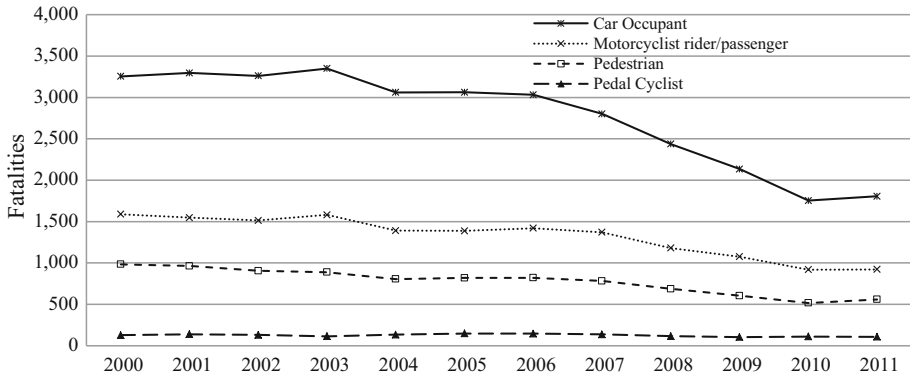


Fig. 1 Road accident fatalities in Great Britain (2000–2011)

accident cause loss of precious lives, but it also damages vehicles and property. Additionally, depending upon the severity of the accident the passengers can be severely injured that may take them years to recover or may result in a permanent disability. An accident is normally considered to involve only car, however, this can involve collision of any two or more vehicles. At times collision of a vehicle with a building or pedestrian can also result in an accident (Akin and Akbas 2010). Many countries maintain the record of traffic accidents through their transportation department, traffic police or hospitals (Imkamon et al. 2008; Suriyawongpaisal and Kanchanasut 2003). These records show that the number of casualties and injuries due to road accidents are huge. Figure 1 shows a graphical representation of fatalities for the period from 2000 to 2011 caused due to road accidents in Great Britain alone (Andersson and Chapman 2011). It is worth mentioning that the number of fatalities in car occupants is way too high as compared to the rest. Based on Fig. 1, for 1 pedestrian fatality there are 2 car occupant fatalities, for 1 motorcyclist there are 2 car occupant fatalities and for 1 pedal cyclist there are 11 car occupant fatalities. It can be seen that from the year 2000 to 2011 there has been a gradual decrease in the number of fatalities. However, these numbers are still high. The presented results in Fig. 1 are that of a developed country, unfortunately, the situation in developing and underdeveloped countries is even worse (Hu et al. 2011).

Based on the aforementioned statistics on the loss of precious lives and property, road accidents have been a major and growing concern worldwide (Hu et al. 2011). There can be multiple approaches to prevent road accidents. Some of these may include traffic rules awareness, better transportation system, affective traffic policing and precautionary systems in vehicles. There are a number of reasons for road accidents, for example, using cell phone while driving, playing music, high acceleration, a sudden brake, responding to a request or congested roads (Dixon et al. 2005; Imkamon et al. 2008). Accident prediction can have a significant influence on traffic flow and safety. There are many existing approaches to detect unsafe driving patterns for accident prediction. Some of these are based on biometric detection and some use facial movement for driving fatigue assessment (Ji et al. 2004; Orazio et al. 2004). However, there is no fixed list of actions or data, based upon which accidents could be predicted accurately. This makes it difficult to consider all the reasons at a time for accident prediction. Due to this reason, a number of techniques have been proposed considering different methods and having diverse datasets for the detection of unsafe driving patterns (Dixon et al. 2005; Imkamon et al. 2008).

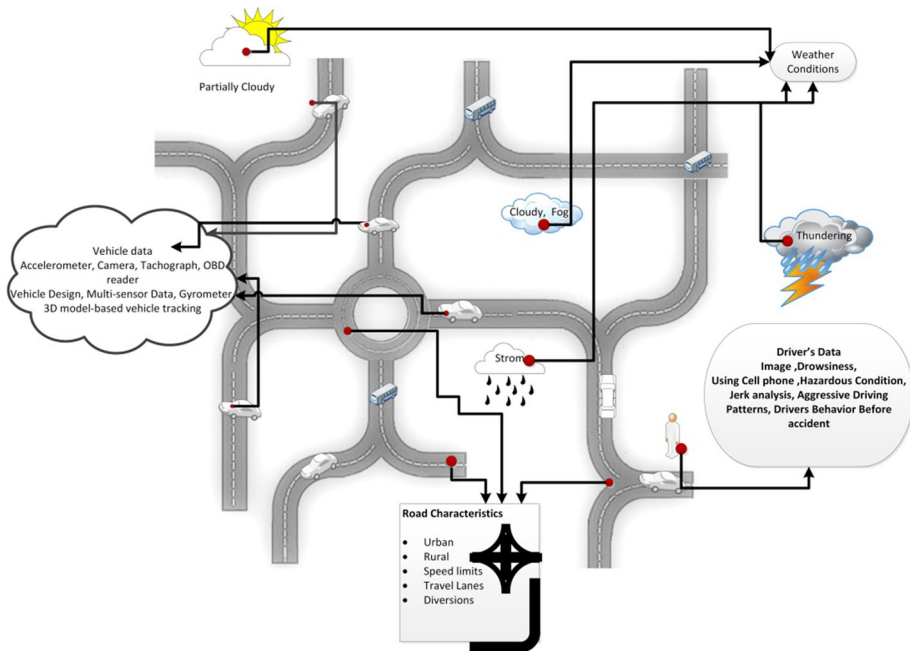


Fig. 2 Data produced by a vehicle and/or traffic system

The choice of a particular technique to be used for detection of unsafe driving patterns is mainly dependent upon the type, size, and format of the data being recorded for this purpose. Figure 2 shows different types, sizes, and formats of data that can be gathered from both the vehicle and the person driving it. As shown in Fig. 2, the data can be collected from the vehicles, road conditions/infrastructure, data related to the drivers and weather conditions.

Previous studies show that AI-based techniques use neural network more often for the detection of unsafe patterns and crash prediction (Akin and Akbas 2010; Yuejing et al. 2010; Moghaddam et al. 2010; Wahab et al. 2009). Other studies are based on the statistical techniques like conditional logit model (Abdel-Aty et al. 2004, 2005; Zheng et al. 2010; Abdel-Aty and Rajashekar 2006; Xu et al. 2012a; Abdel-Aty et al. 2012). Genetic programming, an evolutionary technique, has also been used for the crash prediction systems. Work like Dixon et al. (2005) uses both learning algorithm and genetic algorithm for the experiments. In Xu et al. (2012b), genetic programming model is developed for the real time crash prediction on freeways. There also exists wide literature related to crash prediction based on statistical approaches (Zhou et al. 2007; Singh and Dongre 2012; Ning et al. 2009, 2008; Jeong et al. 2004; Wang et al. 2012; Bruns et al. 2005). Similarly, in Rygula (2009), analysis of speed profile is performed to identify the driving style.

The study based on vehicle's data and/or driving features can be used in diverse ways and for many useful objectives. The key problems that can be addressed in this domain are listed as follows:

- **Driver identification:** The data from the vehicle and also of the driver can be recorded and later matched with the pre-stored datasets. This can be used in minimizing car thefts and for parental control (Miyajima et al. 2007).

- Profiling (grouping) drivers based on their driving features: Profiling the drivers can be done by recording driver's driving features only. Where, a signature can be created for each driver using his/her recorded data. The said signature can then be used to identify the driver and, based on the signature, assign him/her to a particular profile. The profiling can be more conveniently done via clustering (Wahab et al. 2009; Kalsoom and Halim 2013).
- Accident prediction: Accident prediction is a critical aspect in road safety, which can be addressed from four perspectives. An accident can be predicted using a) vehicle data, b) driver's data c) road or traffic condition/history, and e) weather conditions (Yuejing et al. 2010). Any combination of these four perspectives will further increase the accuracy of an accident prediction system.
- Early warning generation: An early warning generation system for vehicles seems to be somewhat similar to accident prediction, however, they are different. An accident prediction system can predict an anticipated accident with a margin of error, whereas; an early warning system may detect unsafe driving states/patterns and intimate the driver to be careful. An early warning system is mainly dependent on the individual driver's style of driving a vehicle (Jabon et al. 2011).
- Modeling individual driver's features: Modeling an AI-based controller to replicate a driver can be done by training the controller using the driver's patterns. This can have many applications like auto drive mode and autonomous vehicles (Ali et al. 2013).
- Accident identification: Accident identification systems are activated once an accident has actually occurred. These systems usually identify an accident based on the jerk caused due to collision and may use additional information recorded from the vehicle sensors. Such post-accident systems are used to alert emergency services (Lv et al. 2009).
- Prediction of driver's suitability for driving: AI techniques can be used to evaluate a driver and propose his/her suitability for the driving of a particular vehicle. This can be done by recording the driver's data generated during a test drive and matching it with one of the predefined clusters. Such techniques can be very useful for traffic police departments before issuing a driving license (Singh and Dongre 2012).
- Postmortem analysis: Study of the reasons for an accident comes under the postmortem analysis. The study can benefit from AI and machine learning (ML) techniques to identify correlation between different events that could have triggered the accident (Akin and Akbas 2010).

There can be many other problem domains, such as study of traffic flow, effects of weather on road condition or appropriate path finding. However, these domains do not strictly take into account the individual driver's driving features, and thus are out of the scope of this study.

Although, there exists a rich literature on different AI techniques used to predict accidents and unsafe driving behaviours but, there lacks a comprehensive survey on the existing methods in the literature. This paper aims at bridging this gap by encapsulating all the literature related to AI-based techniques for accident prediction, unsafe driving pattern identification, methods for profiling drivers, listing the key datasets/sources and identifying the future trends for road safety based on AI.

The rest of the paper is organized as follows. Section 2 lists the previous work on accident prediction/unsafe driving pattern analysis and comparative analysis of various AI techniques. Section 3 explains the available datasets and simulators for road safety studies. Section 4 covers the discussion, Sect. 5 lists the open questions and finally Sect. 6 concludes the paper.

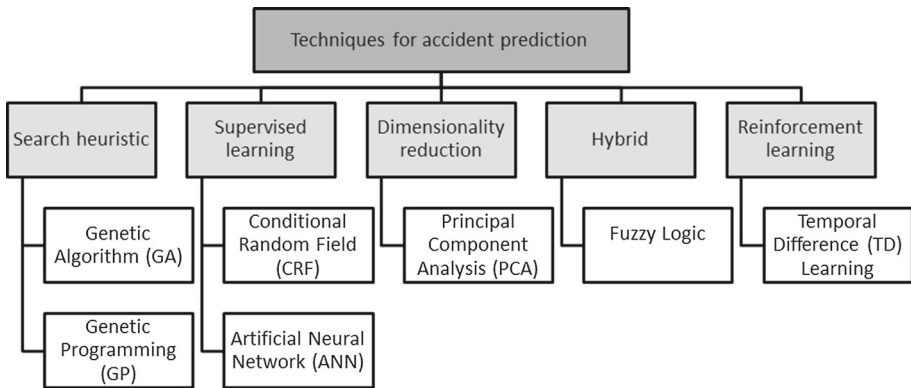


Fig. 3 AI techniques used in literature for accident prediction

2 AI techniques for accident prediction and unsafe driving pattern analysis

This section covers the key techniques from the domain of AI used in literature for accident prediction and to study the driver's driving behavior. The techniques covered can be categorized into five subdomains, namely, search heuristic, supervised learning, dimensionality reduction, hybrid approaches and reinforcement learning. Figure 3 lists the categories and approaches covered in this work. There are many other techniques like, unsupervised learning, association rule learning, deep learning and hierarchical clustering however, this section covers the techniques which have been previously used for accident prediction or for the purpose of unsafe driving pattern analysis.

2.1 Evolutionary algorithms (EA)

Genetic algorithm (GA) (Koza 1998) is a heuristic based evolutionary search approach inspired by the Darwin theory and applies genetic operations of mutation and crossover to find the best solutions. GAs start with a set of candidate solutions known as the population. The candidate solutions are usually represented as a one dimensional array. The initial population is randomly generated using rules of the problem domain. The next population, also known as the next generation, is generated by selecting solutions from the previous population. For the purpose of promoting individuals from one population to the next, a fitness function is used to gauge fitness of the candidate solutions. It is assumed that new population will perform better as compared to the previous population and individuals having better fitnesses have more chances to be selected for the reproduction process. Crossover and mutation are the two reproduction operators used in GAs.

Genetic programming (GP), on the other hand, genetically breeds the computer programs to solve the problem (Koza 1998). Although GPs work like GAs but the candidate solutions in GPs are represented as a tree like structure. The breeding in GP is also based upon the fitness function and the reproduction operators are the same as were in the case of GAs. The main goal of the GP is to find the program from the search space which satisfies the fitness function and gives the best approximation to the objective. GP has two major advantages over the previous techniques. First, without any perspective, GP model can find a better functional solution to a problem. Second, GP can remove the black box effect caused by AI model (Xu

et al. 2012b). In general, GP model is based on the evolution theory which works on the population of mathematical models. In each generation a mathematical model is selected, based on a fitness function of the GP model. When the predetermined number of generations is executed or the best fitness value is achieved, the GP model stops. Equation (1) is a fitness function developed based on the error values predicted by GP model and actual data (Abdel-Aty et al. 2004), where $F(B_j)$ is the fitness value of the j th model, B_j in the population, $B_j(x_i)$ is the value of the j th model, and $C(B_j(x_i))$ is a scaling function that converts the value calculated by the model B_j into either 1 or 0. The functional form of $C(B_j(x_i))$ can be in any shape where values greater than a threshold are converted to 1 and others to 0.

$$F(B_j) = \sum_{i=1}^n \left(\beta^{y_i} \times |y_i - C(B_j(x_i))| + (y_i - B_j(x_i))^2 \right) \quad (1)$$

A GP based model is proposed in Xu et al. (2012b) to predict a crash in real time. The proposed system in Xu et al. (2012b) is used for freeways based on traffic, weather, and crash data obtained from a local authority. Authors have studied two types of traffic conditions; uncongested and congested. A separate GP model is proposed for each type of the traffic conditions where, random forest (RF) technique is employed to select the variables from the dataset that influence a crash the most. The candidate solutions in Xu et al. (2012b) are initialized using the ramped half-and-half method (Koza 1992) by restricting tree depth to six levels. Later, during the evolutionary process the trees are reported to have a maximum depth of 30 levels. For reproduction, models with best fitnesses are selected and priority is given to models with less depth. The fitness function used in Xu et al. (2012b) is based on the number of hits and square errors as shown in Eq. (1). Each state is analyzed using GP model and results in Xu et al. (2012b) show that traffic flow characteristics, which may lead to a crash, are found different in case of congested and uncongested states. Authors in Xu et al. (2012b) also show a comparison between binary logit model and GP model using receiver operating characteristic (ROC) which depicts that GP has better performance as compared to binary logit model (Abdel-Aty and Rajashekar 2006). The population size for experiments in Xu et al. (2012b) is set to 1000, and the stopping condition for the GP is the maximum number of generations set at 100. The number of iterations, however, seems to be set at a lower side. It would, however, have been interesting to see the results of the experiments for 1000 or more iterations. The stopping iteration number having a larger value can also be more convincingly set unless the solution converges or until no further improvement in the best found solution is achieved for many consecutive iterations. Random forest (RF) modeling technique which consists of an ensemble randomized classification and regression trees (Breiman 2001) is used in Xu et al. (2012b) for the selection of contributing factors in crash risk. These candidate variables selected by RF model were used to generate the GP model. GP model was used for the detection of hazardous conditions that lead to a crash under each traffic state. Prediction performance is compared with a binary logit model which was developed for the same dataset.

The work reported in Dixon et al. (2005), Damousis et al. (2007), and Yang (2012) addresses the problem of modeling human recognition, accident prevention and accident prediction, respectively. However, they do not use GAs directly to solve the problem. Instead, the GAs are used to train the actual learning method for better performance. Authors in Dixon et al. (2005) have used GAs to minimize the error between the converged samples and ground-truth labels for prediction. GA is used to train the parameters of the fuzzy expert system in Damousis et al. (2007) where the chromosome is a coded GA to avoid loss of accuracy. Similarly, the work in Yang (2012) utilizes GAs to train the parameters of support vector regression (SVR).

2.2 Conditional random field (CRF)

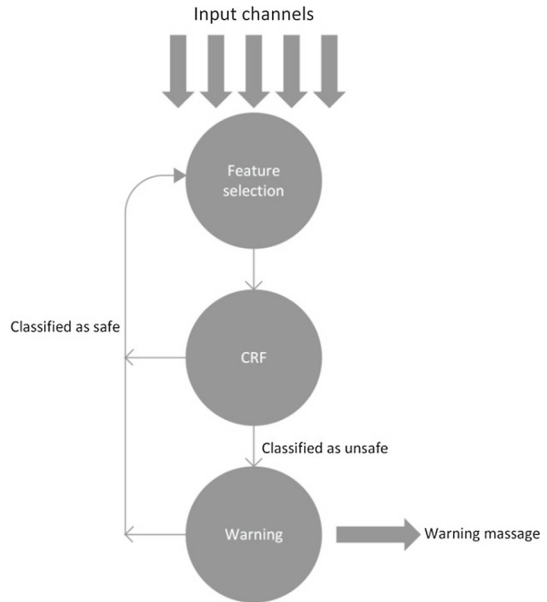
Labeling of sequential data is a problem which arises in many fields; CRF is an undirected probabilistic graphical model that is used for labeling and segmentation of data (Lafferty et al. 2001). A simple classifier predicts the label without considering neighboring samples, while CRF takes all samples under observation in a linear chain to predict the label. Since CRFs use undirected graph, there is no biased value regarding fewer states. CRF is a discriminatively trained undirected graphical model which uses Markov property in hidden nodes for observation and has characteristics of both discriminative and generative models. An ordinary classifier labels the sample data without considering the neighboring samples while CRF can take the whole context into account. To maximize the conditional probability of the label, given observation features are combined using exponential functions. Linear chain CRF feature function, for a node y_i given an input sequence x , (Zhou et al. 2007) is shown in Eq. (2), where, f_k is the feature and i is the node number.

$$F_i(Y, X) = \sum_k \lambda_k \cdot f_k(y_i, y_{i-1}, X) \quad (2)$$

Since, multichannel sequential data cannot be applied directly to a discriminative classifier (like, support vector machine (SVM)) because the data is temporally correlated that is why in Zhou et al. (2007) CRF was used as an inference model. The work in Zhou et al. (2007) aims at combining the data received from multiple channels for the detection of unsafe driving patterns. The dataset is generated using STISIM car driving simulator (Zhou et al. 2007; Ning et al. 2009, 2008; Jabon et al. 2011) and a total of nine channels are used. The channels record data of throttle, brake, steering wheel, position, speed, acceleration, lane position, distance to the same lane vehicle, and distance to incoming vehicle. Sampling frequency of 30 Hz is used to record the data. For each of the channels, authors in Zhou et al. (2007) calculate minimum, maximum, mean, variance and first-order derivative to estimate Gaussian mixture model (GMM) to be used for safe and unsafe driving pattern, respectively. The fusion of multiple channels in Zhou et al. (2007) has an advantage of avoiding the need of labeling each sample in training data since both labeled and unlabeled data are used in the semi-supervised learning algorithms. Figure 4 depicts working of the system. The data from multiple channels is recorded and features are extracted to be forwarded to the CRF. The feature vector is passed to the inference model to estimate the current driving state where CRF classifies a stream to be safe or unsafe and warnings are generated as per desired threshold. While training, CRF finds the best inference model parameters. Training the CRF with labels of the complete dataset is a time consuming and expensive process because if we label some patterns unsafe then it is not necessary that all unsafe patterns may lead to an accident. On the other hand, since it is difficult to label the data manually, therefore semi-supervised learning is used in Zhou et al. (2007).

Results in Yuejing et al. (2010) show that CRF outpaces the SVM and hidden Markov model (HMM) in classification of both labeled and unlabeled data. The work in Wang et al. (2010a) presents an approach to predict driving danger-level based on multiple sensor input data. The data taken into consideration is that of the vehicle dynamic parameters, physiological data of the driver and driver dependent features of the vehicle while driving. CRF is used in Wang et al. (2010a) to model the temporal patterns that can be used to predict unsafe driving states. As was in the case of (Zhou et al. 2007), the data in Wang et al. (2010a) is also generated using STISIM car driving simulator. For the purpose of generating danger level source, a history of 1 min is scanned for the CRF. Authors in Wang et al. (2010a) have compared the performance of CRF with HMM and reinforcement learning where the

Fig. 4 Flow of activities to detect an unsafe driving pattern



later seems to outperform other approaches. Similarly, an approach for dangerous driving warning system is introduced in Wang et al. (2010b) that uses sparsely labeled dataset for training. Nine vehicle dynamic parameters are recorded using STISIM car driving simulator (Wang et al. 2010b) and a semi-supervised learning approach is used for prediction which outperforms CRF.

2.3 Artificial neural network (ANN)

Modern neural networks (NN) are non-linear statistical data modeling tools that are used to model complex relationships between input and output, and to find patterns in the data. ANN analysis is considered to be an alternative approach for the investigation of non-linear relationships in engineering problems (Akin and Akbas 2010). An ANN model is developed in three phases: modeling, training and testing. A typical neural network structure is shown in Fig. 5. Rules, input parameter, and gathering of data are used in the modeling phase. For training phase, preparation of the data and adaption of learning laws are performed. Accuracy and performance evaluation is performed during the testing phase by calculating error between estimated and actual outputs.

During the past decade a number of ANN-based methods have been presented for the detection of accidents and its severity using different datasets and conditions. In Chong (2004) students of Oklahoma State University presented a model for the severity of injury caused by traffic accidents. Work in Chong (2004) used decision tree and ANN using backpropagation with different number of iterations to train the network. To minimize the error, conjugate gradient descent with 500 epochs was implemented. Analysis in Chong (2004) shows driver's seat belt usage, light condition of the roadway and driver's alcohol usage were the most critical features in fatal injuries. The dataset used in the study was taken from the national automotive sampling system (NASS) general estimates system (GES) consisting of traffic accident records from 1995 to 2000 with a total of 417,670 instances. Recently, in another

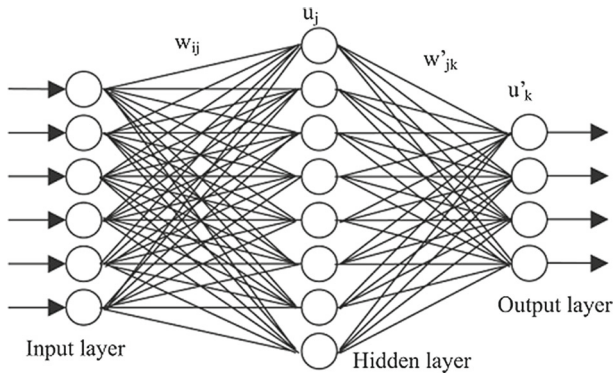


Fig. 5 A typical neural network structure

work (Tambouratzis et al. 2010) for the prediction of accident severity (light, serious or dead), a fused methodology of probabilistic neural network (PNN) and decision tree is used on the data collected by the republic of Cyprus police. The result shows that it enhances the classification accuracy (Tambouratzis et al. 2010). Both the studies in Chong (2004) and Tambouratzis et al. (2010) use historical data for the prediction of an accident. Although, such an approach can be useful but may cause inaccuracy when tested in real time since the data gathered in real time has few attributes and covers a limited history spanning over a few minutes.

It is observed that there are many parameters that affect the occurrence of accidents. Generally, these parameters are related to traffic flow, road section length, infrastructure geometric characteristics, pavement surface conditions, lighting, as well as weather and driver behavior. Decrease in these occurring parameters that leads to an accident can reduce the occurrence and severity of the accidents; several accident prediction models using observational data have been created. Many supervised and unsupervised techniques are implemented for the accident prediction using a combination of neural networks, support vector machine and decision tree. Some of these can be seen in Akin and Akbas (2010); Yuejing et al. (2010); Moghaddam et al. (2010); Wahab et al. (2009); Lv et al. (2009); Qu et al. (2012); Li et al. (2008).

2.4 Principal component analysis (PCA) and hidden Markov model (HMM)

Principal component analysis (PCA) is an unsupervised feature extraction technique mostly used to develop/select smaller number of artificial variables when there are a large number of observed attributes (Zhao and Karypis 2005). These artificial variables are called principal components and they have the best capacity to explain the variance of data and retain the characteristics of data without losing information. HMM is a statistical model based on Markov process (Rabiner 1989). Markov process states are directly visible and observed. While in case of hidden Markov model, states are hidden, but the output state is visible.

A mobile application called crash prediction is reported in Singh and Dongre (2012) which takes various attributes as input like age, gender, disability (if any), vision, date of license expiry or experience in driving. This data is treated as single variable and PCA is applied on it. Afterwards it is converted into driver dependent variables, on which the HMM technique is applied to determine if a driver is fit, unfit or partially fit for driving. For the purpose of prediction and classification, PCA serves as a preprocessing technique to select

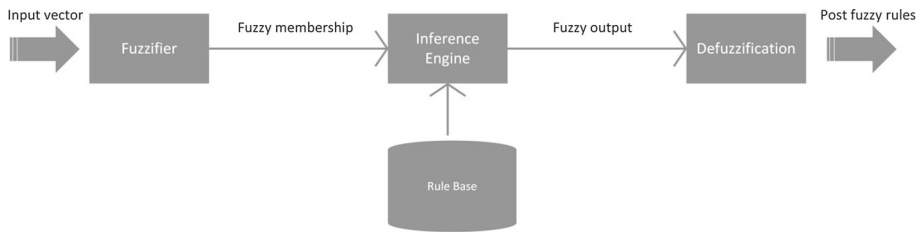


Fig. 6 Fuzzy system overview

minimum key attributes. Road accident prediction is done in [Ali and Bakheit \(2011\)](#) using ANN and for the purpose of selecting input variables for the ANN, PCA is used. PCA separates interrelationships to the independent components which are linear combinations of original items in the data. The approach in [Farah et al. \(2007\)](#) is based on creating an infrastructure coefficient for the freeways to predict an accident. The said coefficient is based on the freeway and its geometric features. PCA is used for the calibration of the coefficient consisting of eleven infrastructure characteristics. For the purposes of automated traffic safety analysis, work in [Yu and Abdel-Aty \(2013\)](#) presents an approach to identify unsafe driving maneuvers at intersections using data from video sensors. HMM is used in [Yu and Abdel-Aty \(2013\)](#) along with k -means clustering for the grouping of vehicle trajectories.

2.5 Fuzzy logic

Fuzzy logic is the representation of possible decisions in natural language, instead of quantitative representation ([Mohan 2011](#)). With fuzzy logic we can express our decision in terms of words instead of numbers and a solution provided by fuzzy logic is basically the reflection of human predicted solution. It enables computerized devices to think and reason like humans. Fuzzy inference system has the ability to map the nonlinear function using fuzzy rules and it can easily map the vector input into scalar output. A fuzzy inference system has four major components: fuzzifier, inference engine, rule base, and defuzzifier. Fuzzifier maps corresponding inputs to fuzzy membership while the rule base contains the rules provided by the human expert. The inference engine converts the fuzzy inputs to the fuzzy outputs using the rule base data. Defuzzifier converts the fuzzy output to a crisp numeric output as shown in [Fig. 6](#).

A detailed study is conducted in [Wahab et al. \(2009\)](#) to extract those features which are effective to profile a driver. A feature extraction technique based on GMM is executed to extract the features from the accelerator and the brake pedal pressure. The extracted features are then used as an input to a fuzzy neural network (FNN) to identify the profile. Results in [Wahab et al. \(2009\)](#) show that FNN has better performance as compared to simple multi-layered perceptron (MLP). An accident prediction framework is presented in [Hu et al. \(2004\)](#) which gets data of moving trajectories using a 3D model vehicle tracking and also records the images in sequence. To learn the vehicle tracking trajectories, a new fuzzy self-learning algorithm, based on fuzzy sets, is proposed in [Hu et al. \(2004\)](#). Accident is predicted by computing the matching degree and monitoring the activity of two moving vehicles. Gaussian field ([Ning et al. 2008](#)) uses semi supervised learning approach which defines a weighted graph based on a similarity function using both labeled and unlabeled data. All the data are sampled using a window size of 1 s and statistics like minimum, maximum, mean, variance and a first order derivative of each sample is calculated. This statistical data is compared with

the GMM to compute the features. CRF features are more expressive and have the ability to capture the dependency between the hidden state and the observations.

In [Imkamon et al. \(2008\)](#) unsafe driving behavior is treated as a subjective quantity and three perspectives are assumed to be used for detection of the unsafe driving behavior. The first perspective is the passenger point of view; second, driver's point of view and the third perspective is vehicle status. For the passenger point of view, the authors detect heavy jolts caused by sudden turns or brakes by using 3-axis accelerometer mounted on the passenger's seat. Video camera is mounted on the car's console to emulate the driver vision and focus on the road as a driver point of view. For vehicle status, velocity and speed of the engine are read directly from engine control unit (ECU) by on-board diagnosis II (OBD II) protocol. Result in [Imkamon et al. \(2008\)](#) shows that the proposed system works according to human opinions. Following are the rules which are used in fuzzy inference system for the output of the system ([Imkamon et al. 2008](#)) which exhibits the probability of hazardous driving behavior.

(i) Rules to minimize the impact to passengers

- If acc_X or acc_Y or acc_Z is HIGH then output = HIGH

(ii) Common rules for driving safety

- If turn rate is HIGH then output = HIGH
- If speed and density of car is HIGH then output = HIGH

(iii) Rules for car protection

- If Speed is HIGH then output = HIGH

The first rule is to discourage high acceleration while the second rule is for inferring about low speed requirement in critical situations. The third rule is for controlling the fuel consumption.

2.6 Temporal difference (TD) learning

Temporal difference learning is a supervised learning technique mostly used for reinforcement-based learning to measure the expected future reward ([Doherty et al. 2003](#)). Main goal of the temporal difference learning is to predict the quantity that depends on the future values. To find that, it uses prediction of successive time steps due to which it is called temporal difference learning.

Previously, TD learning has been used to find the unsafe driving patterns; however, it is still difficult to quantify the list of actions that contribute to unsafe driving. A generic framework is proposed in [Ning et al. \(2008\)](#) to detect unsafe driving patterns at runtime using multiple sensor readings. A danger level function is learned in [Ning et al. \(2008\)](#) and TD is used to approximate the expected function reward. The dataset used in [Ning et al. \(2008\)](#) is recorded using STISIM car driving simulator using 36 subjects where a sample of 20 min is recorded for each of the subjects. A signature is generated for each of the subject using mean, maximum, minimum, and variance which is used for the experiments instead of using raw data. Since the main challenge is labeling of data for training, therefore the authors in [Ning et al. \(2008\)](#) have used TD learning as reinforcement learning to give future expected reward to the danger level. TD learning gives an approximation to penalty the observable at collapse time. The proposed TD-based approach in [Ning et al. \(2008\)](#) is compared with two-category classifier using logistic regression and linear regressor. While in [Wang et al. \(2012\)](#) the authors have also used TD learning to avoid the labeling issue during training of data by giving an approximation to the risk at crash time.

2.7 Support vector machine (SVM)

Support vector machine (SVM) is a primary classifier based on decision boundaries and hyper plane. SVMs are learning machines that plot training vectors in a high dimensional space. An SVM-based approach in [Borges \(1998\)](#) classifies the data by creating support vectors and hyper plane represents the classified data of different classes. If the selected data for classification does not represent proper information, SVM cannot determine hyper planes in correct direction which degrades its performance. SVM is also limited for binary classification; however, it can be extended to multiple classes using transformation on the input data.

In [Li et al. \(2008\)](#) a prediction model based on statistical theory is proposed which uses SVM. Analysis of the model is performed on rural frontage road data of Texas. SVM model results are compared with negative binomial (NB) regression model which shows that SVM predicts with better accuracy as compared to the NB model. A previous work of the authors was based on the backpropagation neural network (BPNN) where the result showed that SVM is much faster as compared to BPNN. Correspondingly, work in [Lv et al. \(2009\)](#) implements SVM to study the vehicle patterns that results in an accident and also those which do not lead towards an accident. This makes it a binary classification problem and thus SVM is a better suited classifier. The data used in [Lv et al. \(2009\)](#) is of a real time traffic scenario using a traffic simulator software tool. The RBF kernel function is used with SVM for the experiments in [Lv et al. \(2009\)](#). The study uses six variables as candidate accident precursors. These variables are the mean and the standard deviation of traffic, headway speed, headway time, precursor values recorded for 5 min duration 50 min prior to an accident and for 5 min duration right before an accident, respectively. The study reports that, for hazardous condition prediction, SVM must be supplied with more than one traffic variables.

In order to evaluate real-time crash prediction the study in [Yu and Abdel-Aty \(2013\)](#) uses SVM. For the purpose of selecting key attributes from the data, classification and regression models have been used. The data are divided into two categories: training and scoring. Two crash datasets consisting of 265 crashes and 1017 non-crash instances from a 15 mile mountainous freeway are used for the experiment. A number of kernels have been evaluated and RBF is found to be performing better as compared to the rest.

2.8 Other techniques

Image processing is used to analyze and manipulate the data by taking images or video as input and applying image processing operations like compression, segmentation, and thinning. An android based application, CarSafe, is reported in [You et al. \(2012\)](#) for driver safety that uses dual cameras and other embedded sensors in a smart phone and fuses all the collected information. Front camera monitors the driver's head pose and other activity to alert the driver in case of drowsiness or fatigue. While the back camera screens the distance between vehicles and other lane change actions to warn the driver if the vehicle is too close to another vehicle. This whole process is performed in four steps of camera switching, frame dispatching, image preprocessing and driver status implication ([You et al. 2012](#)).

Model predictive control (MPC) generally represents the behavior of a dynamic system and controls the process ([García et al. 1989](#)). MPC predicts the change caused in dependent variables due to independent variables using dynamic process model. Mostly these are used in industries for process control to predict and control all the changes in inputs. A key advantage of MPC is the optimization of current time while considering future events. In [Rygula \(2009\)](#); [Murphy et al. \(2009\)](#) driver style identification was analyzed using speed graphs and jerk

analysis within a specific time window. Observation of the speed graph shows that each driver has a unique speed graph and it can be used as a behavioral biometric and deep analysis can also be used to investigate the driver's psychophysical state (Rygula 2009).

Multiple linear regression (MLR) is a statistical technique used to find the relationship between dependent and independent variables. To minimize the error between observed differences and predicted values it uses least square method. It takes a group of random variables as input and finds the mathematical relationship between them by a straight line which approximates all data points. MLR is used in dendroclimatology for developing models to reconstruct climate variables from tree-ring series. MLR cannot be used for time series problem because, in time series data, most of the observations are dependent on each other. Linear prediction is used mostly for speech processing and analysis of speech signals. They are linear mathematical operations, which are used to estimate the future value of time signal by analysis of previous sampled values. MLR is used in Abdel-Aty and Radwan (2000) for modeling traffic accident occurrence.

DS (Driver's Style) classification (Murphey et al. 2009) is an algorithm which was proposed for online driver style identification. It calculates the jerk within a specific time window and uses average jerk for identification of a driver. It also predicts the road type and traffic state whether it is congested or not. Experimental result in Murphey et al. (2009) shows that DS classification performs better than acceleration-based classification. Halim et al. (2016) addresses the problem of profiling drivers based on their driving features only. A dataset is recorded using 50 subjects and is then profiled using clustering techniques. For the purpose of clustering k -means, Fuzzy c -means and model based clustering is used. Results in Halim et al. (2016) shows that average speed, maximum speed, number of times brakes were applied, and number of times horn was used provide the information regarding drivers' driving behavior, which is useful for clustering. The work later trains multiple classifiers for the prediction of drivers' profile. Canale and Stefano (2002) presents an analysis to decide whether it is possible to determine a driver's behavior. Additionally, the work also investigates the suitable signals for identifying the driver's style and also which parameters can be used to describe a driver's style. An assignment procedure is defined to classify a driver's behavior within the stop and go task. A Dynamic Time Warping (DTW) and smartphone-based sensor-fusion system is presented in Johnson and Trivedi (2011) to classify a driver in two categories: non-aggressive and aggressive. The system recognizes and records various actions without external processing. Work in Ly et al. (2013) explores the possibility of using the vehicle's inertial sensors from the CAN bus to build driver's profile. This helps in reducing the dangerous car maneuvers by providing an appropriate feedback. Results show that the braking and turning events contribute more towards characterizing an individual. Shi et al. (2015) present an approach to quantitatively evaluate driving styles by normalizing driving behavior based on personalized driver modeling. A personalized driver model is established for the drivers. The personalized model is evaluated by a neural network. Later the driver model is used to simulate standard driving cycle test. An aggressiveness index is also proposed using the energy spectral density analysis on the normalized behavior. The index is useful to identify abnormal driving behaviors.

Transportation professionals are paying much attention towards the development of real time crash prediction models for freeways. In Veeraraghavan et al. (2005) physical data such as speed and turn signal is used for experiments. To measure the posture of the driver, a pressure sensitive chair and ultrasonic six degree freedom head tracking system was used and have manually labeled the videos of driving according to a list of situations. The main objective is to predict the time series of human recognized situation using the various sensory inputs. Sandia cognitive framework (SCF)

(Dixon et al. 2005) is used to integrate the information of vehicle state and driver posture which gives a nonlinear dynamical pattern for recognition system. To minimize the error between ground truth and estimated parameters using SCF, authors in Dixon et al. (2005) use gradient descent approach and GAs. Using gradient descent estimation to minimize the error, rules are updated which, consequently guarantees to converge and terminate at a local minimum. For the GA, DAKOTA (design analysis kit for optimization and terascale applications) optimization package (Bruns et al. 2005) is used.

It is a common observation that every driver operates the vehicle in his/her own style, some drive slow while some like aggressive driving. Vehicle state normally depends on the speed, steering angle, acceleration, brake usage frequency, inter car distance and driver's mental state. Authors in Rygula (2009) have analyzed the driver speed profile to identify the driver style. Driver speed profile can also be used to get other information like psycho-physiological state recognition, driver sleepiness and tiredness. Amount of points in which the derivative of speed changes with respect to time is called speed profile in Rygula (2009). Equations (3) and (4) are used in Rygula (2009) to estimate maxima and minima of the speed function, respectively, as

$$\frac{dV}{dt} = 0 \quad \text{and} \quad \frac{d^2V}{dt^2} < 0 \quad (3)$$

$$\frac{dV}{dt} = 0 \quad \text{and} \quad \frac{d^2V}{dt^2} > 0 \quad (4)$$

where, V is the velocity, t shows the time, and, $\frac{dV}{dt}$ is the change in the velocity with respect to the time, i.e., acceleration. Equation (3) defines the amount of points in which the speed offshoots with respect to time. This is called intensity change of speed profile. Equations (3) and (4) are used to estimate extrema of the speed function taken as the points in which acceleration change its sign

Table 1 lists the aforementioned related studies on accident prediction models that are based on different AI techniques. These techniques can be divided into three groups, which are supervised or unsupervised learning, statistics-based learning, and other techniques. Table 1 shows that ANN, SVM, GAs and statistics-based techniques are frequently used AI techniques for the prediction of an accident/unsafe driving patterns analysis during the past 10 years.

Various AI-based techniques have been covered in this section for the prediction of an accident and analysis of unsafe driving patterns. To have a comparative view of these techniques, Table 2 lists eight important features of these techniques. The features covered include: (1) learning technique/domain, (2) whether or not the technique is population-based, (3) best reported accuracy of the techniques for accident prediction, (4) whether or not the technique is suitable for real-time deployment, (5) does the techniques support multi-format data?, (6) state-of-the-art work using the technique, (7) key strength and (8) notable weakness of the technique. The values of these eight features are listed keeping in view the accident prediction and AI domain. The values may differ for other problem domains. Based on the results in Table 2 the best reported accuracy is of PNN and decision trees with an accuracy of 95.9307% in Tambouratzis et al. (2010). However, this accuracy does depend upon the type of data, the amount of data and the features taken into consideration. Any comparison made with a new or different technique will only be valid if it is tested on the same sample as was in the baseline approach.

Table 1 Different AI techniques used in literature for driving safety and vehicle crash prediction

Category	Technique	Work
Supervised/unsupervised learning	Artificial Neural Network	Akin and Akbas (2010), Dixon et al. (2005), Yuejing et al. (2010), Moghaddam et al. (2010), Yoon et al. (2013)
	Support vector machine	Qu et al. (2012), Lv et al. (2009), Li et al. (2008)
	Decision trees	Chong (2004)
	Agglomerative clustering	Veeraraghavan et al. (2005)
Statistics-based learning	Conditional Random Field Probability	Bruns et al. (2005), Rygula (2009), Wang et al. (2010a), Wang et al. (2010b)
	Local probability	Lafferty et al. (2001)
	Principle component and hidden Markov analysis	Wang et al. (2012)
	Temporal difference learning	Chong (2004)
	Self-learning algorithm	Wang et al. (2012)
Other techniques	Genetic programming	Moghaddam et al. (2010), Wang et al. (2012), Yang (2012)
	Speed graph analysis	Rygula and Mitas (2007)
	Fuzzy logic	Imkamon et al. (2008)

3 Datasets and simulators

An important issue in crash prediction system and road safety research is the availability of datasets. In literature, different studies have used diverse datasets, some of which are publicly available over the web, such as the CIAR (Wahab et al. 2009), STISIM data (Jabon et al. 2011) and SPD (Lafferty et al. 2001) datasets for the prediction of unsafe driving analysis and accident prediction, while other studies have collected data from local departments in the countries concerned. Table 3 lists the datasets used in related studies, including their size, instances and number of attributes.

In Dixon et al. (2005), data are collected through five different human subjects who were instructed to drive on urban roads for about 200 km and the data of 24 h was sampled at the rate of 4 Hz. For the classification using supervised learning following eight situations were decided as ground truth, (1) leaving intersection, (2) entering on ramp, (3) high speed roadway, (4) being over taken, (5) high acceleration, (6) approaching a slow vehicle, (7) preparing to change the lane, and (8) changing the lanes. Similarly, the proposed approach in Zhou et al. (2007) is evaluated using the car driving simulator STISIM. Work in Zhou et al. (2007) uses nine channels including: throttle, brake, steering wheel, position, speed, acceleration, lane position, distance to same line, and distance to incoming vehicle for driving performance. In Imkamon et al. (2008) Hitachi H48c accelerometer is used for acceleration where, a high value of acceleration in X-axis and Y-axis indicates the heavy jolts to passenger seat. The high value of acceleration in Z-axis indicates the jump and roughness of road experienced by the vehicle. By ignoring polarity of axes, work in Imkamon et al. (2008) has considered only

Table 2 Comparison of various AI techniques for accident prediction

Technique	Learning technique	Is population based?	Best reported accuracy	Suitable for real-time	Is multi-format data supported?	State-of-the-art reported	Strength	Weakness
Genetic algorithm	Search heuristic	Yes	85% in Dixon et al. (2005)	No	Yes	Dixon et al. (2005)	Provides multiple possible solutions	May stuck in local optimum
Genetic programming	Search heuristic	Yes	82% in Xu et al. (2012b)	No	Yes	Xu et al. (2012b)	Identical parents may produce different child solution	May stuck in local optimum
Conditional random field	Supervised	No	0.081 in Zhou et al. (2007)	Yes	Yes	Zhou et al. (2007)	Conditional in nature	linear state transition factors only
Artificial neural network	Supervised	No	95.9307% in Tambouratzis et al. (2010) using PNN and Decision trees	Yes	No	Tambouratzis et al. (2010)	Reprogramming not required	Training dependent
Principal component analysis (PCA)	Dimensionality reduction	No	N/A	Yes	Yes	N/A	Reduces complexity and noise	Invariance explicitly required in training data
Hidden Markov model (HMM)	Supervised	No	0.063 in Zhou et al. (2007)	Yes	No	Zhou et al. (2007)	Captures dependencies	Represents finite static distributions
Fuzzy logic	Hybrid	No	60–70% in Imkamon et al. (2008)	Yes	Yes	(Imkamon et al. 2008)	Simpler to use	Scaling issues with large rulesets

Table 2 continued

Technique	Learning technique	Is population based?	Best reported accuracy	Suitable for real-time	Is multi-format data supported?	State-of-the-art reported	Strength	Weakness
Temporal difference (TD) learning	Reinforcement learning	No	89.46% in Chong (2004)	Yes	No	(Chong 2004)	Less computations required per observation	May get intractable for large space
Support vector machine (SVM)	Supervised	No	0.81 in Yu and Abdel-Aty (2013)	Yes	No	(Yu and Abdel-Aty 2013)	Fewer hyperparameters	May get to be a black box

the forces imposed on a passenger while ignoring the direction (Imkamon et al. 2008). By passing through a high pass filter, DC offset is also eliminated. All driving events of moving car can be measured by the change in acceleration.

In Xu et al. (2012b), for the experiments of GP model, data were obtained from a 21 mile freeway section of the I-880N freeway in United States using 40 loop detector and 3 weather stations. The crash data for the selected freeway were obtained from the statewide integrated traffic record system (SWITRS) and traffic data were obtained from the highway performance measurement system (PeMS) maintained by Caltrans. Traffic data were collected from the nearest upstream and downstream station for each crash location. For a speed graph analysis, tachograph disk is easily accessible and provides 24 h continuous data of driver by reading driving parameters, which include speed graph and road graph. Detailed data description of tachograph is available in Mitas (2007); Ryguła and Mitas (2007). Using tachograph disk, speed graph is generated as a result of up and down movement of the scribe. Tachograph disk information can be converted into digital form by finding its extreme points and converting pole coordinate to the value of temporary speed and proper position on the time axis (Ryguła 2009). Table 3 shows the datasets which are used in the literature. The format of the data and the recorded attributes depend upon the type of problem being solved.

To acquire the driving data, using a simulator is always a good choice since it will reduce time consumed and avoids any injuries during experiments. There are a number of simulators available for this purpose. STISIM is a programmable driving simulator with the support of 3D graphics. It comes with a number of ready to run simulations and is capable of recording user specific driving features. SCANeR is a driving simulation engine that serves to allow training and safety awareness for five types of vehicles, including: cars, trucks, bus, armored vehicles and emergency vehicles. The SCANeR is also supported by a complete development kit and 3D graphics. Carnetsoft driving simulator is a low cost simulator aimed towards the training and education of drivers. Carnetsoft driving simulator also has a software module for studying the driver behavior and collecting related data. VANET is an open source driving simulator with a key focus on security and supports simulation on real traffic scenarios by providing the ability to import maps. VANET also supports micro-simulation where each vehicle can be simulated individually and may take autonomous decisions. PTV Vissim is a specialized software tool developed to analyze public transport and the effect of signaling. PTV Vissim also provides an application programming interface (API) for integrating external applications. Aimsun is another traffic modeling software with capabilities of mesoscopic, microscopic and hybrid simulation. Quadstone Paramics is traffic simulation software used for planning and design of transportation systems. Quadstone Paramics also supports the study of pedestrians to be incorporated in the transportation infrastructure. For the purpose of evaluating infrastructure and policy changes, SUMO (Simulation of Urban Mobility) is developed as a free simulator. TraNS is a simulator for vehicular ad hoc networks to study the mobility of vehicles. TraNS provide a visual interface integrating the SUMO (a traffic network simulator) and ns2. OpenEnergySim is an online simulator that is used to visualize microscopic traffic and CO₂ emissions of vehicles. iTETRIS provides a flexible simulation platform to analyze traffic at city level. Table 4 shows the key features of the aforementioned simulators.

4 Discussion

The work in this paper encapsulates the studies based on AI techniques for accident prediction and identification of unsafe driving states. However, there are many open issues that have

Table 3 List of datasets used in literature

Data source	No. of records/frames	No. of attributes	Related study
Center for integrated acoustic information research (CIAIR)	800	5	Wahab et al. (2009)
STISIM car driving simulator	30	9	Zhou et al. (2007), Ning et al. (2009), Ning et al. (2008), Jabon et al. (2011)
Camera video	10,231–10,380	16–38	Imkamon et al. (2008), Veeraraghavan et al. (2005)
Analogue tachograph disk	15	3	Rygula (2009)
Republic of Cyprus Police (RPC)	1407	33	Tambouratzis et al. (2010)
I-880N freeway (SWITRS, PeMS)	807	21	Xu et al. (2012b)
State Police Department (SPD) of the State of Michigan	115	48	Lafferty et al. (2001)
OBD-II Reader	15	3	Bruns et al. (2005), Shaout and Bodenmiller (2011)
National automotive sampling system (NASS)	160	9	Abdel-Aty et al. (2005)
Using GPS	106	2	Ellison et al. (2012)

either not been comprehensively examined in the literature or reasons for an efficient and accurate system for accident prediction are lacking. These open issues provide opportunities for further research. This section lists the shortcomings of the previous work and discusses how these can be further investigated. The work in Dixon et al. (2005), has used a labeled dataset for classification of driving situations, however, it is not clearly discussed how many physical states and how many video sensor inputs are used for the experiment. Explicit mention of this information can certainly be useful for others to replicate the results and also for the comparison purpose. The study in Dixon et al. (2005) has used human subjects for the creation of ground truth which seems to be a tedious task keeping in view that each subject had to drive for around 200 km. Although the results are promising, yet a modern tool like STISIM car driving simulator can be used in future for recording the dataset. Authors in Dixon et al. (2005) employed GA which is an evolutionary algorithm and needs more time for convergence. Results show that GA has a very high computation time as compared to the gradient descent approach which makes it less efficient and less reliable for real time system. Since GA is search based heuristic, and in order to get better and quick results, multi-population GAs can be employed in the future. The multi-population GA optimizes the multiple candidate solutions simultaneously, which makes them get to the global best solution using local search in each of the population. The efficiency of multi-population GAs can further be enhanced using parallel GAs (Abu-Lebdeh et al. 2014) where each population of the GA will be assigned to a separate processor. For the purpose of identifying unsafe driving patterns or the prediction of an accident, GAs can be utilized to train a classifier using both labeled and unlabeled examples. For the labeled examples, GA can use the class labels from the training dataset to assign fitness to each candidate chromosome: whereas, for the unlabeled training example a fitness

Table 4 Simulators for vehicle and traffic analysis

Simulator	Open source	Supported OS	Allows scripting	Online 3D graphics	graph-ics	Traffic analysis	Data analysis	Visualization	Hardware support	URL	Freeware
STISIM	No	N/A	Yes	No	Yes	Yes	No	No	Yes	http://www.stisimdrive.com/	No
SCANER	No	Windows XP, Vista, Windows 7, RT OS	Yes	No	Yes	Yes	Yes	No	Yes	http://www.scannersimulation.com/	No
Carnetsoft	No	N/A	No	No	Yes	Yes	Yes	Yes	Yes	https://www.rjsschool-simulator.nl/index_EN.html	No
VANET	Yes	Any	Yes	No	Yes	Yes	Yes	Yes	No	http://svs.informatik.uni-hamburg.de/vanet/	Yes
PTV Vissim	No	Microsoft Windows	Yes	No	Yes	Yes	Yes	Yes	No	http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/	No
Aimsun	No	Microsoft Windows	Yes	No	Yes	Yes	Yes	Yes	No	http://www.aimsun.com/wp/?page_id=21	No
Quadstone Paramics	No	Microsoft Windows	Yes	No	Yes	Yes	Yes	Yes	No	http://www.paramics-online.com/	No
SUMO	Yes	Microsoft Windows	Yes	No	Yes	Yes	Yes	Yes	No	http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/	Yes
TraNS	No	Any	No	No	Yes	Yes	Yes	No	No	http://lca.epfl.ch/projects/trans/	Yes
OpenEnergySim	No	N/A	Yes	No	Yes	Yes	Yes	Yes	No	http://54.65.32.237/globallab/projects/openenergysim/	No
iTETRIS	Yes	Ubuntu	Yes	No	Yes	Yes	Yes	Yes	No	http://www.ict-itetris.eu/	Yes

function based on maximizing or minimizing a reward can be utilized. Although GAs can also be directly used to predict an accident or unsafe driving pattern, instead of being used to train a classifier, however, their offline use is always recommended. The reason is the real-time performance requirement of the prediction system and the naturally slow nature of a GA. The GP and other evolutionary search heuristics share the common performance issue as are for the GAs (Table 5).

In view of the detailed study in Zhou et al. (2007), CRF seems to be a better approach as compared to HMM. HMM models the observations independently given the hidden variables, and these assumptions are too strong in many cases. Since, CRF uses a probabilistic approach for labeling the data based on a conditional probability distribution over a sequence of labels, this makes them perform better than HMM. In addition to the study reported in Zhou et al. (2007), CRF has also performed better than HMM in other studies related to bioinformatics and pattern recognition. For the purpose of computing maximum-likelihood values in CRF, a dynamic programming based approach can be used. Traditional CRFs do pose a limitation when it comes to representing nonlinear dependencies in each frame (Maaten et al. 2011). To overcome this, hidden unit CRFs seem to be potentially useful where the hidden units try to learn the latent distributed structures in the underlying data for better classification accuracy. HMM, on the other hand, can still be useful in better compression that allows sequences to be found than a simple Markov model, provided HMM is well-tuned.

Investigating the prediction performance and accuracy of various ML techniques provide a deeper understanding of the current state of prediction methods, enabling future research to take the next step. Different datasets are used in the previous studies nevertheless; it may be worth examining the relationship between selected datasets, prediction accuracy, and number of features used in those experiments. In the literature, a variety of measurements are used to assess the prediction performance. This survey considers accuracy rate for the comparison of previous approaches. Accuracy is considered for comparison because of two key reasons. Firstly, because it is reported in nearly all related studies, and secondly for accident prediction/unsafe driving pattern identification it is of extreme importance as failure of the prediction approach may be fatal. The results of Dixon et al. (2005) show that the gradient descent algorithm predicts the driving situations correctly over 95 % of the time, while GA predicts with an accuracy of 85 % using the test data. Additionally, gradient descent has a very less convergence time as compared to GA's computation time. For the evaluation and training of 18.3 h driving dataset, 1 h is required in case of gradient descent approach and 1611 h for genetic algorithm (Dixon et al. 2005). However, the results and their comparison depend upon the problem being addressed and the dataset used in addition to the particular AI technique employed. Prediction performance comparison of GP and binary logit model in Xu et al. (2012b) shows that, under uncongested traffic conditions, GP model increases the accuracy of crash prediction by an average of 8.2 % and for congested traffic it is increased by 4.9 %. Fuzzy logic system proposed in Imkamon et al. (2008) is tested against ground truth based on a questionnaire prepared by three passengers. The system is tested in four rounds and average error of each round is 0.225.

In Wahab et al. (2009) authors have compared the performance of GMSS, MLP, ANFIS (adaptive network-based fuzzy inference system) and EFuNN (evolving fuzzy neural network) by taking different combinations of the driver static features, dynamic accelerator and brake pedal pressure. Where, the GMM based extracted feature outperforms the wavelet coefficients. Result in Wahab et al. (2009) depicts that EFuNN has 100 % gender identification rate and GMSS has low error for cross validation of drivers. The proposed combination of PNN and DT in Tambouratzis et al. (2010) shows better accuracy as compared to naive Bayes and BNN. For the work presented in Zhou et al. (2007), Ning et al. (2009) and Ning

et al. (2008), the authors have tried to handle the labeling issue and training of labeled and unlabeled data by using CRF and TD learning, respectively. Experimental results show the effectiveness of the technique. In other statistical techniques, analysis of different driver's speed graph (Rygula 2009) shows that each driver has a unique speed graph and long route driving causes increase in acceleration due to the tiredness of the driver. DS-Classification algorithm (Murphey et al. 2009) effectively classifies the driver style by conducting experiments in PSAT simulation environment. While analyzing the speed profile for the driver identification in Rygula (2009), performance measurement experiments were performed by analysis of 10 tachograph disk of 2 drivers. Results in Rygula (2009) show that both drivers have a significant change in their speed profile characteristic. "Appendix" lists the complete details of the literature covered in this paper as a comprehensive table.

5 Open issues

Based on a careful examination of the literature and on the basis of the analysis presented in Table 2 and "Appendix", this section lists the current open issues for future research in accident prediction and detection of unsafe driving patterns. In general, the open problems in the subject domain are as follows:

5.1 Performance markers

Performance markers have been developed for many problem domains where the comparison of various techniques is done using these markers. For instance, clustering is a subjective issue and the performance markers for this are; Davies Bouldin index (DBI), Dunn index (DI) and Silhouette coefficient (SC). In case DBI is used as performance marker, the best value is the one having the lowest index value. For DI, the best has the highest index value and for SC the value closest to 1 is considered to be the best. However, most of the related studies in the domain of accident prediction and unsafe driving pattern analysis have used different datasets for method evaluation; thus it is not feasible to directly compare these models. Performance markers need to be developed for these studies where the said markers should take into consideration the size of data, type of the data, data format and prediction technique's performance. This will not only help in comparison of results, but will also accelerate the pace of research for accident prediction, as the comparison with state-of-the-art approaches will get easier and more logical. To state an example from the domain of automated generation of game contents, a set of entertainment metrics is proposed in Halim et al. (2014) which is used for the measurement of entertainment that the contents of a game may carry. The same can be used for comparison of different techniques for generation of games. A performance marker as in Halim et al. (2014) is also required for the research in the area of traffic accident prediction and road safety approaches. Once a performance marker is introduced, this will also help in conducting studies by further increasing the attributes of dataset. Moreover, the effect on the performance of the accident prediction based on multiple channel data can be studied.

5.2 Benchmarking datasets

Creation of benchmark datasets is a missing item in accident prediction and unsafe driving patterns analysis domain. Benchmark datasets need to be developed upon which algorithms

can be tested. Currently the UCI machine learning repository¹ is the major source of the benchmark dataset from where researchers can download a variety of domains related dataset for testing the results of the proposed approach against the ground truth. However, unfortunately as yet no dataset has been submitted as a benchmark consisting of vehicles data for accident prediction. Since the accident can be predicted from a variety of data, including vehicle's information, weather data, road/traffic conditions, and driver demographic, datasets need to be created for all these perspectives. Although a benchmark tool, STISIM car driving simulator, to record road/traffic data is available which can be used to record datasets, but the cost is at a higher end and to prompt research in road safety using AI, benchmark datasets can play a vital role. To start with, the data sources listed in Table 3 can be combined at a central repository. A very rich source of data, strategic highway research program (SHRP 2)² can be useful in benchmarking the dataset which is publically available. SHRP 2 is authorized by US Congress to address key issues in highway systems. The four key research areas covered by SHRP 2 are; (a) to understand the interaction among various factors involved in highway crashes—driver, vehicle, and infrastructure, (b) renewal, (c) reliability and (d) capacity.

5.3 Driver identification

Identification of the driver is also an interesting research area that needs to be addressed in future (Wu and Ye 2009a). An individual can be identified based upon his/her vehicle driving patterns. Such approaches can be utilized in protecting vehicles from theft and for parental control. There are many existing techniques that use biometric, image processing and voice signals (Wu and Ye 2009b) for the identification of a driver. However, all of these techniques are visible to the vehicle driver and thus will implicitly ping an authorized driver if he/she is driving the vehicle without permission or if the vehicle is being stolen. A driver identification system needs to be developed taking into account the driving patterns of the individual driving the vehicle. In addition to other advantages, from the view point of road safety driver identification will avoid crashes due to rash driving of a stolen car and can also be used to indicate driving while being drunk. Such a system when integrated with the vehicle security features can prove to be very useful. An interesting work in the same direction can be seen in Miyajima et al. (2006) where spectral analysis is used to identify driving behaviors using signals like gas and brake pedal operation while accelerating or decelerating. Algorithm learning based neural network integrating feature selection and classification is a promising approach introduced in Yoon et al. (2013), the same can be applied for driver identification for better accuracy.

5.4 Clustering drivers and/or their features

Clustering is an unsupervised learning technique that is used to group similar objects using a distance function. Clustering can be applied on a set of drivers based on their driving features or based on drivers' demographics. The resultant clustering can be useful for many domains, such as customizing a car design for a specific group of people, advertisement and road safety analysis. Since clustering is an unsupervised approach, a descriptive analysis of the formed clusters will be required that may involve a domain expert. The said clusters can also be useful in linking a group of drivers with a particular physiological profile that can further be used as a recommender system for a specific vehicle type, driving care or

¹ <https://archive.ics.uci.edu/ml/datasets.html>.

² <http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Blank2.aspx>.

assistive technology that can be useful for that cluster. The results of the clustering are highly dependent on the size, type and features in the dataset. Clustering results vary across different types of clustering algorithms. As a proposed example, the driving features that may be taken into consideration for clustering can include: number of times footbrakes are applied, use of horns, maximum speed achieved while driving, driver's average speed, ratio of left turns to use of left indicators, ratio of right turns to use of right indicators, maximum gear used by the driver, driver's average gear and number of times a vehicle gets into reverse gear. However, the accuracy of data that will be recorded is influenced by the data recording tools used and this can be overcome by using a benchmark simulator like STISIM. Some work in this direction can be seen in [Ellison et al. \(2012\)](#); [Kalsoom and Halim \(2013\)](#); [Halim et al. \(2016\)](#).

5.5 Additional considerations

Accident prediction, as already a major focus of this paper, is also an important aspect in road safety. Algorithm further needs to be built that can predict an accident with better accuracy and less prediction time. Real time analysis of accident prediction techniques is worth considering. Currently, most of the studies use simulation data or the simulators for testing their approaches. There can, however, be many factors like environment, road conditions, weather and number of passengers in the vehicle that can have influence over the occurrence of an accident. These factors can cause a delta in the expected and actual performance of an accident prediction technique. The same accident prediction systems can also serve as an early warning generation for unsafe driving states.

The accident prediction techniques, most of the times, uses a classifier; the same classifier can also be used for modeling of the individual driver's features. The modeling of a driver thus becomes a controller training problem. A trained controller can be used in fully autonomous or semi-autonomous vehicles ([Göhring et al. 2013](#)). Accident prediction systems can also be integrated with post-accident systems and services to trigger rescue and emergency activities. False alarm generation is an important issue for these studies and can be minimized by increasing the accuracy of the prediction technique. The secondary problems in post-accident systems include shortest path finding, searching a path which consumes minimum fuel and nearest hospital finding ([Chen et al. 2012](#); [Islam and Rahman 2014](#)), where these problems can take into account the GPS systems of either the vehicle or the cell phone of the driver ([Mathkour 2011](#)).

Predicting driver suitability for driving is also an open problem that not only can benefit from the clustering approaches for grouping the drivers into homogeneous clusters, but can also utilize the prediction systems that can map a driver into one of the predefined classes. Such systems can be utilized by traffic police departments for issue or renewal of driving licenses. From the point of view of human machine interaction (HMI), appropriate interface development is vital for the accident prediction and unsafe driving pattern analysis systems. The driver needs to concentrate on his/her driving and the accident prediction and avoidance systems will only assist the driver for a safe journey. The interface through which the message is to be conveyed to the driver needs to be explored from the HMI perspective ([Peschel and Murphy 2013](#)) since visual cues to the driver may cause distraction. To resolve this issue, sonification of the signals ([Halim et al. 2016](#)) can be explored. Some other interesting directions for road and passenger safety include, anger management, speed management and stress management.

6 Conclusion

This paper presented a comprehensive survey of various artificial intelligence techniques used in the literature for predicting an accident and also to study the unsafe driving patterns. The literature covered is from the past 10 years, i.e., 2004–2014 and is motivated by lack of such survey in the literature. The review reveals that the artificial neural networks, support vector machine and genetic algorithms have been the most frequently used

Table 5 Acronyms and notations used in this review (unless stated otherwise)

Full form/semantics	Abbreviation/parameter
Artificial neural network	ANN
Average vehicle occupancies	AVO
Binary logit model	BLM
Bayesian neural network	BNN
Crash prediction system	CPS
Conditional random field	CRF
Driving style	DS
Decision trees	DT
Discrete wavelet transform	DWT
Engine control unit	ECU
Electro-oculogram	EOG
Fuzzy neural network	FNN
Genetic algorithm	GA
Gaussian mixture models	GMM
Genetic programming	GP
Hidden Markova model	HMM
Kanade–Lucas–Tomasi algorithm	KLT algorithm
Multiple linear regression	MLR
Model predictive control	MPC
Multilayer perceptron	MPL
Neural network	NN
On-board diagnosis	OBD
Principle component analysis	PCA
Probabilistic neural networks	PNN
Random forest	RF
Recursive least-square	RLS
Receiver operating characteristic	ROC
Support vector machine	SVM
Temporal difference learning	TD learning
Intervention strategy	γ

artificial intelligence techniques for the prediction of an accident/unsafe driving pattern analysis during the last ten years. The best reported accuracy for predicting an accident is of probabilistic neural networks and decision trees with an accuracy value of 95.9307%. However, this accuracy does depend upon the type of data, the amount of data and the features taken into consideration by the prediction technique. The paper has also listed the available datasets and the simulators which can be used to conduct research on accident prediction and studying unsafe driving patterns. Accident prediction can be performed using rich format data like images/videos and the same can also be done using data formats that consume less memory and processing power such as, reading vehicle information from the engine. For the purpose of capturing the required data, richer data formats may also require expensive data recording devices to be installed in a vehicle, whereas, the engine related information can be captured using the on-board diagnosis protocol without any additional expense. Nevertheless, irrespective of the data format and the data capturing mechanism, the accuracy of accident prediction is the most critical aspect in road safety where the artificial intelligence has already contributed and this further need to be improved.

The open issues in the subject domain have also been discussed to provide directions for the future work. The performance marker setting is an important aspect in this regard that can be addressed in the future to support various projects for accident prediction, to compare the results and rank various techniques. Benchmark datasets are also needed that can help the research community in testing new algorithms for accident prediction using simulations. Driver identification and profiling of the drivers is an interesting domain that can also be further looked into. From the production point of view, solutions for accident prediction and unsafe driving pattern identification also need to take into account the financial aspect so that the vehicles with such intelligent features are easily affordable by most individuals.

Appendix

Reference	Research direction	Techniques used	Theory or application	Comparison with other techniques
Veeraraghavan et al. (2005)	Monitoring of the driver activities using camera. Analyzing the images from videos for the detection of safe and unsafe actions based on skin-color segmentation	Unsupervised method—agglomerative clustering, supervised method—Bayesian eigen-image classifier	An application in the area of interior vehicle design, which helps to improve the placement of controls in order to reduce unsafe driving behaviors	No comparison with other techniques is performed. However, the accuracy of the classifier is reported using the test data An accuracy of 95.54% is achieved for safe driving activity, and 73.91% accuracy is reported for the unsafe driving activity

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Dixon et al. (2005)	To develop a system that minimizes the impact of untimely interruptions by providing a physical context to the driving conditions	Gradient-descent approach, GA	The result of the supervised-learning algorithm is a step towards building a system that can identify potentially tough driving conditions	No comparison with other techniques is performed. Results of gradient-descent learning and GA are compared with each other Gradient-descent algorithm predicted with an accuracy of 95%, while the GA has an accuracy of 55%
Zhou et al. (2007)	Discriminative learning approach for fusing multichannel sequential data to detect the unsafe driving patterns	CRF	Application to detect unsafe driving patterns from multi-channel data	A comparison is performed with HMM, and SVM with RBF kernel CRF does not require labeling of all data and uses both labeled and unlabeled data for training. It outperforms the simple discriminative classifier (SVM) and generative model (HMM) with an accuracy of 0.081 based on $P(A U)$
Tawari and Trivedi (2011)	To model the individual driving behavior in order to identify features that may be used in grouping the drivers	ANN using the MLP network and a statistical method based on GMM	Application of GMMs, FNN	Comparison is reported with NN and, statistical method based on the GMM MLP is relatively better but the network has significantly longer period needed for training as compared to the GMM

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Rygula (2009)	Analyze the driver's speed profile using techograph and use the same for identifying the driver style	Intensity of speed profile change graph, techograph	Theoretical and application	No comparison with other techniques. However, a comparison of the speed profile changes is done on the common roads
Tambouratzis et al. (2010)	Analysis is performed on the dataset collected by the Republic of Cyprus Police using combination of PNN's and DT's to investigate the potential of predicting accident severity (light, serious or fatal) from the collected parameters	PNN and DT	Application	A comparison is reported between, ANN, DT, and SVM. Severity prediction accuracy of the proposed methodology is superior to all previous techniques with an accuracy of 95.9307 %
Wang et al. (2010b)	Driving danger level prediction is proposed that uses multiple sensory inputs	HMM, CRF, reinforcement learning	Theoretical	A comparison between HMM, CRF and reinforcement learning is performed where, reinforcement learning outperforms other approaches
Wang et al. (2012)	A generic model has been established with a set of parameters to capture individual driver characteristics	A RLS self-learning algorithm has been developed for determining the model parameters	Application	Comparison of typical adaptive cruise control and self-learning adaptive cruise control in manual braking is performed
Singh and Dongre (2012)	Analysis of the driver profile in done for the prediction of driver suitability for driving. Drivers are categorized into following three categories: fit, unfit and partially fit	PCA, HMM	Application named CPS for mobile device with android operating system	No comparison is reported

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Ali et al. (2013)	Predictive approaches to the problem of roadway departure prevention via automated steering and braking	MPC	Application	A comparison is performed between the braking torque applied by the intervention γ_3 and that by the onboard electronic stability control system in combination with the driver
Xu et al. (2012b)	Investigate the applications of the GP model for real-time crash prediction on freeways	RF for the selection of candidate variables and GP	Application of the GP model for real-time crash prediction on freeways	Comparison is done with BLM. The prediction accuracy of the GP mode was found to be greater than that of the BLM
Akin and Akbas (2010)	To assess accidents that occur at intersections with different underlying reasons attributed to time of occurrence, weather and surface conditions, and user and vehicle characteristics	ANN trained by back propagation	Application	No comparison with other approaches is performed. However, a sensitivity analysis of the design parameters is reported
You et al. (2012)	CarSafe fuses information from both cameras and other embedded sensors on the phone—such as the GPS, accelerometer and gyroscope—to detect and alert the driver about dangerous driving conditions in and outside the car	Image processing	Car Safe App for Android operating system based phones	No comparison is reported
Imkamon et al. (2008)	Detection of unsafe driving patterns using data from three different sensors (Accelerometer, Camera and OBD reader)	KLT algorithm, Fuzzy logic	Application	Comparison is done with the ground truth using questionnaire

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Shaout and Bodenmiller (2011)	Primary objective of this research is to capture, measure, and warn users of unsafe and inefficient driving using data from ECU by OBD-II reader	Direct measurements reading and by setting a threshold value to detect the unsafe and inefficient driving	Android based mobile application	No comparison is reported
Liu (2007)	Statistical analysis of vehicle occupancy rates using the accident data with respect to their geographic, temporal, and vehicle coverage design. Investigation and identification of three potential factors namely, accident severity, driver age, and driver gender	Average vehicle occupancies	Application	A comparison of countywide AVOs, countywide AVOs for different facility and AVOs from the field and from accidents is performed
Ning et al. (2008)	Use of a danger level function (expected negative reward) to alert the user in advance about a dangerous situation	TD learning	Application	Hard and soft label approaches are compared with our TD learning method
Jabon et al. (2011)	An active driver-safety framework that captures both vehicle dynamics and the driver's face. A bottom up approach which uses 22 raw facial features, time and frequency domain statistics to determine the most valuable statistics for accident prediction	DWT, Bayesian nets, decision tables, decision trees, SVMs, regressions, and LogitBoost	Application	No comparison with other techniques is reported. However, the ROC curves to analyze the major accident is given. ROC depicts true versus false positives for the classifiers
Li et al. (2008)	To evaluate the application of SVM models for predicting motor vehicle crashes	SVM models based on statistical learning theory	Application	A comparison with BNN is made

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
				SVM model is faster than neural network models The training of neural networks is usually computationally intensive
Ning et al. (2009)	Detection of unsafe system states based on the analysis of multi-sensor data streams	TD learning	Application	A comparison is performed with logistic regression and general linear regression
Chong (2004)	Modeling the severity of injury resulting from traffic accidents using ANNs and DTs	ANN and DT	Application	No comparison is performed with other approaches. However, the ANN and DTs are compared with each other for accuracy
Xie (2007)	Application of BNN models for predicting motor vehicle crashes. A series of models are estimated using data collected on rural frontage roads in Texas	BNN Models	Application	A comparison is performed with negative Binomial regression model Neural network models perform better than the NB regression model in terms of data prediction
Bundele and Banerjee (2009)	A system to monitor the fatigue/drowsiness/stress level of a driver using physiological parameters	ANN	Application	Comparing NN with different number of layers
Manan (2011)	Analysis of the accident data to determine the location of accident at intersection with the highest rank of accident point weightage and to identify the causes of accidents occurred	MLR	Theoretical	No comparison reported

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Guo et al. (2009)	Monitoring and analyzing the fatigue and attention state of driver by using the features of face orientation	Face orientation	Application and theoretical	No comparison reported
Gong and Yang (2011)	Pattern recognition of drivers' behavior before accidents	Fuzzy logic based on multiple regression theory, multi-objective decision theory	Application	No comparison reported
Takatori and Hasegawa (2004)	Influence of prediction methods in the driving assistance system	Linear Prediction	Application	Comparison is done via influence of the system prediction time on the average accident interval
Ma et al. (2012)	Detection of unsafe driving states is presented. The detection is based on the multi sensor approaches, including gyrometer, accelerometer, radar and videos	Unsupervised learning algorithm to perform the unsafe states detection	Application	No comparison reported
Damousis et al. (2007)	Development of physiological algorithms for real-time, unobtrusive, sleepiness-related prediction for time critical operations, such as driving within sensation	GA, Fuzzy expert system	Application	Comparison is reported with a EOG-based sleep prediction algorithm and the proposed approach provides more than 90 % prediction accuracy
Yuejing et al. (2010)	Establishing a functional relationship between accident forms and influencing factors. Provides basis for the screening of safety degree and targeted reasonable reconstruction of the intersection	ANN	Theoretical	No comparison reported

Reference	Research direction	Techniques used	Theory or Application	Comparison with other techniques
Hu et al. (2004)	Probabilistic model for predicting traffic accidents using three-dimensional model-based vehicle tracking is proposed	Fuzzy self-organizing neural network algorithm	Application	No comparison with other approaches is reported. However, a comparison between various structures of NNs is listed
Rujun and Xiuqing (2010)	Study on traffic accident prediction models. The RBF neural network model used to predict and extrapolate the number of fatalities	RBF Neural Network model	Application	RBF NN has a simple structure, concise training, quick convergence study speed and also have a strong advantage in the approximation ability, classification and speed over BP network. Predictive values of the network are closer to the actual one
Jeong et al. (2004)	Using local adaptive threshold and local probability for detecting driving region and the regions where driving is possible	Adaptive threshold method and local-probability	Application	A comparison based on the ability of extension between prevailed contour extension and local probability is reported
Lv et al. (2009)	Identification of traffic conditions leading to traffic accidents based on the data collected from software simulator	SNM	Theoretical	No comparison reported
Quintero et al. (2012)	Modeling of driving behaviors for identification of different types of drivers, and identify high risk areas on the roads	ANN	Application	No comparison reported
Murphey et al. (2009)	Developing a new algorithm for the classification driving style by analyzing the jerk profile of the driver	DS Classification	Theoretical	Comparison of the proposed approach is reported with an acceleration based algorithm

References

- Abdel-Aty M, Hassan H, Ahmed M, Al-Ghamdi A (2012) Real-time prediction of visibility-related crashes. *Transp Res C* 24:288–298. doi:[10.1016/j.trc.2012.04.001](https://doi.org/10.1016/j.trc.2012.04.001)
- Abdel-Aty M, Radwan AE (2000) Modeling traffic accident occurrence and involvement. *Accid Anal Prev* 32(5):633–642
- Abdel-Aty M, Rajashekar P (2006) Calibrating a real-time traffic crash prediction model using archived weather and its traffic data. *IEEE Trans Intell Transp Syst* 7(2):167–174
- Abdel-Aty M, Uddin N, Abdalla F, Pande A, Hsia L (2004) Predicting freeway crashes from loop detector data using matched-case-control logistic regression. *Transp Res Rec* 1897(1):88–95
- Abdel-Aty M, Uddin N, Pande A (2005) Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways. *Transp Res Rec* 1908(1):51–58
- Abu-Lebdeh G, Chen H, Ghanim M (2014) Improving performance of genetic algorithms for transportation systems: case of parallel genetic algorithms. *J Infrastruct Syst.* doi:[10.1061/\(ASCE\)IS.1943-555X.0000206](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000206)
- Akin D, Akbas B (2010) A neural network (NN) model to predict intersection crashes based upon driver vehicle and roadway surface characteristic. *Sci Res Essays* 5(19):2837–2847
- Ali GA, Bakheit CS (2011) Comparative analysis and prediction of traffic accidents in Sudan using artificial neural networks and statistical methods. In: SATC 2011
- Ali M, Falcone P, Olsson C, Sjöberg J (2013) Predictive prevention of loss of vehicle control for roadway departure avoidance. *IEEE Trans Intell Transp Syst* 14(1):56–68. doi:[10.1109/TITS.2012.2206584](https://doi.org/10.1109/TITS.2012.2206584)
- Andersson AK, Chapman L (2011) The impact of climate change on winter road maintenance and traffic accidents in West Midlands. *Accid Anal Prev* 43(1):284–289
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Bruns A et al (2005) EEG- and context-based cognitive-state classifications lead to improved cognitive performance while driving. In: Schmorow DD (ed) *Foundations of augmented cognition, Proceedings of the 2005 human-computer interaction conference*. CRC Press, Boca Raton, p 1065
- Bunde MM, Banerjee R (2009) Detection of fatigue of vehicular driver using skin conductance and oximetry pulse: a neural network approach. In: *Proceedings of the 11th international conference on information integration and web-based applications and services*. ACM, pp 739–744
- Burges CJC (1998) A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov* 2(2):121–167. doi:[10.1023/A:1009715923555](https://doi.org/10.1023/A:1009715923555)
- Canale M, Stefano M (2002) Analysis and classification of human driving behaviour in an urban environment. *Cogn Technol Work* 4(3):197–206
- Chong MM (2004) Traffic accident analysis using decision trees and neural networks. In: *IADIS International conference on applied computing, Portugal*. IADIS Press, Pedro
- Chen BY, Lam WH, Sumalee A, Li ZL (2012) Reliable shortest path finding in stochastic networks with spatial correlated link travel times. *Int J Geogr Inf Sci* 26(2):365–386
- Damoussi I, Cester I, Nikolaou S, Tzovaras D (2007) Physiological indicators based sleep onset prediction for the avoidance of driving accidents. Presented at the Engineering in Medicine and Biology Society, 2007. EMBS 29th annual international conference of the IEEE, Lyon, pp 6699–6704. doi:[10.1109/IEMBS.2007.4353898](https://doi.org/10.1109/IEMBS.2007.4353898)
- Dixon KR, Lippitt CE, Forsythe JC (2005) Supervised machine learning for modelling human recognition of vehicle-driving situations. In: *IEEE/RJS international conference on intelligent robots and systems (IROS)*, pp 604–609
- Doherty JP, Dayan P, Friston K, Critchley H, Raymond JD (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38(2):329–337. doi:[10.1016/S0896-6273\(03\)00169-7](https://doi.org/10.1016/S0896-6273(03)00169-7)
- Ellison AB, Greaves SP, Daniels R (2012) Profiling drivers' risky behavior towards all road users. In: *Australasian College of Road Safety Conference Sydney*, pp 461–472
- Farah H, Polus A, Cohen M (2007) Multivariate analyses for infrastructure-based crash-prediction models for rural highways. *Road Transp Res J Aust N Z Res Pract* 16(4):26
- García CE, Prett DM, Morari M (1989) Model predictive control: theory and practice a survey. *Automatica* 25(3):335–348. doi:[10.1016/0005-1098\(89\)90002-2](https://doi.org/10.1016/0005-1098(89)90002-2)
- Gong J, Yang W (2011) Driver pre-accident operation mode study based on vehicle-vehicle traffic accidents. Presented at the international conference on electric information and control engineering (ICEICE), pp 1357–1361. doi:[10.1109/ICEICE.2011.5777792](https://doi.org/10.1109/ICEICE.2011.5777792)
- Göhring D, Latotzky D, Wang M, Rojas R (2013) Semi-autonomous car control using brain computer interfaces. In: *Intelligent autonomous systems, vol 12*. Springer, Berlin, pp 393–408
- Guo K, Yu G, Li Z (2009) An new algorithm for analyzing driver's attention state. Presented at the IEEE intelligent vehicles symposium, pp 21–23. doi:[10.1109/IVS.2009.5164246](https://doi.org/10.1109/IVS.2009.5164246)

- Halim Z, Kalsoom R, Baig AR (2016) Profiling drivers based on driver dependent vehicle driving features. *Appl Intell*. doi:[10.1007/s10489-015-0722-6](https://doi.org/10.1007/s10489-015-0722-6)
- Halim Z, Baig AR, Zafar K (2014) Evolutionary search in the space of rules for creation of new two-player board games. *Int J Artif Intell Tools* 23(2):1–26
- Halim Z, Baig R, Bashir S (2006) Sonification: a novel approach towards data mining. In: International conference on emerging technologies ICET'06. IEEE, pp 548–553
- Hu G, Baker T, Baker SP (2011) Comparing road traffic mortality rates from police-reported data and death registration data in China. *Bull World Health Organ* 89(1):41–45
- Hu W, Xiao X, Xie D, Tan T, Maybank S (2004) Traffic accident prediction using 3-D model-based vehicle tracking. *IEEE Trans Veh Technol* 53(3):677–694. doi:[10.1109/TVT.2004.825772](https://doi.org/10.1109/TVT.2004.825772)
- Imkamon T, Saensom P, Tangamchit P, Pongpaibool P (2008) Detection of hazardous driving behavior using fuzzy logic. In: Proceedings of ECTI-CON, pp 657–660
- Islam M, Rahman HA (2014) Customer feedback-based healthcare facilities' identification using location based computing technology. In: Global humanitarian technology conference-South Asia Satellite (GHTC-SAS). IEEE, pp 189–194
- Jabon ME, Bailenson JN, Pontikakis E, Takayama L, Nass C (2011) Facial expression analysis for predicting unsafe driving behavior car driving simulator. *IEEE Pervasive Comput* 10(4):84–95. doi:[10.1109/MPRV.2010.46](https://doi.org/10.1109/MPRV.2010.46)
- Jeong P, Nedvtschi S, Daniliuc M (2004) Local probability based safe region detection for autonomous driving. In: IEEE intelligent vehicles symposium, pp 744–749. doi:[10.1109/IVS.2004.1336477](https://doi.org/10.1109/IVS.2004.1336477)
- Ji Q, Zhu Z, Lan P (2004) Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Trans Veh Technol* 53(4):1052–1068
- Johnson DA, Trivedi MM (2011) Driving style recognition using a smartphone as a sensor platform. In: International IEEE conference on intelligent transportation systems (ITSC), pp 1609–1615
- Juárez F, Gayet C (2014) Transitions to adulthood in developing countries. *Annu Rev Sociol* 40:521–538
- Kalsoom R, Halim Z (2013) Clustering the driving features based on data streams. In: 16th international multi topic conference, INMIC13. IEEE, pp 89–94
- Koza R (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge
- Koza JR (1998) Genetic programming on the programming of computers by means of natural selection. The MIT Press, Cambridge
- Lafferty J, McCallum A, Pereira F (2001) Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of 18th international conference on machine learning, pp 282–289
- Liu K (2007) Estimation and prediction of average vehicle occupancies using traffic accident records. FIU Electronic Theses and Dissertations, Florida International University
- Li X, Lord D, Zhang Y, Xie Y (2008) Predicting motor vehicle crashes using support vector machine models. *Accid Anal Prev* 40(4):1611–1618. doi:[10.1016/j.aap.2008.04.010](https://doi.org/10.1016/j.aap.2008.04.010)
- Lv Y, Tang S, Zhao H, Li S (2009) Real-time highway accident prediction based on support vector machines. In: Control and decision conference, CCDC '09, pp 4403–4407. doi:[10.1109/CCDC.2009.5192409](https://doi.org/10.1109/CCDC.2009.5192409)
- Ly MV, Martin S, Trivedi MM (2013) Driver classification and driving style recognition using inertial sensors. In: IEEE intelligent vehicles symposium, pp 1040–1045
- Ma J, Yan X, Chu D, He Y (2012) Online driving states monitoring using fusion of multi-sensor information. *J Theor Appl Inf Technol* 46(1):274–283
- Manan WNBW (2011) Accident prediction model at un-signalized intersections using multiple regression method. Faculty of Civil and Environmental Engineering University Tun Hussein, Malaysia
- Maaten L, Welling M, Saul LK (2011) Hidden-unit conditional random fields. *Int Conf Artif Intell Stat* 15(479):488
- Mathkour HI (2011) A GPS-based mobile dynamic service locator system. *Appl Comput Inform* 9(2):95–106
- Mitas RA (2007) The computing unit for tachometer dataanalysis by means of driving characteristics. In: Proceedings of the 7th international conference on transportation and logistics integrated systems. Cracow 5(3):260–267
- Miyajima C, Nishiwaki Y, Ozawa K, Wakita T, Itou K, Takeda K, Itakura F (2007) Driver modeling based on driving behavior and its evaluation in driver identification. *Proc IEEE* 95(2):427–437
- Miyajima C et al (2006) Cepstral analysis of driving behavioral signals for driver identification. In: Proceedings of IEEE international conference on acoustics, speech and signal processing (ICASSP), vol 5
- Mohan BM (2011) Fuzzy PID control via modified Takagi–Sugeno rules. *Intell Autom Soft Comput* 17(2):165–174
- Moghaddam R, Afandizadeh S, Ziyadi M (2010) Prediction of accident severity using artificial neural networks. *Int J Civ Eng* 9(1):41–48

- Murphey YL et al (2009) Driver's style classification using jerk analysis. In: IEEE workshop on computational intelligence in vehicles and vehicular systems, pp 23–28. doi:[10.1109/CIVVS.2009.4938719](https://doi.org/10.1109/CIVVS.2009.4938719)
- Ning H, Xu X, Zhou Y, Gong Y, Huang TS (2009) A general framework to detect unsafe system states from multisensor data stream. *IEEE Trans Intell Transp* 11(1):4–15. doi:[10.1109/TITS.2009.2026446](https://doi.org/10.1109/TITS.2009.2026446)
- Ning H, Xu X, Zhou Y, Gong Y, Huang TS (2008) Temporal difference learning to detect unsafe system states. In: Proceedings of the international conference on pattern recognition, pp 1–4. doi:[10.1109/ICPR.2008.4761237](https://doi.org/10.1109/ICPR.2008.4761237)
- Orazio TD, Leo M, Spagnolo P, Guaragnella C (2004) A neural system for eye detection in a driver vigilance application. In: The 7th international IEEE conference on, pp 320–325
- Peschel JM, Murphy RR (2013) On the human–machine interaction of unmanned aerial system mission specialists. *IEEE Trans Hum Mach Syst* 43(1):53–62
- Qu X, Wang W, Wang W, Liu P (2012) Real-time prediction of freeway rear-ends crash potential by support vector machine. In: Annual meeting transportation research board, Washington, DC
- Quintero M, Lopez JO, Pinilla ACC (2012) Driver behavior classification model based on an intelligent driving diagnosis system. Presented at the 15th international IEEE conference on intelligent transportation systems, Anchorage, Alaska, USA, pp 894–899. doi:[10.1109/ITSC.2012.6338727](https://doi.org/10.1109/ITSC.2012.6338727)
- Rabiner L (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286. doi:[10.1109/5.18626](https://doi.org/10.1109/5.18626)
- Rujun Y, Xiuqing L (2010) Study on traffic accidents prediction model based on RBF neural network. Presented at the 2nd international conference on information engineering and computer science (ICIECS), Wuhan, pp 1–4. doi:[10.1109/ICIECS.2010.5678126](https://doi.org/10.1109/ICIECS.2010.5678126)
- Rygula A (2009) Driving style identification method based on speed graph analysis. In: International conference on biometrics and kansei engineering, pp 76–79. doi:[10.1109/ICBAKE.2009.51](https://doi.org/10.1109/ICBAKE.2009.51)
- Rygula A, Mitas A (2007) Numeric tolls for tachogram analysis. *Transp Probl Gliw* 2(4):73–81
- Shaout AK, Bodenmiller AE (2011) A mobile application for monitoring inefficient and unsafe driving behavior. In: The international Arab conference on information technology (ACIT), pp 1–8
- Shi B, Xu L, Hu J, Tang Y, Jiang H, Meng W, Liu H (2015) Evaluating driving styles by normalizing driving behavior based on personalized driver modeling. *IEEE Trans Syst Man Cybern Syst* 45(12):1502–1508
- Singh GR, Dongre SS (2012) Crash prediction system for mobile device on android by using data stream mining techniques. In: Sixth Asia modeling symposium, pp 185–190. doi:[10.1109/AMS.2012.16](https://doi.org/10.1109/AMS.2012.16)
- Suriyawongpaisal P, Kanchanasut S (2003) Road traffic injuries in Thailand: trends, selected underlying determinants and status of intervention. *Inj Control Saf Promot* 10(1–2):95–104
- Takatori Y, Hasegawa T (2004) Influence of accident prediction methods in the driving assistance system on safety. In: The 7th international IEEE conference on, pp 396–401. doi:[10.1109/ITSC.2004.1398931](https://doi.org/10.1109/ITSC.2004.1398931)
- Tambouratzis T, Souliou D, Chalikiaris M, Gregoriades A (2010) Combining probabilistic neural networks and decision trees for maximally accurate and efficient accident prediction. In: International joint conference on neural networks (IJCNN), pp 1–8. doi:[10.1109/IJCNN.2010.5596610](https://doi.org/10.1109/IJCNN.2010.5596610)
- Tawari A, Trivedi MM (2011) Audio visual cues in driver affect characterization: issues and challenges in developing robust approaches. Presented at the proceedings of international joint conference on neural networks, California, USA
- Veeraraghavan H et al (2005) Driver activity monitoring through supervised and unsupervised learning. In: IEEE conference on intelligent transportation systems. Vienna, Austria, pp 580–585
- Wahab A, Quek C, Keong T, Takeda K (2009) Driving profile modeling and recognition based on soft computing approach. *IEEE Trans Neural Netw* 20(4):563–582
- Wang J, Xu W, Gong Y (2010) Real-time driving danger-level prediction. *Eng Appl Artif Intell* 23(8):1247–1254
- Wang J, Zhang L, Zhang D, Li K (2012) An adaptive longitudinal driving assistance system based on driver characteristics. *IEEE Trans Intell Transp* 14(1):1–12. doi:[10.1109/TITS.2012.2205143](https://doi.org/10.1109/TITS.2012.2205143)
- Wang J, Zhu S, Gong Y (2010) Driving safety monitoring using semisupervised learning on time series data. *IEEE Trans Intell Transp Syst* 11(3):728–737. doi:[10.1109/TITS.2010.2050200](https://doi.org/10.1109/TITS.2010.2050200)
- Wu JD, Ye SH (2009a) Driver identification using finger-vein patterns with Radon transform and neural network. *Expert Syst Appl* 36(3):5793–5799
- Wu JD, Ye SH (2009b) Driver identification based on voice signal using continuous wavelet transform and artificial neural network techniques. *Expert Syst Appl* 36(2):1061–1069
- Xie Y (2007) Predicting motor vehicle collisions using Bayesian neural network models: an empirical analysis. Zachry Department of Civil Engineering, Texas A&M University
- Xu C, Liu P, Wang W, Li Z (2012a) Evaluation of the impacts of traffic states on crash risks on freeways. *Accid Anal Prev* 47:162–171
- Xu C, Wang W, Liu P (2012) A genetic programming model for real-time crash prediction on freeways. *IEEE Trans Intell Transp* 14(2):1–13. doi:[10.1109/TITS.2012.2226240](https://doi.org/10.1109/TITS.2012.2226240)

- Yang ZQ (2012) Highway traffic accident prediction based on SVR trained by genetic algorithm. *Adv Mater Res* 433:5886–5889
- Yoon H, Park CS, Kim JS, Baek JG (2013) Algorithm learning based neural network integrating feature selection and classification. *Expert Syst Appl* 40(1):231–241
- You CW et al. (2012) CarSafe: a driver safety App that detects dangerous driving behavior using dual-cameras on smartphones. Presented at the ACM 978-1-4503-1224-0/12/09
- Yu R, Abdel-Aty M (2013) Utilizing support vector machine in real-time crash risk evaluation. *Accid Anal Prev* 51:252–259. doi:[10.1016/j.aap.2012.11.027](https://doi.org/10.1016/j.aap.2012.11.027)
- Yuejing L, Jie L, Ming L, Xing-lin Z, Haixia Z (2010) Research on accident prediction of intersection and identification method of prominent accident form based on back propagation neural network. In: International conference on computer application and system modeling (ICCASM), pp V1–V434
- Zhao Y, Karypis G (2005) Prediction of contact maps using support vector machines. *Int J Artif Intell Tools* 14(05):849–865
- Zheng Z, Ahna S, Monsere C (2010) Impact of traffic oscillations on freeway crash occurrences. *Accid Anal Prev* 42(2):626–636
- Zhou Y, Xu W, Ning H, Gong Y, Huang TS (2007) Detecting unsafe driving patterns using discriminative learning. In: IEEE international conference on multimedia and expo ICME, pp 1431–1434. doi:[10.1109/ICME.2007.4284929](https://doi.org/10.1109/ICME.2007.4284929)