

RESEARCH

Open Access



On combining Big Data and machine learning to support eco-driving behaviours

Giovanni Delnevo* , Pietro Di Lena, Silvia Mirri, Catia Prandi and Paola Salomoni

*Correspondence:
giovanni.delnevo2@unibo.it
Department of Computer
Science and Engineering,
University of Bologna,
Bologna, Italy

Abstract

A conscious use of the battery is one of the key elements to consider while driving an electric vehicle. Hence, supporting the drivers, with information about it, can be strategic in letting them drive in a better way, with the purpose of optimizing the energy consumption. In the context of electric vehicles, equipped with regenerative brakes, the driver's braking style can make a significant difference. In this paper, we propose an approach which is based on the combination of big data and machine learning techniques, with the aim of enhancing the driver's braking style through visual elements (displayed in the vehicle dashboard, as a Human–Machine Interface), actuating eco-driving behaviours. We have designed and developed a system prototype, by exploiting big data coming from an electric vehicle and a machine learning algorithm. Then, we have conducted a set of tests, with simulated and real data, and here we discuss the results we have obtained that can open interesting discussions about the use of big data, together with machine learning, so as to improve drivers' awareness of eco-behaviours.

Keywords: Machine learning, Human–Machine Interface, Eco-driving behaviours

Introduction

The automotive industry is currently facing a moment of radical change. According to the 2018 edition of the future automotive industry structure (FAST) study,¹ conducted by Oliver Wyman and the German Automotive Association, there are seven main factors that will drive this sector over the next decade (until 2030), thanks to the exploitation of digitization, artificial intelligence (AI), and machine learning (ML). These identified factors are: (i) connected vehicles, (ii) autonomous vehicles, (iii) electric mobility, (iv) digital industry, (v) new distribution channel pay-per-use, (vi) changes in customer structure, and (vii) new concepts for Human–Machine Interface.

In this emerging context, the Internet of Things paradigm can be exploited, taking advantage of interconnected sensors, which are more and more accurate, smaller, and powerful [1]. Accordingly, vehicles are increasingly being equipped with a variety of sensors, provided with integrated connectivity, that can monitor different components and situations, such as the engine, the driving style or the environmental conditions. Indeed, these sensors lead to the generation of large volumes of data about the vehicle,

¹ <https://www.oliverwyman.com/our-expertise/insights/2018/jun/fast-2030.html>.

that can be analyzed to reveal patterns and trends, referring to the so-called big data [2, 3]. In fact, the relevant impact of big data also applies to the automotive context, where data gathered from the vehicle can be exploited in different ways, ranging over predictive maintenance and advanced real-time analytics [4]. These emerging technologies and paradigms are particularly interesting with regard to the electric vehicles segment, the spread of which is constantly growing thanks to several factors, such as a reduced impact on the environment [5], a faster and more accurate motor torque generation [6], and lower long-term costs of ownership [7]. In this context, the streams of data, collected by exploiting interconnected sensors, could be a unique chance to mitigate the so-called “range anxiety”, defined as the fear of fully depleting a battery electric vehicles (BEVs) battery in the middle of a trip, leaving the driver stranded [8]. In fact, despite the several advantages described above, the use of electric vehicles still presents some limitations, including:

- Limited traveling distance: currently the distance that can be covered by electric cars is limited.
- Lack of recharge points: electric vehicles require charging stations, but there is a significant lack of infrastructures. While this problem is mitigated in an urban context, it becomes critical in long distance travels, since they require a network of charging stations strategically located.
- Long recharge time: unlike gasoline-powered cars, electric ones take hours to get fully charged.

Considering these open issues, it is clear the need for optimizing energy consumption in electric vehicles. With this purpose in mind, we focused on different driving styles, evaluating which one leads to the efficient use of the battery charge. From an empirical analysis, and according to the current literature [9], we found out that the regenerative brake has a significant impact in this sense. In fact, electric vehicles are equipped with two different brakes: the regenerative and the friction ones. In particular, the activation of the regenerative brake implies recharging the battery, on the contrary of the activation of the friction one. Hence, a more aware use of the brake can significantly change the battery level. Starting from these considerations, our aim is to provide drivers with an integrated solution that fosters eco-driving behaviours, and in particular, that enhances an optimized use of the regenerative brake.

This study was conducted within the TIME project² (which stands for “Tecnologie Integrate per la Mobilità Elettrica” in Italian language, meaning “integrated technologies for electric mobility”). The whole TIME project is focused on the development of complete technological systems (powertrains) with integrated solutions for electric vehicles. The main purpose is to identify a methodology to build the integrated components of an efficient electric vehicle, obtained by converting a petrol car (retrofit) or by directly being installed in a new vehicle. While the TIME project involves several research areas, in this paper we are mainly investigating how to combine big data and machine learning in order to predict the activation of the friction brake. This aims to optimize energy consumption in electric vehicles, improving drivers’ consciousness, and then let them

² <https://site.unibo.it/time/en>.

better deal with the “range anxiety”. Then, the in-vehicle Human–Machine Interface can take advantage of such a real-time prediction, equipping the driver with more accurate and complete information about his/her braking styles, thus enhancing eco-driving behaviours.

Hence, here we present the design and implementation of a prototype system, integrated in a real electric vehicle. Such a prototype exploits: (i) vehicles data collection, (ii) a machine learning algorithm, and (iii) the Human–Machine Interface (i.e., the car dashboard). In this scenario, we have taken advantage of the use of the most adequate machine learning algorithm so as to predict the friction brake activation. Such an algorithm has been evaluated on simulated and real datasets of driving cycles.

Summarizing, the main contributions of our study are: (i) the design and implementation of a system that uses big data gathered by in-vehicle sensors and components to provide drivers with real-time visual feedback to optimize the battery energy consumption; (ii) the identification of the machine learning algorithm that best suits our need to predict the friction brake activation. Moreover, we assessed the feasibility of our approach by using both simulated and real data (collected from the field, integrating the system in a real electric vehicle).

The remainder of the paper is structured as follows. “[Background and related work](#)” section briefly presents studies conducted in the field, which analyze and exploit topics and research projects similar to the ones of interest in our work. “[Problem statement](#)” section details the context and the research questions that motivated our study, then, “[Methods](#)” section presents our approach, detailing the dataset used, the system overview, and describing the machine learning module. “[Results and discussion](#)” section illustrates the results obtained with the consequent discussion. Finally, “[Conclusion and future work](#)” section concludes this paper, presenting some final remarks and future work.

Background and related work

In this Section, we present some projects and studies related to our approach, introducing some cases based on sensed data and machine learning techniques and on increasing the user awareness about driving styles and eco-driving behaviours in the automotive context. Finally, some works that exploit gamification strategies to support drivers and users on the move are briefly introduced.

The growing number of sensors integrated in vehicles, in addition to the possibility of having a complete view of the vehicle status, thanks to the messages exchanged on the controller area network (CAN) bus, made possible exploiting big data to obtain useful explanation and/or insights behind different phenomena. The data gathered from in-vehicle sensors can be variegated, ranging from the speed of the vehicle to the percentage of the battery charge. In this context, it emerges the opportunity to feed machine learning algorithms with these data, in order to extract new knowledge, as witnessed by recent studies. For instance, the authors of [10] developed ML_EMO_HEV, a framework that uses big data about motor torque, power, and vehicle speed, together with machine learning algorithms with the aim of predicting roadway types, traffic congestion levels, and driving trends. Then, starting from this prediction, it exploits another algorithm so as to learn optimal energy settings. This way, a real-time quasi-optimal control of energy

flow in a hybrid electric vehicle (HEV) can be obtained. With the same aim in mind (optimizing the engine performance), [11] tried to determine optimal biodiesel ratio using advanced machine learning methods, exploiting a dataset composed by engine torque and fuel blends. The authors of this work compared different techniques to create engine models based on experimental data, in particular: Extreme Learning Machine, Least-Squares Support Vector Machine, and Radial-Basis Function Neural Network.

Different aspects were analyzed by [12, 13] in their studies. In particular, [12] exploited machine learning to model human recognition of vehicle-driving situations. The purpose of such research is to provide physical context to mitigate unnecessary distractions, allowing the driver to maintain focus during periods that require high concentration. While [13], proposed a system to detect early onset of fatigue in drivers using heart rate variability, since fatigue can acutely impair driver's alertness and performance. In particular, the detection was performed by a neural network, trained and tested on a set of electrocardiogram (ECG) data recorded under laboratory conditions. Finally, Ezzini et al. [14] addressed the problem of driver identification using in-vehicle sensors measurements. They investigated both the minimum learning and classification times required, exploiting several algorithms like Extra Tree, Random Forest, KNN, and SVM. Monitoring heart rate is also at the basis of [15], where the authors have assessed the correlation of driving behaviour with heart rate changes. The authors considered that there are situations where the emotional conditions of drivers vary due to different reasons (beyond their control or not). With the aim of controlling such variations, the authors designed and developed an Android application to real-time monitoring both physiological data from the driver and diagnostic data from the vehicle, so as to investigate their correlation. In particular, an OBD-II connector has been used to obtain the vehicle diagnostic data, while non-invasive biomedical sensors packaged in belts and smart watches have been used to gather data about the drivers' heart rate, focusing on the connection between driving aggressiveness and heart rate.

The study and the identification of driving styles are at the basis of [16] and of [17]. The authors integrated both data mining techniques and neural networks with the aim of generating a classification of driving styles by analyzing the driver's behavior along specific routes. The authors implemented a neural network-based algorithm, that can characterize the type of road on which the vehicle is moving, as well as drivers' styles, taking into account parameters such as speed, acceleration, and revolutions per minute of the engine (rpm). The final aim of this work is to assist the driver, so as to avoid or at least to limit bad habits in their driving behavior, while supporting useful tips to improve fuel economy, by exploiting the evolution of mobile terminals and the availability of a standard interface to access car data. The results presented in these papers show that neural networks were able to achieve a high degree of exactitude at classifying both road and driver types based on user traces.

The prediction of consumption in vehicles while supporting drivers in optimizing vehicles consumption are the purpose of [18–20]. In particular, [18] proposes a methodology to real-time calculate the consumption and the environmental impact of spark ignition and diesel vehicles from a set of variables, such as engine fuel rate, speed, mass air flow, absolute load, and manifold absolute pressure. All these data can be collected from the vehicle's Electronic Control Unit (ECU). The methodology

has been defined with the aim of supporting eco-driving behaviors in drivers, by improving fuel economy. This work shows, through data mining, how much the driving style really affects (both negatively or positively) the fuel consumption, as well as the improvement or the reduction of greenhouse gas emissions generated by vehicles. While predicting the consumption of electric vehicles and their electricity demand are at the basis of [20]. In fact, it proposes an approach based on having specific knowledge about information on the number of vehicles that are going to recharge their batteries at a particular point and instant and on the available charging points. Obtained results report that the proposed approach can predict the electricity requirements of the electric vehicles that are expected to recharge their batteries up to 180 minutes in advance [20].

Both in the industrial and academic world, there are a lot of examples of the adoption of Gamification strategies with different purposes, such as fostering eco-driving behaviours. Different car manufacturers developed several Gamification mechanisms and embedded them in their models. Diewald et al. [21] provided a comprehensive review of such applications. From Volkswagen's BlueMotion Roulette to Ford's SmartGauge with EcoGuide, passing through the Kia's ECOdynamics and Fiat's eco:Drive. With regard to academic research, there are plenty of works on this subject, even if with different aims [22]. Corcoba Magaña and Munoz-Organero [23] developed a mobile application to provide a vocal eco-driving assistant. The application employs two types of feedback, one real-time (while the user is driving) and the other one at the end of the trip. The first one is activated when a driver does not comply with predefined eco-driving rules. In this case, the system tries to identify a group of drivers in a similar situation, exploiting information about weather, traffic, time, and vehicle location. Then, using fuzzy logic, each attribute of the user telemetry is compared with the ones of the best driver of the cluster. If the difference exceeds a threshold, the relative eco-advice will be provided. The second type of feedback, instead, is provided at the end of every trip, and generally reports about his/her driving style. In particular, it consists in the unlocking of achievements, that are one of the most common Gamification mechanisms. In this study, achievements have to encourage drivers to save fuel and motivate them to continually use the eco-driving assistant. Vaezipour et al. [24], through a user-centered design approach, evaluated a series of conceptual designs of in-vehicle gamified interfaces. A particular focus has been paid to the ability to increase drivers' acceptance of feedback from such interfaces in order to improve fuel efficiency and promote safe driving.

An example exploiting real-time analysis of data gathered by in-vehicle sensors is the study presented by Steinberg [25], that proposed a game concept to encourage safer driving behaviours, in particular avoiding tailgating, the practice of driving too close to a frontward vehicle. They defined an area around cars and they implemented a real-time data analysis with a zombie narrative as theme [26], so that crossing the area results in the infection of the car.

All these studies differ from the approach we propose and the objectives which are at the basis of our study. In fact, our approach seeks to take advantage of sensed big data and machine learning strategies to understand the user's driving style. Then, real-time feedback through graphical elements is prompted to induce a more efficient driving style and management of the electric vehicle.

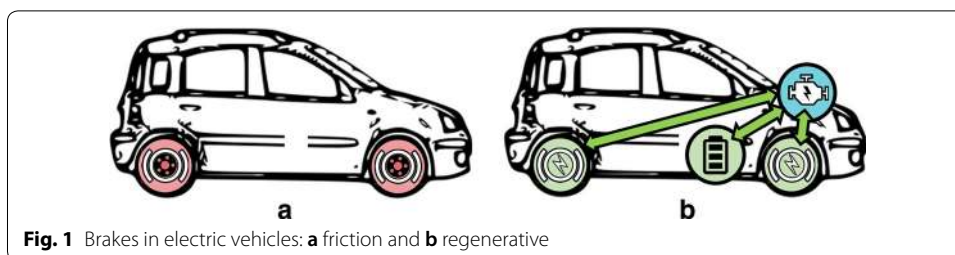


Fig. 1 Brakes in electric vehicles: **a** friction and **b** regenerative

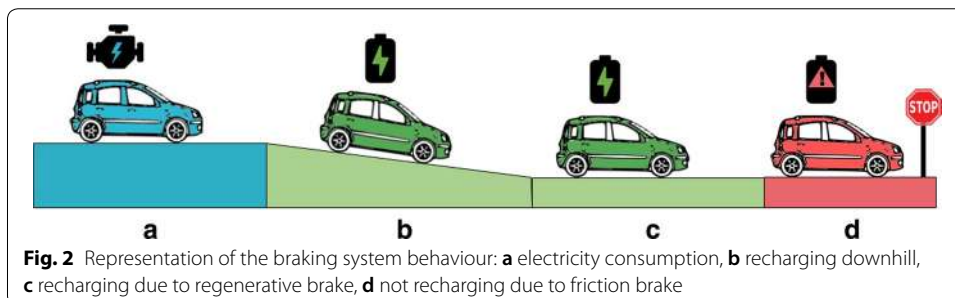


Fig. 2 Representation of the braking system behaviour: **a** electricity consumption, **b** recharging downhill, **c** recharging due to regenerative brake, **d** not recharging due to friction brake

Problem statement

Electric vehicles are equipped with two different brakes. Figure 1 presents how the activation of the two brakes impact the battery in terms of energy consumption optimization. In particular, the case a in Fig. 1 represents the activation of the friction brake, that consists of a pad pushing on a disc to create friction, resulting in a neutral effect on the battery; while case b shows the positive impact on the battery of the activation of the regenerative brake, where the vehicle’s momentum is used to turn the motor into a generator and to recover energy in the battery that would be, otherwise, lost as heat.

Figure 2 describes the behavior of the vehicle. While running, the vehicle simply consumes the electricity stored in the batteries (case a). They can be recharged in two ways. The first one is through the engine brake when the vehicle is downhill (case b). The second is during a brake. In the TIME prototype, as in many electric vehicles, light or moderate braking takes advantage of the regenerative brake (case c) while heavy braking will also engage the friction braking system (case d). In the latter situation, the energy is not fully recovered but it is partially lost as heat. Hence, the braking style that allows maximizing the recovered energy consists of using only and exclusively the regenerative brake (i.e., reducing the red area of Fig. 2, and increasing, consequently, the green one). Since a threshold (e.g., brake pedal position), that defines the friction brake activation, does not exist, we want to predict its activation exploiting in-vehicle sensors and components data, and machine learning techniques. Starting from these predictions, the system should provide real-time suggestions and indications with the aim of improving the driver’s braking style, taking advantage only of the regenerative brake, with a consequent increase of the recovered energy.

Given these premises, with this study, we intend to prove that it is possible to effectively exploit big data gathered from in-vehicle sensors and components to provide drivers with real-time feedback aimed to optimize their braking style, and, accordingly, the energy consumption.

In more detail, the research questions that drove our study are the following ones:

1. What is the machine learning algorithm that best allows to extract meaningful information from our dataset and address our objective to predict the friction brake activation, keeping into account the need of continuously train the model and the limited computational power?
2. Is the identified algorithm able to predict the friction brake activation with *good* performances?
3. Are the performance of the selected algorithm comparable to the ones of other common machine learning algorithms?
4. Is the algorithm, trained on simulated data, able to predict the friction brake activation using sensed real data?

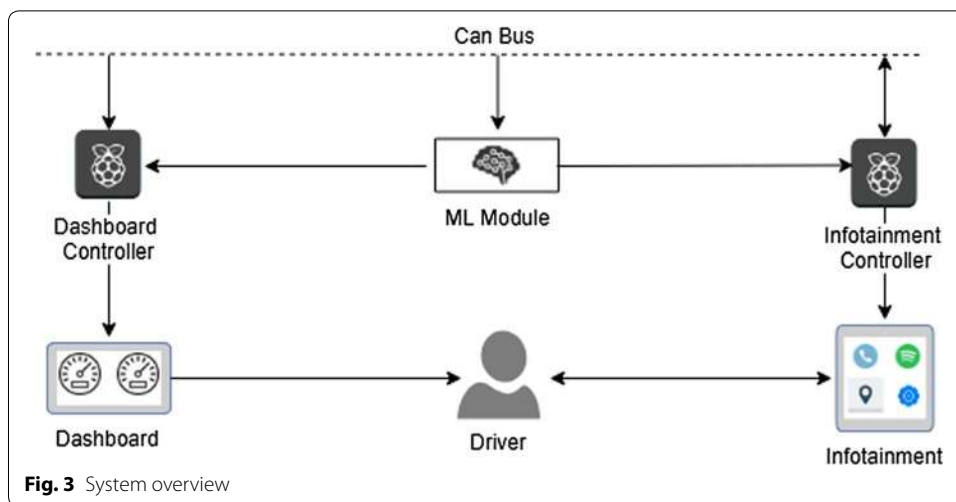
Methods

In this section, as a main contribution of the paper, we describe our approach, detailing the dataset, the system architecture, and the machine learning algorithm.

Dataset description

The dataset is composed of (i) simulated data, used for the training, and (ii) real data, used as a further validation set. The simulated data were derived from Matlab Simulink simulator of the TIME's vehicle. The simulations were conducted on six heterogeneous driving cycles: urban (ECE-15, ARTEMIS urban, Bologna urban, Rome urban), rural (ARTEMIS rural road) and motorway (ARTEMIS motorway). The Urban Driving Cycle ECE-15 was introduced first in 1970 as part of ECE vehicle regulations. The cycle has been designed to represent typical driving conditions of busy European cities and is characterized by low engine load, low exhaust gas temperature, and a maximum speed of 50 km/h. The ARTEMIS cycles were defined within the ARTEMIS research project, supported by the European Commission to set-up and improve the European methods for estimating and inventorying the pollutant emissions from the transports. In particular, we used the driving cycles (urban, rural and motorway ones) developed by [27]. Finally, Bologna and Rome urban are both driving cycles created within the TIME project, starting from a set of GPS data in Bologna and Rome. The first one consists in Bologna avenues, the latter one is mainly an urban cycle, but it also includes a part of the "*Grande Raccordo Anulare*" (orbital motorway that encircles Rome). With regards to the real data, since the vehicle has not been yet type-approved, it was not possible to collect a large amount of data. For this reason, we were able to collect real data related to a cycle of 1.2 km performed on a private road.

The features collected are: (i) gas pedal position (expressed in degrees), (ii) percentage of use of the electric brake, (iii) percentage of use of the mechanic brake, (iv) speed of the vehicle (expressed in km/h), (v) percentage of the battery charge. While the first four features are closely related to the activation of the friction brake, the last one is not. We choose to include it since it may be useful in view of a possible future integration with a different approach based on the prediction of the future state of charge of the battery, as described in [28]. All the features are sampled at a rate of 0.1 s (100 ms).



We aggregated the data according to two main parameters: the input temporal window width (i.e., how many samples are used as input) and the time offset (i.e., the future instant of the prediction). We evaluated different temporal-window widths (from 0.1 to 2.3 s) and different time offsets (from 0.1 to 2.3 s). More details about the motivation behind these intervals value are provided in the “[Prediction of the friction brake activation](#)” subsection. Only temporal windows included within a braking are considered (i.e., the driver is braking). The datasets contain from 1500 to 6000 sample windows, with an average of 4300.

System overview

Figure 3 presents an overview of the proposed system. All the data about the vehicle state are transmitted on the CAN (Controlled Area Network) bus, including the input features of the machine learning module. The vehicle has two tangible-car-infotainment-systems (TCIS). One of them is the car dashboard, and it is directly ahead of a vehicle’s driver (named car dashboard), displaying instrumentation and controls for the vehicle’s operation. The other one is the control panel that delivers entertainment and information content (named infotainment system). It allows calling someone, playing music, etc. Each TCIS has its own controller, that consists of a single-board computer. The dashboard controller can only read from the CAN bus and visualizes the correspondent information on the display. Instead, the infotainment controller can both read and write on the can bus. Thus, its display is touch-screen, since it has to manage also the driver’s input.

Essentially, considering our approach, the primary information to display drivers, in order to affect their driving style, is the real-time prediction of the imminent activation of the friction brake. This visual information can be displayed in both the car dashboard and the infotainment system since the machine learning (ML) module can interact with both the interfaces. Even if the car dashboard has strong constraints imposed by various regulations [29], we decided to display the prediction in the car dashboard, since such information has to be provided real-time to the driver. In this way, the information is immediately visible to the driver, in a clear and not distracting way.

Machine learning module

The main purpose of the ML module is to predict how likely (calculating a value between 0 and 1) the friction brake will be active in a future instant. To address this objective, answering to the first research question, we evaluated the use of different machine learning algorithms. Two main factors drove our final choice.

First of all, the module should be able to continuously train the model with the new data generated while the vehicle is running, without starting from scratch (i.e., without employing all the data used for training until then). This is a crucial aspect since the different components developed within the TIME project are designed to be suitable not only for installation on new vehicles but also on existing vehicles produced by different car manufacturers. So the model, designed to address this constraint, must adapt itself to different vehicles. Thus, each driver has his/her own driving styles, that can also change during time, so the model must also adapt itself to different driving styles. For this reason, we excluded the use of traditional approaches such as linear regression, gradient boosting, etc.. that do not support incremental learning.

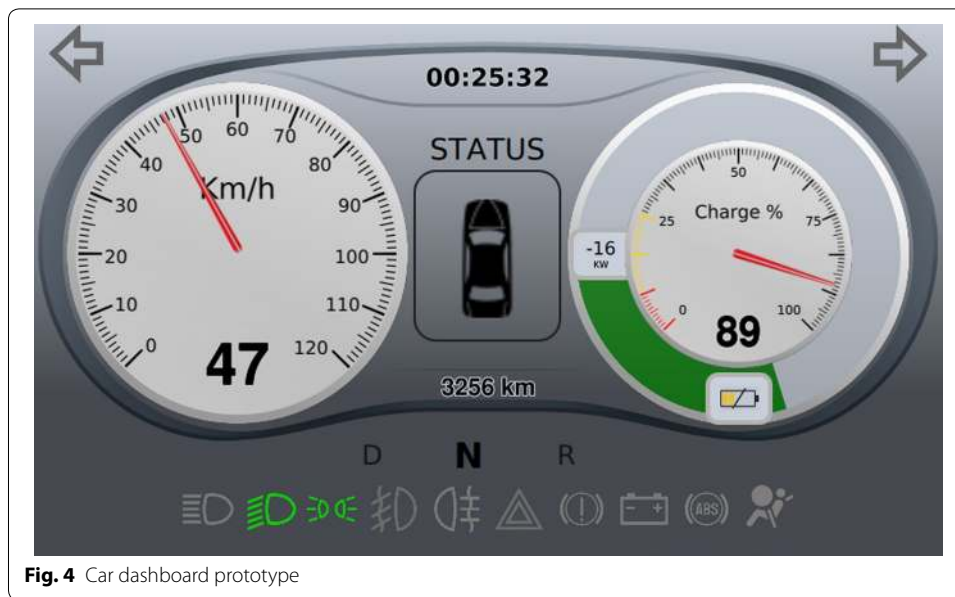
The second factor to take into account when selecting the more suitable algorithm is the available computational power. The Dashboard controller (Fig. 3) runs on a single-board computer, like a Raspberry PI 3 Model B, that has a Quad Core 1.2GHz Broadcom BCM2837 64bit CPU with 1GB RAM. Thus, the sample rate is 0.1 s. Since our system aims to be as real-time as possible, the model has to predict the activation of the friction brake with the lowest sample time (i.e., 0.1 s). These constraints exclude deep learning approaches like Recurrent Neural Networks, that are the best option when working with sequential data [30]. Hence, we decided to employ a multilayer perceptron (MLP) with one hidden layer composed of 20 neurons. Such a neural network can be trained on new samples, adapting itself to different cars and drivers, even if it runs on a single-board computer, and it is able to compute the result every 0.1 s.

Then, we focused on how to visualize the feedback about the activation of the friction brake in the car dashboard. An important aspect, that we took into consideration, is the cognitive load. It is particularly relevant since driving is one of the most complex activity. Different studies [31, 32] investigated the impact of car interfaces in terms of the distraction of the driver. Based on the findings of past studies, we decided to use an icon with a battery that appears only during braking, and its color changes with respect to the neural network prediction. In particular, it can be:

- green (low probability of activation, range [0, 0.33])
- yellow (medium probability of activation, range [0.33, 0.67])
- red (high probability of activation, range [0.67, 1]).

The whole car dashboard was developed using QT,³ standard “de facto” of automotive HMI. A screenshot of the car dashboard is visible in Fig. 4, where braking is simulated. It is possible to notice the yellow icon of the battery, which indicates a medium probability of activation of the friction brake.

³ <https://www.qt.io/qt-automotive-suite/>.



Results and discussion

In this section we evaluate and compare our method to predict the friction brake activation. Finally, we validate our approach using both simulated and real data, collected in the field.

Prediction of the friction brake activation

To answer the second research question, we focused on the simulation data. We evaluated the performances of the model varying both the time-width and the time-offset from 0.1 to 2.3 s. The tests have been performed in cross-validation on the data of the six simulated driving cycles. In the simulations, we employed the same vehicle characteristics and the same driving style. A crucial aspect of our approach is how far in advance the activation of the friction brake can be predicted, so that, through the car dashboard battery icon, the driver can be advised far in advance.

In Table 1 we reported the average AUC on the six driving cycles. As shown, the neural network has the best performances with larger time-width and lower time-offset. However, the neural network has good prediction capabilities (0.85 and 0.82) also when the time-offset is large (e.g., 2.1 s and 2.3 s). It is possible to notice that already with time-width equal to 2.3 s, results get worse compared to the previous time-width in many cases. Since the driver's reaction time is on the average 2 s [33, 34] we decided to fix the time-offset to 2.1 s. With such time-offset, the best model is the one with a temporal width equal to 2.3 s.

Performance comparison

Subsequently, to answer the third research question, we checked if the performances of the selected algorithm were comparable to the ones of other common machine

Table 1 Average AUC varying width and offset

Width/offset	0.1	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9	2.1	2.3
0.1	0.978	0.932	0.894	0.871	0.843	0.830	0.828	0.809	0.810	0.798	0.759	0.774
0.3	0.985	0.943	0.909	0.885	0.859	0.841	0.834	0.835	0.806	0.797	0.799	0.787
0.5	0.986	0.952	0.918	0.886	0.870	0.849	0.822	0.832	0.813	0.794	0.812	0.751
0.7	0.987	0.955	0.922	0.897	0.884	0.853	0.834	0.825	0.815	0.823	0.785	0.776
0.9	0.990	0.959	0.931	0.916	0.897	0.876	0.844	0.839	0.823	0.812	0.786	0.775
1.1	0.994	0.973	0.950	0.924	0.909	0.886	0.850	0.843	0.829	0.824	0.795	0.801
1.3	0.994	0.974	0.947	0.929	0.908	0.854	0.869	0.847	0.830	0.800	0.814	0.797
1.5	0.994	0.974	0.948	0.932	0.908	0.858	0.849	0.848	0.829	0.815	0.819	0.816
1.7	0.993	0.974	0.950	0.933	0.888	0.867	0.860	0.844	0.830	0.800	0.828	0.827
1.9	0.994	0.973	0.948	0.930	0.908	0.873	0.854	0.877	0.808	0.814	0.835	0.801
2.1	0.992	0.969	0.950	0.933	0.913	0.878	0.883	0.833	0.856	0.826	0.841	0.812
2.3	0.993	0.971	0.948	0.925	0.897	0.870	0.878	0.835	0.856	0.801	0.855	0.822

learning algorithms. It is worth to notice that we already discussed the reasons behind our choice in "Machine learning module" section. However, we want to compare the performances of the MLP to be sure that other algorithms do not outperform it.

In particular, we evaluated the performances of the following ones:

- Linear regression (LR)
- K-nearest neighbors (KNN)
- Classification and regression tree (CART)
- Support vector regression (SVR)
- Adaptive boosting (AB)
- Gradient boosting (GB)
- Random forest (RF)

We fixed the time-width and the time-offset respectively to 2.3 s and 2.1 s, as explained in the previous section. The tests have always been performed in cross-validation on the data of the six simulated driving cycles. The results are shown in Fig. 5, in which the box plot for each algorithm is reported, and in Table 2, where the performances on the test path are reported for each algorithm. As shown, our MLP has generally better performances than the other approaches. In particular, it works better than almost all the others except for the adaptive boosting and the gradient boosting which are slightly more accurate on average.

In conclusion, the MLP has proved to be the right choice, since it works better than many of the presented approaches, in addition to presenting all the advantages described in "Machine learning module" section.

Evaluation on real data

To answer the fourth and final question, we also considered the prediction capabilities of the model using real data related to the test cycle directly sampled from the first prototype of the TIME vehicle (i.e. not Simulink samples). Figure 6 shows the ROC curve on the six simulating driving cycles (continuous lines) and on real data sampled from the

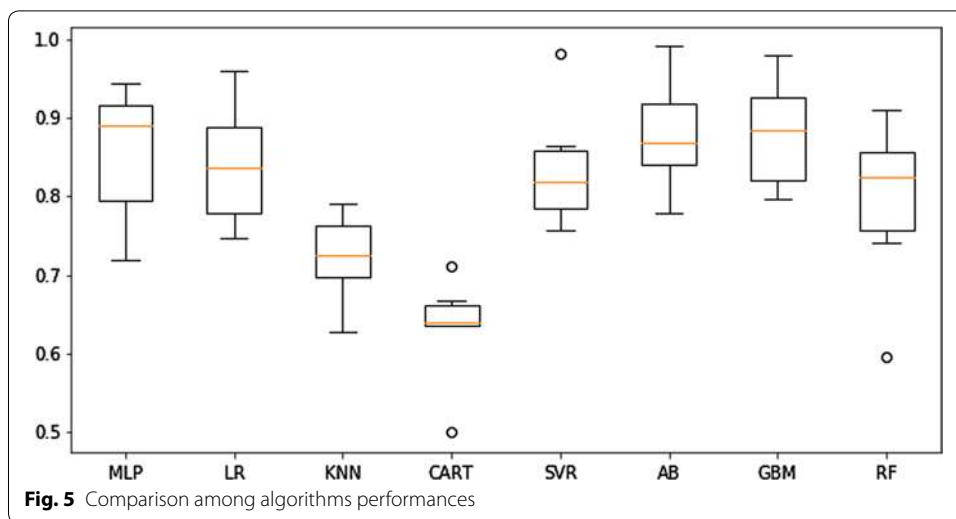


Table 2 Comparison among algorithms performances on the six driving cycles

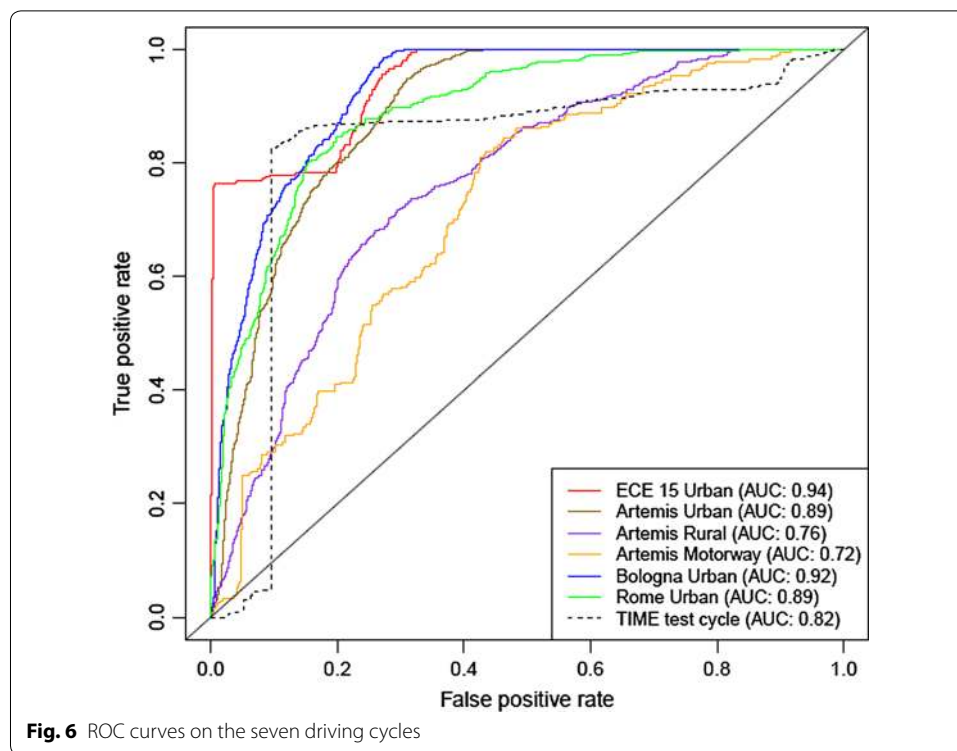
Path	MLP	LR	KNN	CART	SVR	AB	GBM	RF
1	0.944	0.960	0.774	0.639	0.982	0.991	0.980	0.911
2	0.889	0.841	0.718	0.668	0.782	0.861	0.868	0.802
3	0.764	0.746	0.690	0.635	0.756	0.779	0.797	0.741
4	0.719	0.760	0.628	0.499	0.796	0.834	0.805	0.596
5	0.923	0.904	0.790	0.640	0.864	0.932	0.935	0.859
6	0.891	0.830	0.733	0.711	0.840	0.875	0.899	0.847
avg	0.855	0.840	0.722	0.632	0.837	0.879	0.881	0.793

TIME vehicle, named TIME test cycle (dashed line), for temporal-window of 2.3 s and time-offset of 2.1 s. As shown, the prediction performances of the neural network model are also good when applied to real data, with an AUC equal to 0.82. Although our real test case is quite limited (as explained in "Dataset description" section), this confirms that the implemented neural network model can predict the activation of the friction brake with good performances also when fed with real data.

As shown in Fig. 6, the prediction capabilities are overall good on all driving cycles with lower performances on the rural and motorway cycles. Since these two driving cycles are the two more under-represented in our training set, we can expect better prediction results with larger datasets including more data in these driving conditions.

Conclusion and future work

This paper describes a prototype we have designed and deployed on a real electric vehicle, with the aim of optimizing the consumption of energy battery in electric vehicles, exploiting big data gathered from in-vehicle sensors and components. A neural network is employed to predict the activation of the friction brake in order to visualize this information in the HMI, and foster eco-driving behaviours. The tests, conducted using both simulated and real data, highlighted how it is possible to predict the activation of the



friction brake with a good accuracy, in time to allow the driver to correct his/her driving style, avoiding wasting energy and optimizing the battery charge.

Even if the results obtained are quite promising, it is worth mentioning some limitations of this study. The main one is surely the use of a limited set of real data. Another one is the type of driving cycles, since there were only one rural and one motorway, contrasted with four urbans. Such limitations can be overcome once the vehicle will be type-approved, simply conducting extensive tests sessions.

Finally, we plan to combine our approach with some elements and aspects typical of the gamification strategy, to improve the driver's involvement. In fact, we plan to take advantage of the gamification paradigm, that has already proved to be useful and effective applied to the automotive context, to present the information about the energy consumption in a more engaging way.

Abbreviations

HMI: Human–Machine Interface; FAST: future automotive industry structure; BEV: battery electric vehicles; TIME: Tecnologie Integrate per la Mobilità Elettrica (integrated technologies for electric mobility); HEV: hybrid electric vehicle; CAN: controller area network; TCIS: tangible-car-infotainment-systems; ML: machine learning; MLP: multilayer perceptron; LR: linear regression; KNN: K-nearest neighbors; CART: classification and regression tree; SVR: support vector regression; AB: adaptive boosting; GB: gradient boosting; RF: random forest; AUC: area under the curve; ROC: receiver operating characteristic.

Acknowledgements

The authors thank all the TIME researchers, with a special mention to Prof. Claudio Rossi, who is the leader of the project, and Marco Bertoldi. Another special thank goes to Dr. Matteo Bezzi, Dr. Francesco Ceroni, Dr. Michele Paci, Dr. Marco Di Felice, and Prof. Massimo Lanzoni for their precious support.

Authors' contributions

All the authors contributed equally to this manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Emilia Romagna region under Grant 737751 TIME.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 9 May 2019 Accepted: 7 July 2019

Published online: 22 July 2019

References

1. Atzori L, Iera A, Morabito G. The internet of things: a survey. *Comput Netw.* 2010;54(15):2787–805. <https://doi.org/10.1016/j.comnet.2010.05.010>.
2. Lee I. Big data: dimensions, evolution, impacts, and challenges. *Bus Horiz.* 2017;60(3):293–303. <https://doi.org/10.1016/j.bushor.2017.01.004>.
3. Wamba SF, Akter S, Edwards A, Chopin G, Gnanzou D. How 'big data' can make big impact: findings from a systematic review and a longitudinal case study. *Int J Prod Econ.* 2015;165:234–46. <https://doi.org/10.1016/j.ijpe.2014.12.031>.
4. Johanson M, Belenki S, Jalminger J, Fant M, Gjertz M. Big automotive data: leveraging large volumes of data for knowledge-driven product development. In: 2014 IEEE international conference on Big Data (Big Data). New York: IEEE; 2014. <https://doi.org/10.1109/bigdata.2014.7004298>.
5. Choi H, Shin J, Woo J. Effect of electricity generation mix on battery electric vehicle adoption and its environmental impact. *Energy Policy.* 2018;121:13–24. <https://doi.org/10.1016/j.enpol.2018.06.013>.
6. Ehsani M, Gao Y, Longo S, Ebrahimi K. Modern electric, hybrid electric, and fuel cell vehicles. Boca Raton: CRC Press; 2018.
7. Weldon P, Morrissey P, O'Mahony M. Long-term cost of ownership comparative analysis between electric vehicles and internal combustion engine vehicles. *Sustain Cities Soc.* 2018;39:578–91. <https://doi.org/10.1016/j.scs.2018.02.024>.
8. Neubauer J, Wood E. The impact of range anxiety and home, workplace, and public charging infrastructure on simulated battery electric vehicle lifetime utility. *J Power Sources.* 2014;257:12–20. <https://doi.org/10.1016/j.jpowsour.2014.01.075>.
9. Nadeau J, Boisvert M, Micheau P. Implementation of a cooperative strategy between a vehicle's mechanical and regenerative brake system. In: 2014 IEEE vehicle power and propulsion conference (VPPC). New York: IEEE; 2014. p. 1–6.
10. Murphey YL, Park J, Chen Z, Kuang ML, Masrur MA, Phillips AM. Intelligent hybrid vehicle power control—part i: machine learning of optimal vehicle power. *IEEE Trans Veh Technol.* 2012;61(8):3519–30.
11. Wong KI, Wong PK, Cheung CS, Vong CM. Modeling and optimization of biodiesel engine performance using advanced machine learning methods. *Energy.* 2013;55:519–28. <https://doi.org/10.1016/j.energy.2013.03.057>.
12. Dixon KR, Lippitt CE, Forsythe JC. Supervised machine learning for modeling human recognition of vehicle-driving situations. In: 2005 IEEE/RSJ international conference on intelligent robots and systems, IROS; 2005. p. 474–9. <https://doi.org/10.1109/IROS.2005.1545026>.
13. Patel M, Lal SKL, Kavanagh D, Rossiter P. Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Syst Appl.* 2011;38(6):7235–42. <https://doi.org/10.1016/j.eswa.2010.12.028>.
14. Ezzini S, Berrada I, Ghogho M. Who is behind the wheel? Driver identification and fingerprinting. *J Big Data.* 2018; <https://doi.org/10.1186/s40537-018-0118-7>.
15. Meseguer JE, Calafate CT, Cano JC. DrivingStyles: assessing the correlation of driving behavior with heart rate changes. Smart objects and technologies for social good. Cham: Springer; 2017. p. 21–30.
16. Meseguer JE, Calafate CT, Cano JC, Manzoni P. DrivingStyles: a smartphone application to assess driver behavior. In: 2013 IEEE symposium on computers and communications (ISCC). New York: IEEE; 2013. <https://doi.org/10.1109/iscc.2013.6755001>.
17. Meseguer Anastasio JE, Tavares De Araujo Cesariny Calafate CM, Cano Escrivá JC, Manzoni P. Characterizing the driving style behavior using artificial intelligence techniques; 2013.
18. Meseguer JE, Calafate CT, Cano JC, Manzoni P. Assessing the impact of driving behavior on instantaneous fuel consumption. In: 2015 12th annual IEEE consumer communications and networking conference (CCNC). New York: IEEE; 2015. <https://doi.org/10.1109/ccnc.2015.7158016>.
19. Beloufa S, Vaillau B, Boucheix J-M, Kemeny A, Merienne F. Eco-driving with electric cars: effect of continuous and on-demand feedback on driving behavior and safety. Reston: American Society of Civil Engineers; 2014.
20. Torres-Sanz V, Sanguesa JA, Garrido P, Martinez FJ, Calafate CT, Marquez-Barja JM. On the prediction of electric vehicles energy demand by using vehicular networks. In: 2017 wireless days. New York: IEEE; 2017. <https://doi.org/10.1109/wd.2017.7918143>.
21. Diewald S, Möller A, Stockinger T, Roalter L, Koelle M, Lindemann P, Kranz M. Gamification-supported exploration and practicing for automotive user interfaces and vehicle functions. Gamification in education and business. Cham: Springer; 2015. p. 637–61. https://doi.org/10.1007/978-3-319-10208-5_32.
22. Ecker R, Holzer P, Broy V, Butz A. Ecochallenge: a race for efficiency. In: Proceedings of the 13th international conference on human computer interaction with mobile devices and services; 2011. p. 91–4. <https://doi.org/10.1145/2037373.2037389>.
23. Magana VC, Munoz-Organero M. GAFU: using a gamification tool to save fuel. *IEEE Intell Transp Syst Mag.* 2015;7(2):58–70. <https://doi.org/10.1109/mits.2015.2408152>.

24. Vaezipour A, Rakotonirainy A, Haworth N. Design of a gamified interface to improve fuel efficiency and safe driving. In: Jini J, editor. Design, user experience, and usability: novel user experiences. Cham: Springer; 2016. p. 322–32. https://doi.org/10.1007/978-3-319-40355-7_31.
25. Steinberger F, Schroeter R, Lindner V, Fitz-Walter Z, Hall J, Johnson D. Zombies on the road: a holistic design approach to balancing gamification and safe driving. In: Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications—AutomotiveUI '2015;15:320–7. <https://doi.org/10.1145/2799250.2799260>.
26. Prandi C, Rocchetti M, Salomoni P, Nisi V, Nunes NJ. Fighting exclusion: a multimedia mobile app with zombies and maps as a medium for civic engagement and design. *Multimed Tools Appl*. 2017;76(4):4951–79.
27. André M. The ARTEMIS European driving cycles for measuring car pollutant emissions. *Sci Tot Environ*. 2004; <https://doi.org/10.1016/j.scitotenv.2004.04.070>.
28. Di Lena P, Mirri S, Prandi C, Salomoni P, Delnevo G. In-vehicle human machine interface: an approach to enhance eco-driving behaviors. In: SmartObject 2017—Proceedings of the 2017 ACM workshop on interacting with smart objects, co-located with IUI 2017; 2017. p. 7–12. <https://doi.org/10.1145/3038450.3038455>.
29. Road vehicles—ergonomic aspects of transportation and control systems—dialogue management principles and compliance procedures. Standard, International Organization for Standardization, Geneva, CH (February 2017).
30. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44. <https://doi.org/10.1038/nature14539>.
31. Wilfinger D, Meschtscherjakov A, Perterer N, Tscheligi M. On the track: comparing distraction caused by interaction with tertiary interfaces with cars on a test track. In: Adjunct proceedings of the 4th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '12), October 17–19, 2012, Portsmouth, NH, USA; 2012. p. 99–102.
32. Kountouriotis GK, Merat N. Leading to distraction: driver distraction, lead car, and road environment. *Accid Anal Prev*. 2016;89:22–30. <https://doi.org/10.1016/j.aap.2015.12.027>.
33. McGehee DV, Mazzae EN, Baldwin GS. Driver reaction time in crash avoidance research: validation of a driving simulator study on a test track. In: Proceedings of the human factors and ergonomics society annual meeting, vol. 44. Los Angeles: SAGE Publications Sage CA; 2000. p. 3–320.
34. Marfia G, Rocchetti M, Amoroso A, Pau G. Safe driving in LA: report from the greatest intervehicular accident detection test ever. *IEEE Trans Veh Technol*. 2013;62(2):522–35.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
