Engineering Applications of Artificial Intelligence 87 (2020) 103312

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review☆

Zouhair Elamrani Abou Elassad[a,*], Hajar Mousannif[a], Hassan Al Moatassime[b], Aimad Karkouch[b]

[a] LISI Laboratory, Computer Science Department, FSSM, Cadi Ayyad University, Morocco
[b] OSER Research Team, Computer Science Department, FSTG, Cadi Ayyad University, Morocco

ARTICLE INFO

Keywords:
Machine learning
Driving behavior analysis
Conceptual framework
Systematic literature review

ABSTRACT

Driving Behavior (DB) is a complex concept describing how the driver operates the vehicle in the context of the driving scene and surrounding environment. Recently, DB assessment has become an emerging topic of great importance. However, in view of to the stochastic nature of driving, measuring and modeling, DB continues to be a challenging topic today. As such, this paper argues that to move forward in understanding the individual and organizational mechanisms influencing DB, a conceptual framework is outlined whereby DB is viewed in terms of different dimensions established within the Driver–Vehicle–Environment (DVE) system. Moreover, DB assessment has been approached by various machine learning (ML) models. Still, there has been no attempt to analyze the empirical evidence on ML models in a systematic way, furthermore, ML based DB models often face problems and raise questions that must be resolved. This article presents a systematic literature review (SLR) of the DB investigation concept; In the first phase, a framework for conceptualizing a holistic approach of the different facets in DB analysis is presented, as well as a scheme to guide the future development and implementation of DB assessment strategies. In the second phase, an overview of the literature on ML is designed, revealing a premier and unbiased survey of the existing empirical research of ML techniques that have been applied to DB analysis. The results of this study identify an interpretive framework incorporating multiple dimensions influencing the driver's conduct, in an attempt to achieve a thorough understanding of the DB concept within the DVE system in which the drivers operate. Additionally, 82 primary studies published during the last decade and eight broadly used ML models were identified. The findings of this review prove the performance capability of the ML techniques for assessing DB. The models using the ML techniques outperform other conventional approaches. However, the application of ML models in DB analysis is still limited and more effort is needed to obtain well-formed and generalizable results. To this end, and based on the outcomes obtained in this work, future guidelines have been provided to practitioners and researchers to grasp the major contributions and challenges in the state-of-the-art research.

1. Introduction

The phenomenon concerning road traffic crashes has become a major concern worldwide. According to the global status report on road safety conducted by the World Health Organization (WHO) in 2015, 1.25 million traffic-related fatalities occur annually worldwide, with millions more sustaining serious injuries and living with long-term adverse health consequences; road traffic injuries are currently estimated to be the leading cause of death among young people, and the main cause of death among those aged 15–29 years. ("WHO | Global status report on road safety 2015", 2015). Road safety perception cannot be detached from the analysis of the driver behavior (DB) as the major part of traffic accidents is caused by human factors as it was inferred that they took part in the manifestation of 95% of all accidents (Evans, 1991, 1996). For these purposes, analyzing DB can aid assessing driver performance, enhance traffic safety and, furthermore, endorse the development of intelligent and resilient transportation systems.

Human driving behavior is a complex concept that, in general terms, delineate how the driver manipulates the vehicle in the context of the driving scene and surrounding environment (Martinez et al., 2017). In recent years, DB has become a topic of interest among the public and researchers; it is generally considered to be one of the most important factors in crash occurrence, yet due to the stochastic nature of driving, measuring and modeling, driving behavior continues to be a challenging topic today (Sagberg et al., 2015). For instance,

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.engappai.2019.103312.

* Corresponding author.
E-mail address: z.elamrani@uca.ma (Z. Elamrani Abou Elassad).

https://doi.org/10.1016/j.engappai.2019.103312
Received 4 March 2019; Received in revised form 25 August 2019; Accepted 18 October 2019
Available online 25 October 2019

many researchers have studied the effects of driver's physiological and psychological characteristics on DB from different standpoints (Ariën et al., 2013; Faure et al., 2016; Dahlen and White, 2006; Li et al., 2017a; Jacobé de Naurois et al., 2017a). On the other hand, taking driver's operational maneuvers as the research objective, some scholars explored various behaviors such as lane change, acceleration and turning events (Ghasemzadeh and Ahmed, 2018; Kim et al., 2017; Paefgen, Kehr, Zhai, & Michahelles, 2012); Within this context, significant correlations were found between driving performance and the driver's profile (e.g. gender, age and experience amongst others).

While the number of traffic accidents and their outcomes, mainly human injuries and fatalities emphasize the prominence of investigating the factors which contribute to their occurrence; principal factors in accidents include human, vehicle, and environment factors. From this perspective, DB has been considered a direct outcome of the impulse encountered from the road infrastructure, surrounding environment and atmosphere inside the vehicle (Waard, 1996). As such, in order to conduct a more reliable analysis of the DB, diverse types of parameters have been taken into account; The Driver–Vehicle–Environment (DVE) is based on the concept of the "joint" cognitive system, where the dynamic interactions between driver, vehicle and environment are represented in a harmonized and integrated way (Hollnagel, 2005). We have adopted an approach with the purpose of investigating a set of parameters needed for estimation of the DB, consequently, various DVE characteristics have been considered in an effort to analyze DB to improve driver safety and reduce accident rates. Although it is generally assumed that the characterization of the DB is related to the aforementioned triptych: the DVE philosophy, there has been only limited investigation focusing on the details of this relationship and how the DVE characteristics influence the evaluation of the DB. However, perhaps most importantly, there is still a lack of a common underlying conceptual framework to guide this research and clearly distinguish the relationship between the different dimensions of the DB and the DVE model, as well as a commonly accepted scheme to instruct future implementations of DB assessment strategies.

As the DB field evolved, researchers studied the use of algorithms from Machine Learning (ML), an area of artificial intelligence (AI) that has been studied since the late 1950s (Martens, 1959); numerous ML-based models were proposed to recognize or predict the DB in an attempt to improve the safety and comfort of drivers, as well as other road users. ML addresses the question of how to build a computer system that improves automatically through experience (Jordan and Mitchell, 2015). ML techniques have been characterized by (i) their autonomously surmounting major non-linear problems using data sets from multiple sources; and (ii) their ability to easily incorporate newly data in an attempt to improve estimation performance (Begg and Kamruzzaman, 2005; Chlingaryan et al., 2018). They are being used in the context of studying DB to provide users with better recommendations and help constructing powerful ML models. However, the ML discipline does not have a definite classification scheme for its algorithms, mostly due to the number of paradigms and the uncertainties introduced in the literature (Lv and Tang, 2011). Subsequently, it becomes difficult and confusing to choose an ML algorithm that fits one's need when developing a DB computational model.

Since the volume of research in the field of DB is expanding constantly, researchers may find it challenging to track the use and the trends of ML algorithms in analyzing DB. In addition, it is becoming more and more difficult to evaluate critically and to provide an overview of the related empirical studies. One way to assist academics and practitioners in using ML techniques in the assessment of the DB is to lay the foundations for grasping the major contributions and shortcomings in the state-of-the-art research. Existing literature reviews of DB investigation can be divided into two categories: traditional literature reviews (TLR) and systematic literature reviews (SLR). The traditional literature reviews (Halim et al., 2016b; Meiring and Myburgh, 2015); mainly cover the state-of-the-art and contemporary beliefs, whereas more methodologically explicit approaches (Vilac et al., 2017) aim to answer various research questions pertaining to the addressed problem. To the best of authors' knowledge, there is no existing SLR that focuses on ML models for DB analysis, which motivates our work in this paper. Further to this, little to no research has directly constructed a thorough treatment of the DB concept accounting for the variety of its dimensions within the DVE structure. This article provides a systematic literature review (SLR) to identify the applications of ML techniques in DB assessment domain. In the first phase, A framework for conceptualizing DB is outlined, which illustrates a holistic approach of the different facets in perceiving DB as well as a scheme to instruct future development and implementation of the DB assessment strategies. It is expected that more information can be obtained about the investigation of the DB in the context of the DVE approach, and make better implementation or research decisions. In the second phase, an overview of the literature on ML is designed, revealing a premier and unbiased survey of the existing empirical research of ML techniques that have been applied to DB evaluation. Specifically, we performed an SLR on ML models published during the last decade (2009–2019). We further provide future guidelines to DB analysis practitioners and researchers regarding the application of the ML techniques in the field.

The work presented on this paper is an extension of our conference paper originally presented at the International Conference on Advanced Intelligent Systems for Sustainable Development - AI2SD' 2018 (Elamrani Abou Elassad and Mousannif, 2019) which proposed unconventional taxonomies based on the nature of the conducted study, measurement patterns and supervision motives underlying the assessment models of driving behavior. Additional material has been included in order to create a more in-depth research paper. In comparison to our original work, progress has been made regarding the following aspects:

(i) propose a theoretical framework to conceptualize driving behavior within the broader organizational context, providing an overview of the mechanisms thought to be involved;
(ii) summarize and clarify the available evidence regarding the ML techniques for constructing DB estimation models;
(iii) identify performance accuracy and capability of ML techniques for constructing DB assessment models in reference to the DB dimensions;
(iv) identify the modeling metrics adopted in the analysis of each of the DB dimensions;
(v) illustrate comparisons of estimation accuracy between ML models and non-ML models;
(vi) illustrate comparisons of estimation accuracy between different ML models;
(vii) summarize the strength and weakness of the ML techniques;
(viii) assist academics to position new research activity in this domain appropriately.

The rest of this paper is organized as follows: Section 2 summarizes the features of related work. Section 3 the DB conceptual framework is presented, followed by Section 4 which presents our results, while Section 5 serves the conclusion and discusses research implications, limitations and suggestions for future research.

2. Related work

In the past few years, substantive research has been conducted regarding the DB examination. Conventionally a driver-centric approach has been applied to understand the precise individual features that contribute to the increased risk experienced by drivers, with a particular focus on their DB. Level in the physiological state of the driver, Bergasa and Nuevo (2005), Karkouch et al. (2018), Lee and Chung (2012) and Lee et al. (2015) developed a flexible alertness monitoring system to capture driver physiological condition and driving aptitude under real-time situations. Different physiological metrics have been collected to detect stress responses or predict its upcoming values

(Ge et al., 2014; Kaiseler et al., 2015; Munoz-Organero and Corcoba-Magana, 2017). Other characteristics such as motion sickness (Domeyer et al., 2013) and different psychological disorders namely Autism Spectrum Disorders (Reimer et al., 2013), Attention-Deficit/Hyperactivity Disorder (Zheng et al., 2014) and Parkinson's Disease (Ranchet et al., 2013) were assessed to elucidate their impact on driving safety. Drivers' psychology is a widespread and perplexing topic as it is considered be a key to an effective perception and right judgments making; road safety practitioners have begun to consider the nature and breadth of the driver's psychosocial characteristics including the personality traits to better comprehend the risky behavior which takes part in the crash involvement and offenses. That is, the relationship between anxiety and fear has been developed substantially in an attempt to provide interesting implications for how they differentially impact DB (Taylor, 2018). On the other hand, a research of Lu et al. (2013) explores how and why anger and fear influence driving risk perception; the findings highlight the necessity to differentiate anger and fear in road safety management. Another study that explores the self-reported risky DB of the young novice in the frame of their sensation seeking , reward sensitivity, depression, and anxiety in a longitudinal methodology have been performed; the research does not, however, examine why the novice who is depressed and/or anxious drives in a different manner (Scott-parker et al., 2013).

Studying DB also means analyzing driving profiles such as age, gender and other demographic traits. Much attention has been paid to older drivers in research on DB and safety particularly in the countries that knows an increased aging of population. Relevant scholars aimed to better understand the contributions of readily measurable age-related factors that may provide a basis for developing evaluations and interventions in order to mitigate driving impairments (Blanchard et al., 2010; Dawson et al., 2010; Lucidi et al., 2014; Molnar et al., 2018). Conversely, an overarching emphasis has been placed on investigating young drivers' risky behavior (Gheorghiu et al., 2015; Zicat et al., 2018). Another aspect relatively frequently studied in driving profiles is the impact of texting while driving; A meta-analysis was conducted to synthesize the effects of text messaging on driving performance and to provide convergent evidence that such attitude compromises the safety of the driver, passengers and other road users (Caird et al., 2014). Various demographics and personality traits have been examined in Ben-ari and Yehiel (2012) adopting a multi-dimensional approach to integrate sociodemographic and motivational factors in order to provide a fuller picture of DB. Several efforts have also been made to examine the driver's operational decisions that could result in unintended risky behaviors such as errors in vehicle handling or in traffic maneuvering. A driver foot gesture modeling and prediction framework was proposed based on vision sensors; despite the fact that embedded vehicle sensor parameters from the Controller Area Network (CAN-bus) like brake or acceleration pedal states tell us something about the foot behaviors, the foot movement before and after a pedal press detected from vision-based sensors can provide valuable information for better semantic understanding of the driver states and to predict when a pedal is pressed before it actually happens (Tran et al., 2012). A substantial amount of research has been dedicated to evaluate the driver behavior in dilemma zones as they become very skeptical and may take improper decision might lead to right angle collision or rear end collision (Lavrenz et al., 2014; Pathivada and Perumal, 2017).

Research consistently demonstrates the contributions provided by the traditional driver-centric approach in improving road safety that emerges from the driver and their immediate driving environment. Within this context, the work presented by Amditis et al. (2010) investigated the formulations of the DVE three parameters, namely the driver, the vehicle and the environment model in order to design an inserted driver–vehicle interface with the purpose of increasing the effectiveness and the safety gains of advanced driver assistance systems and to reduce the level of workload and omission inflicted by in-vehicle information systems and nomad devices. Many estimation methods have been proposed in the driver behavior domain based on the DVE measures comprising the driver's characteristics (voice activity, head movement, gazing direction, etc.), the vehicle parameters (throttle position, steering angle, etc.) and the environment features (rainy weather, road type, etc.). With the growing data quantity and the boosted computational power, ML modeling is more efficient and accurate than ever before.

Driver characteristics such as head, eye and hand cues were leveraged to detect the driver's activity state using Support Vector Machine (SVM) classifier to evaluate the driver's performance in on-road settings (Ohn-bar et al., 2014). The drowsiness recognition research has examined a variation of methods including: Artificial Neural Networks (Correa et al., 2013; Jacobé de Naurois et al., 2017a), Logistic regression (Murata, 2016) and Dynamic Bayesian Networks (McDonald et al., 2018). Drivers' gaze behavior was measured prior to and during the execution of different driving maneuvers performed in a dynamic driving simulator using (ANNs), Bayesian Networks (BNs), and Naive Bayes Classifiers (NBCs) (Lethaus et al., 2013). Driving events like acceleration/deceleration, breaking, lane change and turning amongst others have been assessed using different ML models combining various sets of DVE parameters. Wang et al. (2010) evaluated three types of sequential supervised learning algorithms for the aforementioned events, the Hidden Markov model (HMM), the Conditional Random Field (CRF), and the Reinforcement Learning (RL). There are also studies where the driver behavior profiling was examined based on smartphones; Tchankue et al. (2013) discussed how the driving context with the distraction level of the driver can be predicted using mobile phone sensors, many ML algorithms have been applied to determine which one is the most effective.

With the many studies already done on DB and notwithstanding the importance of these findings and their contribution to the literature, what appears to be lacking is an effort to present a fuller and more holistic picture of the harnessed ML techniques, and the measures accounted for DB evaluation. We retain that this kind of work is the most essential to be explored as it outlines a conceptual framework incorporating multiple dimensions influencing the driver's conduct, in an attempt to achieve a comprehensive understanding of the DB within the DVE system in which the drivers operate.

3. A driving behavior conceptual framework

Theoretical or conceptual frameworks are essential contributions in grasping insight about the mechanisms influencing behavior, and in the development of interventions to enhance safety outcomes (Newnam and Watson, 2011). In order to structure the present review, there is a need for a more precise conceptualization of the DB analysis within the broader organizational context pertaining to the different influencing factors; therefore, a theoretical framework is outlined to guide this research and clearly distinguish the relationship between the different dimensions of the DB and the DVE model. Accordingly, practitioners and researchers will be better directed in the future development of intervention strategies designed to improve safety outcomes with regard to the driving conduct. Based on this initial analysis, the following research questions are detailed (see Table 1).

3.1. Driving behavior concept

In spite of the agreement on the prominence of traffic-specific human factors in understanding the complex concept of DB, there is little agreement on their conceptualization due to the influencing factors which complicates its description (Martinez et al., 2017; Taubman-ben ari et al., 2004; Yanagihara et al., 2015). However, the present section along with the DVE triptych (Section 3.2) are believed to explain a large portion of the concept and provide elucidation regarding how the DB should best be modeled and measured.

Table 1
DB conceptual framework research questions.

ID	Research question	Motivation
RQ1	What dimensions best convey the concept of the DB within the DVE system?	Identify the components delineating the DB along with the influencing factors with reference to the DVE system.
RQ2	What are the DVE data measures that underpin the empirical studies investigating the DB?	Develop driving data-collection taxonomies based on the dominant DVE measures thought to be involved.
RQ3	How can a grounded interpretive framework be outlined in an attempt to conceptualize driving behavior analysis within the broader organizational context?	Propose a theoretical framework whereby the DB is viewed in a holistic perspective in terms of dimensions and influencing aspects.

Relying on whether driving data is context relevant, DB can be deemed as contextual or non-contextual. Analyzing DB based on the vehicle kinematics such as speed and acceleration/deceleration or the driver's state namely the physiological and psychological conditions could be acknowledged as non-contextual DB since it does not consider the surrounding environment of the driver that may contribute to crash risk such as road and weather conditions. Research adopting non-contextual DB approach has its limitation. For instance, if a driver depicts stop-and-go or abrupt accelerate/decelerate behavior, without additional information on traffic conditions, it is usually difficult to tell if this is simply due to the driver's behavior, or caused by heavy traffic conditions. Conversely, context-aware models can improve the driver's conduct by augmenting the awareness of the vehicle's state (e.g. headway distance), the environment (e.g. weather conditions) and the physiological and psychological status of the driver (e.g. available alertness level) (Rakotonirainy, 2005).

Following Elander et al. (1993) and Taubman-ben ari et al. (2004), DB is composed of two separate components, driving skills and driving styles. Driving skills concern attitudes and characters of the driver, therefore constitute a driver's maximum level of performance on elements of the driving task. Driving styles, on the other hand, referring to the style in which individuals choose to drive or habitually drive. Although driving styles and driving skills are argued to be independent and to have different mechanism of development, they sometimes interact with each other (Zhang et al., 2018). While drivers with better vehicle maneuvering skills have been reported to be more likely to adopt a careless driving style by involving in risky and distracting activities (Pöysti et al., 2005); it has been depicted that driving styles interact with driving skills by influencing their acquisition as well as their actual execution (Catherine et al., 2016).

3.2. The driver–vehicle–environment model

The ostensible driver, vehicle and environment model (i.e. DVE model) is based on the concept of the "joint" cognitive system, where the dynamic interactions between the driver, vehicle and environment are represented in a harmonized and integrated way (Hollnagel, 2005); such a model can be seen as a closed loop system in which the perception of a given driving scenario and its impact on the driver and other road users is observed. Martinez et al. (2017) and Waard (1996) presented DB as a direct consequence of the stimuli received from the road infrastructure, surrounding environment and atmosphere inside the vehicle; therefore, a realistic illustration of the DB must take into considerations the interactions of the driver within the DVE model. In our framework, information about the characteristics of the vehicle (e.g. velocity, acceleration, action on steering wheel, etc.), the driving environment (e.g. weather and road conditions, etc.) and the driver's properties (e.g. HRV, head movement and voice activity, etc.) are captured. As such, the DB is established as a transference system between the states of the vehicle operated by the driver and the surrounding environment. On this basis, we make an attempt to capture most of the common elements in the DVE system in an effort to describe and analyze its facets' modeling to efficiently investigate DB.

3.2.1. Driver module

In the DVE system, the driver basically performs a set of actions on the vehicle commands and controls and makes decisions based on the perceived environment. Of all the factors influencing traffic safety, the driver is a main factor and the vehicle–road–environment influences traffic through drivers (Cai and Lin, 2011). As reported by Sayer et al. (2018) and Wen et al. (2011), the driver state can be psychological or physiological which can be correlated to environmental conditions. Cacciabue and Carsten (2010) selected five major categories of factors influencing driver capability, performance and behavior within the DVE system:

- Experience/competence (EXP): knowledge or aptitudes aggregation resulting from direct participation in the driving process;
- Attitudes/personality (ATT): a complex mental state involving beliefs and feelings and values and dispositions to act in certain ways;
- Task Demand (TD): the demands of the process of achieving a specific and measurable goal using a prescribed method;
- Driver State (DS): driver physical and mental ability to drive (fatigue, sleepiness.);
- Situation Awareness/Alertness (SA): perception of the features in the environment within a volume of time and space along with the apprehension of their significance and the projection of their status in the near futures.

The driver's module has been developed taking into account some of the aforementioned parameters (directly or indirectly), and focused on a certain number of basic components affecting variability in driver's conduct and performance identified and listed as follows:

Physiological State: Refers to the condition or state of a body or bodily function of a driver when driving the vehicle such as fatigue, stress, position amongst others (Sayer et al., 2018). The physiological status can be detected or predicted by supervising driver's positions and actions; for instance, eyes activity, facial expressions, head nodding, body sagging posture and others are often reasoned to be vital signs of the driver's perception and concentration. Lanata et al. (2015) have presented meaningful results on the relationship between the physiological changes in drivers and their driving behaviors, such as steering-wheel angle corrections, velocity changes, and time responses under incremental stress conditions. Moreover, physiological characteristics are deemed to be the essential material that underlies the psychological factors, which are ultimately reflected in driver's psychological load (Wang et al., 2018b). According to international and national research, driver's physiological and psychological characteristics are closely related to traffic safety, and psychological characteristics that arise from physiological characteristics are important factors influencing DB (Scott-Parker et al., 2015).

Psychological State: Refers to the mental or emotional state of the driver such as being happy, angry, sad, distracted and so on (Sayer et al., 2018). The driver's psychological situation can be known by monitoring the heart rate, breathing rate and other physiological parameters and when it changes drastically, the driver cannot control his thinking activity and the aptitude to judge (Zhang et al., 2014). The different psychological properties have been characterized as driving tendencies or propensities; these latter refer to the car driver's psychological experience of actual traffic conditions and the preference of decision-making under dynamic influence of various factors, they also depict the emotional state of automobile operators in time-varying dynamic environment (Wang et al., 2018a). To capture the affective factors which are used to describe a state in which people deeply involved in an activity that nothing else seems to matter (e.g. speeding behavior), the psychological flow theory was outlined (Chen and Chen, 2011). It is a complex concept and researchers often measure it through multiple dimensions, such as perceived enjoyment, concentration, perceived control, and curiosity (Ghani and Deshpande, 1994; Koufaris, 2008; Lu et al., 2009; Moon and Kim, 2001).

Driver profile: According to the present study, driver profile is affiliated to the driver's demographic characteristics (e.g. gender, age, education, income) and driving history (e.g. years on the road, record of violations). Accumulate research (Aberg and Rimmo, 1998; Blockey and Hartley, 1995; Parker et al., 2000; Reason et al., 1990) reported that men and young drivers have a tendency to commit violations repeatedly comparing to women and older drivers, and the more time spent on the road by drivers the more they violate traffic regulations. Conversely, female and older drivers committed more errors than male and young drivers. Increased annual mileage has been proved to increase crash risk (Davey et al., 2007; Lourens et al., 1999). Furthermore, educational level and marital status have been proved to imperil indirectly driving performance and driving safety (Mehdizadeh et al., 2018).

Driving events: Pertain to fundamental driving operations performed by the driver while controlling the vehicle and leading to changes in the car motion/state (e.g. turning, tailgating, flashing headlights). Undesirable events frequency (EF) is a useful safety surrogate as such events were found to be related to crash involvement and driver safety (Musicant et al., 2010). For instance, over 2.5 million rear-end collisions are reported every year, making them the most common type of automobile crash (42%) (Center for Disease Control, 2009). Also, it has been reported, based on traffic police data for 2010, that in 85% of cases, accidents are caused by violations of traffic rules by vehicle drivers; from them, 25% of road accidents are caused by the non-observation of speed regulations by drivers on roads; 15%, by violations of the rules of passage through a crossroad; 10%, by moving into an oncoming traffic lane (Prokhorov and Shmakov, 2013). It is important to detect these typical driving events as they are fundamental to the evaluation of driver behavior and it would be highly beneficial to many application domains in the road safety perspective such as an automated advanced warning system (Saiprasert et al., 2017).

3.2.2. Vehicle module

The driver of a vehicle generally executes a variety of tasks comprising those that directly involve driving and those that do not. Typically, the primary task is to steer the vehicle and to avoid potential road hazards. Secondary tasks refer to those in-vehicle functions that require from the driver a separate attention to perform. These latter include visual-related tasks, such as using vehicle displays or looking for items, auditory-related tasks like listening to radio or a passenger, and haptic-related tasks namely eating/drinking or using a cell phone. Naturalistic driving studies have found that both complex visual-haptic tasks and looking away from the road for more than two seconds in a six-second window increase crash and near-crash risk (Klauer et al., 2006; Olson et al., 2009). Moreover, Ranney et al. (2001) highlighted that visual, auditory and haptic distractions can have individual or joint effects on drivers' performance. Secondary tasks have been considered the reason for 30% of the traffic collisions caused by driver distraction (Nevile and Haddington, 2010); they will affect the driver's response time to danger, braking time and will lead to higher acceleration (Nowosielski et al., 2018). Thus, further investigation of the synergy between the driver and the vehicle can offer fundamental guidance for a safer DB, and as Koo et al. (2015) stated, "By running controlled studies of driver response to vehicle interactions, we are better able to predict user behavior".

3.2.3. Environment module

The environmental conditions have been identified to hold a substantial impact on the assessment of the DB (Derbel and Landry, 2017; Hamdar et al., 2016). Most scholars investigating the environmental effects are directly assessing performance and are pursuing appropriate metrics which represents the adjustments attached to imposed behavior (Hancock, 1997). With regard the environment context, many characteristics influence the performance of the vehicle and the driver, thus, the complexity of the module raises rapidly with the number of factors considered. These characteristics cover the detailed information relating to road geometry (e.g. curvature, gradient, lane width, shoulder width), road condition (e.g. dry, wet, snowy, icy), road type (e.g. asphalt, concrete, gravel, earth), weather condition (e.g. clear, rain, fog/snow/severe wind), light condition (e.g. daylight, night time) and traffic condition (traffic lights, traffic flow, traffic disturbance). Road geometry is often correlated to risky driving; Kantowitz and Simsek (2001) considered that odd road geometry contributes to driver workload, while road condition was proven to cause nervous and cautious behavior (Hamdar et al., 2016). As for road type, even though some scholars concluded that the pavement nature indicator becomes statistically insignificant comparing to different safety attributes (Anastasopoulos et al., 2008); other authors found that road type factors are strongly related to the accidents rate (Lee et al., 2008). Empirical evidence suggests that the likelihood of unsafe conduct increases during abnormal weather conditions (Hamdar et al., 2016) and driving during midnight and rush hours (Zhang et al., 2016). On the account of the above studies, it is demonstrated that different driving environments being uncontrollable, continuously impose different changes to the DB of an individual which involves transformations in the DVE system.

3.3. Modeling metrics

Numerous metrics are used for estimating DB which can be assorted into intrusive and non-intrusive categories as presented in our original work (Elamrani Abou Elassad & Mousannif, in press). Defining these variables could be a sensitive process since they affect the robustness of the ML model. Still, there is no common agreement of a suggested set of metrics in the literature (Ericsson, 2000; Taubman-ben ari et al., 2004). Vehicle-based, behavioral, physiological and subjective metrics are the most common groups of measures used in assessing DB described as follows:

Vehicle-based measures: mainly provide an analysis of driving performance by examining the capabilities of the driver while operating the vehicle, they include vehicle speed, acceleration, steering wheel movement, lane position deviation, gear changes and others (Hori et al., 2016; Kim et al., 2017). These metrics are easy to obtain as they only require the vehicle to be equipped with necessary sensors, yet they are strongly dependent on the type of vehicle, the conduct capability of drivers and surrounding environment (Chen et al., 2017a; Sahayadhas et al., 2012). The vehicle-based measures have been categorized into three major classes summarizing some of the recent findings regarding their potential relationship with driver's rendering level (Aghaei et al., 2016): (i) driver input to the vehicle (e.g. steering, braking), (ii) vehicle response to driver input (e.g. Velocity/acceleration, jerk) and (iii) vehicle state relative to the environment (e.g. Headway distance, time to lane change); the first two categories can be directly measured by sensors mounted inside the vehicle, whereas the third category requires information regarding the driving environment. These metrics have the advantage to be real-time, continuous, non-intrusive and reliable (Li et al., 2017b). However, they manifest only after the driver state changes which is often too late to prevent an accident.

Physiological-based measures: record the driver's physiological state during the drive and considered being a fundamental material that underlies the psychological factors reflected in driver's psychological status as described earlier. Physiological signals such as Electrocardiogram (ECG), electroencephalogram (EEG), galvanic skin response

(GSR), Electrooculogram (EOG), respiration, electromyogram (EMG) and Heart Rate Variability and others, display relatively high recognition accuracy and offer a reliable insight of the operator functional state (Healey and Picard, 2005; Jacobé de Naurois et al., 2017a; Jinjun et al., 2010; Nilsson et al., 2018; Vicente et al., 2016). Some of these metrics require sensors and cables attached on the body, which constrains the behavior of drivers in some extent, despite that, due to miniaturization of the measurement equipment, many physiological measures can be collected continuously and relatively unobtrusively (Waard, 1996). Albeit the physiological-based measures are largely used by practitioners, they often increase the complexity of the experimental setup and require costly and particular sensors to capture the signals. Additionally, environmental factors or motion artifact noises could extremely affect the sensor readings which would imply the use of sophisticated signal conditioning and filtering operations.

Behavioral-based measures: mostly identified by image and audio processing techniques. They examine the drivers' attitude and visual alertness in order to detect behaviors such as head movements, facial expression, voice activity and so on (Hori et al., 2016; Jacobé de Naurois et al., 2017a; Lee and Chung, 2012). These metrics are not only technically hard to get as they demand more processing steps, but also because the required computer methods are computationally expensive. The driver behavioral measures are non-intrusive, real-time and fast at detecting the onset of abnormal driver states. It is perceived that image processing techniques by watching driver's facial expressions have reached quite a high level of sophistication, especially used of deep learning techniques (Schmidhuber, 2015). Nevertheless, behavioral-based measures have been found to return unreliable results because their detection capabilities are strongly affected by variations in environmental factors and driving conditions (Chen et al., 2017a). Therefore, they cannot always return reliable results especially in the changes of lighting conditions inside or outside the vehicle during the day/night and while wearing glasses. In addition, driver facial signs are susceptible to false interpretation when the driver's perception is influenced by certain self-contained aspects during the process of driving.

Subjective-based measures: In which drivers themselves delineate their own assessment of DB using questionnaires and self-reports, which are easy to use in most work environments. Some of the frequently used examples are the Karolinska Sleepiness Scale (KSS) (Akerstedt and Gillberg, 1990), the NASA Task Load Index (NASA-TLX) (Hart and Staveland, 1988), the Manchester Driving Behavior Questionnaire (DBQ) (Reason et al., 1990), the Aggressive Driving Behavior Scale (Houston et al., 2003) and others which have been widely utilized and validated in various application domains. Subjective-based measures have been deemed to perform well in simulated environments (Mittal et al., 2016). They are also provided directly by the driver, assembled before or after the task is done, and are relatively inexpensive to obtain. Although much useful information can be collected by the use of subjective-based measures, these metrics have their limitations; there have been some concerns with this type of measurements as they may be subjects to errors of recall and reporting like social desirability bias (Hatfield et al., 2017). Furthermore, subjective measures are missing the potential to record in real-time and to continuously monitor the DB in order detect sudden changes.

3.4. Framework design

Theoretical development in the field of DB has been difficult to pin down. As previously mentioned, there is only a little agreement on the DB analytical scheme which complicates its description. Sagberg et al. (2015) presented a tentative framework to distinguish the concept of driving style from other constructs and to discuss key terms and definitions commonly used in this research area. However, it was inferred that further research is clearly needed to understand more precisely how the influencing factors shape driving style and how they may interact. Sarma et al. (2013) considered that the majority of studies that have examined attitudes towards driving have done so through the lens of the Theory of Planned Behavior (TPB; Ajzen, 1985), which is a theoretical framework offering a model of human behavior that integrates beliefs, attitudes and intentions. Other notable research within this scope have predominantly adopted theoretical frameworks based on the tenets of social-cognitive theory, which emphasize self-efficacy and psychometric experience as prerequisites for adaptive behavior change (Vetter et al., 2018; Windsor and Anstey, 2006). In terms of ML models, Morris et al. (2011) addressed the problem of perceiving driver's intention to change lanes seconds before it occurs using a real-time intent detection framework. To build a consistent stochastic model of driver's interactions with the environment, the road network and other traffic participants, Gindele et al. (2015) designed a probabilistic framework for context-sensitive state estimation and prediction describing driver's behaviors and plans based on unlabeled traffic observations. Bahram et al. (2016) examined the motion intention of highway drivers regarding the modeled interaction with the driver's surrounding road users in a traffic scene, combined with a supervised maneuver-based classifier. The developed framework achieves a robust prediction against possibly improper model assumptions in the case of unusual style of driving.

Although all of these research papers provide a significant contribution to furthering our understanding of the DB concept, there is still a lack of broad overviews based on relevant theoretical frameworks. In particular those that conceptualize DB within the broader organizational context, providing an overview of the mechanisms thought to be involved. The structure of the proposed framework is illustrated in Fig. 1.

Table 2
ML based DB models research questions.

ID	Research question	Motivation
RQ4	Which ML techniques have been used for DB dimensions' analysis?	Identify the ML techniques that have been used to estimate each of the driving behavior dimensions. Practitioners can take the identified ML techniques as candidate solutions in their practice.
RQ5	What is the overall estimation accuracy of ML models in reference to DB dimensions?	Identify estimation accuracy of ML models as it is considered to be the primary performance metric for ML models.
RQ6	Do ML models outperform non-ML models?	Illustrate performance comparison between ML models and non-ML models to verify which ones are superior.
RQ7	Are there any ML models that distinctly outperform other ML models?	Identify the ML models with relatively excellent performance.
RQ8	What are the strengths and weaknesses of different of ML models?	Identify the strengths and weaknesses of different ML models. With fully grasping the properties of the candidate ML models, researchers can make a sound conclusion on choosing the ML models that favor the defined estimation contexts.
RQ9	What are the existing related challenges and open research areas?	Highlight the opportunities and the raised concerns with regard to the driving behavior.

4. Machine learning techniques for driving behavior analysis

Machine Learning models use computers to simulate human learning process in order to improve the performance of specific tasks based on the knowledge identified and acquired from the real world. They operate by building algorithms that are guided by data, rather than

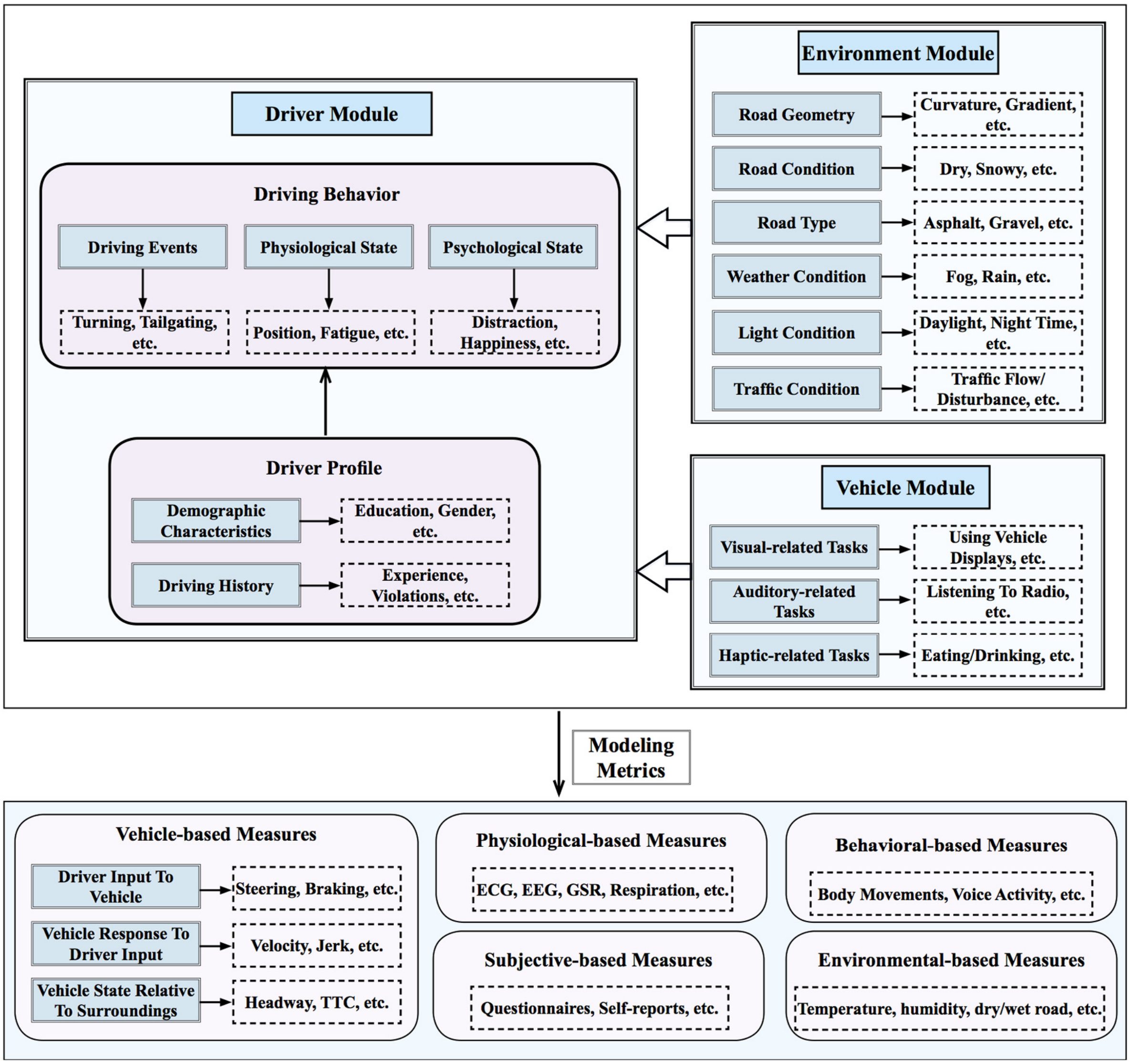


Fig. 1. Conceptual framework for driving behavior analysis within the driver–vehicle–environment system.

relying on human programmers to provide hard-coded instructions and can initially be classified as supervised, semi-supervised, unsupervised, or reinforcement learning. With the major influx of attention devoted to the ML-based analysis of driving behavior and the high activity of progress in this field, remaining knowledgeable of the existing progress can be a considerable work.

4.1. Methodology

In this review, an overview of the literature on ML is designed, revealing a premier and unbiased survey of the existing empirical research of ML techniques that have been applied to DB evaluation. In order to stimulate this work and guarantee its integrity, Table 2 highlights the adopted research questions that are better suited for the remainder of this review.

The steps involved in conducting this SLR include the identification of resources and selection of studies, along with data extraction and synthesis. The improper identification of the reviewed articles might lead to inconsistent and inaccurate conclusions. Therefore, recognizing the relevant keywords plays a pivotal role in identifying the proper resources. Thus, an extensive search using all possible permutations of the ML techniques and the driver behavior keywords which derived from the conceptual framework formerly established were performed within the most prominent scientific digital libraries. The search is filtered to show only the articles that were written in English and published during the last decade in a reliable peer-reviewed journal or conference. In order to select the relevant studies, we excluded the duplicate articles or the articles that were not clearly related to the research questions. Then, the remaining articles were subjected to a quality screening to evaluate their reliability, integrity and pertinence.

We identified 82 studies (Table A.5 in Appendix A) in the field of ML based driving behavior analysis which were published during the time period 2010–2019. Among them, 68 (83%) papers were published in journals, 14 (17%) papers appeared in conference proceedings. The publication venues and distribution of the selected studies along with the journals' impact factors are shown in Table 3.

From the selected studies, we categorized the ML techniques used for DB estimation as follows:

- Neural networks (NN)
- Support vector machines (SVM)
- Fuzzy & Neuro Fuzzy based (NF)
- Clustering (CL)
- Inductive Rule Based (IR)
- Instance Based (IB)
- Decision trees (DT)

Table 3
Publication venues and distribution of selected studies.

Publication venue	Type	Number of studies	Impact factor (2018)
IEEE Transactions on Intelligent Transportation Systems	Journal	12	5.744
Accident Analysis and Prevention	Journal	9	3.058
Transportation Research Part C	Journal	5	5.775
Expert Systems with Applications	Journal	4	4.292
Transportation Research Part F	Journal	3	2.360
Safety Science	Journal	2	3.619
IEEE Transactions on Human–Machine Systems	Journal	2	3.332
Frontiers in Human Neuroscience	Journal	2	2.870
Plos One	Journal	2	2.766
Sensors	Journal	2	3.031
Journal of Safety Research	Journal	1	2.401
Biomedical Signal Processing and Control	Journal	1	2.943
Pattern Recognition Letters	Journal	1	2.810
Medical Engineering & Physics	Journal	1	1.785
Knowledge-Based Systems	Journal	1	5.101
Journal of Transport & Health	Journal	1	2.583
IEEE Transactions on Mobile Computing	Journal	1	3.822
International Journal of Automotive Technology	Journal	1	1.523
IEEE Transactions on Multimedia	Journal	1	5.452
IEEE Sensors Journal	Journal	1	3.076
Cognitive Systems Research	Journal	1	1.384
Advances in Mechanical Engineering	Journal	1	1.024
Journal of Intelligent Transportation Systems	Journal	1	2.568
Science China Technological Sciences	Journal	1	2.180
IEEE Transactions on Vehicular Technology	Journal	1	4.066
Applied Intelligence	Journal	1	2.882
Neural Computing and Applications	Journal	1	4.664
Engineering Applications of Artificial Intelligence	Journal	1	3.526
IEEE Pervasive Computing	Journal	1	3.813
Fuzzy Sets and Systems	Journal	1	2.907
Applied Informatics	Journal	1	–
Advances in Body Area Networks I	Journal	1	–
Journal of Modern Transportation	Journal	1	–
International Journal of Intelligent Transportation Systems Research	Journal	1	–
Genetic and Evolutionary Computing	Journal	1	–
International Journal of Transportation Science and Technology	Journal	1	–
IEEE Intelligent Vehicles Symposium (IV)	Conference	3	
International Federation of Automatic Control (IFAC)	Conference	2	
International Workshop on Connected & Intelligent Mobility (CIM)	Conference	1	
International Conference Image Analysis and Recognition (ICIAR)	Conference	1	
IEEE International Conference Intelligent Systems (IS)	Conference	1	
EURO Working Group on Transportation (EWGT)	Conference	1	
Conference on Human Factors in Computing Systems (CIH)	Conference	1	
International Conference on Intelligent Transportation Systems (ITSC)	Conference	1	
World Conference on Transport Research (WCTR)	Conference	1	
EURO Working Group on Transportation Meeting (EWGT)	Conference	1	
Total		**82**	

- Bayesian learners (BL)
- Ensemble learners (EL)
- Evolutionary algorithms (EA)
- Miscellaneous (Misc)

Overall, SVM, NN, BL and EL are the four most frequently used ones; they together were adopted by 72% of the selected studies, as illustrated in Fig. 2 presenting the amount of research attention that each type of ML technique received during the last decade; in addition to Fig. 2. The identified ML techniques were used to evaluate DB usually in two forms: (1) studies comparing ML techniques (2) studies comparing statistical and ML techniques; Fig. 3 is plotted to further outline the distribution of research interest in each publication year. As can be seen, the activity of publications in this field is growing at an explosive rate. Note that some studies contain more than one ML technique.

As shown in Fig. 3, for one thing, an apparent publication peak is shown around the years 2017 and 2018; Moreover, compared to other ML techniques, SVM, NN, BL and EL seem to have received dominant research attention in many years. It is noteworthy that in neural network studies, both traditional Artificial Neural Networks (ANNs) and Deep Neural Networks (DNNs) have been considered for this review. Deep learning methods can be seen as an improved extension of the classic ANNs consisting of more layers that enable higher degrees of abstraction and enhanced assessment from data. Of all the selected research that adopted NN for DB analysis, about 57% employed DNNs whereas 43% used classical ANNs.

4.2. Performance estimation of ML models

One of the main goals of this SLR is to give insight into the scope of using of ML algorithms in DB evaluation that can pave the way for future researchers and practitioners in their studies. An in-depth examination has been considered to perceive how the algorithms are being applied to every DB dimension by inspecting the performance measures that researchers adopt to describe ML techniques as the reliability of a method needs of its performance assessment in order to validate the stability of the developed ML models and examine their efficiency in practical applications. With respect to validation methods, K-fold Cross-Validation, Holdout and Leave-One-Out Cross-Validation (LOOCV) are the three dominant ones used in the selected studies. Specifically, the numbers of the studies that used these three validation methods are 30 for k-fold Cross-Validation, 31 for Holdout and 18 for LOOCV.

Beside validation methods a number of metrics are used for gauging the performance of ML approaches in DB analysis. These performance

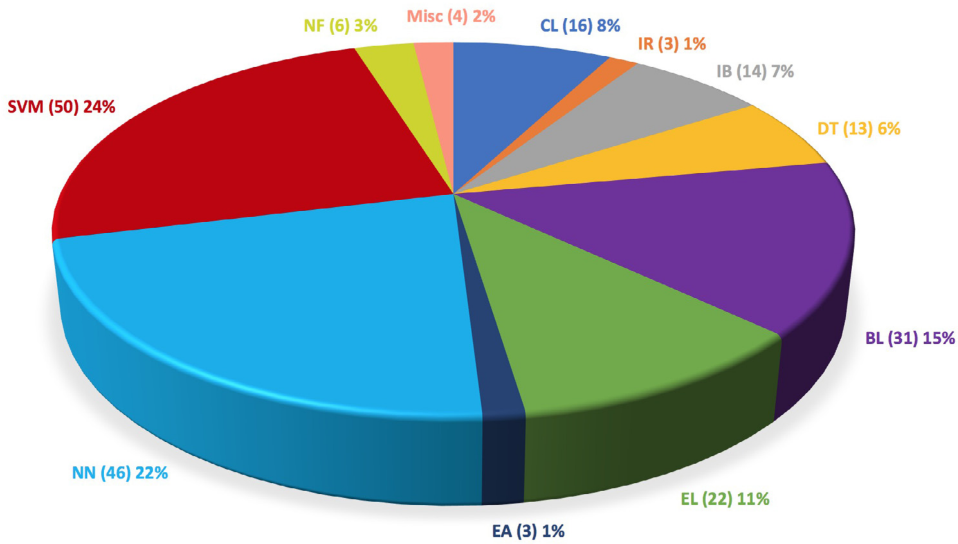


Fig. 2. Distribution of the studies over type of ML technique.

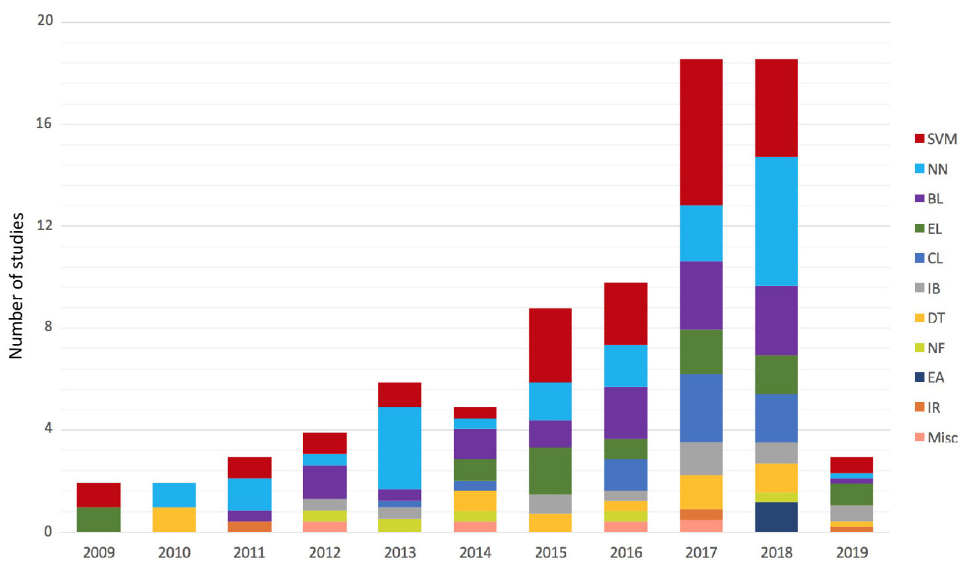


Fig. 3. Distribution of the studies over publication year.

metrics are used for comparing and evaluating models developed using various ML and statistical techniques. Note that an algorithm may use one or more metrics to describe its performance. Thus, it is crucial to adopt appropriate accuracy metrics in evaluating estimation accuracy. Table 4 presents the performance metrics that were employed and the studies in which they are used. An overview of the number and the proportion of studies using each performance metric is illustrated in Fig. 4. According to the figure, it is found that Accuracy, Recall and Specificity, are among the most popular performance metrics considered in the selected studies as 65.85% of them adopted Accuracy, while about 35.36% and 32.92% used Recall and specificity respectively; followed by Precision, F1 measure, FPR, AUC, Confusion Matrix and MSE measures. OOB Error, RMSE, MAPE, Kappa, G-mean are other typically adopted metrics. Some less commonly measures are grouped in the miscellaneous category namely MAE, NMSE, MPE, Youden's Index and Hausdorff Distance.

Developing a more in-depth analysis requires a characterization of ML models on the basis of each of the DB dimensions previously defined (see Fig. 1). Among the three DB domains beforehand, Driving events have been actively investigated when assessing DB using ML techniques, followed by the Physiological and Psychological states. Fig. 5 shows the amount of research attention that each of the DB dimensions received in the selected studies.

The occurrence analysis of the evaluation measures depicts their popularity and presents other metrics to researchers. However, in view of the dominance of Accuracy, Recall and Specificity metrics, we adopted them in this review to evaluate the performance of ML models. The higher values of the selected performance measures indicate more accurate estimate. We have used box plots to gain insight about the distributions of these metrics' values of ML techniques. Considering the small number of observations of the three chosen metrics for some ML categories, we did not plot them in the figures to avoid the insignificant presentation of insufficient points in a box plot. The following subsections provide a comprehensive analysis of the use of ML techniques for each of the DB dimensions.

4.2.1. Driving events performance estimation

On the performance of ML models for Driving events' analysis measured in Recall and Specificity (see Fig. 6a and b), DT and IB are the most accurate one (with median around 0.97). For Recall, NN came

Table 4
Performance metrics.

Performance metric	Studies
Accuracy	Aksjonov et al. (2018), Ba et al. (2017), Barua et al. (2019), Bejani and Ghatee (2018), Bundele and Banerjee (2010), Campilho and Kamel (2014), Chen et al. (2018a,b, 2017a, 2015), Chuang et al. (2015), Correa et al. (2013), Darzi et al. (2018), Deshmukh and Dehzangi (2019), Elhenawy et al. (2015), Gwak et al. (2018), Halim et al. (2016a), Henni et al. (2018), Hong et al. (2014), Hou et al. (2014, 2015), Ihme et al. (2018), Jeon et al. (2017), Khushaba et al. (2013), Kim et al. (2017), Lee et al. (2017), Liang et al. (2017), Liu et al. (2016), Manawadu et al. (2018), Masood et al. (2018), Memory et al. (2011), Min et al. (2017), Minhad et al. (2017), Miyajima et al. (2010), Munoz-Organero and Corcoba-Magana (2017), Osafune et al. (2017), Osman et al. (2019), Qi and Fries (2018), Rodriguez Gonzalez et al. (2014), Scenarios et al. (2018), Tango and Botta (2013), Wang et al. (2018c, 2017b, 2016, 2017a), Xie et al. (2017), Xiong et al. (2018), Xuan et al. (2010), Yang et al. (2018), Yeo et al. (2009), Yu et al. (2017), Yuan et al. (2018), Zhang and Kumada (2018), Zhao et al. (2012) and Zhao et al. (2017)
Recall	Ba et al. (2017), Barua et al. (2019), Bejani and Ghatee (2018), Bundele and Banerjee (2010), Chen et al. (2018b), Chen et al. (2017a), Gwak et al. (2018), Ihme et al. (2018), Khushaba et al. (2013), Li et al. (2016), Liang et al. (2017), Liu et al. (2016), McDonald et al. (2018), Memory et al. (2011), Min et al. (2017), Miyajima et al. (2010), Okamoto et al. (2017), Osafune et al. (2017), Qi and Fries (2018), Ragab et al. (2014), Riccardo et al. (2012), Scenarios et al. (2018), Singh et al. (2013), Sysoev et al. (2017), Tango and Botta (2013), Vlahogianni and Barmpounakis (2017), Wang et al. (2016), Yu et al. (2017) and Zhang and Kumada (2018)
Specificity	Ba et al. (2017), Barua et al. (2019), Bejani and Ghatee (2018), Bundele and Banerjee (2010), Chen et al. (2018b), Chen et al. (2017a), Gwak et al. (2018), Ihme et al. (2018), Khushaba et al. (2013), Liang et al. (2017), Liu et al. (2016), Memory et al. (2011), Min et al. (2017), Miyajima et al. (2010), Okamoto et al. (2017), Osafune et al. (2017), Qi and Fries (2018), Ragab et al. (2014), Riccardo et al. (2012), Scenarios et al. (2018), Singh et al. (2013), Sysoev et al. (2017), Tango and Botta (2013), Vlahogianni and Barmpounakis (2017), Wang et al. (2016), Yu et al. (2017) and Zhang and Kumada (2018)
Precision	Bejani and Ghatee (2018), Gwak et al. (2018), Khushaba et al. (2013), Memory et al. (2011), Okamoto et al. (2017), Qi and Fries (2018), Ragab et al. (2014), Riccardo et al. (2012), Singh et al. (2013), Sysoev et al. (2017), Vlahogianni and Barmpounakis (2017), Yu et al. (2017); Zhang and Kumada (2018), Zhao et al. (2012) and Zhu et al. (2017a)
F-measure	Bejani and Ghatee (2018), Chen et al. (2018a), Ihme et al. (2018), Jeon et al. (2017), Li and Busso (2016), Memory et al. (2011), Okamoto et al. (2017), Qi and Fries (2018), Ragab et al. (2014), Riccardo et al. (2012), Scenarios et al. (2018), Singh et al. (2013), Sysoev et al. (2017) and Vlahogianni and Barmpounakis (2017)
False positive rate	Bejani and Ghatee (2018), Elhenawy et al. (2015), Ihme et al. (2018), Jahangiri et al. (2018), Li and Busso (2016), McDonald et al. (2018), Miyajima et al. (2010), Riccardo et al. (2012), Vlahogianni and Barmpounakis (2017), Wu et al. (2013) and Yu et al. (2017)
AUC (Area under the Curve)	Barua et al. (2019), Bundele and Banerjee (2010), Chen et al. (2018b), Chen et al. (2017a), Jahangiri et al. (2018), Jeon et al. (2017), Li et al. (2013), Liang et al. (2017), McDonald et al. (2018), Riccardo et al. (2012) and Singh et al. (2013)
Confusion Matrix	Barua et al. (2019), Bundele and Banerjee (2010), Chen et al. (2018a), Jeon et al. (2017), Khushaba et al. (2013), Manawadu et al. (2018), Masood et al. (2018), McDonald et al. (2018), Rodriguez Gonzalez et al. (2014), Yang et al. (2017) and Zhao et al. (2012)
MSE	Bejani and Ghatee (2018), Bundele and Banerjee (2010), Halim et al. (2016a), Huang et al. (2018), Jacobé de Naurois et al. (2017b), Li et al. (2013), Singh et al. (2013) and Tango and Botta (2013)
OOB Error	Hou et al. (2015), Jahangiri et al. (2018), Jahangiri et al. (2016) and Wang et al. (2016)
RMSE	Jacobé de Naurois et al. (2018), Kim et al. (2017), Li et al. (2018) and Tang et al. (2018)
MAPE	Kim et al. (2017), Tang et al. (2018), Vlahogianni and Golias (2012) and Wang et al. (2015)
Kappa	Chen et al. (2018b), Jabon et al. (2011) and Zhang and Kumada (2018)
G-mean	Liu et al. (2016), Ragab et al. (2014) and Singh et al. (2013)
Other miscellaneous measures (MAE, NMSE, MPE, Youden's Index, Hausdorff Distance)	Chen et al. (2018a), Kim et al. (2017), Riccardo et al. (2012), Shimosaka et al. (2015) and Tang et al. (2018)

next (with median around 0.92), followed by BL and SVM (with median around 0.88). With regard to Specificity, DT and BL were followed by BL (with median around 0.95) then lastly NN (with median around 0.82). In terms of Accuracy, NN performed the best with value 0.97, IB, BL and SVM came next (with median around 0.91), followed by EL then DT. NF models performed the worst with 0.73 value of accuracy. In addition to the aforementioned observations, according to Fig. 6, the NN, DT models for Specificity along with SVM for recall and DT with BL for Accuracy all have their median nearly in the centers of the boxes, which implies that the values of these models are symmetrically distributed around the medians; conversely, apart from IB and NF for accuracy which indicate a positive skewness, all the others depict lopsided boxplots skewed to the left. Moreover, the Specificity values in general have less variation than Recall and Accuracy, since they have proportionately shorter boxes and tighter ranges of values.

Further, we also aimed to deeply investigate the estimation accuracy for Driving events; we provided in Table B.6 (Appendix B) the detailed statistics of Recall, Specificity and Accuracy for each type of ML model. Note that the outliers identified in Fig. 6 were eliminated before computing the statistics. It can be observed that all the ML models for all the metrics, except for NF, have their means approximately ranging from 0.82 to 0.96. This reflects that the ML techniques have reasonable estimation capability as the values of the selected metrics indicate acceptable levels.

4.2.2. Physiological state performance estimation

As can be seen in Fig. 7, the Physiological state assessment for Recall demonstrates higher values of performance in terms of NN and IB (with median around 0.91), followed by SVM and EL with 0.89 value, while BL was the least performant with 0.76 value. As for specificity, approximately all the four models (SVM, NN, EL and IB) have their medians around 90%. With regard to Accuracy, NN is more accurate with 0.9 value, followed by SVM and EL (with median around 0.85) next IB and CL with 0.82 and 0.79 values respectively, then lastly BL with a median around 0.74. In addition to the aforementioned observations, (EL, BL, IB) for Recall and (SVM, NN, EL) for accuracy all have their median nearly in the centers of the boxes, which implies that the values of these models are symmetrically distributed around the medians. Additionally, once again the Specificity values typically have less variation than Recall and Accuracy, since they have relatively smaller boxes.

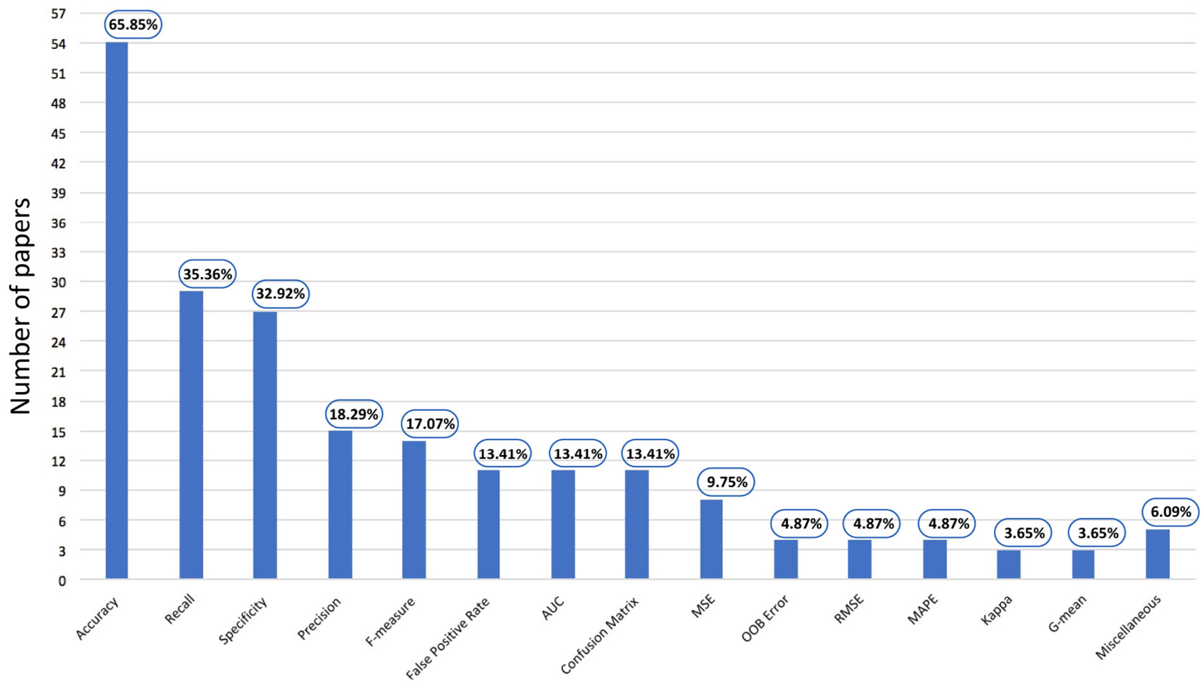


Fig. 4. Distribution of the studies over performance metric.

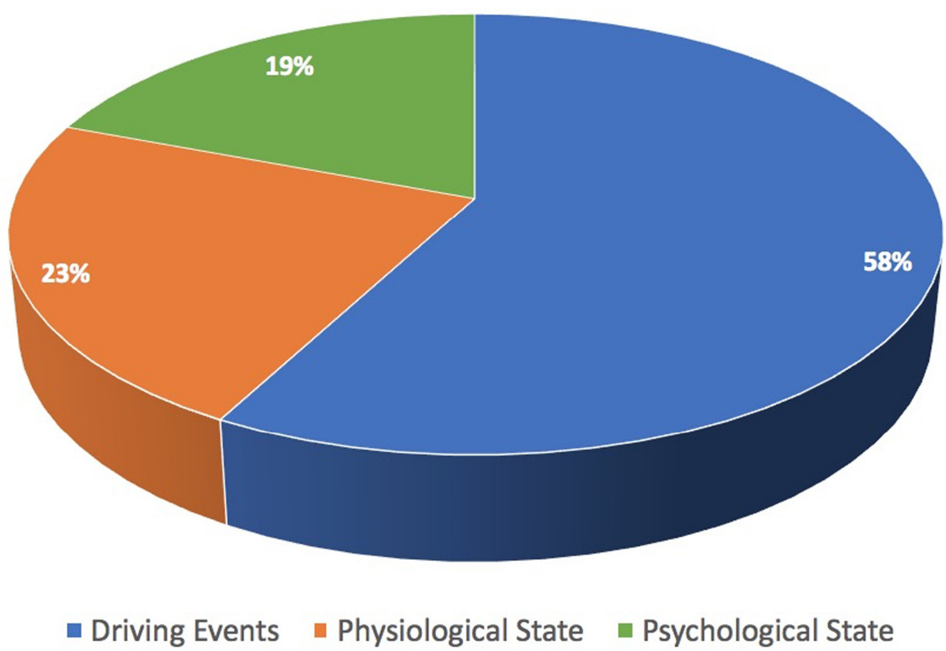


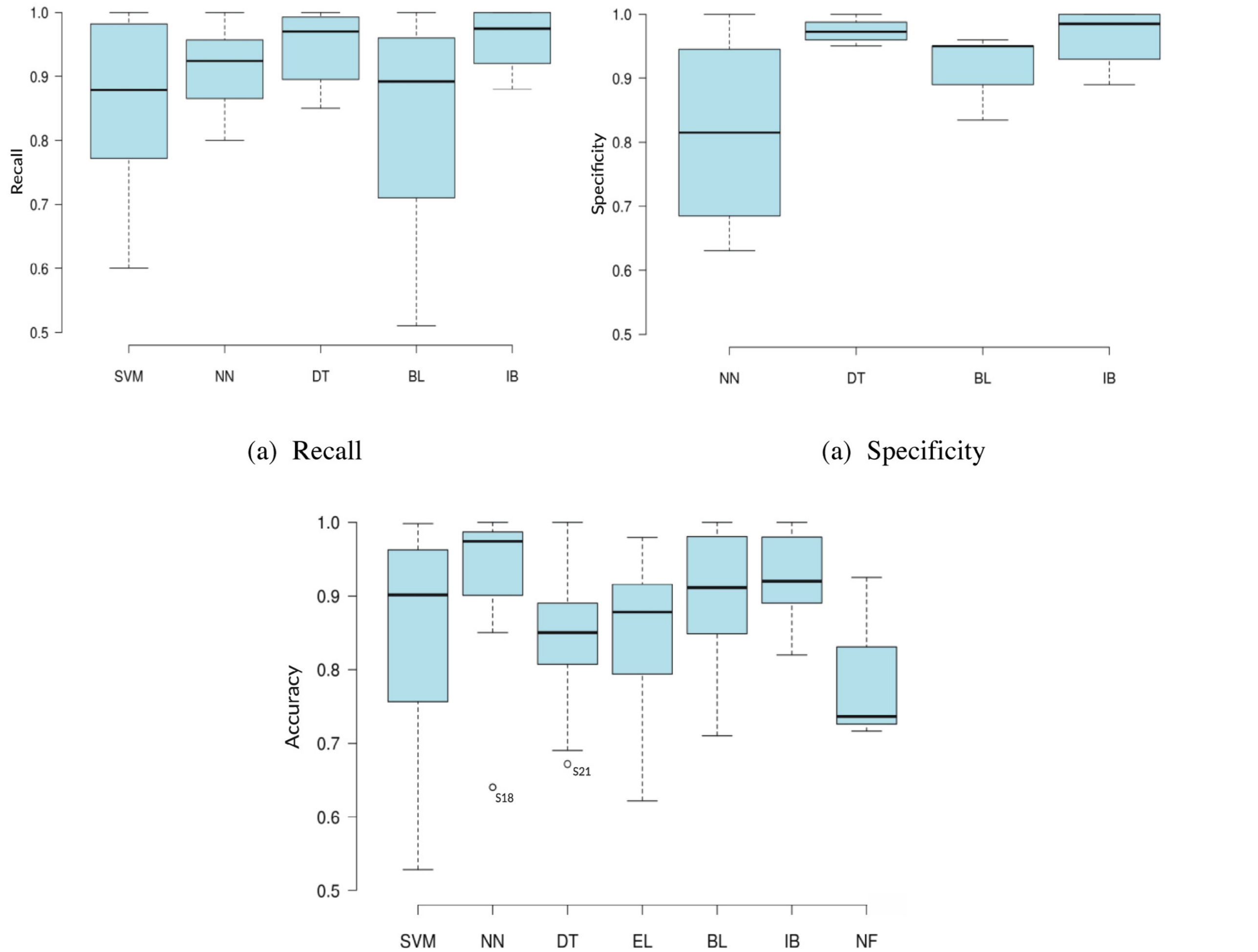
Fig. 5. Distribution of studies over DB dimensions.

More performance statistics for the DB Physiological state can be found in Table B.7. As shown, apart from the BL model that has a mean around 0.76, all the other ML models have their means ranging from 0.8 to 0.91 which indicates an adequate effectiveness in evaluating the Physiological state.

4.2.3. Psychological state performance estimation

As concerns the Psychological state, Fig. 8 depicts that for Recall and Specificity SVM and NN are more efficient (with median around 0.87 for Recall and 0.94 for Specificity), followed by EL with 0.77 and 0.9 values for Recall and Specificity respectively, whereas BL was last for specificity with 0.84 value. In reference to Accuracy, EL achieved the highest performance with value 0.83, followed by NN, DT and SVM (with median around 0.88). BL and IB represented the worst accuracy with 0.66 value. On top of that, only EL for all the three metrics and BL for Accuracy have their median closely to the middle of the boxes, and both Recall with Specificity have less deviation displaying fairly smaller boxes.

In the same way, Table B.8 presents detailed statistics of Recall, Specificity and Accuracy for analyzing psychological-based ML models. As can be seen, SVM, NN and EL for Accuracy and Recall achieved good capability by having their means higher than 0.85, while the rest of results reflect acceptable estimation ability by having their means between 67% and 77%.



(a) Recall

(a) Specificity

(a) Accuracy

Fig. 6. Box plots of Recall, Specificity and Accuracy (outliers are labeled with associated study IDs) for the Driving events.

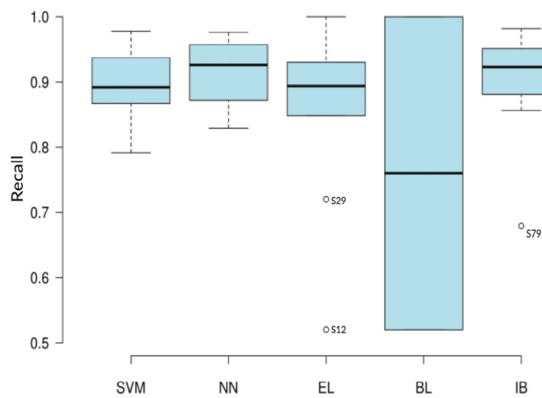
4.2.4. Modeling metrics modalities for DB dimensions

In addition to the detailed ML performance analysis for each of the DB dimensions, here we highlight the characterization of the adopted modeling measures in evaluating these dimensions as we provide a deeper insight into capturing their main underlying patterns. Fig. 9 presents the focus of research activity, based on the DB dimensions, that each type of modeling measures received in the selected studies. First, we can notice that all of the publications covered more than one type of modeling metrics. It can be observed that subjective-based measures are used in different aspects of analyzing DB since they have been known to solely delineate the driver's own assessment, thus they are being reasonably acquired by the three DB dimensions. Clearly, the vast majority of the studies that analyzed the Driving events employed vehicle-based measures in the process. Followed by environmental-based and subjective-based measures. whereas behavioral and physiological input have not been actively adopted in studying Driving events. Indeed, such a behavior could be highly examined by considering the impact of driver inputs to the vehicle and its responses to these actions taking into account the influence of the surroundings. This may be one explanation for endorsing these metrics in evaluating the most investigated DB dimension in the selected publications. Not surprisingly, the physiological-based measures are considerably used in examining the driver's Physiological state. Environmental and vehicle based measures came next, whereas behavioral measures were last. Finally, we have the Psychological state inspection that counted for an acceptable level of these metrics considering that this DB dimension is the least explored in the selected studies. Strangely enough, we can see that environmental-based inputs are not intensively engaged in the examining of the Psychological behavior. This is probably due to the fact that the adoption and use of these metrics in this processing is at its early stages.

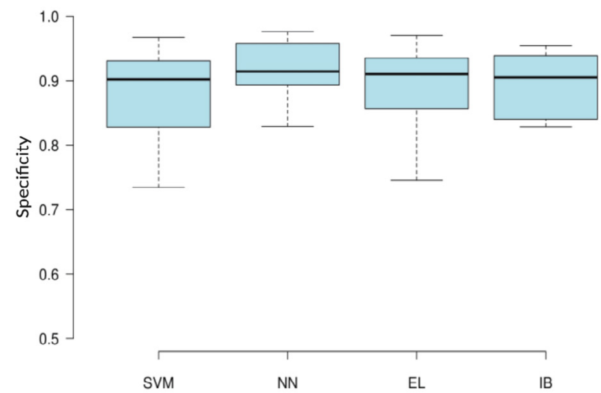
4.3. ML models vs. non-ML models

The ML models have been compared with three conventional non-ML models: Regression model, Discriminant Analysis (DA), AFVD (Gong et al., 2008). One model is alleged to outperform another in an experiment if the Accuracy value of the first model achieves at least 2% improvement comparing to the second one. The authors of S5 and S14 employed error metrics to measure the accuracy of the adopted models and thus we have decided to include them in the analysis. Most of the ML techniques have been compared to the Regression model which is considered to be the most widely used statistical technique in the literature (Malhotra, 2015). The details of the comparisons between ML models and non-ML models are provided in Table C.9 in Appendix C.

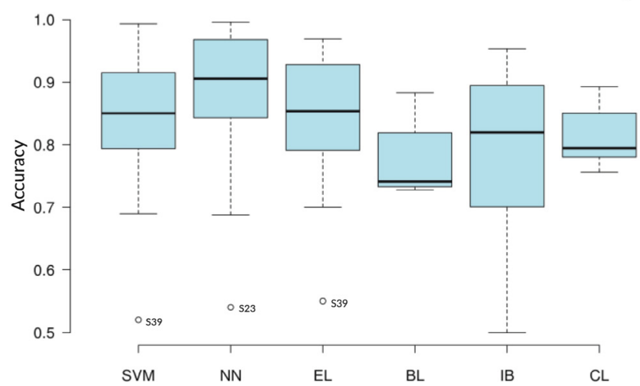
Fig. 10 indicates the overall results of the comparisons between ML models and non-ML models, where all the comparisons were in terms of Accuracy metric. The bars above zero line indicate that ML models are more accurate, whereas the bars below zero line indicate that non-ML models are more accurate. Every bar is comprised of three proportions



(b) Recall



(b) Specificity



(b) Accuracy

Fig. 7. Box plots of Recall, Specificity and Accuracy (outliers are labeled with associated study IDs) for the Physiological state.

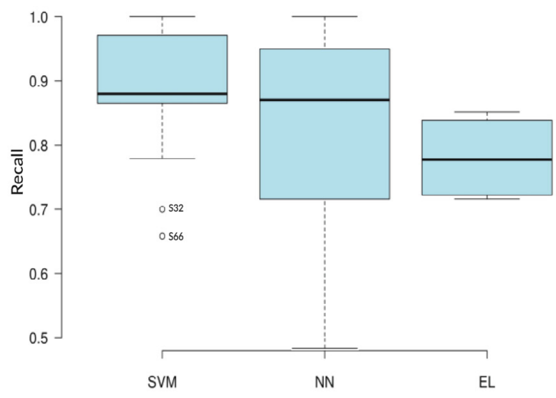
representing the distribution of the three DB dimensions. As can be seen in Fig. 10, the majority of the experiments indicate that ML models outperform non-ML models. The results show that 74% (92 of 124) of experiment results reveal the superiority of ML models whereas only 26% (32 of 79) of experiment results exhibit the superiority of non-ML models (note that some studies carry out more than one trial). Furthermore, the larger part of the studies comparing these two model types depicts the analysis of the Driving events dimension. The comparison results show that NN, SVM and BL outperform the Regression model (Reg) in 96%, 94% and 83% of experiments respectively, and these observations are sustained by a number of experiments, followed by EL in 80% of trials. NN and SVM are more accurate than Regression model in the investigation of all the three DB dimensions while Regression model outperforms BL and EL in the analysis of the Physiological state and Psychological state respectively.

For IB, NF and DT, no more than three evaluations reported the comparisons with Regression model indicating higher accuracy for the IB and NF in terms of assessing the Driving events behavior, as for DT, it is difficult to determine whether it is more accurate than regression model or not, since the number of experiments reporting that DT outperformed Regression model is identical to the ones depicting the opposite results. In addition to Regression model, other non-ML models have also been compared with some ML models; EL, SVM and DT outperformed Discriminant Analysis in the experiments where the most part was the evaluation of Driving events, whilst DA provided better overall results than NN and BL and similar accuracy comparing to IB in both Psychological state and Driving events behaviors, but in smaller number of observations. With regard to the AFVD model, although there are seven experiments in total which include four trails showing that AFVD is less accurate than NN model in examining Driving events, these evaluations actually come from an identical study (S14).

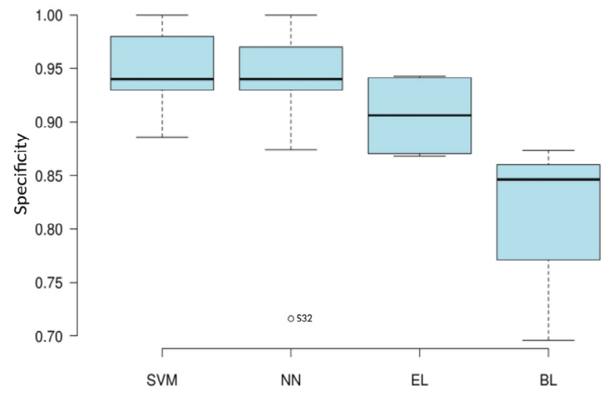
In compliance with the above-mentioned results and analysis, we may conclude that even though non-ML models show a slight upper hand in analyzing the Physiological state over some ML models, in general, the ML models outperform the non-ML. However, there is a threat to the validity to this conclusion as this argument holds only for the three ML models (NN, SVM and BL) vs. Regression and (SVM and EL) vs. DA, which have been supported by a sufficient number of experiments; while it is difficult to fully validate and generalize the conclusions for the other evaluations due the small number of experiments comparing the performance of ML models with the non-ML models. Thus, more number of studies comparing ML models and other models for DB estimation should be conducted in order to procure satisfactory and generalized results.

4.4. ML models vs. ML models

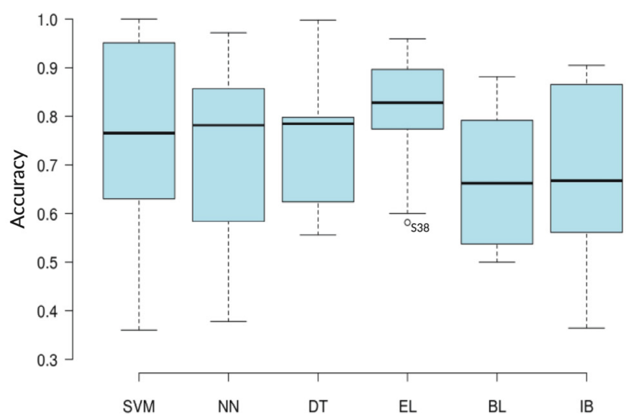
With the aim of comparing the results of ML models with respect to other ML models we adopted the same methodology as the one used in the comparisons between ML models and non-ML models. That is, the comparisons were conducted on the same experiments in terms of Accuracy metric on the basis of each of the DB dimensions. Fig. 11 shows the overall results of the comparisons between different ML models as well as the distribution of the DB dimensions and the corresponding number of supporting experiments; the bars on top of zero line display that models in horizontal axis are more accurate, whereas the bars below zero line indicate that models in horizontal axis are less accurate. On top of that, for every comparison between two models, DB dimensions are illustrated to display the performance



(c) Recall



(c) Specificity



(c) Accuracy

Fig. 8. Box plots of Recall, Specificity and Accuracy (outliers are labeled with associated study IDs) for the Psychological state.

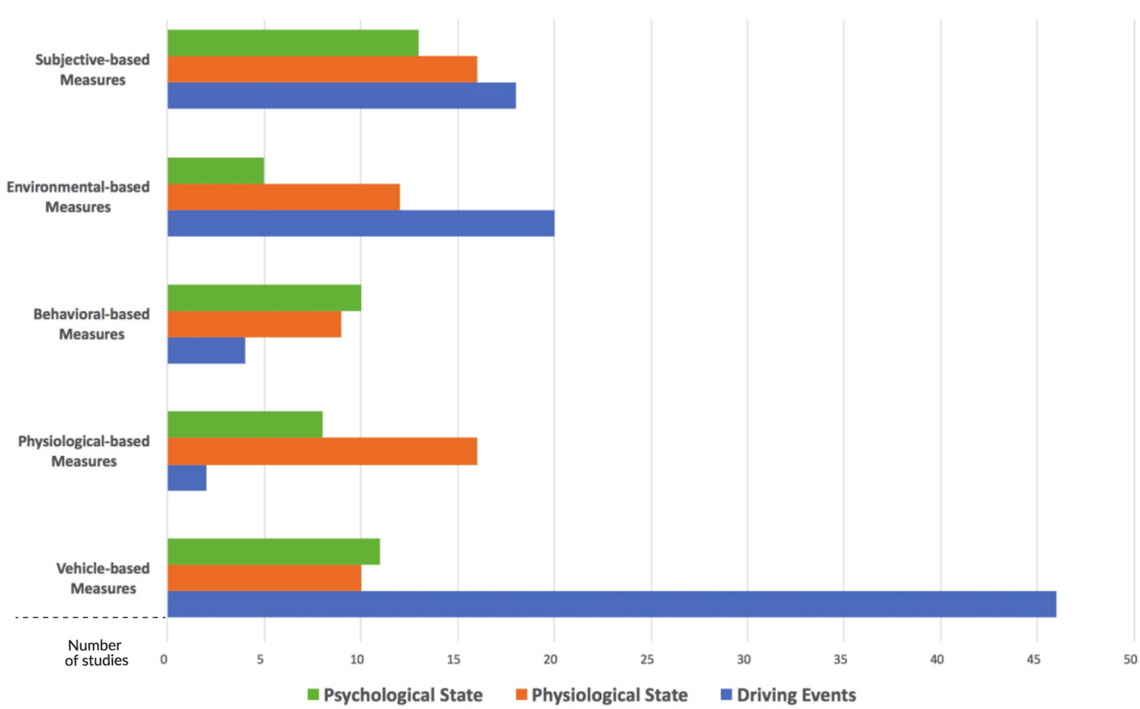


Fig. 9. Distribution of the DB dimensions' modeling metrics over number of studies.

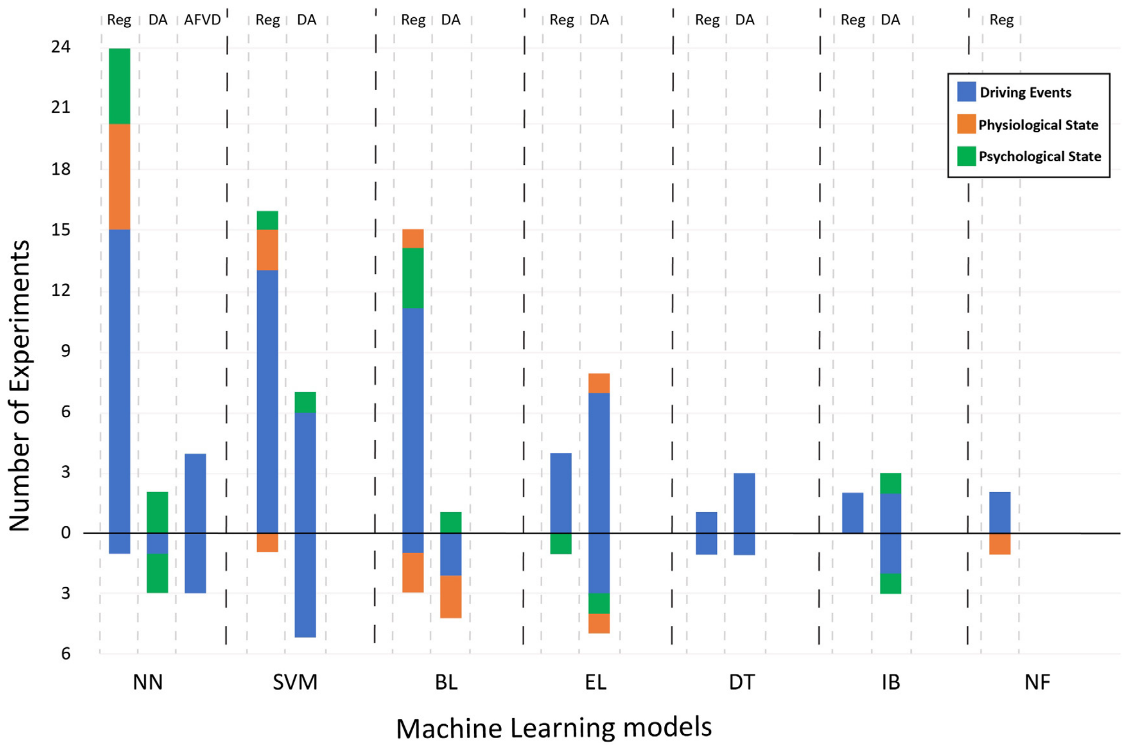


Fig. 10. Comparisons of accuracy between ML models and non-ML models in reference to DB dimensions.

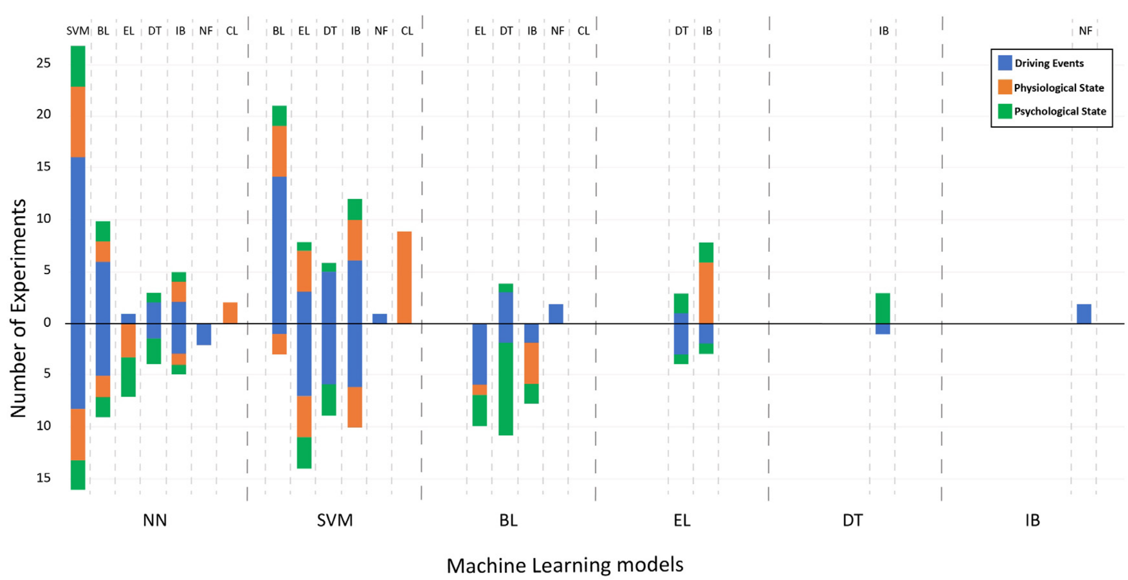


Fig. 11. Comparisons of accuracy between ML models in reference to DB dimensions.

of ML model by dimension. More details of the comparison results can be found in Table C.10 in Appendix B.

Several significant comparison results can be found in Fig. 11. First, NN, EL and DT are more accurate than SVM for the investigation of all the three DB dimensions, and this is supported by an adequate number of experiments. On the other hand, SVM, EL, DT and IB are more accurate than BL. More precisely, SVM and EL are highly performant than BL for the assessment of all DB dimensions. Further, EL is more efficient than IB and NN, while NN was found to perform equally to BL on every DB dimension. However, some of these findings are inconsistent with what we have found in (Sections: 4.2.1, 4.2.2, 4.2.3), where SVM was found to perform better than EL in evaluating Driving events and BL was shown to outperform SVM for the same DB dimension. Also, NN was proven to outperform EL in the examination of the Physiological state and NN was found to outperform BL in all DB dimensions. These contradictions may be resulting from: (1) all the experiments demonstrating (i) the superiority of EL over SVM and SVM over BL in the analysis of Driving events, (ii) the advantage of DT over SVM in the Psychological state and (iii) the upper hand of EL over NN and over BL in the analysis of the Physiological state and Driving events respectively, come from studies that are likely to have a bias towards the superior models; (2) these studies barely performed trials to compare the superior models and the inferior ones, so that the majority of their studies depicting high estimation accuracy contribute nothing to this comparison; and (3) the number of experiments in these

Table A.5
List of the studies inspected in the ML based SLR.

ID	Paper	Title	Citation
S1	Zhu et al. (2017b)	A Bayesian network model for contextual versus non-contextual driving behavior assessment	https://doi.org/10.1016/j.trc.2017.05.015
S2	Vlahogianni and Barmpounakis (2017)	Driving analytics using smartphones: algorithms, comparisons and challenges	https://doi.org/10.1016/j.trc.2017.03.014
S3	Das et al. (2009)	Using conditional inference forests to identify the factors affecting crash severity on arterial corridors	https://doi.org/10.1016/j.jsr.2009.05.003
S4	Elhenawy et al. (2015)	Modeling driver stop/run behavior at the onset of a yellow indication considering driver run tendency and roadway surface conditions	https://doi.org/10.1016/j.aap.2015.06.016
S5	Tang et al. (2018)	Lane-changes prediction based on adaptive fuzzy neural network	https://doi.org/10.1016/j.eswa.2017.09.025
S6	Yeo et al. (2009)	Can SVM be used for automatic EEG detection of drowsiness during car driving?	https://doi.org/10.1016/j.ssci.2008.01.007
S7	Singh et al. (2013)	A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals	https://doi.org/10.1016/j.bspc.2013.06.014
S8	Liu et al. (2014)	Preempt or yield? An analysis of driver's dynamic decision making at unsignalized intersections by classification tree	https://doi.org/10.1016/j.ssci.2013.12.009
S9	Jahangiri et al. (2016)	Red-light running violation prediction using observational and simulator data	https://doi.org/10.1016/j.aap.2016.06.009
S10	Masood et al. (2018)	Detecting distraction of drivers using convolutional neural network	https://doi.org/10.1016/j.patrec.2017.12.023
S11	Correa et al. (2013)	Automatic detection of drowsiness in EEG records based on multimodal analysis	https://doi.org/10.1016/j.medengphy.2013.07.011
S12	McDonald et al. (2018)	A contextual and temporal algorithm for driver drowsiness detection	https://doi.org/10.1016/j.aap.2018.01.005
S13	Yang et al. (2017)	Driving behavior recognition using EEG data from a simulated car-following experiment	https://doi.org/10.1016/j.aap.2017.11.010
S14	Huang et al. (2018)	A car-following model considering asymmetric driving behavior based on long short-term memory neural networks	https://doi.org/10.1016/j.trc.2018.07.022
S15	Jacobé de Naurois et al. (2017a)	Detection and prediction of driver drowsiness using artificial neural network models	https://doi.org/10.1016/j.aap.2017.11.038
S16	Qi and Fries (2018)	Real-time detection of drivers' texting and eating behavior based on vehicle dynamics	https://doi.org/10.1016/j.trf.2018.06.027
S17	Xie et al. (2017)	Modeling discretionary cut-in risks using naturalistic driving data	https://doi.org/10.1016/j.trf.2017.11.022
S18	Bejani and Ghatee (2018)	A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data	https://doi.org/10.1016/j.trc.2018.02.009
S19	Pariota et al. (2016)	Longitudinal control behavior: analysis and modeling based on experimental surveys in Italy and the UK	https://doi.org/10.1016/j.aap.2016.01.007
S20	Jeon et al. (2017)	A deterministic feedback model for safe driving based on nonlinear principal analysis scheme	https://doi.org/10.1016/j.procs.2017.08.301
S21	Wang et al. (2017b)	Modeling the various merging behaviors at expressway on-ramp bottlenecks using support vector machine models	https://doi.org/10.1016/j.trpro.2017.05.157
S22	Li et al. (2016)	Lane changing intention recognition based on speech recognition models	https://doi.org/10.1016/j.trc.2015.11.007
S23	Chuang et al. (2015)	An EEG-based perceptual function integration network for application to drowsy driving	https://doi.org/10.1016/j.knosys.2015.01.007
S24	Minhad et al. (2017)	Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness	https://doi.org/10.1016/j.jth.2017.11.001
S25	Kocatepe et al. (2017)	Analysis of speed patterns on inter-urban parallel highways: a case study in the southeast Florida	https://doi.org/10.1016/j.trpro.2017.03.064
S26	Tango and Botta (2013)	Real-time detection system of driver distraction using machine learning	https://doi.org/10.1109/TITS.2013.2247760
S27	Yu et al. (2017)	Fine-grained abnormal driving behaviors detection and identification with smartphones	https://doi.org/10.1109/TMC.2016.2618873
S28	Liu et al. (2016)	Driver distraction detection using semi-supervised machine learning	https://doi.org/10.1109/TITS.2015.2496157
S29	Wang et al. (2016)	Drowsy behavior detection based on driving information	https://doi.org/10.1007/s12239-016-0016-y
S30	Hou et al. (2015)	Situation assessment and decision making for lane change assistance using ensemble learning methods	https://doi.org/10.1016/j.eswa.2015.01.029
S31	Li et al. (2013)	Modeling of driver behavior in real world scenarios using multiple noninvasive sensors	https://doi.org/10.1109/TMM.2013.2241416
S32	Ragab et al. (2014)	A visual-based driver distraction recognition and detection using random forest	https://doi.org/10.1007/978-3-319-11758-4_28
S33	Wu et al. (2013)	Reasoning-based framework for driving safety monitoring using driving event recognition	https://doi.org/10.1109/TITS.2013.2257759

(*continued on next page*)

comparisons is higher in favor of the superior models' studies comparing to the inferior ones. Second, other results such as the advantage of SVM over IB and CL, the performance of EL and SVM is nearly the same in investigating the physiological state, also DT was found to

Table A.5 (*continued*).

ID	Paper	Title	Citation
S34	Wang et al. (2017a)	Driving style classification using a Semisupervised support vector machine	https://doi.org/10.1109/THMS.2017.2736948
S35	Li and Busso (2016)	Detecting drivers' mirror-checking actions and its application to maneuver and secondary task recognition	https://doi.org/10.1109/TITS.2015.2493451
S36	Hou et al. (2014)	Modeling mandatory lane changing using Bayes classifier and decision trees	https://doi.org/10.1109/TITS.2013.2285337
S37	Lee et al. (2017)	Stress events detection of driver by wearable glove system	https://doi.org/10.1109/JSEN.2016.2625323
S38	Osman et al. (2019)	A hierarchical machine learning classification approach for secondary task identification from observed driving behavior data	https://doi.org/10.1016/j.aap.2018.12.005
S39	Barua et al. (2019)	Automatic driver sleepiness detection using EEG, EOG and contextual information	https://doi.org/10.1016/j.eswa.2018.07.054
S40	Chen et al. (2018b)	Electroencephalography based fatigue detection using a novel feature fusion and extreme learning machine	https://doi.org/10.1016/j.cogsys.2018.08.018
S41	Wang et al. (2018c)	Risky driver recognition based on vehicle speed time series	https://doi.org/10.1109/THMS.2017.2776605
S42	Yuan et al. (2018)	Lane-change prediction method for adaptive cruise control system with hidden Markov model	https://doi.org/10.1177/1687814018802932
S43	Chen et al. (2018a)	Driver behavior formulation in intersection dilemma zones with phone use distraction via a logit-bayesian network hybrid approach	https://doi.org/10.1080/15472450.2017.1350921
S44	Jahangiri et al. (2018)	Application of real field connected vehicle data for aggressive driving identification on horizontal curves	https://doi.org/10.1109/TITS.2017.2768527
S45	Li et al. (2018)	Research on optimized GA-SVM vehicle speed prediction model based on driver–vehicle–road-traffic system	https://doi.org/10.1007/s11431-017-9213-0
S46	Xiong et al. (2018)	A new framework of vehicle collision prediction by combining SVM and HMM	https://doi.org/10.1109/TITS.2017.2699191
S47	Ihme et al. (2018)	Recognizing frustration of drivers from face video recordings and brain activation measurements with functional near-infrared spectroscopy	https://doi.org/10.3389/fnhum.2018.00327
S48	Darzi et al. (2018)	Identifying the causes of drivers' hazardous states using driver characteristics, vehicle kinematics, and physiological measurements	https://doi.org/10.3389/fnins.2018.00568
S49	Scenarios et al. (2018)	Learning and inferring a driver's braking action in car-following scenarios	https://doi.org/10.1109/TVT.2018.2793889
S50	Min et al. (2017)	Driver fatigue detection through multiple entropy fusion analysis in an EEG-based system	https://doi.org/10.1371/journal.pone.0188756
S51	Sysoev et al. (2017)	Estimation of the driving style based on the users' activity and environment influence	https://doi.org/10.3390/s17102404
S52	Osafune et al. (2017)	Analysis of accident risks from driving behaviors	https://doi.org/10.1007/s13177-016-0132-0
S53	Munoz-Organero and Corcoba-Magana (2017)	Predicting Upcoming Values of Stress While Driving	https://doi.org/10.1109/TITS.2016.2618424
S54	Kim et al. (2017)	Prediction of driver's intention of lane change by augmenting sensor information using machine learning techniques	https://doi.org/10.3390/s17061350
S55	Chandrasiri et al. (2016)	Driving skill classification in curve driving scenes using machine learning	https://doi.org/10.1007/s40534-016-0098-2
S56	Halim et al. (2016a)	Profiling drivers based on driver dependent vehicle driving features	https://doi.org/10.1007/s10489-015-0722-6
S57	Xuan et al. (2010)	Identification of driver's braking intention based on a hybrid model of GHMM and GGAP-RBFNN	https://doi.org/10.1007/s00521-018-3672-1
S58	Deshmukh and Dehzangi (2019)	Characterization and Identification of Driver Distraction During Naturalistic Driving: An Analysis of ECG Dynamics	https://doi.org/10.1007/978-3-030-02819-0_1
S59	Chen et al. (2015)	Driving behavior analysis of multiple information fusion based on Adaboost	https://doi.org/10.1007/978-3-319-12286-1_28
S60	Chen et al. (2017a)	Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers	https://doi.org/10.1016/j.eswa.2017.01.040
S61	Wang et al. (2015)	Short term prediction of freeway exiting volume based on SVM and KNN	https://doi.org/10.1260/2046-0430.4.3.337
S62	Jacobé de Naurois et al. (2018)	Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness	https://doi.org/10.1016/j.aap.2018.08.017
S63	Zhao et al. (2012)	Recognition of driving postures by multiwavelet transform and multilayer perceptron classifier	https://doi.org/10.1016/j.engappai.2012.09.018
S64	Vlahogianni and Golias (2012)	Bayesian modeling of the microscopic traffic characteristics of overtaking in two-lane highways	https://doi.org/10.1016/j.trf.2012.02.002
S65	Jabon et al. (2011)	Facial-expression analysis for predicting unsafe driving behavior	https://doi.org/10.1109/MPRV.2010.46
S66	Memory et al. (2011)	Online driver distraction detection using Long Short-Term Memory	https://doi.org/10.1109/TITS.2011.2119483
S67	Miyajima et al. (2010)	Analysis of real-world driver's frustration	https://doi.org/10.1109/TITS.2010.2070839
S68	Elmitiny et al. (2010)	Classification analysis of driver's stop/go decision and red-light running violation	https://doi.org/10.1016/j.aap.2009.07.007
S69	Bundele and Banerjee (2010)	Roc analysis of a fatigue classifier for vehicular drivers	https://doi.org/10.1109/IS.2010.5548362

(*continued on next page*)

Table A.5 (*continued*).

ID	Paper	Title	Citation
S70	Riccardo et al. (2012)	Comparative analysis of random utility models and fuzzy logic models for representing gap-acceptance behavior using data from driving simulator experiments	https://doi.org/10.1016/j.sbspro.2012.09.799
S71	Hong et al. (2014)	A smartphone-based sensing platform to model aggressive driving behaviors	https://doi.org/10.1145/2556288.2557321
S72	Shimosaka et al. (2015)	Predicting driving behavior using inverse reinforcement learning with multiple reward functions towards environmental diversity	https://doi.org/10.1109/IVS.2015.7225745
S73	Zhao et al. (2017)	Modeling driver behavior at roundabouts: results from a field study	https://doi.org/10.1109/IVS.2017.7995831
S74	Okamoto et al. (2017)	Driver intention-based vehicle threat assessment using random forests and particle filtering	https://doi.org/10.1016/j.ifacol.2017.08.2231
S75	Khushaba et al. (2013)	Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition	https://doi.org/10.1016/j.fss.2012.12.003
S76	Aoude et al. (2012)	Driver behavior classification at intersections and validation on large naturalistic data set	https://doi.org/10.1109/TITS.2011.2179537
S77	Rodriguez Gonzalez et al. (2014)	Modeling and detecting aggressiveness from driving signals	https://doi.org/10.1109/TITS.2013.2297057
S78	Henni et al. (2018)	Feature selection for driving fatigue characterization and detection using visual- and signal-based sensors	https://doi.org/10.1186/s40535-018-0054-9
S79	Gwak et al. (2018)	Early detection of driver drowsiness utilizing machine learning based on physiological signals, behavioral measures, and driving performance	https://doi.org/10.1109/ITSC.2018.8569493
S80	Zhang and Kumada (2018)	Automatic detection of mind wandering in a simulated driving task with behavioral measures	https://doi.org/10.1371/journal.pone.0207092
S81	Aksjonov et al. (2018)	A novel driver performance model based on machine learning	https://doi.org/10.1016/j.ifacol.2018.07.044
S82	Manawadu et al. (2018)	Multiclass classification of driver perceived workload using long short-term memory based recurrent neural network	https://doi.org/10.1109/IVS.2018.8500410

Table B.6
Descriptive statistics of recall, specificity and accuracy for driving events.

Model	Recall						Specificity						Accuracy					
	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max
SVM	17	0.86	0.88	0.12	0.6	1.00	–	–	–	–	–	–	130	0.86	0.90	0.12	0.53	1.00
NN	17	0.92	0.92	0.06	0.8	1.00	4	0.84	0.84	0.19	0.63	1.00	19	0.96	0.98	0.05	0.85	1.00
BL	11	0.82	0.88	0.17	0.51	1.00	5	0.94	0.95	0.06	0.83	1.00	35	0.90	0.91	0.09	0.71	1.00
DT	7	0.94	0.97	0.07	0.85	1.00	4	0.97	0.97	0.02	0.95	1.00	20	0.86	0.85	0.09	0.69	1.00
IB	6	0.96	0.97	0.05	0.88	1.00	4	0.97	0.98	0.05	0.89	1.00	7	0.93	0.92	0.07	0.82	1.00
EL	–	–	–	–	–	–	–	–	–	–	–	–	14	0.86	0.88	0.10	0.62	0.98
NF	–	–	–	–	–	–	–	–	–	–	–	–	4	0.79	0.73	0.12	0.72	0.93

Table B.7
Descriptive statistics of recall, specificity and accuracy for physiological state.

Model	Recall						Specificity						Accuracy					
	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max
SVM	27	0.89	0.89	0.05	0.79	0.98	11	0.88	0.90	0.08	0.73	0.97	37	0.86	0.85	0.08	0.69	0.99
NN	13	0.91	0.93	0.05	0.83	0.98	15	0.92	0.91	0.05	0.83	0.98	30	0.88	0.90	0.10	0.54	1.00
BL	4	0.76	0.76	0.28	0.52	1.00	–	–	–	–	–	–	5	0.76	0.75	0.07	0.73	0.88
IB	22	0.91	0.92	0.06	0.68	0.98	6	0.90	0.91	0.05	0.83	0.95	28	0.80	0.82	0.11	0.50	0.95
EL	14	0.82	0.89	0.22	0.19	1.00	9	0.89	0.91	0.07	0.75	0.97	14	0.86	0.86	0.09	0.70	0.97
CL	–	–	–	–	–	–	–	–	–	–	–	–	6	0.81	0.80	0.05	0.76	0.89

Table B.8
Descriptive statistics of recall, specificity and accuracy for psychological state.

Model	Recall						Specificity						Accuracy					
	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max	Number of values	Mean	Median	Std. dev.	Min	Max
SVM	23	0.92	0.93	0.07	0.78	1.00	22	0.95	0.94	0.03	0.89	1.00	24	0.77	0.77	0.19	0.36	1.00
NN	37	0.85	0.88	0.13	0.48	1.00	28	0.95	0.94	0.03	0.87	1.00	27	0.72	0.78	0.19	0.38	0.97
EL	4	0.80	0.78	0.07	0.72	0.85	4	0.91	0.91	0.04	0.87	0.94	10	0.83	0.83	0.10	0.60	0.96
BL	–	–	–	–	–	–	4	0.78	0.78	0.09	0.70	0.87	11	0.67	0.66	0.14	0.50	0.88
IB	–	–	–	–	–	–	–	–	–	–	–	–	9	0.68	0.67	0.18	0.37	0.90
DT	–	–	–	–	–	–	–	–	–	–	–	–	6	0.76	0.79	0.15	0.56	0.99

outperform IB in assessing psychological state while IB outperformed DT in the driving events, these findings, amongst others, are coherent with the results of Section 4.2. Third, for the comparisons between the other ML models, the number of evaluations is relatively insufficient, particularly when dealing with every dimension separately, and some comparisons are even found to be inconsistent. Thus, it is hard to decide which ML model is more accurate.

Table C.9
Comparisons of Accuracy between ML models and non-ML models ("+" indicates ML model outperforms non-ML model, "−" indicates non-ML model outperforms ML model; the number following study ID is Accuracy improvement in percentage, the study ID in bold indicates the improvement exceeds 2%).

ML model		Non-ML model		
		Regression	DA	AFVD
NN	+	**S5**(59.3), **S5**(58.8), **S5**(76.3), **S5**(57.5), **S5**(51.9), **S5**(58.1), **S5**(57.5), **S5**(54.4), **S5**(52.8), **S5**(43.02), **S5**(39.7), **S5**(32.3), **S5**(38.1), **S5**(26.3), **S5**(20), **S5**(36.4), **S5**(28.3), **S5**(26.1), **S26**(18.6), **S26**(8), **S33**(12.2), **S33**(7), S33(0.63), **S33**(7.9), **S33**(8.12)	**S38**(9.2), **S38**(2.7)	**S14**(24.1), **S14**(21.1), **S14**(34.1), S14(0.69), **S14**(26.6)
	−	**S33**(2.7)	S38(0.3), **S38**(7.4), **S38**(181.8), **S38**(500)	**S14**(18.9), **S14**(2.24), **S14**(14)
SVM	+	**S4**(7.05), **S5**(51.4), **S5**(50.6), **S5**(42.3), **S5**(44.4), **S5**(32.9), **S5**(35.6), **S5**(43), **S5**(34.4), **S5**(41.2), **S16**(10.9), **S16**(11.3), **S26**(26.6), S33(0.76), S33(0.89), **S33**(2), S33(1.66), **S79**(8.7), **S79**(3)	**S16**(13), **S16**(17), **S38**(5.5), **S38**(15.8), **S38**(171.6), **S38**(132.2), **S38**(548.3),	NA
	−	**S33**(4.5)	**S38**(34.6), **S38**(12.5), **S38**(51), **S38**(12.9), S38(0.2), **S38**(4.3)	NA
BL	+	**S5**(25), **S5**(24.7), **S5**(21.5), **S5**(27.1), **S5**(20.9), **S5**(16.3), **S5**(26.4), **S5**(19.6), **S5**(25.1), **S33**(10.8), **S33**(24.7), S33(1.79), **S33**(2), **S33**(10), S33(0.63), S33(1.26), S33(0.64), **S33**(9.36), S33(1.66), **S36**(280.3)	**S38**(3.33)	NA
	−	S33(0.63), **S33**(6.1), **S33**(10.5), **S36**(3.6)	**S38**(4.03), **S38**(6.7), **S38**(86.2), **S38**(47.6), S38(0.9)	NA
EL	+	**S16**(14.1), **S16**(15.9), **S79**(12.5), **S79**(3.4)	**S16**(16.2), **S16**(21.8), **S38**(7.3), **S38**(34.3), **S38**(2390.3), **S38**(1270), **S38**(4.6), **S38**(2.1)	NA
	−	S4(1.22), **S4**(2.3)	**S38**(4.5), **S38**(4), **S38**(4.5), **S38**(10.1), **S38**(100)	NA
DT	+	**S36**(303.3)	**S38**(42.6), **S38**(0.3), S38(0.4), **S38**(2254.8), **S38**(1180)	NA
	−	**S36**(18.4)	**S38**(11.7)	NA
IB	+	**S79**(4.1), **S79**(2)	**S38**(18.6), **S38**(1100), **S38**(453.3)	NA
	−	NA	**S38**(12.9), **S38**(14.9), **S38**(3.9)	NA
NF	+	**S26**(20), **S36**(252.1)	NA	NA
	−	**S36**(33.6), **S70**(0.54)	NA	NA

4.5. Strengths and weaknesses of ML models

Selecting the appropriate ML models for the DB estimation contexts can be addressed by investigating the candidate ML models from their characteristics, which are generally outlined by the strengths and weaknesses of the ML techniques as recorded by researchers. Hence, the strengths and weaknesses of the ML techniques supported by more than one study are introduced in this section. The SVM has been applauded for its excellent ability to deal with redundant features, non-linear, high-dimensional and small samples, this makes the SVM suitable for classification problems with redundant data sets. NN works reliably with noisy data and has been proven to hold strong generalization and learning ability as well as adaptability. BL are capable of explicitly modeling the time-dependent nature of driver state and allow the inclusion of contextual factors that influence DB, they can also infer the state of an unobserved variable given the state of observed variables. Moreover, the reviewed research have proven that BL can be more analytically demanding when augmented-density attributes are involved, but still tractable and frugal when using flexible inputs; in the author's view, models such as Naïve Bayes and Baseline Bayesian are computationally efficient, while other, more elaborate methods such as Bayesian Networks and Tree Augmented Bayesian Networks can be more computationally expensive for high-dimensional data. EL is extensively used in behavioral modeling because they can model any time series, they are also robust against overfitting and relatively robust to outliers and noise. DT have been acclaimed for their effectiveness to construct classifications of DB through segmenting the data set into smaller and more homogeneous groups and can avoid overfitting by pruning. The detailed information on the synthesized strengths and weaknesses of different ML techniques is presented in Table D.11 in Appendix D.

Although ML techniques have been proved in some studies to be efficient for DB estimation, they do not always perform well on all DB tasks. For that matter, an utterly "supreme" estimation method does not seem to exist, and the ML model performance relies profoundly on the contexts it applies to. Therefore, researchers need to comprehend not only the characteristics of the candidate ML techniques, but also the DB contexts in order to choose ML techniques properly and apply them to real-world DB tasks efficiently.

5. Conclusion, research implications and limitations

This paper presents the findings of a systematic review that academics and practitioners can use to unlock the immense potential of understanding driving behavior analysis in reference to its dimensions along the cross cutting themes identified in the study. Although our work cannot claim to be all-embracing, we believe that it will prove a beneficial resource for anyone interested in driving behavior research, and will help induce further interest in the field. The first section of the paper presented a conceptual framework to outline the relationship between the different dimensions of the DB and the Driver–Vehicle–Environment system. Conceptual frameworks have been proven to be important contributions in understanding the mechanisms influencing behavior, and in the development of interventions to improve safety outcomes (Newnam and Watson, 2011). Therefore, a crucial part of our research is the construction of a broader theoretical framework that provides a direction for the future development and implementation of DB intervention strategies along with the designation of a reference collection of relevant literature. The second section introduced an extensive overview of academic articles in the sense of evaluating the performance of various ML techniques amongst themselves and with the non-ML techniques used for DB analysis. Although the importance of ML techniques in assessment of DB has been recognized, a systematic literature review of their application in DB research studies is lacking. We have within the confines of the formulated research questions completed a rigorous analysis by following a systematic series of steps

Table C.10
Comparisons of Accuracy between different ML models and non-ML models ("+" indicates the model given in the row outperforms the model given in the column, "−" indicates the model given in the column outperforms the model given in the row; the number following study ID is Accuracy improvement in percentage, the study ID in bold indicates the improvement exceeds 2%).

ML model		ML model						
		SVM	BL	EL	DT	IB	CL	NF
NN	+	**S18**(2.4), **S18**(4.8), **S18**(2.3), S27(1.4), S27(0.3), S27(1.5), S28(2.3), S28(1.7), **S33**(11.3), **S33**(4.9), **S33**(12.7), **S33**(6.3), **S38**(3.4), **S38**(41.1), **S38**(7.8), **S40**(3.5), **S40**(3.5), **S40**(5.2), **S40**(6.8), **S40**(6), **S40**(9.4), **S40**(6.3), **S40**(3.1), **S40**(3.9), S50(0.6), S50(1.6), **S56**(79.1), **S58**(18.2), **S63**(8.6), **S66**(3.3), **S66**(5.2), S66(1.4), **S66**(2.9), **S66**(4.6), S75(1.4),	**S18**(2), **S18**(19.7), **S18**(22.5), **S32**(16.2), S33(1.2), **S33**(14.5), S33(1.3), **S33**(10.9), **S38**(13.6), **S38**(42.8), **S38**(3.7), **S58**(26.5)	S38(1.7), **S38**(3.7), S50(1.9), S50(0.7), S58(1.3)	**S18**(2), **S18**(3.5), S18(1.1), **S58**(4)	**S18**(2), **S38**(9.3), **S50**(2.4), **S50**(4.9), **S63**(3.2)	**S28**(14), **S28**(15.9)	NA
	−	**S23**(51.6), **S23**(9.07), **S23**(4.6), **S23**(4.8), **S23**(5.9), **S26**(16.5), **S26**(14.6), S28(1.7), **S33**(3.7), **S38**(6), **S38**(11.8), **S38**(4.3), **S66**(2), **S75**(2), **S75**(2.1), **S75**(2.4), **S75**(2.3)	**S23**(34.8), **S23**(4.7), **S23**(6.54), **S23**(6.3), **S23**(5.4), **S33**(11.1), **S33**(2.7), **S33**(4.6), **S33**(2.8)	**S18**(51.5), **S32**(21), **S32**(21.3), **S32**(2.4), **S38**(22.9), S38(1.3), **S38**(4.1), **S38**(42.7)	**S18**(53.1), **S38**(30.5), **S38**(5.3), **S38**(38.3)	**S18**(53.1), **S18**(3.4), **S18**(5.8), **S38**(8.6), **S38**(15.9), S75(0.6), S75(1), S75(0.7)	NA	**S26**(9.2), **S26**(11),
SVM	+		**S18**(16.9), **S18**(19.7), **S21**(3.3), **S21**(4.9), **S21**(5.4), **S21**(6.7), **S23**(12.5), **S23**(4.1), **S24**(39.7), **S24**(14.6), **S24**(102.8), **S24**(100), S33(0.9), **S33**(4.9), S33(1.5), **S33**(12.3), **S38**(9.7), **S38**(20.5), **S38**(32.4), S38(1.2), **S38**(16), **S38**(8.2), **S46**(13), **S58**(7)	**S4**(9.5), **S4**(8.3), **S4**(8.4), **S4**(8.7), **S38**(7.9), **S38**(7.3), **S39**(2.5), **S39**(2.1), S39(1.1), S39(1.4)	**S17**(6.9), **S17**(17.4), S18(1.1), **S21**(6.3), **S21**(16.6), **S21**(3.4), **S21**(19.2)	**S24**(26.4), **S24**(13.8), **S31**(5.3), **S31**(3.6), S38(1.3), S39(1.2), S39(1), S39(1.1), **S39**(4), **S39**(2.8), S50(1.7), **S50**(3.2), **S55**(8.5), **S55**(9), S75(1.3), S75(1.6), S75(1.7), S75(1.6), **S75**(2.7), **S75**(5.4), **S79**(4.3), S79(0.9)	**S28**(16.1), **S28**(13.9), **S28**(13.3), **S65**(7.9), **S65**(6.3), **S65**(5.9), **S65**(7.4), **S65**(7.8), **S65**(10.1)	**S26**(4.9),
	−		S21(0.3), S23(1.8), **S33**(9.9), **S33**(23.8), S33(0.8), **S33**(7.8), S33(1.2), **S33**(18),	**S16**(2.8) **S16**(4), **S32**(2.1), S38(1.7), **S38**(27.2), **S38**(15.9), **S38**(36), **S38**(43), **S38**(3.9), **S38**(9.3), **S38**(36.7), **S38**(27.6), **S39**(5.7), S39(1.4), **S58**(16.6), **S79**(3.5)	S18(1.2), **S21**(8.7), **S21**(4.9), **S38**(35.1), **S38**(23), **S38**(48.7), **S38**(13.6), **S38**(32.8), **S38**(23.7), **S58**(13.6)	**S18**(8.4), **S18**(5.8), **S38**(12.4), **S38**(2.3), **S38**(29), **S38**(11.3), **S38**(3.7), **S39**(7.2), **S63**(5.1), **S75**(2.5), **S75**(1)	NA	NA
BL	+			NA	**S21**(11.7), **S21**(2.8), **S36**(6.3), **S36**(14.2)	NA	NA	**S36**(8), **S36**(28.9)
	−			**S32**(4.1), **S32**(4.3), **S32**(2.6), **S38**(11.7), **S38**(39.7), **S38**(37.7), **S38**(44.7), **S38**(8.1), **S38**(48.1), **S58**(24.9)	**S18**(18.3), **S21**(2.5), **S21**(7), **S21**(8.3), **S21**(4.9), S21(1.9), **S36**(8.2), **S36**(8.2), **S38**(48.3), **S38**(50.5), **S38**(43.6), **S58**(21.6)	**S18**(26.7), **S24**(26.4), **S24**(14.6), **S24**(78.2), **S24**(75.72), **S38**(23.4), **S38**(30.6), **S38**(20.3)	NA	NA
EL	+				**S38**(9.3), **S38**(3.1), **S58**(2.6),	**S38**(13.2), **S38**(5.4), **S38**(10.8), **S38**(23), **S39**(10), **S39**(4.3), S50(1.6), **S50**(2.8), **S79**(8.1),	NA	NA
	−				**S38**(32.8), **S38**(6.1), **S38**(3.9), **S38**(32.8)	**S38**(10.4), **S38**(11.3), S39(1.2), S39(1), **S39**(5.7)	NA	NA

(*continued on next page*)

and analyzing the quality of the studies we identified 82 primary studies published in the last decade (2009–2019). However, with the dynamic nature of DB research, we cannot fully guarantee to have taken into account all the available studies in this research domain. The investigation of ML based DB estimation was conducted following four standpoints: the type of ML techniques, the estimation accuracy of ML models, the comparison between different models (including ML model vs. non-ML model and ML model vs. other ML model), and the strengths and weaknesses of the ML models. The main findings obtained from the selected primary studies are:

Table C.10 (*continued*).

ML model		ML model						
		SVM	BL	EL	DT	IB	CL	NF
DT	+					**S38**(20.2), **S38**(15.2), **S38**(19.3)	NA	NA
	−					**S18**(7.1)	NA	NA
IB	+						NA	**S36**(12.8), **S36**(9.7)
	−						NA	NA
CL	+							NA
	−							NA

Table D.11
Strengths and weaknesses of ML techniques (the unlisted models are not supported by more than one study).

Strength		Weakness	
Items	Supporting Studies	Items	Supporting Studies
NN			
Hold strong generalization and learning ability as well as adaptability	S5, S7, S63	Require diversified training data set to train the model effectively	S10, S23, S69
Can do fast real-time computation with better computational efficiency	S54, S82	Cost large computational resource	S41, S60
Can avoid over-fitting	S15, S62		
Capable of dealing with noisy data	S7, S15		
SVM			
Good tolerance for redundant features	S6, S39	Has very limited success when applied to imbalanced data sets	S35, S39
Can handle complex non-linear problems	S6, S17, S21, S23, S26, S39, S45, S61, S63, S79	Choosing an adequate kernel function	S26, S55, S61, S76
A remarkable property of SVM is its good generalization capacity independent of the input space dimension	S6, S17, S21, S23, S26, S39, S45, S54, S56, S61, S63		
Simultaneously minimizes the empirical classification error and maximizes the geometric margin in classification	S4, S6, S13, S16, S17, S78, S79		
Robust in nature	S26, S44, S55, S73		
DT			
Intuitive and easy to understand	S3, S17, S68		
Can deal with categorical features	S3, S68		
Can avoid overfitting by pruning	JS17, S68		
BL			
Can infer unobservable states from observable actions	S5, S46, S67, S76	Can be computationally expensive for high-dimensional data	S1, S43, S49
Allow the inclusion of contextual information	S1, S12		
Can mimic the complex nonlinear states more realistically	S36, S64		
Capable of learning causal relationships	S43, S76		
Computationally efficient	S43, S49		
Can easily integrate qualitative and quantitative information, and/or erroneous or missing data in the modeling process	S24, S43, S64		
EL			
Likely to have smaller misclassification error compared to the base classifier	S28, S30	The description of relationships between the variables can be challenging due to the presence of several individual classifiers	S28, S30
Produces insights about factor importance	S9, S79		
Robust against overfitting	S9, S29		
Reduce the bias of the learning algorithm	S12, S30		
Robust against noisy and missing data	S30, S39		
Runs efficiently on large databases	S16, S79		
IB			
Intuitive and easy to understand	S55, S61, S63, S58	Suffer from computational complexity	S63, S75
Robust in nature	S58, S61, S63, S75		
CL			
Capture the complex distribution of multimodal data	S25, S32	Influenced by random initialization	S41, S56

- The ML techniques were broadly categorized into Neural networks (NN), Support vector machines (SVM), Clustering (CL), Instance Based (IB), Decision trees (DT), Bayesian learners (BL), Ensemble learners (EL), Fuzzy & Neuro Fuzzy based (NF), Inductive Rule Based (IR), Evolutionary algorithms (EA) and Miscellaneous. Among them, SVM, NN, EL and BL are used most frequently.
- The Driving events dimension have been actively investigated when analyzing DB using ML techniques, followed by the Physiological and Psychological states.
- Subjective-based measures have been fairly adopted in different aspects of analyzing DB. The vast majority of the studies employed vehicle-based measures in the process, followed by environmental-based measures, then physiological and behavioral inputs.
- Accuracy, Recall and Specificity are the most commonly used performance measures in the primary studies. The overall estimation accuracy of most ML models is close to the acceptable level.
- The assessment of the Driving events dimension showed that ML models have their arithmetic means of Accuracy approximately ranging from 73% to 98%, Recall from 82% to 96%, and Specificity from 84% to 97%.
- The Physiological state analysis showed that ML models have their arithmetic means of Accuracy approximately ranging from 76% to 88%, Recall from 76% to 91%, and Specificity from 88% to 92%.
- The Psychological state analysis depicted that ML models have their arithmetic means of Accuracy approximately ranging from 67% to 83%, Recall from 80% to 92%, and Specificity from 78% to 95%.
- NN, IB, BL and SVM are the most accurate ML models for the analysis of the Driving events dimension, while it was found that NN, SVM and EL in one hand, along with EL, SVM, NN and DT on the other hand, are the most effective ones in terms of accuracy for the Physiological and Psychological state respectively.
- ML model outperformed non-ML model in general, which is supported by most of the studies. Regression model is the non-ML model that is most often compared with ML models. Moreover, the driving events dimension was the evaluated one of all the DB dimensions.

When interpreting the results of this SLR, some limitations should be captured. Of note, only about 24% of the primary studies compare ML techniques with non-ML techniques. Thus, the results of these comparisons are not definitely conclusive. Also, while comparing ML techniques, various experimental settings are likely to be used by each study, which include data sets to construct model, validation methods, feature selection methods and pre-processing methods to remove outliers Wen et al. (2012). Furthermore, this review considered only the metrics of Accuracy, Recall and Specificity when evaluating the performance of ML models or comparing ML models with other models as they are among the most important ones and were used by most of the studies. However, considering other performance metrics such Precision and Classification Errors, which were ignored in this review, could be a valuable addition to the analysis strategy. Generally, academics will exclude a model which does not accomplish the minimum accuracy threshold (Mair et al., 2000). Moreover, some inconsistencies have been displayed throughout the comparisons between ML models and non-ML models and between different ML models is assessing each of the DB dimension, despite the fact these incompatibilities were addressed, it is difficult to determine which model outperform the other due to the insufficient amount of the studies relating the desired comparisons that could have been the cause for such inconsistencies.

We have presumed that all the selected studies are unbiased, however, this is regarded as threat if that is not the case. Also, this review does not include any work in progress, unpublished or non-peer reviewed publications which may be able to answer any of the research questions. Lastly, the strengths and weaknesses of ML techniques were extracted directly from the selected studies. That is, some of them may just stand for the authors' opinions and then prove to be unreliable. For that reason, we only considered the ones that are supported by two or more selected studies. Notwithstanding that these strengths and weaknesses are in general trustworthy coming from studies proven to be with acceptable quality throughout quality assessment process, it is recommended to further validate the veracity of the summarized strengths and weaknesses of ML techniques in this review.

This review provides recommendations for researchers as well as guidelines for practitioner to carry out future research on DB analysis using the ML techniques:

- Additional studies for DB analysis should be conducted using the ML techniques in order to obtain generalizable results and gain more proof on the viability of ML models. As there are few studies that compare the ML techniques with the non-ML techniques, more studies should compare the performance of the ML techniques with non-ML techniques.
- Only a few studies examined the efficiency of Evolutionary Algorithms such as GP, GP and Inductive Rule Based namely CR and M5R for DB assessment. Hence, future work may emphasize the importance on this matter.
- Practitioners should comprehend the rationale of ML models before using them in DB evaluation and collaborate with experienced researchers in the application of ML models.
- ML models should be adopted in parallel with existing conventional models at the early stage in order to unlock the true potential of a given ML technique.
- Sharing proprietary project data sets with research community would highly enable comparative studies and inspire amateurs in the field.

Acknowledgment

This research received funding from the Moroccan Ministry of Equipment, Transport and Logistics and was supported by the Moroccan National Center for Scientific and Technical Research (CNRST).

Appendix A

See Table A.5.

Appendix B

See Tables B.6, B.7, B.8.

Appendix C

See Tables C.9, C.10.

Appendix D

See Table D.11.

References

Aberg, L., Rimmo, P.-A., 1998. Dimensions of aberrant driver behaviour. Ergonomics 41 (1), 39–56. http://dx.doi.org/10.1080/001401398187314.

Aghaei, A.S., Donmez, B., Liu, C.C., He, D., Liu, G., Plataniotis, K.N., … Sojoudi, Z., 2016. Smart driver monitoring: When signal processing meets human factors: In the driver's seat. IEEE Signal Process. Mag. 33 (6), 35–48. http://dx.doi.org/10.1109/MSP.2016.2602379.

Ajzen, I., 1985. From intentions to actions: A theory of planned behavior. In: Action Control. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 11–39. http://dx.doi.org/10.1007/978-3-642-69746-3_2.

Akerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2265922.

Aksjonov, A., Nedoma, P., Vodovozov, V., Petlenkov, E., Herrmann, M., 2018. A novel driver performance model based on machine learning. IFAC-PapersOnLine 51 (9), 267–272. http://dx.doi.org/10.1016/j.ifacol.2018.07.044.

Amditis, A., Pagle, K., Joshi, S., Bekiaris, E., 2010. Driver – Vehicle – Environment monitoring for on-board driver support systems : Lessons learned from design and implementation. Applied Ergon. 41 (2), 225–235. http://dx.doi.org/10.1016/j.apergo.2009.03.002.

Anastasopoulos, P.C., Tarko, A.P., Mannering, F.L., 2008. Tobit analysis of vehicle accident rates on interstate highways. Accid. Anal. Prev. 40 (2), 768–775. http://dx.doi.org/10.1016/j.aap.2007.09.006.

Aoude, G.S., Desaraju, V.R., Stephens, L.H., Jonathan, P.H., 2012. Driver behavior classification at intersections and validation on large naturalistic data set. IEEE Trans. Intell. Transp. Syst. 13 (2), http://dx.doi.org/10.1109/TITS.2011.2179537.

Taubman-ben ari, O., Mikulincer, M., Gillath, O., 2004. The multidimensional driving style inventory — scale construct and validation. 36, pp. 323–332. http://dx.doi.org/10.1016/S0001-4575(03)00010-1.

Ariën, C., Jongen, E.M.M., Brijs, K., Brijs, T., Daniels, S., Wets, G., 2013. A simulator study on the impact of traffic calming measures in urban areas on driving behavior and workload. Accid. Anal. Prev. 1–11. http://dx.doi.org/10.1016/j.aap.2012.12.044.

Ba, Y., Zhang, W., Wang, Q., Zhou, R., Ren, C., 2017. Crash prediction with behavioral and physiological features for advanced vehicle collision avoidance system. Transp. Res. C 74, 22–33. http://dx.doi.org/10.1016/j.trc.2016.11.009.

Bahram, M., Hubmann, C., Lawitzky, A., Aeberhard, M., Wollherr, D., 2016. A combined model- and learning-based framework for interaction-aware maneuver prediction. IEEE Trans. Intell. Transp. Syst. 17 (6), 1538–1550. http://dx.doi.org/10.1109/TITS.2015.2506642.

Barua, S., Ahmed, M.U., Ahlström, C., Begum, S., 2019. Automatic driver sleepiness detection using EEG, EOG and contextual information. Expert Syst. Appl. 115, 121–135. http://dx.doi.org/10.1016/j.eswa.2018.07.054.

Begg, R., Kamruzzaman, J., 2005. A Machine Learning Approach for Automated Recognition of Movement Patterns using Basic, Kinetic and Kinematic Gait Data, Vol. 38. pp. 401–408. http://dx.doi.org/10.1016/j.jbiomech.2004.05.002.

Bejani, M.M., Ghatee, M., 2018. A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data. Transp. Res. C 89 (February), 303–320. http://dx.doi.org/10.1016/j.trc.2018.02.009.

Ben-ari, O.T., Yehiel, D., 2012. Driving styles and their associations with personality and motivation. Accid. Anal. Prev. 45, 416–422. http://dx.doi.org/10.1016/j.aap.2011.08.007.

Bergasa, L.M., Nuevo, J., 2005. Real-time system for monitoring driver vigilance. IEEE Int. Symp. Ind. Electron. III (1), 1303–1308. http://dx.doi.org/10.1109/ISIE.2005.1529113.

Blanchard, R.A., Myers, A.M., Porter, M.M., 2010. Correspondence Between Self-Reported and Objective Measures of Driving Exposure and Patterns in Older Drivers, Vol. 42. pp. 523–529. http://dx.doi.org/10.1016/j.aap.2009.09.018.

Blockey, P.N., Hartley, L.R., 1995. Aberrant driving behaviour: errors and violations. Ergonomics 38 (9), 1759–1771. http://dx.doi.org/10.1080/00140139508925225.

Bundele, M.M., Banerjee, R., 2010. ROC Analysis of a Fatigue Classifier for Vehicular Drivers. pp. 1–6. http://dx.doi.org/10.1109/IS.2010.5548362.

Cacciabue, P.C., Carsten, O., 2010. A simple model of driver behaviour to sustain design and safety assessment of automated systems in automotive environments. Applied Ergon. 41 (2), 187–197. http://dx.doi.org/10.1016/j.apergo.2009.03.008.

Cai, H., Lin, Y., 2011. Modeling of operators' emotion and task performance in a virtual driving environment. Int. J. Hum.-Comput. Stud. 69 (9), 571–586. http://dx.doi.org/10.1016/J.IJHCS.2011.05.003.

Caird, J.K., Johnston, K.A., Willness, C.R., Asbridge, M., Steel, P., 2014. A meta-analysis of the effects of texting on driving. Accid. Anal. Prev. 71, 311–318. http://dx.doi.org/10.1016/j.aap.2014.06.005.

Campilho, A., Kamel, M., 2014. Image analysis and recognition: 11th international conference. In: ICIAR 2014 Vilamoura, Portugal, October (2014) 22-24 Proceedings, Part I. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8814, pp. 256–265. http://dx.doi.org/10.1007/978-3-319-11758-4.

Catherine, C., Sommers, M.S., Winston, F.K., 2016. Novice Teen Driver Crash Patterns. pp. 47–58, (n.d.).

Chandrasiri, N.P., Nawa, K., Ishii, A., 2016. Driving skill classification in curve driving scenes using machine learning. J. Mod. Transp. 24 (3), 196–206. http://dx.doi.org/10.1007/s40534-016-0098-2.

Chen, C.-F., Chen, C.-W., 2011. Speeding for fun? Exploring the speeding behavior of riders of heavy motorcycles using the theory of planned behavior and psychological flow theory. Accid. Anal. Prev. 43 (3), 983–990. http://dx.doi.org/10.1016/J.AAP.2010.11.025.

Chen, C., Chen, Y., Ma, J., Zhang, G., Walton, C.M., 2018a. Driver behavior formulation in intersection dilemma zones with phone use distraction via a logit-Bayesian network hybrid approach. J. Intell. Transp. Syst. Technol. Plann. Oper. 22 (4), 311–324. http://dx.doi.org/10.1080/15472450.2017.1350921.

Chen, S.-H., Pan, J.-S., Lua, K., Xu, H., 2015. Driving behavior analysis of multiple information fusion based on adaboost. Adv. Intell. Syst. Comput. 329, http://dx.doi.org/10.1007/978-3-319-12286-1_28.

Chen, J., Wang, H., Hua, C., 2018b. Electroencephalography based fatigue detection using a novel feature fusion and extreme learning machine. Cogn. Syst. Res. 52, 715–728. http://dx.doi.org/10.1016/j.cogsys.2018.08.018.

Chen, L. lan, Zhao, Y., Ye, P. fei, Zhang, J., Zou, J. zhong, 2017a. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. Expert Syst. Appl. 85, 279–291. http://dx.doi.org/10.1016/j.eswa.2017.01.040.

Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture : A review. Comput. Electron. Agric. 151 (2017), 61–69. http://dx.doi.org/10.1016/j.compag.2018.05.012.

Chuang, C.H., Huang, C.S., Ko, L.W., Lin, C.T., 2015. An EEG-based perceptual function integration network for application to drowsy driving. Knowl.-Based Syst. 80, 143–152. http://dx.doi.org/10.1016/j.knosys.2015.01.007.

Correa, A.G., Orosco, L., Laciar, E., 2013. Automatic detection of drowsiness in EEG records based on multimodal analysis. Med. Eng. Phys. http://dx.doi.org/10.1016/j.medengphy.2013.07.011.

Dahlen, E.R., White, R.P., 2006. The Big Five factors, sensation seeking, and driving anger in the prediction of unsafe driving. Pers. Individ. Differ. 41 (5), 903–915. http://dx.doi.org/10.1016/j.paid.2006.03.016.

Darzi, A., Gaweesh, S.M., Ahmed, M.M., Novak, D., 2018. Identifying the causes of drivers' hazardous states using driver characteristics, vehicle kinematics, and physiological measurements. Front. Neurosci. 12 (AUG), 1–13. http://dx.doi.org/10.3389/fnins.2018.00568.

Das, A., Abdel-Aty, M., Pande, A., 2009. Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. J. Saf. Res. 40 (4), 317–327. http://dx.doi.org/10.1016/j.jsr.2009.05.003.

Davey, J., Wishart, D., Freeman, J., Watson, B., 2007. An application of the driver behaviour questionnaire in an Australian organisational fleet setting. Transp. Res. F 10 (1), 11–21. http://dx.doi.org/10.1016/J.TRF.2006.03.001.

Dawson, J.D., Uc, Ã.E.Y., Anderson, S.W., Johnson, A.M., 2010. Neuropsychological Predictors of Driving Errors in Older Adults. pp. 1090–1096. http://dx.doi.org/10.1111/j.1532-5415.2010.02872.x.

Derbel, O., Landry, R., 2017. Driver behavior assessment in case of critical driving situations. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. E100A (2), 491–498. http://dx.doi.org/10.1587/transfun.E100.A.491.

Deshmukh, S.V., Dehzangi, O., 2019. Characterization and identification of driver distraction during naturalistic driving: An analysis of ECG dynamics. Adv. Body Area Netw. I http://dx.doi.org/10.1007/978-3-030-02819-0_1.

Center for Disease Control, 2009. Nonfatal, Motor Vehicle-Occupant Injuries (2009) and Seat Belt Use (2008) Among Adults-United States (Reprinted from MMWR, 59(51/52), 1681) (Vol. 305).

Domeyer, J.E., Cassavaugh, N.D., Backs, R.W., 2013. The use of adaptation to reduce simulator sickness in driving assessment and research. Accid. Anal. Prev. 53, 127–132. http://dx.doi.org/10.1016/j.aap.2012.12.039.

Elamrani Abou Elassad, Z., Mousannif, H., 2019. Understanding driving behavior: Measurement, modeling and analysis. Adv. Intell. Syst. Comput. 5, http://dx.doi.org/10.1007/978-3-030-11928-7_41.

Elander, J., West, R., French, D., 1993. Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. In: Psychological Bulletin. American Psychological Association, US, http://dx.doi.org/10.1037/0033-2909.113.2.279.

Elhenawy, M., Jahangiri, A., Rakha, H.A., El-Shawarby, I., 2015. Modeling driver stop/run behavior at the onset of a yellow indication considering driver run tendency and roadway surface conditions. Accid. Anal. Prev. 83, 90–100. http://dx.doi.org/10.1016/j.aap.2015.06.016.

Elmitiny, N., Yan, X., Radwan, E., Russo, C., Nashar, D., 2010. Classification analysis of driver's stop/go decision and red-light running violation. Accid. Anal. Prev. 42 (1), 101–111. http://dx.doi.org/10.1016/j.aap.2009.07.007.

Ericsson, E., 2000. Variability in urban driving patterns. Transp. Res. D 5 (5), 337–354. http://dx.doi.org/10.1016/S1361-9209(00)00003-1.

Evans, L., 1991. Traffic safety and the driver. In: Traffic Safety and the Driver. Van Nostrand Reinhold Co, New York, NY, US.

Evans, L., 1996. Comment: The dominant role of driver behavior in traffic safety. Am J Public Health 86 (6), 784–786. http://dx.doi.org/10.2105/AJPH.86.6.784.

Faure, V., Lobjois, R., Benguigui, N., 2016. The Effects of Driving Environment Complexity and Dual Tasking on Drivers ' Mental Workload and Eye Blink Behavior, Vol. 40. pp. 78–90. http://dx.doi.org/10.1016/j.trf.2016.04.007.

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., Zhang, K., 2014. The effect of stress and personality on dangerous driving behavior among Chinese drivers. Accid. Anal. Prev. 73, 34–40. http://dx.doi.org/10.1016/j.aap.2014.07.024.

Ghani, J.A., Deshpande, S.P., 1994. Task characteristics and the experience of optimal flow in human—computer interaction. J. Psychol. 128 (4), 381–391. http://dx.doi.org/10.1080/00223980.1994.9712742.

Ghasemzadeh, A., Ahmed, M.M., 2018. Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches. Transp. Res. C 90 (2017), 379–392. http://dx.doi.org/10.1016/j.trc.2018.03.018.

Gheorghiu, A., Delhomme, P., Line, M., 2015. Peer pressure and risk taking in young drivers ' speeding behavior. Transp. Res. F 35, 101–111. http://dx.doi.org/10.1016/j.trf.2015.10.014.

Gindele, T., Brechtel, S., Dillmann, R., 2015. Learning driver behavior models from traffic observations for decision making and planning. IEEE Intell. Transp. Syst. Mag. 7 (1), 69–79. http://dx.doi.org/10.1109/MITS.2014.2357038.

Gong, H., Liu, H., Wang, B.-H., 2008. An asymmetric full velocity difference car-following model. Physica A 387 (11), 2595–2602. http://dx.doi.org/10.1016/J.PHYSA.2008.01.038.

Gwak, J., Shino, M., Hirao, A., 2018. Early detection of driver drowsiness utilizing machine learning based on physiological signals. Behav. Meas. Driving Perform. 179, 4–1800. http://dx.doi.org/10.1109/ITSC.2018.8569493.

Halim, Z., Kalsoom, R., Baig, A.R., 2016a. Profiling drivers based on driver dependent vehicle driving features. Appl. Intell. 44 (3), 645–664. http://dx.doi.org/10.1007/s10489-015-0722-6.

Halim, Z., Kalsoom, R., Bashir, S., Abbas, G., 2016b. Artificial intelligence techniques for driving safety and vehicle crash prediction. Artif. Intell. Rev. 46 (3), 351–387. http://dx.doi.org/10.1007/s10462-016-9467-9.

Hamdar, S.H., Qin, L., Talebpour, A., 2016. Weather and road geometry impact on longitudinal driving behavior : Exploratory analysis using an empirically supported acceleration modeling framework. Transp. Res. C 67, 193–213. http://dx.doi.org/10.1016/j.trc.2016.01.017.

Hancock, P.A., 1997. Fatigue, workload and adaptive driver systems'. 29, 495–506. http://dx.doi.org/10.1016/S0001-4575(97)00029-8.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. Adv. Psychol. 52, 139–183. http://dx.doi.org/10.1016/S0166-4115(08)62386-9.

Hatfield, J., Williamson, A., Kehoe, E.J., Prabhakharan, P., 2017. An examination of the relationship between measures of impulsivity and risky simulated driving amongst young drivers. Accid. Anal. Prev. 103 (2016), 37–43. http://dx.doi.org/10.1016/j.aap.2017.03.019.

Healey, J.A., Picard, R.W., 2005. Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans. Intell. Transp. Syst. 6 (2), 156–166. http://dx.doi.org/10.1109/TITS.2005.848368.

Henni, K., Mezghani, N., Gouin-Vallerand, C., Ruer, P., Ouakrim, Y., Vallières, É., 2018. Feature selection for driving fatigue characterization and detection using visual- and signal-based sensors. Appl. Inform. 5 (1), 1–15. http://dx.doi.org/10.1186/s40535-018-0054-9.

Hollnagel, E., 2005. Review and synthesis of models for Joint Driver-vehicle Interaction Design. 0(March).

Hong, J.-H., Margines, B., Dey, A.K., 2014. A smartphone-based sensing platform to model aggressive driving behaviors. pp. 4047–4056. http://dx.doi.org/10.1145/2556288.2557321.

Hori, C., Watanabe, S., Hori, T., Harsham, B.A., Hershey, J.R., 2016. Driver Confusion Status Detection using Recurrent Neural Networks Mitsubishi Electric Research Laboratories. Mitsubishi Electric Corporation Information Technology R & D Center, http://dx.doi.org/10.1109/ICME.2016.7552966.

Hou, Y., Edara, P., Sun, C., 2014. Modeling mandatory lane changing using Bayes classifier and decision trees. IEEE Trans. Intell. Transp. Syst. 15 (2), 647–655. http://dx.doi.org/10.1109/TITS.2013.2285337.

Hou, Y., Edara, P., Sun, C., 2015. Situation assessment and decision making for lane change assistance using ensemble learning methods. Expert Syst. Appl. 42 (8), 3875–3882. http://dx.doi.org/10.1016/j.eswa.2015.01.029.

Houston, J.M., Harris, P.B., Norman, M., 2003. The Aggressive Driving Behavior Scale: Developing a self-report measure of unsafe driving practices. N. Am. J. Psychol. 5 (2), 269–278.

Huang, X., Sun, J., Sun, J., 2018. A car-following model considering asymmetric driving behavior based on long short-term memory neural networks. Transp. Res. C 95 (July), 346–362. http://dx.doi.org/10.1016/j.trc.2018.07.022.

Ihme, K., Unni, A., Zhang, M., Rieger, J.W., Jipp, M., 2018. Recognizing frustration of drivers from face video recordings and brain activation measurements with functional near-infrared spectroscopy. Front. Human Neurosci. 12 (August), http://dx.doi.org/10.3389/fnhum.2018.00327.

Jabon, M., Bailenson, J., Pontikakis, E., Takayama, L., Nass, C., 2011. Facial-expression analysis for predicting unsafe driving behavior. IEEE Perv. Comput. 10 (4), 84–95. http://dx.doi.org/10.1109/MPRV.2010.46.

Jahangiri, A., Berardi, V.J., MacHiani, S.G., 2018. Application of real field connected vehicle data for aggressive driving identification on horizontal curves. IEEE Trans. Intell. Transp. Syst. 19 (7), 2316–2324. http://dx.doi.org/10.1109/TITS.2017.2768527.

Jahangiri, A., Rakha, H., Dingus, T.A., 2016. Red-light running violation prediction using observational and simulator data. Accid. Anal. Prev. 96, 316–328. http://dx.doi.org/10.1016/j.aap.2016.06.009.

Jeon, M., Yang, E., Oh, E., Park, J., Youn, C.H., 2017. A deterministic feedback model for safe driving based on nonlinear principal analysis scheme. Procedia Comput. Sci. 113, 454–459. http://dx.doi.org/10.1016/j.procs.2017.08.301.

Jinjun, W., Xu, W., Gong, Y., 2010. Engineering applications of artificial intelligence real-time driving danger-level prediction. Eng. Appl. Artif. Intell. 23 (8), 1247–1254. http://dx.doi.org/10.1016/j.engappai.2010.01.001.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. 349 (6245).

Kaiseler, M., Cunha, J.P., Cunha, P.S., Member, S., 2015. A mobile sensing approach to stress detection and memory activation for public bus drivers a mobile sensing approach to stress detection and memory activation for public bus drivers. 16 (2016), 3294–3303.

Kantowitz, B.H., Simsek, O., 2001. SECONDARY-TASK MEASURES of DRIVER WORKLOAD. STRESS, WORKLOAD AND FATIGUE, p. 682, Retrieved from https://trid.trb.org/view.aspx?id=683360.

Karkouch, A., Mousannif, H., Al, H., 2018. Sciencedirect sciencedirect CADS : A connected assistant for driving safe. Procedia Comput. Sci. 127, 353–359. http://dx.doi.org/10.1016/j.procs.2018.01.132.

Khushaba, R.N., Kodagoda, S., Lal, S., Dissanayake, G., 2013. Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition. Fuzzy Sets and Systems 221, 90–111. http://dx.doi.org/10.1016/j.fss.2012.12.003.

Kim, I.-H., Bong, J.-H., Park, J., Park, S., 2017. Prediction of driver's intention of lane change by augmenting sensor information using machine learning techniques. Sensors 17 (6), 1350. http://dx.doi.org/10.3390/s17061350.

Klauer, S., Dingus, T.A., Neale, V.L., Sudweeks, J., Ramsey, D., 2006. The impact of driver inattention on near crash/crash risk: An analysis using the 100-Car naturalistic driving study data. Dot Hs 810 594 (April), 226, https://doi.org/DOT HS 810 594.

Kocatepe, A., Ozguven, E.E., Vanli, A., Moses, R., 2017. Analysis of speed patterns on inter-urban parallel highways: A case study in the southeast florida. Transp. Res. Procedia 22 (2016), 479–488. http://dx.doi.org/10.1016/j.trpro.2017.03.064.

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., Nass, C., 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. Int. J. Interact. Des. Manuf. 9 (4), 269–275. http://dx.doi.org/10.1007/s12008-014-0227-2.

Koufaris, M., 2008. Applying the technology acceptance model and flow. Cyberpsychol. Behav. 11 (3), 205–223. http://dx.doi.org/10.1089/cpb.2007.0117.

Lanata, A., Valenza, G., Greco, A., Gentili, C., Bartolozzi, R., Bucchi, F., . Scilingo, E.P., 2015. How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving. IEEE Trans. Intell. Transp. Syst. 16 (3), 1505–1517. http://dx.doi.org/10.1109/TITS.2014.2365681.

Lavrenz, S.M., Pyrialakou, V.D., Gkritza, K., 2014. Analytic methods in accident research modeling driver behavior in dilemma zones : A discrete / con- tinuous formulation with selectivity bias corrections. Anal. Methods Accident Res. 3–4, 44–55. http://dx.doi.org/10.1016/j.amar.2014.10.002.

Lee, D.S., Chong, T.W., Lee, B.G., 2017. Stress events detection of driver by wearable glove system. IEEE Sens. J. 17 (1), 194–204. http://dx.doi.org/10.1109/JSEN.2016.2625323.

Lee, B., Chung, W., 2012. Driver alertness monitoring using fusion of facial features and bio-signals. IEEE Sens. J. 12 (7), 2416–2422.

Lee, J.Y., Chung, J.H., Son, B., 2008. Analysis of traffic accident size for Korean highway using structural equation models. Accid. Anal. Prev. 40 (6), 1955–1963. http://dx.doi.org/10.1016/j.aap.2008.08.006.

Lee, B., Lee, B., Chung, W., 2015. Smartwatch-Based Driver Alertness Monitoring with Wearable Motion and Physiological Sensor *, Vol. 1. pp. 6126–6129.

Lethaus, F., Baumann, M.R.K., Köster, F., Lemmer, K., 2013. Neurocomputing A comparison of selected simple supervised learning algorithms to predict driver intent based on gaze data. Neurocomputing 121, 108–130. http://dx.doi.org/10.1016/j.neucom.2013.04.035.

Li, N., Busso, C., 2016. Detecting drivers' mirrorchecking actions and its application to maneuver and secondary task recognition. IEEE Trans. Intell. Transp. Syst. 17 (4), 980–992. http://dx.doi.org/10.1109/TITS.2015.2493451.

Li, Y.F., Chen, M.N., Lu, X.D., Zhao, W.Z., 2018. Research on optimized GA-SVM vehicle speed prediction model based on driver-vehicle-road-traffic system. Sci. China Technol. Sci. 61 (5), 782–790. http://dx.doi.org/10.1007/s11431-017-9213-0.

Li, Z., Chen, L., Peng, J., Wu, Y., 2017a. Automatic detection of driver fatigue using driving operation information for transportation safety. Sensors 17 (6), http://dx.doi.org/10.3390/s17061212.

Li, Z., Li, S., Li, R., Cheng, B., Shi, J., 2017b. Online detection of driver fatigue using steering wheel angles for real driving conditions. Sensors 17 (3), 495. http://dx.doi.org/10.3390/s17030495.

Li, N., Member, S., Jain, J.J., Busso, C., 2013. Modeling of driver behavior in real world scenarios using multiple noninvasive sensors. IEEE Trans. Multimed. 15 (5), 1213–1225. http://dx.doi.org/10.1109/TMM.2013.2241416.

Li, K., Wang, X., Xu, Y., Wang, J., 2016. Lane changing intention recognition based on speech recognition models. Transp. Res. C 69, 497–514. http://dx.doi.org/10.1016/j.trc.2015.11.007.

Liang, Y., Horrey, W.J., Howard, M.E., Lee, M.L., Anderson, C., Shreeve, M.S., . Czeisler, C.A., 2017. Prediction of drowsiness events in night shift workers during morning driving. Accid. Anal. Prev. (May), 0–1. http://dx.doi.org/10.1016/j.aap.2017.11.004.

Liu, M., Lu, G., Wang, Y., Wang, Y., Zhang, Z., 2014. Preempt or yield? An analysis of driver's dynamic decision making at unsignalized intersections by classification tree. Saf. Sci. 65, 36–44. http://dx.doi.org/10.1016/j.ssci.2013.12.009.

Liu, T., Yang, Y., Huang, G., Bin, Yeo, Y.K., Lin, Z., 2016. Driver distraction detection using semi-supervised machine learning. IEEE Trans. Intell. Transp. Syst. 17 (4), 1108–1120. http://dx.doi.org/10.1109/TITS.2015.2496157.

Lourens, P.F., Vissers, J.A.M., Jessurun, M., 1999. Annual mileage, driving violations, and accident involvement in relation to drivers' sex, age, and level of education. Accid. Anal. Prev. 31 (5), 593–597. http://dx.doi.org/10.1016/S0001-4575(99)00015-9.

Lu, J., Xie, X., Zhang, R., 2013. Focusing on appraisals : How and why anger and fear in fl uence driving risk perception. J. Saf. Res. 45, 65–73. http://dx.doi.org/10.1016/j.jsr.2013.01.009.

Lu, Y., Zhou, T., Wang, B., 2009. Exploring Chinese users' acceptance of instant messaging using the theory of planned behavior, the technology acceptance model, and the flow theory. Comput. Hum. Behav. 25 (1), 29–39. http://dx.doi.org/10.1016/J.CHB.2008.06.002.

Lucidi, F., Mallia, L., Lazuras, L., Violani, C., 2014. Personality and attitudes as predictors of risky driving among older drivers. Accid. Anal. Prev. 72, 318–324. http://dx.doi.org/10.1016/j.aap.2014.07.022.

Lv, H., Tang, H., 2011. Machine learning methods and their application research. http://dx.doi.org/10.1109/IPTC.2011.34.

Mair, C., Kadoda, G., Lefley, M., Phalp, K., Schofield, C., Shepperd, M., Webster, S., 2000. An investigation of machine learning based prediction systems. J. Syst. Softw. 53 (1), 23–29. http://dx.doi.org/10.1016/S0164-1212(00)00005-4.

Malhotra, R., 2015. A systematic review of machine learning techniques for software fault prediction. Appl. Soft Comput. J. 27, 504–518. http://dx.doi.org/10.1016/j.asoc.2014.11.023.

Manawadu, U.E., Kawano, T., Murata, S., Kamezaki, M., Muramatsu, J., Sugano, S., 2018. Multiclass classification of driver perceived workload using long short-term memory based recurrent neural network. In: IEEE Intelligent Vehicles Symposium, Proceedings, 2018– June(Iv). pp. 2009–2014. http://dx.doi.org/10.1109/IVS.2018.8500410.

Martens, H., 1959. Two Notes on Machine Learning, Vol. 379. pp. 364–379.

Martinez, C.M., Heucke, M., Wang, F., Gao, B., Cao, D., 2017. Driving style recognition for intelligent vehicle control and advanced driver assistance : A survey. pp. 1–11.

Masood, S., Rai, A., Aggarwal, A., Doja, M.N., Ahmad, M., 2018. Detecting distraction of drivers using convolutional neural network. Pattern Recognit. Lett. 1–7. http://dx.doi.org/10.1016/j.patrec.2017.12.023.

McDonald, A.D., Lee, J.D., Schwarz, C., Brown, T.L., 2018. A contextual and temporal algorithm for driver drowsiness detection. Accid. Anal. Prev. 113 (January), 25–37. http://dx.doi.org/10.1016/j.aap.2018.01.005.

Mehdizadeh, M., Shariat-Mohaymany, A., Nordfjaern, T., 2018. Accident involvement among Iranian lorry drivers: Direct and indirect effects of background variables and aberrant driving behaviour. Transp. Res. F 58, 39–55. http://dx.doi.org/10.1016/j.trf.2018.05.029.

Meiring, G.A.M., Myburgh, H.C., 2015. A review of intelligent driving style analysis systems and related artificial intelligence algorithms. Sensors (Switzerland) 15 (12), 30653–30682. http://dx.doi.org/10.3390/s151229822.

Memory, L.S., Wöllmer, M., Blaschke, C., Schindl, T., Schuller, B., Färber, B., Trefflich, B..., 2011. Online driver distraction detection using. Transportation 12 (2), 574–582. http://dx.doi.org/10.1109/TITS.2011.2119483.

Min, J., Wang, P., Hu, J., 2017. Driver fatigue detection through multiple entropy fusion analysis in an EEG-based system. PLoS One 12 (12), 1–19. http://dx.doi.org/10.1371/journal.pone.0188756.

Minhad, K.N., Ali, S.H.M., Reaz, M.B.I., 2017. Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness. J. Transp. Health 7 (xxxx), 75–89. http://dx.doi.org/10.1016/j.jth.2017.11.001.

Mittal, A., Kumar, K., Dhamija, S., Kaur, M., 2016. Head movement-based driver drowsiness detection: A review of state-of-art techniques. In: 2016 IEEE International Conference on Engineering and Technology (ICETECH). IEEE, pp. 903–908. http://dx.doi.org/10.1109/ICETECH.2016.7569378.

Miyajima, C., Kitaoka, N., Takeda, K., Member, S., 2010. Analysis of real-world driver ' s frustration. IEEE Trans. Intell. Transp. Syst. 12 (1), 1–10. http://dx.doi.org/10.1109/TITS.2010.2070839.

Molnar, L.J., Eby, D.W., Bogard, S.E., Leblanc, D.J., Jennifer, S., Molnar, L.J., . Zakrajsek, J.S., 2018. Using naturalistic driving data to better understand the driving exposure and patterns of older drivers patterns of older drivers. http://dx.doi.org/10.1080/15389588.2017.1379601, 9588.

Moon, J.-W., Kim, Y.-G., 2001. Extending the TAM for a world-wide-web context. Inform. Manage. 38 (4), 217–230. http://dx.doi.org/10.1016/S0378-7206(00)00061-6.

Morris, B., Doshi, A., Trivedi, M., 2011. Lane change intent prediction for driver assistance : On-road design and evaluation, (Iv). pp. 895–901.

Munoz-Organero, M., Corcoba-Magana, V., 2017. Predicting upcoming values of stress while driving. IEEE Trans. Intell. Transp. Syst. 18 (7), 1802–1811. http://dx.doi.org/10.1109/TITS.2016.2618424.

Murata, A., 2016. Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures. IEEE Trans. Intell. Transp. Syst. 7323 (March), 1802–1811. http://dx.doi.org/10.1080/21577323.2016.1164765.

Musicant, O., Bar-Gera, H., Schechtman, E., 2010. Electronic records of undesirable driving events. Transp. Res. F 13 (2), 71–79. http://dx.doi.org/10.1016/J.TRF.2009.11.001.

Jacobé de Naurois, C., Bourdin, C., Bougard, C., Vercher, J.L., 2018. Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. Accid. Anal. Prev. 121 (April), 118–128. http://dx.doi.org/10.1016/j.aap.2018.08.017.

Jacobé de Naurois, C., Bourdin, C., Stratulat, A., Diaz, E., Vercher, J.L., 2017a. Detection and prediction of driver drowsiness using artificial neural network models. Accid. Anal. Prev. (October), 0–1. http://dx.doi.org/10.1016/j.aap.2017.11.038.

Jacobé de Naurois, C., Bourdin, C., Stratulat, A., Diaz, E., Vercher, J.L., 2017b. Detection and prediction of driver drowsiness using artificial neural network models. Accid. Anal. Prev. (July), 0–1. http://dx.doi.org/10.1016/j.aap.2017.11.038.

Nevile, M., Haddington, P., 2010. In-Car Distractions and their Impact on Driving Activites. Australian Road Safety Report RSGR 2010-001, http://dx.doi.org/10.1021/acs.cgd.5b01175, Available online: http://www.infrastructure.gov.au/roads/safety/publications/2010/pdf/rsgr_2010001.pdf.

Newnam, S., Watson, B., 2011. Work-related driving safety in light vehicle fleets: A review of past research and the development of an intervention framework. Saf. Sci. 49 (3), 369–381. http://dx.doi.org/10.1016/j.ssci.2010.09.018.

Nilsson, E.J., Aust, M.L., Engström, J., Svanberg, B., Lindén, P., 2018. Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). Transp. Res. Part F: Traffic Psychol. Behav. 59, 463–474. http://dx.doi.org/10.1016/J.TRF.2018.09.026.

Nowosielski, R.J., Trick, L.M., Toxopeus, R., 2018. Good distractions: Testing the effects of listening to an audiobook on driving performance in simple and complex road environments. Accid. Anal. Prev. 111, 202–209. http://dx.doi.org/10.1016/J.AAP.2017.11.033.

Ohn-bar, E., Martin, S., Tawari, A., Trivedi, M., 2014. Head, eye, and hand patterns for driver activity recognition. http://dx.doi.org/10.1109/ICPR.2014.124.

Okamoto, K., Berntorp, K., Di Cairano, S., 2017. Driver intention-based vehicle threat assessment using random forests and particle filtering. IFAC-PapersOnLine 50 (1), 13860–13865. http://dx.doi.org/10.1016/j.ifacol.2017.08.2231.

Olson, R.L., Hanowski, R.J., Hickman, J.S., Safety, V. T. T. I. C. for T. and B., 2009. Driver Distraction in Commercial Vehicle Operations. Federal Motor Carrier Safety Administration, United States, Retrieved from https://rosap.ntl.bts.gov/view/dot/17715.

Osafune, T., Takahashi, T., Kiyama, N., Sobue, T., Yamaguchi, H., Higashino, T., 2017. Analysis of accident risks from driving behaviors. Int. J. Intell. Transp. Syst. Res. 15 (3), 192–202. http://dx.doi.org/10.1007/s13177-016-0132-0.

Osman, O.A., Hajij, M., Karbalaieali, S., Ishak, S., 2019. A hierarchical machine learning classification approach for secondary task identification from observed driving behavior data. Accid. Anal. Prevent. 123 (May), 274–281. http://dx.doi.org/10.1016/j.aap.2018.12.005.

Pariota, L., Bifulco, G.N., Galante, F., Montella, A., Brackstone, M., 2016. Longitudinal control behaviour: Analysis and modelling based on experimental surveys in Italy and the UK. Accid. Anal. Prev. 89, 74–87. http://dx.doi.org/10.1016/j.aap.2016.01.007.

Parker, D., McDonald, L., Rabbitt, P., Sutcliffe, P., 2000. Elderly drivers and their accidents: the Aging Driver Questionnaire. Accid. Anal. Prev. 32 (6), 751–759. http://dx.doi.org/10.1016/S0001-4575(99)00125-6.

Pathivada, B.K., Perumal, V., 2017. Sciencedirect modeling driver behavior in dilemma zone under mixed traffic conditions. Transp. Res. Procedia 27, 961–968. http://dx.doi.org/10.1016/j.trpro.2017.12.120.

Pöysti, L., Rajalin, S., Summala, H., 2005. Factors influencing the use of cellular (mobile) phone during driving and hazards while using it. 37, pp. 47–51. http://dx.doi.org/10.1016/j.aap.2004.06.003.

Prokhorov, B.B., Shmakov, D.I., 2013. Causes of people's death in peacetime and economic assessment of the value of losses. Stud. Russ. Econ. Dev. 24 (4), 394–399. http://dx.doi.org/10.1134/S1075700713040096.

Qi, Y., Fries, R., 2018. Real-time detection of drivers ' texting and eating behavior based on vehicle dynamics. Transp. Res. Part F: Psychol. Behav. 58, 594–604. http://dx.doi.org/10.1016/j.trf.2018.06.027.

Ragab, A., Craye, C., Kamel, M.S., Fakhri, K., 2014. A visual-based driver distraction recognition and detection using random forest. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8814, pp. 256–265. http://dx.doi.org/10.1007/978-3-319-11758-4_28.

Rakotonirainy, A., 2005. Design of context-aware systems for vehicles using complex system paradigms.

Ranchet, M., Paire-ficout, L., Uc, E.Y., Bonnard, A., Sornette, D., Broussolle, E., 2013. Impact of specific executive functions on driving performance in people with Parkinson's disease. 28, (14), pp. 1941–1948. http://dx.doi.org/10.1002/mds.25660.

Ranney, T.A., Garrott, W.R., Goodman, M.J., 2001. Nhtsa driver distraction research: past, present, and future. Retrieved from https://www.sae.org/publications/technical-papers/content/2001-06-0177/.

Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K., 1990. Errors and violations on the roads: a real distinction?. Ergonomics 33 (10–11), 1315–1332. http://dx.doi.org/10.1080/00140139008925335.

Reimer, B., Fried, R., Mehler, B., Zhao, N., Goldin, R., Biederman, J., 2013. Brief report : Examining driving behavior in young adults with high functioning autism spectrum disorders : A pilot study using a driving simulation paradigm. http://dx.doi.org/10.1007/s10803-013-1764-4.

Riccardo, R., Massimiliano, G., Gregorio, G., Claudio, M., 2012. Comparative analysis of random utility models and fuzzy logic models for representing gap-acceptance behavior using data from driving simulator experiments. Procedia - Soc. Behav. Sci. 54, 834–844. http://dx.doi.org/10.1016/j.sbspro.2012.09.799.

Rodriguez Gonzalez, A.B., Wilby, M.R., Vinagre Diaz, J.J., Sanchez Avila, C., 2014. Modeling and detecting aggressiveness from driving signals. IEEE Trans. Intell. Transp. Syst. 15 (4), 1419–1428. http://dx.doi.org/10.1109/TITS.2013.2297057.

Sagberg, F., Selpi, Bianchi Piccinini, G.F., Engström, J., 2015. A review of research on driving styles and road safety. Hum. Factors 57 (7), 1248–1275. http://dx.doi.org/10.1177/0018720815591313.

Sahayadhas, A., Sundaraj, K., Murugappan, M., Sahayadhas, A., Sundaraj, K., Murugappan, M., 2012. Detecting driver drowsiness based on sensors: A review. Sensors 12 (12), 16937–16953. http://dx.doi.org/10.3390/s121216937.

Saiprasert, C., Pholprasit, T., Thajchayapong, S., 2017. Detection of driving events using sensory data on smartphone. Accid. Anal. Prevent. Res. 15 (1), 17–28. http://dx.doi.org/10.1007/s13177-015-0116-5.

Sarma, K.M., Carey, R.N., Kervick, A.A., Bimpeh, Y., 2013. Psychological factors associated with indices of risky, reckless and cautious driving in a national sample of drivers in the Republic of Ireland. Accid. Anal. Prev. 50, 1226–1235. http://dx.doi.org/10.1016/j.aap.2012.09.020.

Sayer, B., Arbor, A., Us, M.I., Domeyer, J.E., Arbor, A., Ci, U.S., 2018. Driver and vehicle monitoring feedback system for an autonomous vehicle. US Patent 20180118219A1, 1.

Scenarios, C., Wang, W., Member, S., Xi, J., Zhao, D., 2018. Learning and inferring a driver ' s braking action in car-following scenarios. IEEE Trans. Veh. Technol. XX (Xx), 1–13. http://dx.doi.org/10.1109/TVT.2018.2793889.

Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural Netw. 61, 85–117. http://dx.doi.org/10.1016/J.NEUNET.2014.09.003.

Scott-Parker, B., King, M.J., Watson, B., 2015. The psychosocial purpose of driving and its relationship with the risky driving behaviour of young novice drivers. Transp. Res. Part F: Traffic Psychol. Behav. 33, 16–26. http://dx.doi.org/10.1016/j.trf.2015.06.004.

Scott-parker, B., Watson, B., King, M.J., Hyde, M.K., 2013. A further exploration of sensation seeking propensity, reward sensitivity, depression, anxiety, and the risky behaviour of young novice drivers in a structural equation model. Accid. Anal. Prev. 50, 465–471. http://dx.doi.org/10.1016/j.aap.2012.05.027.

Shimosaka, M., Nishi, K., Sato, J., Kataoka, H., 2015. Predicting driving behavior using inverse reinforcement learning with multiple reward functions towards environmental diversity. In: IEEE Intelligent Vehicles Symposium, Proceedings, 2015– Augus(Iv). pp. 567–572. http://dx.doi.org/10.1109/IVS.2015.7225745.

Singh, R.R., Conjeti, S., Banerjee, R., 2013. A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals. Biomed. Signal Process. Control 8 (6), 740–754. http://dx.doi.org/10.1016/j.bspc.2013.06.014.

Sysoev, M., Kos, A., Guna, J., Pogačnik, M., 2017. Estimation of the driving style based on the users' activity and environment influence. Sensors (Switz.) 17 (10), 00. http://dx.doi.org/10.3390/s17102404.

Tang, J., Liu, F., Zhang, W., Ke, R., Zou, Y., 2018. Lane-changes prediction based on adaptive fuzzy neural network. Expert Syst. Appl. 91, 452–463. http://dx.doi.org/10.1016/j.eswa.2017.09.025.

Tango, F., Botta, M., 2013. Real-time detection system of driver distraction using machine learning. IEEE Trans. Intell. Transp. Syst. 14 (2), 894–905. http://dx.doi.org/10.1109/TITS.2013.2247760.

Taylor, J.E., 2018. The extent and characteristics of driving anxiety. Transp. Res. Part F: Psychol. Behav. 58, 70–79. http://dx.doi.org/10.1016/j.trf.2018.05.031.

Tchankue, P., Elizabeth, P., Africa, S., Elizabeth, P., Africa, S., Elizabeth, P., . Janetwessonnmmuacza, P.T., 2013. Using machine learning to predict the driving context whilst driving, (figure 1). pp. 47–55.

Tran, C., Doshi, A., Trivedi, M.M., 2012. Modeling and prediction of driver behavior by foot gesture analysis. Comput. Vis. Image Underst. 116 (3), 435–445. http://dx.doi.org/10.1016/j.cviu.2011.09.008.

Vetter, M., Schünemann, A.L., Brieber, D., Debelak, R., Gatscha, M., Grünsteidel, F., . Ortner, T.M., 2018. Cognitive and personality determinants of safe driving performance in professional drivers. Transp. Res. Part F: Traffic Psychol. Behav. 52, 191–201. http://dx.doi.org/10.1016/j.trf.2017.11.008.

Vicente, J., Laguna, P., Bartra, A., Bailón, R., 2016. Drowsiness detection using heart rate variability. Med. Biol. Eng. Comput. 54 (6), 927–937. http://dx.doi.org/10.1007/s11517-015-1448-7.

Vilac, A., Cunha, P., Car, B., 2017. Systematic literature review on driving behavior.

Vlahogianni, E.I., Barmpounakis, E.N., 2017. Driving analytics using smartphones: Algorithms, comparisons and challenges. Transp. Res. Part C: Emerg. Technol. 79, 196–206. http://dx.doi.org/10.1016/j.trc.2017.03.014.

Vlahogianni, E.I., Golias, J.C., 2012. Bayesian modeling of the microscopic traffic characteristics of overtaking in two-lane highways. Transp. Res. Part F: Traffic Psychol. Behav. 15 (3), 348–357. http://dx.doi.org/10.1016/j.trf.2012.02.002.

Waard, D. De, 1996. The measurement of drivers ' mental workload.

Wang, X., An, K., Tang, L., Chen, X., 2015. Short term prediction of freeway exiting volume based on SVM and KNN. Int. J. Transp. Sci. Technol. 4 (3), 337–352. http://dx.doi.org/10.1260/2046-0430.4.3.337.

Wang, Xi J., Chong, A., Li, L., 2017a. Driving style classification using a semisupervised support vector machine. IEEE Trans. Hum.-Mach. Syst. 47 (5), 650–660. http://dx.doi.org/10.1109/THMS.2017.2736948.

Wang, M.S., Jeong, N.T., Kim, K.S., Choi, S.B., Yang, S.M., You, S.H., . Suh, M.W., 2016. Drowsy behavior detection based on driving information. Int. J. Automot. Technol. 17 (1), 165–173. http://dx.doi.org/10.1007/s12239-016-0016-y.

Wang, X., Liu, Y., Guo, Y., Xia, Y., Wu, C., 2018a. Transformation mechanism of vehicle cluster situations under dynamic evolution of driver's propensity. Transp. Res. Part F: Traffic Psychol. Behav. http://dx.doi.org/10.1016/J.TRF.2018.08.011.

Wang, X., Liu, Y., Wang, J., Zhang, J., 2018b. Study on influencing factors selection of driver's propensity. Transp. Res. D (June), 0–1. http://dx.doi.org/10.1016/j.trd.2018.06.025.

Wang, D., Pei, X., Li, L., Yao, D., 2018c. Risky driver recognition based on vehicle speed time series. IEEE Trans. Hum.-Mach. Syst. 48 (1), 63–71. http://dx.doi.org/10.1109/THMS.2017.2776605.

Wang, E.G., Sun, J., Jiang, S., Li, F., 2017b. Modeling the various merging behaviors at expressway on-ramp bottlenecks using support vector machine models. Transp. Res. Procedia 25, 1327–1341. http://dx.doi.org/10.1016/j.trpro.2017.05.157.

Wang, J., Xu, W., Gong, Y., 2010. Real-time driving danger-level prediction. Eng. Appl. Artif. Intell. 23 (8), 1247–1254. http://dx.doi.org/10.1016/j.engappai.2010.01.001.

Wen, J., Li, S., Lin, Z., Hu, Y., Huang, C., 2012. Systematic literature review of machine learning based software development effort estimation models. Inf. Softw. Technol. 54 (1), 41–59. http://dx.doi.org/10.1016/j.infsof.2011.09.002.

Wen, D., Yan, G., Zheng, N.N., Shen, L.C., Li, L., 2011. Toward cognitive vehicles. IEEE Intell. Syst. 26 (3), 76–80. http://dx.doi.org/10.1109/MIS.2011.54.

WHO | Global status report on road safety 2015. 2015. Retrieved from http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/.

Windsor, T.D., Anstey, K.J., 2006. Interventions to reduce the adverse psychosocial impact of driving cessation on older adults. Clin. Interv. Aging 1 (3), 205–211, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18046872.

Wu, B.F., Chen, Y.H., Yeh, C.H., Li, Y.F., 2013. Reasoning-based framework for driving safety monitoring using driving event recognition. IEEE Trans. Intell. Transp. Syst. 14 (3), 1231–1241. http://dx.doi.org/10.1109/TITS.2013.2257759.

Xie, G., Qin, H., Hu, M., Ni, D., Wang, J., 2017. Modeling discretionary cut-in risks using naturalistic driving data. Transp. Res. Part F: Psychol. Behav. http://dx.doi.org/10.1016/j.trf.2017.11.022.

Xiong, X., Chen, L., Liang, J., 2018. A new framework of vehicle collision prediction by combining SVM and HMM. IEEE Trans. Intell. Transp. Syst. 19 (3), 699–710. http://dx.doi.org/10.1109/TITS.2017.2699191.

Xuan, Z., Shu, W., Ma, J., Yu, Q., Gao, Q., Yu, M., 2010. Identification of driver's braking intention based on a hybrid model of GHMM and GGAP-RBFNN. Neural Comput. Appl. 6 (1), 63–70. http://dx.doi.org/10.1007/s00521-018-3672-1.

Yanagihara, M., Uno, N., Nakamura, T., 2015. Latent class analysis for driving behavior on merging section. Transp. Res. Procedia 6, 259–271. http://dx.doi.org/10.1016/j.trpro.2015.03.020.

Yang, L., Ma, R., Zhang, H.M., Guan, W., Jiang, S., 2017. Driving behavior recognition using EEG data from a simulated car-following experiment. Accid. Anal. Prev. (October), 1–11. http://dx.doi.org/10.1016/j.aap.2017.11.010.

Yang, L., Ma, R., Zhang, H.M., Guan, W., Jiang, S., 2018. Driving behavior recognition using EEG data from a simulated car-following experiment. Accid. Anal. Prev. 116 (April), 30–40. http://dx.doi.org/10.1016/j.aap.2017.11.010.

Yeo, M.V.M., Li, X., Shen, K., Wilder-Smith, E.P.V., 2009. Can SVM be used for automatic EEG detection of drowsiness during car driving? Saf. Sci. 47 (1), 115–124. http://dx.doi.org/10.1016/j.ssci.2008.01.007.

Yu, J., Chen, Z., Zhu, Y., Jennifer Chen, Y., Kong, L., Li, M., 2017. Fine-grained abnormal driving behaviors detection and identification with smartphones. IEEE Trans. Mob. Comput. 16 (8), 2198–2212. http://dx.doi.org/10.1109/TMC.2016.2618873.

Yuan, W., Li, Z., Wang, C., 2018. Lane-change prediction method for adaptive cruise control system with hidden Markov model. Adv. Mech. Eng. 10 (9), 1–9. http://dx.doi.org/10.1177/1687814018802932.

Zhang, T., Hajiseyedjavadi, F., Wang, Y., Samuel, S., Qu, X., 2018. Training interventions are only effective on careful drivers, not careless drivers. Transp. Res. Part F: Psychol. Behav. 58, 693–707. http://dx.doi.org/10.1016/j.trf.2018.07.004.

Zhang, Y., Kumada, T., 2018. Automatic detection of mind wandering in a simulated driving task with behavioral measures. PLoS One 13 (11), e0207092. http://dx.doi.org/10.1371/journal.pone.0207092.

Zhang, Q., q. Xu, G., Wang, M., Zhou, Y., Feng, W., 2014. Webcam based non-contact real-time monitoring for the physiological parameters of drivers. In: The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent. pp. 648–652. http://dx.doi.org/10.1109/CYBER.2014.6917541.

Zhang, G., Yau, K.K.W., Zhang, X., Li, Y., 2016. Traffic accidents involving fatigue driving and their extent of casualties. Accid. Anal. Prev. 87, 34–42. http://dx.doi.org/10.1016/j.aap.2015.10.033.

Zhao, C., Gao, Y., He, J., Lian, J., 2012. Recognition of driving postures by multiwavelet transform and multilayer perceptron classifier. Eng. Appl. Artif. Intell. 25 (8), 1677–1686. http://dx.doi.org/10.1016/j.engappai.2012.09.018.

Zhao, M., Kathner, D., Jipp, M., Soffker, D., Lemmer, K., 2017. Modeling driver behavior at roundabouts: Results from a field study. In: IEEE Intelligent Vehicles Symposium, Proceedings, (Iv). pp. 908–913. http://dx.doi.org/10.1109/IVS.2017.7995831.

Zheng, C., Lichtenstein, P., Brian, M.D., 2014. Serious transport accidents in adults with attention-deficit/hyperactivity disorder and the effect of medication a population-based study. 71, (3), pp. 319–325. http://dx.doi.org/10.1001/jamapsychiatry.2013.4174.

Zhu, X., Yuan, Y., Hu, X., Chiu, Y.C., Ma, Y.L., 2017a. A Bayesian network model for contextual versus non-contextual driving behavior assessment. Transp. Res. Part C: Emerg. Technol. 81, 172–187. http://dx.doi.org/10.1016/j.trc.2017.05.015.

Zhu, X., Yuan, Y., Hu, X., Chiu, Y., Ma, 2017b. A Bayesian network model for contextual versus non-contextual driving behavior assessment. Transp. Res. C 81, 172–187. http://dx.doi.org/10.1016/j.trc.2017.05.015.

Zicat, E., Bennett, J.M., Chekaluk, E., Batchelor, J., 2018. Cognitive function and young drivers : The relationship between driving, attitudes, personality and cognition. Transp. Res. Part F: Psychol. Behav. 55, 341–352. http://dx.doi.org/10.1016/j.trf.2018.03.013.