Corcoba-Magaña, Víctor, Muñoz-Organero, Mario, and Pañeda, Xabiel G. 'Prediction of Motorcyclist Stress Using a Heartrate Strap, the Vehicle Telemetry and Road Information'. Journal of Ambient Intelligence and Smart Environments, vol. 9, no. 5, pp. 579-593, 2017

DOI: 10.1039/c3lc51419f

Prediction of motorcyclist stress using the vehicle telemetry and road information

Corcoba-Magaña Víctor[a,*], Muñoz-Organero Mario and Garcia-Pañedo Xicu Xaviel[a]
[a] *Telematics Department, University of Oviedo, Campus de Gijon, Calle Fojanes, s/n, 33204 Gijon, Spain*
[b] *Telematics Department, Av. de la Universidad, 30, 28911 Leganes, Spain*

Abstract. The number of motorcycles on the road has increased in recent years. Although the total number of motorcycles is lower than the number of cars, the accident rate is much higher. A large number of these accidents are due to human errors. Stress is one of the main reasons behind human errors while driving. In this paper, we present a novel mechanism to predict upcoming values for stress levels based on current and past values for both the driving behavior and environmental factors. First, we analyze the relationship between stress levels and different variables that model the driving behavior (accelerations, decelerations, positive kinetic energy, standard deviation of speed, and road shape). Second, we study the accuracy of several machine learning algorithms (Support Vector Machine, Multilayer Perceptron, Naïve Bayes and Deep learning) when used to estimate the stress based on our input data. Finally, an experiment was conducted in a real environment. We considered four different scenarios: normal traffic, heavy traffic, tired motorcyclist, and rested motorcyclist. The results show that the proposal can estimate the upcoming stress with high accuracy. This algorithm could be used to develop driving assistants that recommend actions to smooth the current driving behavior in order to reduce the workload of the motorcyclist when high levels of upcoming stress are predicted.

Keywords: stress prediction, motorbike driving, machine learning, road safety, recommender systems.

[*]Corresponding author. E-mail: editorial@iospress.nl.

1. Introduction

Motorcyclists are drivers who have accidents more frequently. The percentage of these vehicles on the roads is small compared with the number of cars. However, they suffer a high proportion of fatal accidents. In Queensland in 2008 [1], the motorcycles were involved in a 20% of the road accidents. In Australia, motorcycle accidents account for 15% of all accidents.

For this type of drivers, factors such as weather and road conditions, driver stress and the road topology have an important impact on safety. Motorcyclists have to pay attention to more risky situations than car drivers. These specific hazards can be classified into:

- Roadway hazards such as debris, unexpected speed bumps, uneven roadways and potholes.
- Motorcycle-specific hazards such as traffic appearing in a sharp turn.

Although motorcyclists are more exposed to accidents, the industry has not invested as much money in driving assistants that improve safety for them in comparison with those for automobiles. The majority of the proposals are still in an experimental stage and there is a lack for real tests.

On the other hand, The Motorcycle Accidents In-Depth Study (MAIDS) [3] concluded that human factors are the primary accident contributing factor in approximately 87.9% of all cases. Stress is one of the causes of human errors. Stress degrades the cognitive capabilities of driver.

Riding a motorbike is a task that requires a great physical and mental effort. Cognitive errors appear in highly cognitive demanding situations in which the cognitive load, as perceived by the driver, is high and the actions taken by the driver to handle those situations are in many occasions not appropriate.

In the literature, there are many proposals on measuring and quantifying the driver's cognitive load and stress levels. There are two major types of proposals to measure the stress:

- Questionnaires: They allow us to assess a large part of the population. However, the result is based on the subjective perception of the participant.
- Physiological signals: They allow us to objectively quantify the stress levels. However, they require the use of sensors, increasing the cost and reducing the number of possible participants. In addition, these solutions can cause discomfort if they are intrusive or heavy.

Many or the challenges for developping driving assistants for motorcyclists have been partially solved in recent years. One of them is how to present the information to the driver during a ride without distracting them. Devices such as smartphones, GPS or heads-up displays have been successfully used to mitigate these barriers.

In addition, driving assistants require information about the driver and the driving in order to provide appropriate recommendations. This information is extracted from the data obtained from user and environmental sensors. In this context, the wereable sensors have managed to minimize the driver's discomfort.

Currently, a single wristband can monitor several physiological signals. For example, the Empatica E4[4] wristband is an unobtrusive, wearable, lightweight, wireless, multisensory signal acquisition device. It has four inbuilt sensors for continuously reporting Galvanic Skin Response (GSR), Photoplethysmograph (PPG) data, Skin Temperature (ST), and Tri-Axial Acceleration (ACC). It also reports Inter-Beat Interval (IBI) at discrete intervals.

In conclusion, motorcycling is a new sector for the driving assitants. Existing solutions are very scarce and its state is very premature. These proposals are limited to provide generic recommendations without taking into account the mental state of the motorcyclist. An example is the proposed system in [5], where the recommended speed is indicated by speed limit signs. However, the proposal does not consider the current cognitive ability of the driver.

In this work we propose an algorithm to predict stress based in Deep Learning and the motorbyke's telemetry. The proposal is evaluated under different conditions regarding the road state, the road type (urban or motorway), and the previous stress levels. This algorithm can be used to build a driving assistant that recommends an optimal vehicle speed. This recommended speed will minimize the driving workload and the speed fluctuations.

1.1. State of the art

In this subsection, we are going to describe the most relevant work relating to the stress detection and solutions to improve the driving of the motorcyclists.

Stress Detection

Sely [6] was the first researcher to refer to the term "stress" in a biological context. According to this author, the stress includes an inappropriate physiological response to any kind of demand [7]. The term "stress" refers to the human condition and the term "stressor" to the stimulus causing it. Currently, there are many definitions of stress. However, all authors agree in that the stress can have a negative effect in the intelligence, health and making decisions hability [8] [9] [10].

In [11], Itoh et al., data from electrocardiogram (ECG) signals as well as head rotational angles, pupil diameters, and eye blinking, measured with a faceLAB device installed in a driving simulator, are used to calculate the driving workload. In the study captured in [12], the driver's workload was estimated from meassured lane changing behaviour, and the measurements were taken through simulation test driving. In [13], the authors proposed a multiple linear regression equation to estimate the driving workload. The model employs variables such as: speed, steering angle, turn signal, and acceleration.

In [14] the authors propose a method for detecting stress based on facial expressions. They employed a near-infrared (NIR) camera to capture the near frontal view of the driver's face. Tracking face is made using a supervised descent method (SDM) [15]. In [16], the research analyzed the suitability of the heart rate variability (HRV) to measure the driving workload. The results conclude that the HRV could be used as a good workload indicator, although it is also affected by many other factors that may have an influence on it.

In [17], the authors presented a solution to evaluate the emotional states (high stress, low stress, disappointment, and euphoria) of car-racing drivers. Support vector machines (SVMs) and adaptive neuro-fuzzy inference system (ANFIS) were used for the classification The proposed approach performs an assessment of the emotional states using facial electromyograms, electrocardiogram, respiration, and electro dermal activity. The system was validated by using data obtained from ten subjects in simulated racing conditions. The maximum predictive rate was 79.3% using support vector machines (SVM).

As we have seen above, the proposals to detect stress are getting very promising results. However, the stress prediction is a more complex task. Firstly, it is difficult to label stress because the prerceived effects of events change depending on the user profile and their current state. In addition, drivers tend to forget about stress situations when manually labeling stressful regions after driving. On the other hand, the physiological signals are very sensitive to noise and are only proxies to assess human mental states. Finally, combining all the data collected from different sensors is not a trivial task.

Improving the motorcyclist safety

We find different solutions to improve safety in the motorbikes' field. Many of these proposals are based on systems developed inside the automobile industry. However, the motorcycles present some special features that hinder its implementation. According to [18], the advanced rider assistance systems could reduce accidents up to 40%.

Authors in [19] present an Intersection Support System for motorcycles. Warning feedback is given to the rider by an appropriate combination of human–machine interface elements, such as a haptic throttle, a vibrating glove, and a visual display. The proposal was validated in a simulator with 20 riders. The simulated test track had a length of about 10 km, consisting of 6.5 km on a rural road and 3.5 km in an urban scenario, and including a total of 26 intersections. The traffic volume was kept low. The results were positive. Drivers, however, highlighted the need to improve the design of the system.

An Intelligent Curve Warning System was proposed in [20]. The frequency of curve crashes using a motorcycle is very high. In order to solve this problem, the authors designed a system that gives the riders support when approaching a curve. The solution was tested in a simulator by 20 riders. The drivers performed three rides: one without the system (baseline) and two experimental rides using a version of the Curve Warning system, one providing the warnings by a force feedback throttle and one by a haptic glove. The conclusions were that riding with the Curve Warning system with the haptic glove further reduces the critical curve events. Moreover, the force feedback throttle required an increased attention.

In [21], researchers presented the experimental results on comfort and safety aspects of two advanced rider assistance systems: the Front Collision Warning (FCW) and the Lane Change Support (LCS). They used three different machine learning models (Hidden Markov Models, Support Vector Machines and Artificial Neural Networks) in order to get the riders' behavior patterns according to the reaction time needed to avoid front collisions. Finally, the paper describes the implementation of the warning delivery strategy, that was implemented in a HMI (Human Machine Interface) installed on

motorbikes. The authors conclude that the interface should show understandable recommendations and nonintrusive.

A model to compute the minimum distance needed to swerve and avoid a collision against a fixed obstacle was proposed in [22]. It is important to compute and estimation for this distance in order to develop autonomous emergency brake systems. These systems help to reduce the severity of those accidents. The proposal was validated by 12 volunteers riding a scooter equipped with a prototype for autonomous emergency braking. These type of solutions are called motorcycle autonomous emergency braking system (MAEB). The results show that the last-second swerving model represents the lower limit for a non-professional rider performing a maneuver with a large scooter swerving.

The feasibility and quantitative potential benefits of a motorcycle autonomous emergency braking (MAEB) system in crashes was assessed in [23]. They analyzed seven cases from the Swedish national in-depth fatal database crash. They simulated accidents taking into account the vehicles involved and their precrash trajectories. The MAEB proved to be beneficial in a large number of cases. However, they highligthed that the solution was still not mature for the motorciclye industry. MAEB comes from automobiles. In the motorcycle is difficult define a deceleration pattern due to the loss of stability in the motobike.

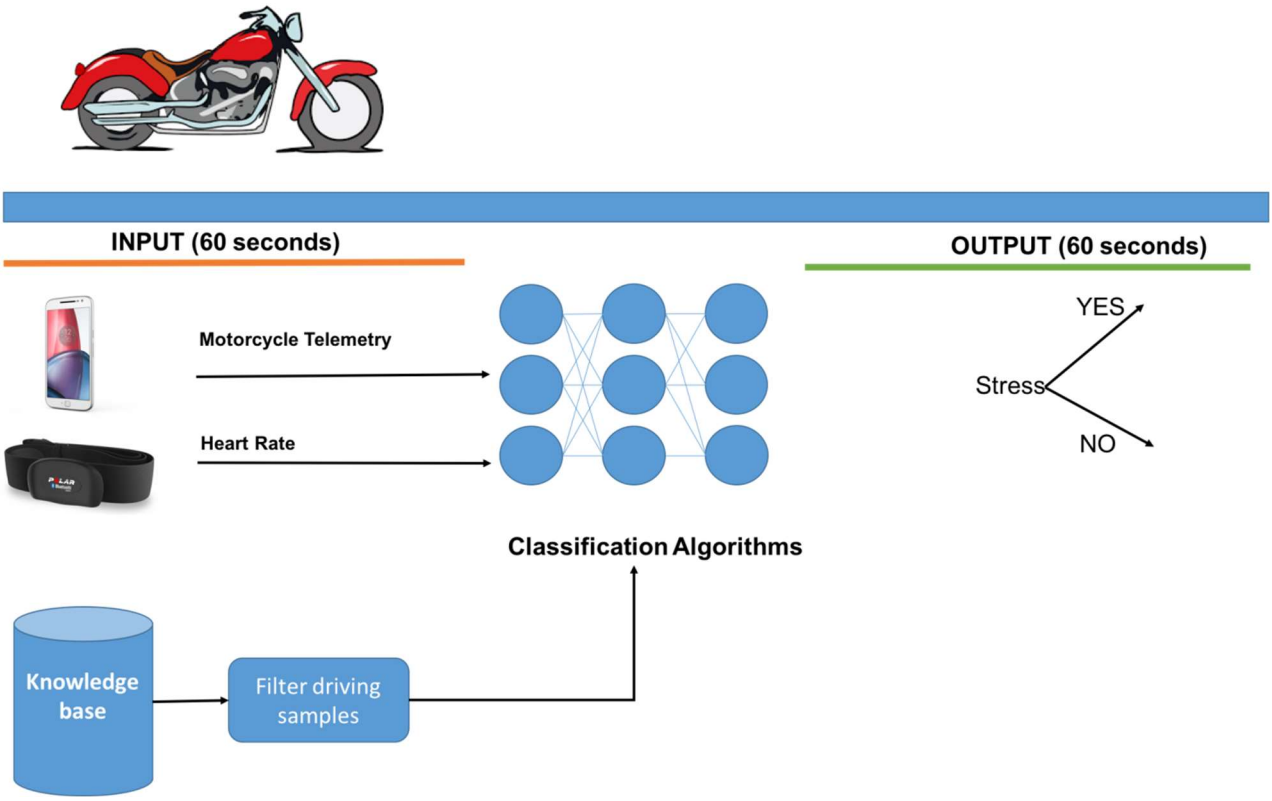
2. Estimation of stress level on motorcyclists

2.1. Objective

Our objective is to build a model to predict the upcomming values for the motorcyclist stress. The proposal uses telemetry data from the motorcycle and stress-related measures from physiological signals in order to estimate near future stress. The prediction of the upcoming stress levels in advance is the basis to make recommendations in order to reduce stress or avoid it (and therefore to mitigate stress induced human errors while drivng a motorbike). On the other hand, stress levels depend on many factors. We will analyze our proposal under different traffic conditions and levels of tiredness of the driver.

2.2. Description of proposal

Figure X shows a schema of the solution. We use a mobile device to obtain the motorcycle telemetry. The smartphone has a GPS sensor that allows us to estimate variables such as: acceleration, deceleration, positive



kinetic energy, standard deviation of speed and intensity of turning. This type of devices are ideal due to their multiple sensors (GPS, accelerometer, and gyroscope), network connections (WIFI, Bluetooth, and LTE) and low cost.

On the other hand, we monitor the stress using a heart rate strap. This device is connected to the phone through bluetooth in order to send the heart rate signal. Using this signal, we employ different computations in order to estimate the level of stress on the driver.

In addition, we have a knowledge base that contains driving samples previously obtained by other drivers or by the own motorcyclist . Stress is a state that depends on many factors. We choose the driving samples from the knowledge base taking into account the traffic state and time. The aim is to predict the upcoming stress accurately for a particular driver in a particular situation.

Finally, a classification algorithm determines if the driver is going to suffer stress in the near future using as input the motorbike's telemetry, the heart rate signal and the filtered driving samples.

Input variables

Fig. 1. Schema of the proposal to anticipate the stress leve of motorcyclist

In this work, the input variables can be classified into two groups: variables related to the stress level and variables associated with the vehicle's telemetry.

Measurements to estimate the stress levels

Heart Rate signals are employed as an indicator for the Autonomic Nervous System (ANS) neuropathy for normal, fatigued and drowsy states because the ANS is influenced by the sympathetic nervous system and parasympathetic nervous systems. This indicator is not intrusive.

Among the different variables presented in the existing literature we have used the Heart Rate Variability (HRV) since it has been assessed as one having a higher correlation with stress levels together with Skin Conductivity (SC) [27].

One major limitation of the HRV signal in order to estimate the level of stress and cognitive load is that there are other factors such as the physical exercise that also impact the measured values. Our experiment has been designed to minimize the impact that factors outside the study have on the measurements. In this way, only data from drivers driving alone to work and back home in similar situations each day (same hour, same traffic conditions, with moderated previous walking to get into the motorbike and a relaxation period of 30 seconds before driving, with the mobile phone muted, and without using any navigation system) have been taken.

In addition, we analyze the driving behavior. The combination of these two groups of variables allows us to build a model to predict the stress on drivers and passengers accurately.

We can consider Heart Rate Variability (HRV) as a proxy variable to estimate the stress level using computations from two different domains (as captured in previous research studies such as [27]): Time and Frequency domains. A time domain analysis of the HRV signal implicates quantifying the mean or standard deviation of RR intervals. Frequency domain analysis means calculating the power of the respiratory-dependent high frequency and low frequency components of the HRV signal. In our case, we are going to use measures on the time domain. There are many HRV features that can be defined on this domain such as: mean RR interval (mRR), mean heart rate (mHR), standard deviation of RR interval (SDRR) or standard deviation of heart rate (SDHR).

We have chosen the following variables extracted from data from real tests:

- Average HeartRate (b.p.m): This variables has a high value when the driver or passenger experiences high levels of stress.
- Average RR (ms): It measures the time between beat-beat (consecutive heartbeats). Its value decreases when there is an event that causes stress on the driver. On the contrary, a high value means that the driver is relaxed.
- Standard deviation of RR intervals (ms): the variation between beat and beat (inter-beat period) increases when the driving workload is high.
- RR50: It is the number of pairs of successive RRs that differ by more than 50 ms. A high number allows us to detect stress situations.

Average acceleration (positive and negative): The accelerations and decelerations capture reactions to different stressors that have an impact on stress levels and imply changes in the HRV signal such as a decrease in the time between consecutive heartbeats. The percentage depends on the intensity of these accelerations. The sudden accelerations significantly increase the driving workload. Figure 2 captures the RR and the motobike speed when motorcyclist is braking. The values have been normalized between 0 and 1 using the following equation:

$$N = \frac{a - \min(A)}{\max(A) - \min(A)} \quad (1)$$

The RR value decreased by 18.23%. Figure 3 shows the results when the motorcyclist is accelerating after a stop. In that case, the RR value is also reduced by a 8.36%.

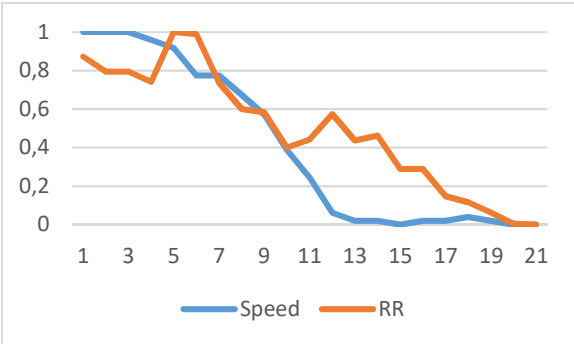


Fig. 2. RR values while motorcyclist was braking.

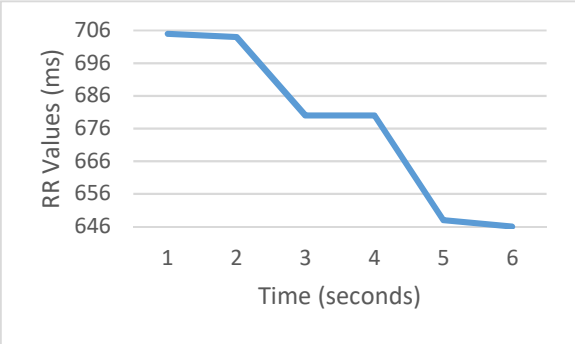


Fig. 3. RR values while motorcyclist was speeding up.

The current acceleration of the vehicle is calculated based on the measured speed as follows:

$$a_i = \frac{v_i - v_{i-1}}{t_i - t_{i-1}} \quad (2)$$

In which v_i represents the speed at the sample number i, a_i the estimated acceleration at that sample and the derivative of the speed is estimated by dividing the increment in speed by the time elapsed between the consecutive samples i-1 and i. Motorcycle speed was obtained using the GPS data from the system's smartphone.

Standard deviation of motorbike speed: The workload decreases when the motorcyclist is driving at steady speed. High deviations of speed capture reactions of the suject to different stressor that cause stress to the driver. He or she has to do several tasks at the same time. Figure 3 shows the RR values obtained in two different cases by the same motorcyclist. In the first case, the standard deviation of motorbike speed was 1.97. In the second case, the value was 16.38. We can observe that in the first case the RR values are higher than in the second case. In addition, the standard deviation of RR values in the second scenario is higher than in the first scenario. In conclusion, we can see that there is a strong relationship between standard deviation of motorcycle speed and the inter-beats time.

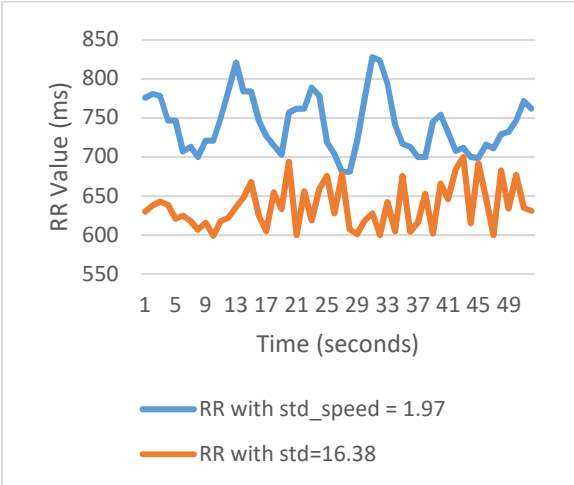


Fig. 4. Comparison of RR values according to the standard deviation of motorbike speed.

Positive Kinetic Energy: This variable measures the aggressiveness of driving. Its value depends on the intensity and frequency of the accelerations. If it is high it means that the driver accelerated sharply and frequently. This driving style has a negative impact on the stress level. He or she has to make faster decisions in order to avoid accidents.

The PKE is estimated over a period of time as follows:

$$PKE = \frac{\sum(v_i - v_{i-1})^2}{d};\ v_i > v_{i-1} \quad (3)$$

Where the sum is performed for the period considered and d is the cumulated distance traveled during this time

The intensity of turning: We detected during testing that the tension increased in the majority of the motorists when there were curves on the road. The degree of impact depends on the road angle (intensity of turning required).

The intensity of turning is estimated using the following equation:

$$TI_i = \cos^{-1} \frac{\overline{l_i} \cdot \overline{l_{i-1}}}{\|\overline{l_i}\| \|\overline{l_{i-1}}\|};\ v_i > th \quad (4)$$

Where the numerator represents the dot product between the average direction vectors in the last 5

seconds and the average direction vectors in the next 5 seconds and the denominator captures the norm of such averaged vectors. The direction vectors are calculated from the GPS coordinates. The average over a period of 5 seconds is used to minimize the impact of random errors in the GPS signal. In order to eliminate the errors introduced at low speeds, a threshold in the speed is used. This threshold has been empirically evaluated and a value of 1 m/s has been found to perform well and therefore selected for the experiment.

2.3. Output

Table 1

RRMSS and pNN50 correlation coefficients

Previous Values	**RRMSS**	**pNN50**
Acceleration	0.6911	0.6180
Deceleration	0.7021	0.6678
PKE	0.6501	0.7180
std_speed	0.4451	0.7099
std_rr	0.3470	0.5612
pNN50	0.2125	0.4391

The output is the stress state in the next minute. In order to label the driving samples, the HRV signal has been translated into 2 different levels of cognitive load: stressed or not stressed. As we saw previously, there are several methods for measuring the stress using the HRV signal.

Table I captures the correlation indexes between the upcoming stress levels and the current values for the independent variables that we will use for their prediction. Upcoming stress levels are computed by using two of the previously captured variables (the RMSSD and the pNN50).

We can observe that the motorcyclist behavior significantly affects the values of future stress levels. On the other hand, the previous stress levels do not have a strong relationship with the upcoming stress. The reason is that an unexpected action causes an increment in the stress level. However, after this increase the motorcyclist relaxes.

The PKE is the variable most correlated with upcoming levels of stress. Analyzing the table, we also observed that we get the best results when we calculate the correlation in relation to the pNN50. Therefore, we have used this measure to classify the driving samples.

3. Classification algorithms

3.1. Deep Learning

A deep-belief network (DBN) [21] is defined as a stack of restricted Boltzmann machines (RBM), in which each RBM layer communicates with both the previous and subsequent layers. The nodes of any single layer don't communicate with each other laterally. The end of DBN is a classifier. We employ gradient-descent algorithm to revise the weight matrix of the whole network. The error is propagated in the opposite direction. Therefore, the parameters of RBMs change slightly. DBN has the following steps:

Layer-wise Unsupervised Learning: We train the first RBM using the original data without the labels (unsupervised) and fixing up the parameters of this RBM. Then, the first layer configuration is frozen. We train the second layer using the output of the first layer. Finally, we get a DBN with several layers, whose parameters are appropriate to extract the features of data. This method avoids the overfitting. In addition, we can take advantage of unlabelled data.

Fine-Turning: We unfreeze all weights, and train full DBN with supervised model (SoftMax classifier) to fine-tune weights. Gradient-descent algorithm is employed to update the weight matrix of the whole network. This solution avoids drastic changes because the error is propagated in the opposite direction.

Restricted Boltzmann Machine (RBM)

Restricted Boltzmann Machine are bipartite graph with a layer of "hidden" neurons and a layer of "visible" neurons, without connections between neurons in the same layer. Each node represents a random variable and each edge a dependency between variables that connects.

We employ energy function (E) and probability distribution to describe a RBM. The energy of a configuration (pair of boolean vectors) (v,h) is defined as:

$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j v_i w_{i,j} h_j \quad (5)$$

where a_i is the bias weight (offset) for the visible unit v_i, b_j is the bias weight for the hidden unit h_j, and $w_{i,j}$ the weight associated with the connection between hidden unit h_j and visible unit v_i.

Probability Distribution:

$$p(v,h) = \frac{1}{Z} e^{-E(v,h)} \quad (6)$$

where Z is a partition function defined as the sum of $e^{-E(v,h)}$ over all possible configurations. The aim is to ensure the probability distribution sums to 1.

Through summation we can get the marginal distribution of visible layer $\boldsymbol{v}$:

$$p(v) = \frac{1}{Z} \sum_h e^{-E(v,h)}$$

SoftMax Classifier

There are a set of training samples such as: $\{(x^1, y^1), (x^2, y^2),\ldots,(x^m,y^m)\}$, $y^i \in \{1,2,\ldots m\}$. The classifier is used in order to estimate the probability that x is a sample of j class. The activation function is:

$$h_\theta(x) = \begin{bmatrix} P(y=1|x,\theta) \\ P(y=2\,|x,\theta) \\ \ldots \\ P(y=k\,|x,\theta) \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} e^{\theta_1 \times X} \\ e^{\theta_2 \times X} \\ \ldots \\ e^{\theta_k \times X} \end{bmatrix} \quad (8)$$

Where $Z = \sum e^{\theta_j \times X}$ is normalization factor.

The cost function to train the classifier is:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{n} 1\{y^{(i)}\} log \frac{e^{\theta_j^T x(i)}}{\sum_{s=0}^{n} \theta_s^T x(i)} \right] \quad (9)$$

In the equation, $1\{y(i) = j\}$ is indicative function, whose value is 1 when $y(i) = j$ and 0 when not. The aim is to minimize the cost function adjusting the parameters. We employ gradient-descent algorithm to make this task.

3.2. MultiLayer Perceptron

A multi-layer perceptron (MLP) [29] and the vehicle telemetry obtained in real time are employed in order to predict the driver behavior. This algorithm is an artificial neural network that has multiple layers and whose main advantage is to allow non-linearly-separable problems. This type of algorithms can be generalized. We can classify an element based on other elements that have been previously classified and which have the same characteristics.

Neural networks were proposed in the 1940, when Warren McCulloch (a psychiatrist and neuroanatomist) and Walter Pitts (a mathematician) explored the computational capabilities of networks made of very simple neurons [30].

Later, in 1943, [31] introduced the perceptron (figure 10), the simplest form of a neural network. The perceptron consists of a single neuron with adjustable synaptic weights and a threshold activation function. This proposal guaranteed the convergence only if the problem was linearly separable due to the basic properties of the perceptron which separate entries into two outputs.

(7)

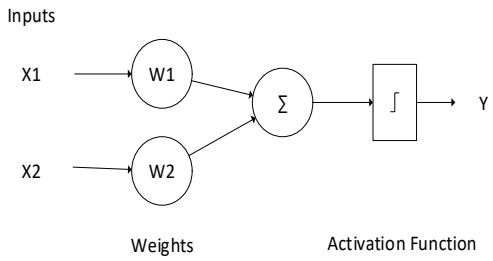


Fig. 5. Perceptron Model

A multi-layer perceptron overcome many of limitations of single-layer perceptron. However, there were no effective training algorithms. This problem was solved by [29]. They called the method "backward propagation" and it is based on least means square algorithm (LMS) [32].

The basic MLP structure consists of an input layer, output layer and one or more hidden layers. The number of layers determines the kind of problem that we can solve. The single layer perceptron is limited to calculating a single line of separation between classes. On the other hand, a three layer perceptron can produce arbitrarily shaped decision regions. The single layer perceptron is limited to calculating a single line of separation between classes. On the other hand, a three layer perceptron can produce arbitrarily shaped decision regions (Kolmogorov theorem), and are capable of separating any classes. Each layer has a set of neurons. The number of neurons depends on the type of problem to be solved. The neurons are connected with other neurons using weighted connections.

The multilayer perceptron has the following steps:

- Initialization of weights and bias. Usually it is randomly done.
- Calculate the output of all neurons of the neuronal network according to equation (9). The output value of a neuron in layer

n is part of the input value from neurons in layer n + 1.

$$y = \sum_i w_i x_i - \theta \quad (9)$$

where w_i is the synaptic weight of the connection, x_i is the input value, θ is a bias term that regulates the degree of an activation to induce firing.

- The third step is to calculate the error in order to minimize it. The training stage on this type of algorithms is supervised. It defines a set of pairs (Xi, Yi) training patterns and an error function (10) (difference between the desired output and the value obtained). Once retrieved the error, the connection weights are updated (11) to minimize them.

$$E_r = \sum_{p=1}^{P} E_p = \frac{1}{2}\sum_{p=1}^{P}(d_p - O_s^p)^2 \quad (10)$$

where E_r is the total output error, E_p is the output error from neuron p, P is the number of neurons from the last layer, O_s^p is the output value from neuron p on layer s, and d_p is the expected output

$$w_{ij}^L(k+1) = w_{ij}^L(k) - \mu \frac{dE_T}{dw_{ij}^L} \quad (11)$$

where w_{ij}^L is the connection weight between neuron i and j in layer L, E_T is the total error ,k is the current iteration, and μ is the learning factor.

The algorithm can be improved by introducing a momentum term focused on accelerating the process, changing the size of the neuronal network or initializing the weight connections in a non-random way.

3.3. Naïve Bayes

Naïve Bayes algorithm [106] is a probabilistic classifier based on Bayes theorem assuming independence between variables. This independence is that gives the name of "Naive". For example, a computer consists of CPU, GPU, RAM and hard disk. A Naïve Bayes Classifier considers each of these features contributes independently to be describing a computer, regardless of the presence or absence of other features. This assumption makes that this algorithm has a great performance, even when the number of instances is very large.

To classify a driving sample in one of the two classes (stress or without stress), we build a probabilistic model. This model allows us to estimate the posterior probability $P(c|x)$ of different classes. Applying Bayes Theorem [10], we obtain:

$$P(c|x) = \frac{P(X|C)P(c)}{P(x)} \quad (12)$$

where $P(c)$ can be estimated by counting the proportion of class c in the training set and $P(x)$ can be ignored since we are comparing different classes on the same driving sample. Therefore, we have only to calculate $P(x|c)$. However, this probability is not easy to be estimated because this requires the estimation of the parameters of the joint-probabilities of the features. As we see previously, Naïve Bayes considers that the variables are independent in order to reduce the complexity. Although it is not always right, it simplifies the classification task dramatically because Creating a new document with the Word templat

3.4. Support Vector Machines

Support Vector Machines (SVM) [28] constitute a different approach for learning from data based on the use of hyperplanes to split the data samples into different classes. By using different kernels, samples can be classfied using non-linear regions [29]. The hyperplanes are defined by the following equation:

$$\sum_i w_i x_i - \theta = 0 \quad (13)$$

A hard margin and a soft margin approach can be used to calculate the paramenters in the previous equation. Soft margin techniques allow the algorithm to find a solution in the case of misclassfied samples. The idea is to find a hyperplane which makes the distance to the nearnest samples from each class as big as possible (highest gap). The generic expression to minimize for the linear soft margin classification is captured in the following equation:

$$\left[\frac{1}{N}\sum_i \max\left(0, 1 - y_i\left(\sum_j w_j x_{ij} - \theta\right)\right)\right] + C\sum_j w_j w_j \quad (14)$$

Where the sub-index “i” has been used to capture the different data samples and the sub-index “j” to

implment the dot product for each m-dimensional sample and the m-dimensional vector of coefficients $\vec{w}$. The value of y_i represents 0 or 1 for the class of each of the samples. The value of C controls the importance of minimizing the misclassfied samples.

4. Validation of the proposal

In this section, the proposed prediction algorithm is validated. We used a Nexus 6 in order to get the driving samples. Nexus 6 has a 2.7 GHz quad-core Snapdragon 805 processor with 3 GB of RAM, GPS chipset Qualcomm IZat Gen8B, and Android 6.0.

A Polar H7 band was employed to record the HRV signal. The band was paired with the smartphone running an application implemented for the experiment which recorded the HRV together with GPS data and telemetry data such as the driving speed.

We have analyzed three different scenarios. In the first scenario, we show how accurately upcoming levels of stress can be assessed based on recent levels of stress, driving behavior and road shape when the motorcyclist drives to their daily commute from home to the workplace. In this case, the driver has rested. However, the stress can be high if the rider does not arrive on time to work.

In the second scenario, we capture the algorithm response when rider drove from workplace to home. In this case, the fatigue alters the heart rate signal. Finally, the third scenario is dedicated to show how the traffic conditions affects the stress level.

In order to conduct the tests, a driver used a motorbike (kawasaki vn900 classicf) to travel an urban route in Madrid (Spain). 22 different drives on the same route have been used. Table 2 captures the features of the motorcycle.

To make the stress prediction, we have analyzed 4 different algorithms to capture different families of machine learning techniques: Naïve Bayes, Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and Deep Learning. 10-fold cross technique ws employed in order to make the validation of the proposal.

Table 2

Features of a kawasaki vn900 classic

Feature	Value
Engine	0.6911
Displacement	0.7021
Bore x Stroke	0.6501
std_speed	0.4451
Overall Length	0.3470
Overall Width	0.2125
Overall Heigh	

Fig. 4. Motorbike used on the tests.

4.1. Scenario 1 (Driving from home to workplace)

Tables 1, 2, and 3 capture the confusion matrix obtained using the different classification algorithms. In order to build the model, we employed a training database that contained driving samples acquired on the daily commute to work.

We can observe how the four algorithms obtained similar results. Deep Learning algorithm is that it is able to predict more accurately stress. This algorithm works especially well for predicting stress when really there is stress. However, it also predicts false positives by 14%. Naive bayes is the algorithm that gets worse results. In that case, the false positives rate is 16%. On the other hand, both MLP and SVM presented results similar. The total error rate (false positives and false negatives) is 24% in both cases.

Table 3

Naïve Bayes

Actual/Predicted	**Yes**	**No**
Yes	0.9	0.1
No	0.16	0.84

Table 4

Support Vector Machine

Actual/Predicted	**Yes**	**No**
Yes	0.88	0.12
No	0.12	0.88

Table 4

Multilayer Perceptron

Actual/Predicted	Yes	No
Yes	0.86	0.14
No	0.10	0.90

Table 5

Deep Learning

Actual/Predicted	Yes	No
Yes	0.92	0.08
No	0.14	0.86

Figure X shows the time to build the classification model using the different algorithms. The time was obtained running 10 times the algorithms on a PC with the features shown in table 6. We can see that Naïve Bayes is the fastest. Deep Learning was the algorithm that obtained more accurate results. However, it is an slower (450 ms) than Naïve Bayes (1 ms). In this case, the algorithm Naïve Bayes is offering a better balance between time and accuracy of the results. Multilayer Perceptron is faster (19 ms) than Deep Learning and slower than Naïve Bayes.

Table 6

PC Specification

Feature	Value
Processor	Intel Core i7-3520M (2.9 GHz
RAM	8 GB 1600 MHz DDR3
Hard Disk	SSD SanDisk Ultra II 240GB
OS	macOs 10.12

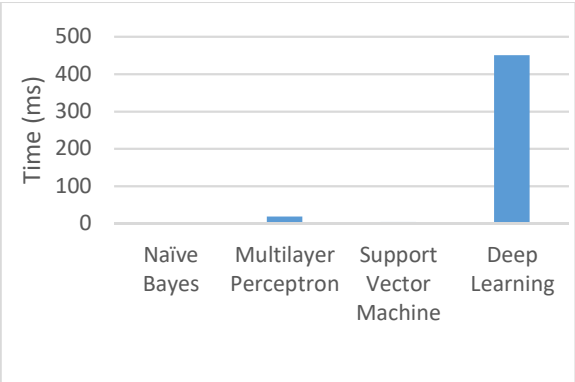


Fig. 2. Time to build the models using the classification algorithms when rider drives from home to workplace

4.2. Scenario 2 (Driving from workplace to home)

Tables 12, 13, 14, and 15 capture the results obtained using driving samples of the commute between workplace and home. In this case, there were significant differences between the tested algorithms.

Naïve Bayes presents a very low success rate. This algorithm is only able to predict that the motorcyclist will suffer stress by 72%. On the other hand, the Deep Learning algorithm improves the accuracy compared to the results obtained in the first scenario.

Deep Learning is a method that introduces a new way of training multilayer networks. This new technique allows us to discover the complex relationships between variables. On the contrary, Naïve Bayes considers that the variables are independent between itself, simplifying the model. However, the tiredness of the motorcyclist after the work increases the complexity for predicting correctly the stress. The consequences are that NaÏve Bayes is not an algorithm suitable in this case.

Finally, the MLP and SVM algorithms get similar results. SVM anticipates the motorcyclist stress by 16% when there is no stress. Moreover, the false positives rate is 12%. For the MLP algorithm, the total error rate (false positives + false negatives) is 24%.

Table 7

Naïve Bayes

Actual/Predicted	Yes	No
Yes	0.72	0.28
No	0.16	0.84

Table 8

Support Vector Machine

Actual/Predicted	Yes	No
Yes	0.84	0.16
No	0.12	0.88

Table 9

Multilayer Perceptron

Actual/Predicted	Yes	No
Yes	0.88	0.12
No	0.12	0.88

Table 10

Deep Learning

Actual/Predicted	Yes	No
Yes	1	0

No	0.08	0.92

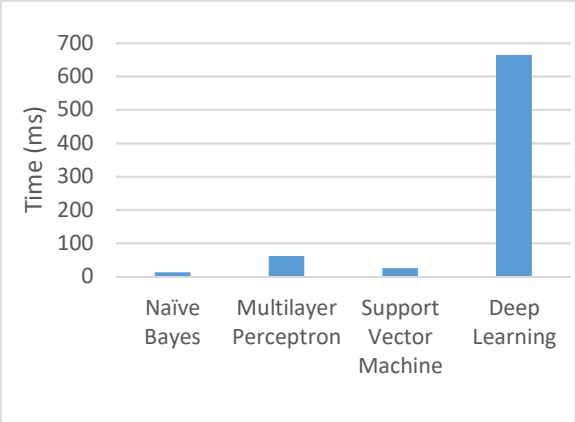


Fig. 2. Time to build the models using the classification algorithms when rider drives from workplace to house

Figure X captures the time required to build models using the proposed algorithms. Naïves Bayes is the fastest (13 ms). However, the success rate is lower than the rest of the algorithms. SVM is presents a more balanced results taking into account the total error rate and execution time. Deep Learning is the algorithm more slowly as it happened on the first stage. This method is 94.87% slower than the rest of algorithms

4.3. Scenario 3 (Driving under heavy traffic)

Tables 14, 15, 16 and 17 show the results obtained under heavy traffic. The speed is decreased due to this external factor. In this case, the sudden accelerations and sudden decelerations also are minimized, but their number increase. Figure X compares the speed profile with and without traffic heavy in the same road section. We can observe that the number of accelerations and decelerations increased by X% when the traffic was dense. In addition, the sudden accelerations (positive and negative) decreased by X %.

Analyzing the results of the confusion matrix, we can see that Naïve Bayes classified correctly the stress level by 94%. In this situation, a significant change in the value of any variable means stress in the near future. For this reason Naïve Bayes is the most suitable algorithm. The success rate is high. In addition, it is the algorithm that takes less time to build the model. MLP predicted the stress level correctly by 94%. However, this algorithm is 78.46% slower.

SVM has a high false negative rate (18%). On the contrary, Deep Learning got a high false positive rate (16%). Furthermore, the time to build a classification model is 91.33% greater than the others. Figure X captures the execution times for each algorithm.

Table 11

Naïve Bayes

Actual/Predicted	Yes	No
Yes	0.94	0.06
No	0.06	0.94

Table 12

Support Vector Machine

Actual/Predicted	Yes	No
Yes	0.82	0.18
No	0	1

Table 13

Multilayer Perceptron

Actual/Predicted	Yes	No
Yes	0.92	0.08
No	0.04	0.96

Table 14

Deep Learning

Actual/Predicted	Yes	No
Yes	0,98	0.02
No	0.16	0.84

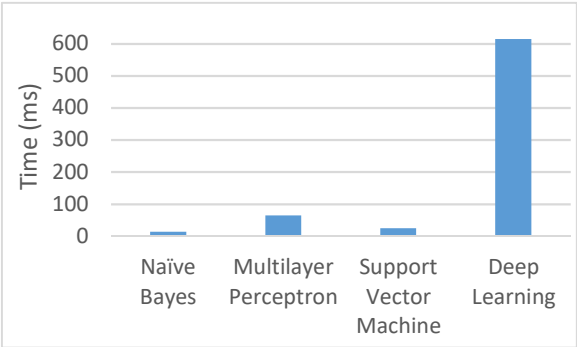


Fig. 2. Time to build the models using the classification algorithms when rider drives from workplace to house

5. Conclusions

Acknowledgements

The research leading to these results has received funding from the "HERMES-SMART DRIVER" project TIN2013-46801-C4-2-R funded by the Spanish MINECO, from the grant PRX15/00036 from the Ministerio de Educación Cultura y Deporte and from a sabbatical leave by the Carlos III of Madrid University.

References

[1] Filtness, Ashleigh J. & Rudin-Brown, Christina M. (2012) Drinking and riding: is subjective workload related to performance? In Australasian Road Safety Research, Policing and Education Conference, 4 - 6 October 2012, Wellington, New Zealand.

[2] M. P. Jenkins and D. Young, "BARRACUDA: An augmented reality display for increased motorcyclist en route hazard awareness," 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Diego, CA, 2016, pp. 68-72.

[3] ACEM (Association des Constructeurs Européens de Motocycles): "In-depth investigations of accidents involving powered two-wheelers (MAIDS)" (Brussels, Belgium, 2009), pp. 1–179.

[4] Empatica E31. URL: https://www.empatica.com, [Last access: 11-07/2016].

[5] Simpkin, B., Lai, F., Chorlton, K., Fowkes, M.: 'ISA-UK, Intelligent Speed Adaptation. Results of motorcycle trial' (University of Leeds, Leeds, UK, 2007).

[6] H. Selye, "A Syndrome Produced by Diverse Nocuous Agents," in Nature, vol. 138, p. 32, Jul. 1936.

[7] H. Selye, "Stress and disease," in Science, vol. 122, p. 625-631, Oct. 1955.

[8] J.E. McGrath, "Stress and behavior in organizations," in Handbook of industrial and organizational psychology, M.D. Dunnette, Ed. Chicago: Rand McNally, 1976, pp. 1351–1395.

[9] D. I. Tepas and J. M. Price, "What is stress and what is fatigue?" in Stress, workload, and fatigue, P.A. Hancock and P.A. Desmond, Eds. Mahwah, NJ: L. Erlbaum, 2001, pp. 607-622.

[10] R. Vecchio, "Theoretical and empirical examination of cognitive resource theory," in J. Applied Psychology, vol.75 (2), 1980, pp. 141-147.

[11] M. Itoh, E. Kawakita, and K. Oguri, "Real-time estimation of driver's mental workload using physiological indexes," in Proc. ITS World Congr., 2010, pp. 1–11.

[12] E. T. T. Teh, S. Jamson, and O. Carsten, "How does a lane change performed by a neighboring vehicle affect driver workload?" in Proc. ITS World Congr., 2012, pp. 1–8.

[13] S. Sega and H. Iwasaki, "Verification of driving workload using vehicle signal data for distraction-minimized systems on ITS," in Proc. ITS World Congr., 2011, pp. 1–12.

[14] H. Gao, A. Yüce and J. P. Thiran, "Detecting emotional stress from facial expressions for driving safety," *2014 IEEE International Conference on Image Processing (ICIP),* Paris, 2014, pp. 5961-5965. doi: 10.1109/ICIP.2014.7026203.

[15] X. Xiong and F. De la Torre, "Supervised Descent Method and Its Applications to Face Alignment," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 532–539.

[16] B. Eilebrecht, "The relevance of HRV parameters for driver workload detection in real world driving," 2012 Computing in Cardiology, Krakow, 2012, pp. 409-412.

[17] C. D. Katsis, N. Katertsidis, G. Ganiatsas and D. I. Fotiadis, "Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach," in IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 38, no. 3, pp. 502-512, May 2008. doi: 10.1109/TSMCA.2008.918624.

[18] Rakotonirainy, A., Haworth, N.: 'Institutional challenges to ITS deployment and adoption'. 29th Australasian Transport Research Forum, Australia, Gold Coast, September 2006

[19] Biral, F., Lot, R., Rota, S., et al.: 'Intersection support system for powered two-wheeled vehicles: Threat assessment based on a receding horizon approach', IEEE Trans. Intell. Transp. Syst., 2012, 13, (2), pp. 805–816

[20] Véronique Huth, Francesco Biral, Óscar Martín, Roberto Lot, Comparison of two warning concepts of an intelligent Curve Warning system for motorcyclists in a simulator study, Accident Analysis & Prevention, Volume 44, Issue 1, January 2012, Pages 118-125, ISSN 0001-4575, http://dx.doi.org/10.1016/j.aap.2011.04.023.

[21] Valtolina, S., Vanzi, S., Montanari, R., Minin, L., Marzani, S.: 'Design of warning delivery strategies in advanced rider assistance systems'. ASME 2011 World Conf. on Innovative Virtual Reality, Italy, Milan, June 2011, pp. 41–50.

[22] Federico Giovannini, Giovanni Savino, Marco Pierini, Niccolò Baldanzini, Analysis of the minimum swerving distance for the development of a motorcycle autonomous braking system, Accident Analysis & Prevention, Volume 59, October 2013, Pages 170-184, ISSN 0001-4575, http://dx.doi.org/10.1016/j.aap.2013.05.020.

[23] Savino, G., Giovannini, F., Baldanzini, N., Pierini, M., Rizzi, M.: 'Assessing the potential benefits of the motorcycle autonomous emergency braking using detailed crash reconstructions', Traffic Injury Prev., 2013, 14, pp. 40–49.

[24] L.L. Di Stasi, D. Contreras, A. Cándido, J.J. Cañas, A. Catena, Behavioral and eye-movement measures to track improvements in driving skills of vulnerable road users: First-time motorcycle riders, Transportation Research Part F: Traffic Psychology and Behaviour, Volume 14, Issue 1, January 2011, Pages 26-35, ISSN 1369-8478.

[25] Charles Goldenbeld, Divera Twisk, Saskia de Craen, Short and long term effects of moped rider training: a field experiment, Transportation Research Part F: Traffic Psychology and Behaviour, Volume 7, Issue 1, January 2004, Pages 1-16, ISSN 1369-8478, http://dx.doi.org/10.1016/j.trf.2003.09.003.

[26] Charles C. Liu, Simon G. Hosking, Michael G. Lenné, Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment, Transportation Research Part F: Traffic Psychology and

Behaviour, Volume 12, Issue 4, July 2009, Pages 325-334, ISSN 1369-8478.

[27] Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on intelligent transportation systems, 6(2), 156-166.

[28] Cortes, C.; Vapnik, V. (1995). "Support-vector networks". Machine Learning. 20 (3): 273–297. doi:10.1007/BF00994018.

[29] Crammer, Koby & Singer, Yoram (2001). "On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines" (PDF). Journal of Machine Learning Research. 2: 265–292.