

PAPER • OPEN ACCESS

## Driver drowsiness detection using different classification algorithms

To cite this article: N. S. Nor Shahrudin and K.A. Sidek 2020 *J. Phys.: Conf. Ser.* **1502** 012037

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Driver drowsiness detection using different classification algorithms

N. S. Nor Shahrudin<sup>1</sup>, K.A. Sidek<sup>2</sup>

<sup>1,2</sup>Department of Electrical and Computer Engineering, International Islamic University Malaysia, P.O. Box 10, Jalan Gombak, 50728 Kuala Lumpur

*shahirah.ciera@gmail.com<sup>1</sup>, azami@iium.edu.my<sup>2</sup>*

**Abstract.** Capability of electrocardiogram (ECG) signal in contributing to the daily application keeps developing days by days. As technology advances, ECG marks the possibility as a potential mechanism towards the drowsiness detection system. Driver drowsiness is a state between sleeping and being awake due to body fatigue while driving. This condition has become a common issue that leads to road accidents and death. It is proven in previous studies that biological signals are closely related to a person's reaction. Electrocardiogram (ECG) is an electrical indicator of the heart, provides such criteria as it reflects the heart activity that can detect changes in human response which relates to our emotions and reactions. Thus, this study proposed a non-intrusive detector to detect driver drowsiness by using the ECG. This study obtained ECG data from the ULg multimodality drowsiness database to simulate the different stages of sleep, which are PVT1 as early sleep while PVT2 as deep sleep. The signals are later processed in MATLAB using Savitzky-Golay filter to remove artifacts in the signal. Then, QRS complexes are extracted from the acquired ECG signal. The process was followed by classifying the ECG signal using Machine Learning (ML) tools. The classification techniques that include Multilayer Perceptron (MLP), k-Nearest Neighbour (IBk) and Bayes Network (BN) algorithms proved to support the argument made in both PVT1 and PVT2 to measure the accuracy of the data acquired. As a result, PVT1 and PVT2 are correctly classified as the result shown with higher percentage accuracy on each PVTs. Hence, this paper present and prove the reliability of ECG signal for drowsiness detection in classifying high accuracy ECG data using different classification algorithms.

## 1. Introduction

Driver's safety needs to be focused, especially when it relates to the road activities that include the behavioural and condition of a person while driving. The condition of the drivers, including their health condition, should be monitored due to the common occurrence of road accidents causes by drowsiness issues. Earlier, a drowsy bus driver was reported that he accidentally crashes a trailer that leads him to death [1]. According to Dungun District Police Chief, Superintendent Baharudin Abdullah mention several times to avoid drowsy driving in preventing the occurrence of accidents.

This is due to the Ops Sikap's report that stated over the last year, which Dungun district marks 52 deaths while 15 deaths were recorded within the first five months of this year [2]. According to the Head of Traffic Department in the RMP, Dato' Mohd Fuad Abd Latif, the common causes that lead to road accidents is the attitude of the driver itself. In addition, according to the Road Accident Statistics Report 2014 by RMP, age range between 11 and 25 years old are the major contributors of road accidents, which is 47.88 per cent of the total deaths [3]. The index on road death based on the range of age is shown in Figure 1. Thus, this scenario indirectly causes the nation to lose potential future leaders.





**Figure 1.** Index of road death (2004-2014) due to road accidents based on the range of age

Another case in Penang also brings a rider to death when he rides in a drowsy condition [4]. This case should be monitored deeply as microsleep at the early stage has the potential to end with road accidents. Based on the incidents happen, it can be summarized that the statistics for road deaths due to drowsiness will keep increasing if no effective solution were made. Thus, this study proposed a driver drowsiness monitoring system to overcome these burgeoning issues.

### 1.1 Open issues

Previous researches proposed many techniques in order to measure the drowsiness level of a driver. Traditionally, ECG used electrodes to measure the electrical signal of the human body. Unfortunately, this method involves invasive contact due to the wires which could make the driver feel uncomfortable with the measurement of the system. For example, the design of a smartwatch was proposed to measure the drowsiness level of the driver. If the driver wears it too loose, it might affect the measurement of the system such as having errors and the values are inaccurate. In another situation, if the driver wears the smartwatch too tight, it might lead to an inconvenient situation for the driver. In order to have a better performance system in monitoring the drowsiness level of a driver, these are the open issues that can be improved based on contributing a convenient method with high accuracy of ECG measurement by proposing potential classification techniques and simultaneously able to differentiate between normal and drowsy conditions of a driver.

### 1.2 Drowsiness detection

Previous studies summarized that there are three types of the drowsiness detection technique. Those techniques include vehicle-based measures, behavioural and psychological measures [5, 6]. The first technique, vehicle-based method calculates the deviation of lane position, angular steering wheel movements [7], and force towards the acceleration pedal. By marking a threshold of continuous data acquired, this technique summarized that any changes that cross the threshold value were considered as an abnormal driving condition and might be caused by drowsy driving.

In recent years, the electrocardiogram (ECG) signal was introduced as a potential biological signal metric to overcome biomedical issues. Preliminary investigations on the validity of using ECG based driver drowsiness have been manifested with different techniques to support its usability and reliability. This also will show that even ECG is well-known in clinical used, such as identifying the heart rate abnormality and any critical disease, ECG also can be used to measure the drowsiness level. Hence, this study proposed a non-invasive method that will analyze with a simple new improvement and innovative driver drowsiness monitoring techniques to highlight the practicality of ECG based driver drowsiness system. The experimentation results will be compared with related and existing driver drowsiness techniques to evaluate the effectiveness and robustness of the proposed approach. The outcome of this research will enable us to measure driver drowsiness level. It is hoped that by implementing the driver drowsiness monitoring system using ECG will help in reducing the number of road accidents and death cases in Malaysia.

### 1.3 Rest of the paper

The remaining paper is organized as follows: Section 2 reviews literature on the background of the studies and related papers relating to ECG for drowsiness detection. Section 3 explains the proposed

methodology, and Section 4 describes the experimental results and their performance. Finally, Section 5 concludes the whole research.

## 2. Literature review

In this section, the study will divide into two subsections which are background studies and related works. Most of the previous works propose techniques which include the driver's safety and convenience way to detect drowsiness. This section briefly discusses related literature on the driver drowsiness monitoring system.

### 2.1 Background studies

Physiological measure for drowsiness detection was the most implementation for the drowsiness research area. This is due to human reaction reflects the body response, which can be measured by PPG, EMG, EEG and ECG. ECG signal is obtained from the electrical signal of the heart, which responds to the physiological signal in the human body. Essentially, ECG works on amplifying the small amount of electrical signal that change as the heart pumps the blood in the body that cause by the heart muscle depolarizes in each heartbeat in any human response whether in emotions and reactions.

A cycle of ECG waveform consists of P-Q-R-S-T wave is well known as PQRST morphology [10]. Many studies will carefully select the wave to analyze data that relates to the application of the researcher. Some of the researchers will measure the amplitude of the wave and the peak-to-peak in another cycle to get the information from the signal. Thus, this study proposed to use ECG signal as the indicator of the drowsiness measurement since it is one of the most reliable physiological signals that can interpret data in corresponding with a varying wave in PQRST morphology.

Nowadays, the development of ECG contributes too many implementations in a monitoring system. It can be seen when ECG have been beyond in health monitoring system which been explained on a study of ECG biometric with abnormal cardiac conditions [11] and implementation of ECG for health and sports [12]. Since the development of ECG is rarely focused on drowsiness, this study decides to have a monitoring system which leads to a better system in order to reduce the number of road accident. Thus, this research proposes a monitoring system which can be summarized to monitor the conditions of a driver using the ECG signal. Next, this research proceeds with reviewing related paper from the previous researcher to identify the limitations of the current drowsiness detection system to be improved.

### 2.2 Related works

This section discusses related works regarding drowsiness detection, which include vehicle-based measures, behavioral and psychological measures.

Lee et al. [13] designed a smartwatch that is used to collect motion data such as linear acceleration and the radial velocity from the human response for driver drowsiness detection. This is due to the design of the smartwatch that contains gyroscope and accelerometer that is capable as motion sensors to acquire motion data. The motion data is used as a feature extraction to obtain the magnitude of the hand movements. Then, this research proceeds to the calculation of phase, spectral and time-domain features using the motion data. There are eight features that have been chosen as an input to support vector machine (SVM) classifier to measure drowsiness level. The output shows that the highest accuracy from the SVM training and testing was 98.15% that is corresponding to the Karolinska Sleepiness Scale (KSS). The smartwatch also can be used for both left-handed and right-handed users since different hands were used by different SVM models. However, the smartwatch must be closely to wrist to obtain less error and accurate result, but if the driver wears the smartwatch too tight, it might lead to an invasive method to detect driver drowsiness.

Hwang et al. [9] suggested one of the highest demand technologies to detect driver drowsiness is Electroencephalogram (EEG) and several signals in order to reduce the number of road accidents. In this study, The EEG is designed in-ear in order to evaluate the performance of the driver alertness classification. EEG is traditionally well known as an intrusive method in order to measure driver drowsiness, so the researcher takes the opportunity to test EEG simultaneously with other peripheral signals such as Electrocardiogram (ECG), Photoplethysmogram (PPG) and Galvanic Skin Response (GSR). This method used simulation by asking the subject to drive his/her car with one hand and keep

the car on track as much as possible. The EEG will be on the ear-canal, scapha and conventional method of EEG on the head of the driver while PPG and GSR on armrest and ECG on the chest are applied through the simulation process. The result shows that the performance of EEG in-ear has the highest accuracy rather than EEG in the conventional method. However, the researcher finds that there is competitive performance between ECG, PPG and GSR if the peripheral signal is combining together in pairwise on detecting driver drowsiness. Yet, this simulation still needs to be in the normal driving simulation since this study use only one hand to be on the steering wheel that can be inconvenient for the driver throughout this driver drowsiness detection.

Colic<sup>4</sup> et al. [8] used visual input to design and implement the driver drowsiness detection system. This study is roughly used a novel way which includes the combination off-the-shelf software components for face detection, eye detection in opened and closed condition and human skin colour detection. This detection is basically referred to behavioural of a driver that includes the eyes closure, sodding on the head and yawning extraction by the observation of a video camera. The drowsiness detection begins with face detection, followed by the tracking stage and warning stage. If sleep detected, the system would continue to alert stage which produces sounding an audible alarm as to wake up the driver. The face detection will extract the skin colour of the driver to create a skin model for the system and identifying a set of eye samples. The different to distinguish both eyes are classified using the Support Vector Machine (SVM) classifier. The outcome shows that this method achieved 96% success rate in the preliminary result by using available SVM MATLAB version. However, this method is not relevant to continuously drive, especially in a tunnel since it might affect the light condition of the proposed system in detecting driver drowsiness.

Babaeian et al. [14] designed a system that can detect driver drowsiness at the early stage by computing heart rate variation using a logistic regression-based machine learning algorithm. This research tested the capability of two sensors: one sensor in contact and another one is a non-contact sensor to detect and measure the gain of the ECG signal, by testing with two different micro-controllers that is Beagle Bone Black and Arduino. Then, this signal is converted to discrete data that prove that Arduino's sampling frequency is much lower than Beagle Bone Black that needs higher space and low-speed performance which developed in MATLAB 2015. Three subjects have been tested in two different methods, that is logistic regression and Naïve Bayes. The result shows that logistic regression has higher accuracy (over 92%) rather than Naïve Bayes (never exceeds 85%) which relies on independent future, rather than in the research proposal that depends in the frequency domain. The probability of drowsiness is predict occurred in a range of 12 and 24 seconds that shows the result of three subjects has over 90% accuracy from the experimental result. However, this research doesn't mention how the data being collected, such as using wires that attach to the chest or only by having a simulation of the ECG database.

These literature are the basis of conducting this research as this study associates the measurement of the ECG signal in order to detect driver drowsiness such as on motor vehicle, behavioural and physiological methods and mitigate the limitation of previous works. Yet, there are still lacking to be a better system in a driver monitoring system. According to Healey et al. in [15], the physiological measure is the best measurement in measuring body signal. One of the advantages of ECG is that it can detect the internal signal more accurate than using other measures. Thus, due to this advantage of ECG, this study proposed the development of driver drowsiness in monitoring system using ECG.

### 3. Research methodology

In this section, the proposed system consists of data acquisition, pre-processing, feature extraction and classification that will be explained briefly in the next subsection as in Figure 2.



**Figure 2.** Methodology in driver drowsiness detection using different classification algorithms

### 3.1. Data acquisition

ECG data used in this study were collected from the ULg multimodality drowsiness database [16]. The database consists of 10 polysomnographic subjects which being collected intentionally based on the psychomotor vigilance tests (PVTs) stages which are PVT1 and PVT2. Both stages refer to the sleep stages from early sleep to deep sleep. The duration of each ECG signals recording is 10 seconds.

### 3.2. Pre-processing

The raw signal collected might not be in a smooth waveform due to the occurrence of the noise, which is caused by the surrounding environment during the data collection. In order to remove the unwanted signal, data processing is performed by using the Savitzky-Golay filter (also known as Sgolay filter) into the raw signal since the selected filter is broadly utilized in smoothing signals. The main advantage of Sgolay filter in ECG is this filter tends to preserve the original signal such that P wave will not be affected during the noise removal [17]. Although the amplitude of the QRS complex is shrinking, a high amount of noise was successfully denoised, and all the key features are preserved.

### 3.3. Feature extraction

The extracted signal includes PR interval, PR segment, ST segment, QT interval and the commonly used feature extractor is the QRS complex. One of the most prevalent QRS detection algorithms is the Pan Tompkins algorithm. Pan and Tompkin introduced this method to extract the QRS complex from the ECG signal. Pan Tompkins proved to gain higher sensibility in QRS detection rather than based wavelet transform, which are 99.81% and 98.28% respectively [18]. Thus, this research used the QRS complex to extract ECG signal, simultaneously act as the input for the classification process in the next subsection.

### 3.4. Classification

The classification algorithms that have been proposed in this study includes the Multilayer Perceptron (MLP), k-Nearest Neighbor (IBk) and Bayesian Network. These techniques were classified using WEKA software. As it is the last step of the proposed method, this process is needed to validate the accuracy of the proposed data.

**3.4.1. Multilayer Perceptron (MLP).** Multilayer Perceptron (MLP) is one of the frequently used types of networks applied in the task of classification. The MLP architecture consists of three layers; an input layer, a hidden layer, and an output layer. The MLPs ability to learn from examples makes the neural network very powerful and flexible. There is no necessity to understand the internal mechanism of a task, or even the need to formulate an algorithm to carry out any specific tasks [19].

**3.4.2. K-Nearest Neighbour (IBk).** This classifier is a simple algorithm that is based on a similarity measure. It functions by acquiring the majority class or nearest neighbour among k-neighbours. It detects the similar element among k nearest neighbours, and the similarity can be defined using distance function, which is Euclidean distance,  $ed(i)$  [20].

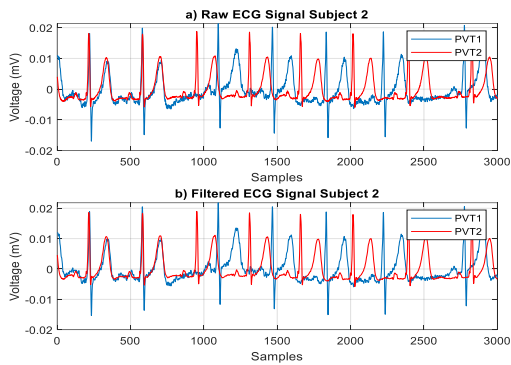
$$ed(i) = \left[ \sqrt{(c(y) - y(i))^2 + (c(x) - x(i))^2} \right] \quad (1)$$

**3.4.3. Bayesian Network (BN).** A Bayesian network consists of a set of local distributions and a directed acyclic graph. The advantages that can be obtained from the Bayesian Network is that it saves a lot of table space and computation since it only relates nodes that are probabilistically related by some dependency. Secondly, the Bayesian network is useful because of its adaptability. It requires only a small amount of data, and from there, it can evaluate essential parameters on its own [21].

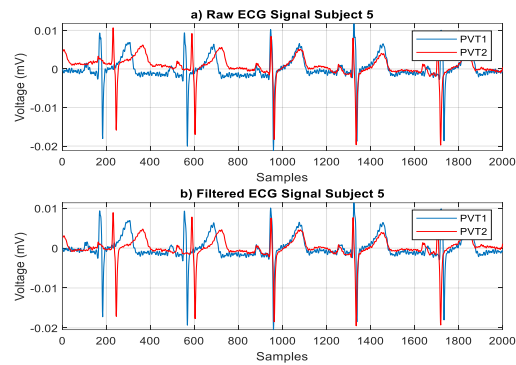
## 4. Experimentation and results

This section describes the proposed research methodology. In the data acquisition stage, two examples out of 10 subject's raw ECG signal were collected as in Figure 3(a) and 4(a). Next, the raw ECG signals

were filtered in the preprocessing stage. Figure 3(b) and 4(b) represents the filtered ECG signals for each subject. Generally, the graphs show changes in ECG signal from PVT1 to PVT2.

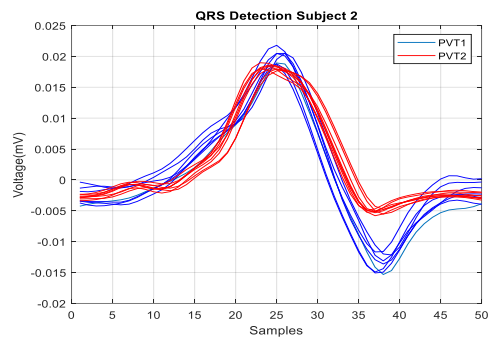


**Figure 3.** (a) Raw and (b) Filtered ECG Signal Subject 2

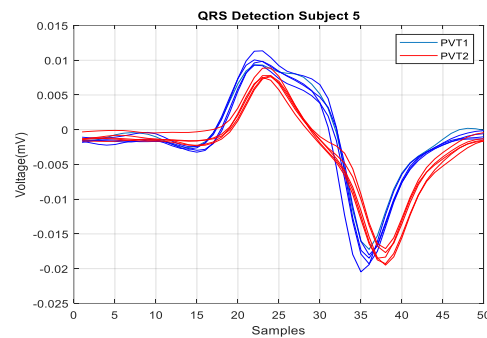


**Figure 4.** (a) Raw and (b) Filtered ECG Signal Subject 5

As can be seen, the figures show changes in ECG signal from both PVTs for each cycle. This graph also presents the capability of Sgolay filter, which tends to preserve the characteristics of the ECG signal. Then, the study continued with the feature extraction stage by extracting QRS complexes using Pan Tompkins algorithm from each ECG cycle. The illustration of this stage from all subjects is shown as in Figure 5 and 6.



**Figure 5.** QRS Complex Subject 2



**Figure 6.** QRS Complex Subject 5

After obtaining QRS complexes, this study proposed a machine learning technique for the classification process. In order to validate the correct data classification, the ML technique was applied using WEKA software. The first dataset contains the readings of QRS data with PVT1 signals. The second dataset contains the readings of QRS data with PVT2 signal. While the third dataset contains the readings of QRS data of the combination between PVT1 and PVT2 signal, these dataset acts as the input for the software, and selected classifiers will produce the results that differentiate between PVT1 and PVT2 signals. The results produced by the classifiers are shown in Table 1.

**Table 1.** Result of classifiers in WEKA software

Classification Technique	Percentage of PVT1	Percentage of PVT2	Percentage of combination of PVT 1 and PVT2
MLP	92.5 %	97.5 %	77.5 %
IBk	92.5 %	97.5 %	70.0 %
BN	92.0 %	95.0 %	70.0 %

From the result obtained, the percentage of each classification techniques for PVT1, PVT2 and combination between both PVTs are shown. All three proposed classifiers which are MLP, IBk and BN algorithms produced high percentage for the PVT1 and PVT2 signals, separately. This result clarifies that both PVTs are correctly classified in differentiating both PVTs since both PVTs are not in the same environment which PVT1 were taken in the morning, while PVT2 was taken on the next day, during night time without sufficient amount of light that makes them prompt to deep sleep. However, the percentage for the combined signals between PVT1 and PVT2 is significantly low. The MLP and IBk seemed to produce better results than the Bayes Network. This is because the performance of neural networks that inhibit additional hidden layers appear to be much better than networks with lesser hidden layers [22]. Apart from that, neural networks work better with continuous data as compared to the Bayesian network that works better with discrete data. It can be concluded that sleep stage in PVT2 can be detected when the ECG signal of a subject appears to be in a deep sleep from ECG signal in early sleep, which is in PVT1.

## 5. Conclusion

As a conclusion, the objectives of this study have been successfully achieved. This study was established as an efficient method on driver drowsiness monitoring system using ECG and been conducted with the deep understanding of all the stages in this study which are data acquisition, pre-processing, feature extraction and classification such as RRI and Cardioid graph.

In order to check the accuracy of PVT1 and PVT2 data, ML tools were applied with the aid of MLP, IBk and BN in WEKA software. The result proved that PVT1 and PVT2 were correctly classified in their own categories. However, the percentage of PVT1 and PVT2 decreases as the research combined both PVTs. This happens due to both PVTs were incorrectly classified and acknowledge future research to be aware of incorrectly classified of ECG signal based on both sleep stages. Thus, the experimentation results suggest that the data is constantly stable throughout this study.

Even though the vital step is crucial as it determines the driver's condition, there will be some room for improvement for this study by having real-time ECG data in order to verify the accuracy of the suggested technique; it is preferred to acquire real-time data. This is due to some parameters that can't be changed if we used an online database. By having our own data, we might be able to increase our accuracy of this proposed technique. In addition, this study might suggest selecting potential data acquisition device due to the current device limitation can't record ECG signal more than 30 seconds, so by having a device that can record ECG signal with a longer period might help on the improvement of the effectiveness of the proposed system.

## Acknowledgements

This work was funded by the Fundamental Research Grant Scheme (FRGS) from the International Islamic University Malaysia. (Project ID: FRGS19-056-0664)

## References

- [1] N. H. Bahaudin, "[UPDATE] Cuai, mengantuk punca kemalangan," *Harian Metro*, 07-Feb-2019
- [2] R. Ilham, "Padah jika terlelap," *Harian Metro*, 31-May-2019.
- [3] Road Safety Department of Malaysia, Message from the Director General RMP, <http://www.jkjr.gov.my/en/about-us/message-from-the-director-general.html>, Retrieved on 15th February 2017.
- [4] "Akibat mengantuk, lelaki maut langgar pokok," *Astro Awani*, 23-Sep-2018.
- [5] M. Rezaei and R. Klette, "Driver Drowsiness Detection," *Computer Vision for Driver Assistance Computational Imaging and Vision*, pp. 95–126, 2017.
- [6] Arun Sahayadhas, Kenneth Sundaraj, "Detecting Driver Drowsiness Based on Sensors A Review", pp.16937-16953, ISSN 1424-8220, Malaysia 2012
- [7] G. Zhenhai, L. Dinhdatt, H. Hongyu, Y. Ziwen, and W. Xinyu, "Driver Drowsiness Detection Based on Time Series Analysis of Steering Wheel Angular Velocity," *2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 2017.



- [8] A. Čolić, O. Marques, and B. Furht, "Design and Implementation of a Driver Drowsiness Detection System - A Practical Approach," *Proceedings of the 11th International Conference on Signal Processing and Multimedia Applications*, 2014.
- [9] T. Hwang, M. Kim, S. Hong, and K. S. Park, "Driver drowsiness detection using the in-ear EEG," *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016.
- [10] Basic Waveform of ECG Signal in PQRST Morphology, (Retrieved on 7th Feb 2017) from <https://thephysiologist.org/study-materials/the-normal-ecg/>
- [11] Sidek, K. A., Khalil, I., & Jelinek, H. F. (2014). ECG Biometric with Abnormal Cardiac Conditions in Remote Monitoring System. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*,44(11), 1498-1509. doi:10.1109/tsmc.2014.2336842
- [12] Valchinov, E., Antoniou, A., Rotas, K., & Pallikarakis, N. (2014). Wearable ECG System for Health and Sports Monitoring. *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming Healthcare through Innovations in Mobile and Wireless Technologies"*. doi:10.4108/icst.mobihealth.2014.257236
- [13] B.-L. Lee, B.-G. Lee, and W.-Y. Chung, "Standalone Wearable Driver Drowsiness Detection System in a Smartwatch," *IEEE Sensors Journal*, vol. 16, no. 13, pp. 5444–5451, 2016.
- [14] M. Babaeian, N. Bhardwaj, B. Esquivel, and M. Mozumdar, "Real time driver drowsiness detection using a logistic-regression-based machine learning algorithm," *2016 IEEE Green Energy and Systems Conference (IGSEC)*, 2016.
- [15] Healey J.A.; Picard R.W., "Detecting Stress during Real-World Driving Tasks Using Physiological Sensors", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 6, NO. 2, pp. 156- 166, 2005.
- [16] "The ULg Multimodality Drowsiness Database (called DROZY) and Examples of Use", by Quentin Massoz, Thomas Langohr, Clémentine François, Jacques G. Verly, *Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV 2016)*, Lake Placid, NY, March 7-10, 2016.
- [17] W. Liang, S. Hu, Z. Shao and J. Tan, "A real-time cardiac arrhythmia classification system with wearable electrocardiogram," *2011 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems*, Kunming, 2011, pp. 102-106.
- [18] J. Fajardo, D. Astudillo, K. Palacio-Baus, L. Solano-Quinde and S. Wong, "Evaluation of two QRS detection algorithm on ECG stress test database," *2016 IEEE ANDESCON*, Arequipa, 2016, pp. 1-4.
- [19] Multilayer Perceptron, "DeepLearning 0.1 Documentation", Accessed on 27th May 2019, <http://deeplearning.net/tutorial/mlp.html>
- [20] Gilbert, D.; "Nearest Neighbour Classifier", Accessed on 19th April 2019, <http://www.robots.ox.ac.uk/~dclaus/digits/neighbour.htm>
- [21] Bayesian Network, Accessed on 12th January 2019, <http://www.pr-owl.org/basics/bn.php>.
- [22] Neal, Radford M. Bayesian Learning in Neural Networks. Vol. 118. N.p.: Springer Science & Business Media, 2012. p.p 126-33.