EMPIRICAL STUDY OF MUSCLE FATIGUE FOR DRIVER'S ERGONOMIC ANALYSIS DURING PROLONGED DRIVING

BY

NOOR AZLYN BINTI AB GHAFAR

A thesis submitted in fulfillment of the requirement for the degree Master of Science in Mechatronics Engineering

Kulliyyah of Engineering

International Islamic University Malaysia

FEBRUARY 2023

ABSTRACT

Driving has become essential in transporting people from one place to another. However, prolonged driving could cause muscle fatigue, leading to drowsiness and microsleep. Electromyography is an important type of electro-psychological signal that is used to measure electrical activity in muscles. This work classifies and predicts muscle fatigue from trapezius muscle of 10 healthy subjects. The EMG signals and the time when muscle fatigue was experienced by the subjects were recorded. The mean frequency and median frequency of the EMG signals were extracted. For classification of muscle fatigue in non-fatigue and fatigue condition, six machine learning models were used: Logistic Regression, Support Vector Machine, Naïve Bayes, k-nearest Neighbour, Decision Tree and Random Forest. From the value of median frequency and slope coefficient of median frequency, mathematical model was developed with respect to driver's physical factors. The results show that both the median and mean frequency are lower when fatigue conditions exist. In term of the classification performance, the highest accuracy for classifying muscle fatigue due to prolonged driving was obtained by the Random Forest classifier with 85.00%, using both the median and mean frequency of the EMG signals. This method of using the mean and median frequency will be useful in classifying driver's non-fatigue and fatigue conditions and predict muscle fatigue during prolonged driving. The significant factor influencing muscle fatigue of the driver was BMI. This study successfully developed mathematical model of second order polynomial of muscle fatigue and BMI (p<0.05 and the R^2 = 0.85). The model was successfully validated where the residual errors compared between predicted values and actual values were less than 10%.

ملخص البحث

أصبحت قيادة المركبات امراً ضرورياً في نقل الأشخاص من مكان إلى آخر، إلا أن القيادة لفترات طويلة قد تسبب تعبًا في العضلات مما يؤدي إلى النعاس او الغفوة القصيرة. يعد تخطيط كهربية العضل (EMG) نوعًا مهمًا للإشارة النفسية الكهربائية المستخدمة في قياس النشاط الكهربائي في العضلات. تصنف هذه الدراسة وتتنبأ بجُهد عضلة من عضلات شبه منحرفة لعشرة أشخاص أصحاء. تم تسجيل إشارات تخطيط كهربية العضل (EMG) والوقت الذي عانى فيه الأشخاص من التعب العضلي، واستخراج متوسط التردد ووسيط التردد للإشارات تخطيط كهربية العضل (EMG). ولتصنيف تعب العضلات في حالة عدم الإرهاق والتعب تم استخدام ستة نماذج للتعلم الآلي وهي: الانحدار اللوجستي Logistic Regression، وآلة المتجهات الداعمةSupport Vector Machine ، وبيز ساذج Naïve Bayes والجار الأقرب لنقطة الاختبار KNN وشجرة القرارات Decision Tree ومصنف الغابة العشوائية Random Forest. من قيم متوسط التردد ومعامل الانحدار لمتوسط التردد تم تطوير نموذج رياضي متعلق بالعوامل الفيزيائية للسائق. أظهرت النتائج أن كلاً من متوسط التردد ووسيط التردد يكونوا أقل عند وجود حالات التعب. ومن حيث أداء التصنيف، تم الحصول على أعلى دقة في تصنيف التعب العضلي بسبب القيادة الطويلة بنسبه 85% ذلك بواسطة نموذج Random Forest ونتائج متوسط التردد و الوسيط لإشارات تخطيط كهربية العضل(EMG) . أن هذه الطريقة سوف تكون مفيدة في تصنيف حالات عدم إرهاق السائق وتعبه، وستمكن من التنبؤ بإجهاد العضلات أثناء القيادة لفترات طويلة. في هذه الدراسة كان العامل المهم هو مؤشر كتلة الجسم (BMI) الذي يؤثر على تعب عضلات السائق. نجحت هذه الدراسة في تطوير نموذج الرياضي من الدرجة الثانية متعدد الحدود لإرهاق العضلات ومؤشر كتلة الجسم بمقياس (P <0.05) و (R^2 = 0.85). تم التحقق من صحة النموذج بنجاح حيث كان معامل الخطأ المتبقي مقارنة بين القيم المتوقعة والقيم الفعلية اقل من 10٪.

APPROVAL PAGE

I certify that I have supervised and read this study and that in my opinion, it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science in Engineering.

DR. NUR LIYANA AZMI
Assistant Professor
Department of Mechatronics Engineering
Kulliyyah of Engineering
International Islamic University Malaysia

…………………………………
Nur Liyana Azmi
Supervisor

DR. KHAIRUL AFFENDY MD NOR
Assistant Professor
Department of Mechatronics Engineering
Kulliyyah of Engineering
International Islamic University Malaysia

…………………………………
Khairul Affendy Md Nor
Co-Supervisor

DR. NOR HIDAYATI DIYANA NORDIN
Assistant Professor
Department of Mechatronics Engineering
Kulliyyah of Engineering
International Islamic University Malaysia

…………………………………
Nur Hidayati Diyana Nordin
Co-Supervisor

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science in Engineering.

PROF IR DR SITI FAUZIAH TOHA
Department of Mechatronics
Faculty of Engineering
International Islamic University Malaysia
PE NO. P117184

…………………………………
Siti Fauziah Toha @ Tohara
Internal Examiner

…………………………………

External Examiner

This thesis was submitted to the Department of Mechatronics and is accepted as a fulfilment of the requirement for the degree of Master of Science in Engineering.

……………………………….
Ali Sophian
Head, Department of Mechatronics

This thesis was submitted to the Kulliyyah of Engineering and is accepted as a fulfillment of the requirement for the degree of Master of Science in Engineering.

……………………………….
Sany Izan Ihsan
Dean, Kulliyyah of Engineering

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Noor Azlyn Ab Ghafar

Signature…………………………… Date……10-Feb-2023………

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

EMPIRICAL STUDY OF MUSCLE FATIGUE FOR DRIVER'S ERGONOMIC ANALYSIS DURING PROLONGED DRIVING

I declare that the copyright holder of this thesis are jointly owned by the student and IIUM.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Noor Azlyn Ab Ghafar

………………………………	10-Feb-2023 …………………..
Signature	Date

This thesis is dedicated to my late father, mother and husband for laying the foundation of what I turned out to be in life.

ACKNOWLEDGEMENTS

All glory is due to Allah, the Almighty, whose Grace and Mercies have been with me throughout the duration of my programme. Although, it has been tasking, His Mercies and Blessing on be ease the herculean task of completing this thesis.

I am most indebted to by supervisor, Dr Liyana Azmi whose enduring disposition, kindness, promptitude, thoroughness and friendship have facilitated the successful completion of my work. I put on record and appreciate her detailed comments, useful suggestion and inspiring queries which have considerably improved this thesis. Her brilliant grasp of the aim and content of this work led to her insightful comment, suggestions and queries which helped me in great deal. Despite her commitments, she took time to listen and attend to me whenever requested. The moral support she extended to me is in no doubt a boost that helped in building and writing the draft of this research work. I am also grateful to my co-supervisor, Dr Khairul Affendy Md Nor and Dr Nor Hidayati Diyana Nordin whose support and cooperation contributed to the outcome of this work.

Lastly, my gratitude goes to my beloved husband and my loving mother; for their prayers, understanding and endurance while away.

Once again, I glorify Allah for His endless mercy on me of which is enabling me to successfully round off the efforts of writing this thesis. Alhamdulillah.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EMG	Electromyography
MNF	Mean Frequency
MDF	Median Frequency
BMI	Body Mass Index
YOD	Years of Driving
SENIAM	Surface EMG for Non-Invasive assessment of Muscle
H_0	Null Hypothesis
H_a	Alternative Hypothesis
ANOVA	Analysis of Variance
MVC	Maximum Voluntary Contraction
UAV	Unmanned Arial Vehicle
SVM	Support Vector Machine
kNN	k-Nearest Neighbor

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

Road transportation is one of the major modes of transportation used by Malaysians. Driving has become more important because it a is fast, cheap, and practical way of moving people from one place to another (Kamat et al. 2020). According to a Ministry of Transport Malaysia (MOTM) report, the number of register vehicle recorded in 2020 was 32.28 million. In 2021, the number had increased to 33.57. In addition, car have been recorded as the type of vehicle used most frequently by Malaysian with a rate of 47.10% followed by motorcycle at a rate of 46.19 in 2021 (MOTM 2022).

As a developing country, Malaysia gains income from greater productivity, and at the same time requires people to move faster and further (Sanjaya, Lee, and Katsuura 2016). Therefore, the transportation system of roads and highways should be greatly improved, which will enable Malaysians to experience better infrastructure, facilities and comfort. Although the increase in driving activity offers major benefits, it also has negative effects due to the increasing number of road accidents (Ani, Kamat, and Husin 2017). Malaysia has one of the highest rates of road accidents worldwide in relation to its population. Since 2012 to 2018, Malaysia has been ranked as the seventh-highest country in the world for the overall number of traffic accident. Additionally, Malaysia has had the greatest global mortality per 100,00 people since 1995 (Wan Husin et al. 2021).

In 2019, the Road Safety Department of Malaysia recorded 5764 cases of fatal accidents (Mahat, Jamil, and Sarah Raseli 2020). The three main causes of traffic accidents are human, environmental, and technical factors (Hawa Harith and Mahmud

2018). According to Malaysian Institute of Road Safety Research (MIROS) reported that the main contributor to road accident is human factor as much as 80% (Shariff et al. 2022). Mahat et al. (2020) categorized human factor into subfactor and according to their finding, the first ranking is drunk driving while drowsiness or microsleep rank as second factor contributing to road accident (Mahat et al. 2020). Fatigue is one of the factors leading to microsleep or drowsiness of the drivers besides prolonged driving, road condition, environment and health (Zaleha et al. 2021).

Research defines fatigue as a lack of ability to exert additional force or power (Al-Mulla, Sepulveda, and Colley 2011). Fatigue detection is important in many areas such as the health sector to monitor health and welfare of the patients. For example, electromyography (EMG) is implemented in the use of prosthetic control devices. Muscle tiredness detection and classification are also crucial in the fields of human-computer interactions, sports injuries and performance, and ergonomics. Muscle tiredness is one of the most common causes of injuries in athletes, and it is usually identified after the muscle has already been injured. (Freitas 2008). When muscular exhaustion is not diagnosed early enough, it can cause pain and also financial hardship. In addition, the most expensive therapy in this world according to Tlili et al., (2021) is spinal therapy (Tlili et al. 2021). As a result, detecting muscular exhaustion before it becomes obvious is critical.

Driving on the highway involves a monotonous driving environment because of the wide and flat pavement, fewer spatial references and high volume of traffic (Fu, Wang, and Zhao 2016). Prolonged driving in this type of environment requires drivers to sustain attention over long a period which decreases their alertness performance and lead to fatigue.

An important measure in the ergonomics of car seats during driving is the selection of the seat inclination angle to increase the driver’s comfort, reduce fatigue, and avoid musculoskeletal disorders (Ferrari and Croft 2001). Selection of seat inclination angle will affect the spine posture of the driver especially during prolonged driving. The weight distribution supported by the seat-pan and backrest, as well as the boundary condition of upper body vibration and the spine’s curvature, are expected to change when sitting with an inclined backrest (Liu and Qiu 2021).

This study classifies EMG signal of non-fatigue and fatigue condition of the driver during prolonged driving using Machine Learning technique. In addition this work also develops a mathematical model to find the relationship between driver’s physical factors and muscle fatigue during prolonged driving. According to J. Barbosa (2003), the definition of a mathematical model is the behavior of real devices and objects is represented mathematically. Modeling a device or system is essential for both engineers and scientists. The mathematical model developed needs to be validated to ensure that the model is accepted.

This study is significant in detecting and predicting muscle fatigue. In addition, this study will also reveal the relationship between physical factor of the driver and muscle fatigue so that drowsiness and microsleep can be prevented. This study focuses to monitor the muscle activity of the subject using EMG signal which it aimed to prevent musculoskeletal disorder in a longer time. In directly, the risk and number of accidents associated with driving fatigue can be minimized. In addition, this study will also reveal the relationship between physical factor of the driver and muscle fatigue.

1.2 PROBLEM STATEMENT

In the recent year, several advanced signal processing algorithms and machine learning methods have been used in the researches (Karthick, Ghosh, and Ramakrishnan 2018). For the classification of non-fatigue and fatigue condition of the driver, several methods had been proposed previously using different classification techniques and different psychophysical signal (Bhardwaj, Natrajan, and Balasubramanian 2018). To date, Machine Learning classification of muscle fatigue using EMG has mainly focused on the areas of rehabilitation, sports science, human-computer interaction and medical research. However limited research had been conducted in the field of driving.

In addition, it is important to relate the physical factor affecting muscle fatigue of the driver during prolonged driving in order to prevent musculoskeletal injury and accident due to fatigue. Ani et al., (2017) developed and validated a mathematical model of driver fatigue using driving duration, road type, gender, the relation between gender and road type, as well as the relation between driving duration and road type as the input parameters (Ani et al. 2017). Meanwhile Fu et al. (2016) developed a mathematical model based on the Hidden Markov Model (HMM) that used EMG, Electroencephalograms (EEG), and respiration signals, as well as contextual information such as the driver's sleep quality, driving conditions, and circadian rhythm (Fu et al. 2016). Lastly, Wang et al. (2017) developed a model for driver fatigue based on ECG and EMG data using non-contact sensors (Wang, Wang, and Jiang 2017). Currently, less research has been undertaken to develop a mathematical model based on the physical factors of body mass index (BMI), age, and years of driving (YOD) to determine muscle fatigue during driving.

1.3 OBJECTIVE OF THE RESEARCH

The objectives of this research are:

1. To classify non-fatigue and fatigue conditions of the driver during prolonged driving using electromyography (EMG) signal.
2. To identify significant physical factors (body mass index (BMI), age, and years of driving) related to muscle fatigue of the driver during prolonged driving.
3. To formulate and validate the mathematical model of muscle fatigue with respect to the driver's physical information.

1.4 SIGNIFICANCE OF THE RESEARCH

To date, muscle fatigue classification has mainly focused on the areas of rehabilitation, sports science, human-computer interaction and medical research. However, limited research on muscle fatigue classification has been conducted in the field of driving using EMG and Machine Learning. This is an important topic as driving fatigue leads to accidents and loss of life. In terms of modeling muscle fatigue, less research is done on modeling muscle fatigue with respect to driver's physical factors. This research is able to classify the muscle fatigue during prolonged driving. This research has also developed and validated the mathematical model of muscle fatigue and drivers' physical factors during prolonged driving.

1.5 SCOPE OF THE RESEARCH

This research is divided into two parts: experimental and mathematical modeling. In the experimental part, an EMG sensor was used in this research to study the muscle activity of the driver. The EMG signal was measured using BITalino biosignal acquisition board. The EMG signal was further pre-processed using MATLAB software. The targeted muscle was the trapezius (shoulder) muscle. Ten healthy subjects with age between 20 and 40 years old were recruited. The subjects needed to drive a car for 2 hours using a highway route. The road condition is monotonous and the experiment took place at the East Coast Expressway Phase 2, Malaysia. The type of car used in this experiment was Perodua Axia with automatic transmission. The seat inclination angle was set to 10°. All subjects needed to maintain a driving speed of 90km/h during the experiment.

The EMG signal was used to classify non-fatigue and fatigue condition of the driver using six Machine Learning algorithms namely: Logistic Regression, Support Vector Machine, Naïve Bayes, k-nearest Neighbour, Decision Tree and Random Forest. In the mathematical modeling part, the classification of muscle fatigue was carried out using Analysis of Variance (ANOVA) analysis. The most significant parameter was identified and the mathematical model developed was validated by calculating the residual error.

1.6 CONTRIBUTION OF THE RESEARCH

This research successfully classified the muscle fatigue condition of the trapezius (shoulder) muscle during prolonged driving. The relationship between muscle fatigue and the parameters was analyzed. Based on the results, the BMI of the subject contributed the most to muscle fatigue. A mathematical model of second-order polynomial using BMI and muscle fatigue of the driver during prolonged driving was successfully developed. The model was validated by calculating the residual error. Two sets of journal articles were successfully published throughout this work. First, a journal article entitled "Mathematical Model of Physical Factor for Driver Fatigue during Prolonged Driving" was published by Journal of Engineering and Technology (JET). Secondly, a journal article entitled "Classification of Muscle Fatigue during Prolonged Driving" was published by ELEKTRIKA, Journal of Electrical Engineering.

1.7 ORGANIZATION OF THE THESIS

This thesis consists of five main chapters.

Chapter 1 explains the background of this research, problem statement, objectives, and scope of the research. This chapter also explains the significance of this research and the publications that have been produces so far.

Chapter 2 presents the literature review of previous research to obtain important technical and scientific knowledge related to this research.

Chapter 3 explains the methodology of this research in detail which consists of participants, experimental procedure, data processing, classification and regression with the aid of a flowchart.

Chapter 4 presents and discusses all results of EMG data processing, classification of non-fatigue and fatigue conditions of the drivers and regression to develop a mathematical model based on the significant physical parameter of the drivers. All findings of this research are highlighted in this chapter.

Chapter 5 presents the conclusion of the research and recommendations for future work.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter presents the extensive literature review of the optimum seat inclination angle of the driver, driving duration for muscle fatigue, the study on electromyography (EMG) sensors, EMG sensor placement on targeted muscle during prolonged driving, EMG signal pre-processing and the study of muscle fatigue. In addition, this chapter explains the classification of fatigue in machine learning and development and validation of mathematical model of muscle fatigue based on the previous research.

2.2 OPTIMUM SEAT INCLINATION ANGLE

An important feature in an ergonomics study of the driver is the selection of seat inclination angle. According to Ferrari & Croft (2001), previous study stated that ideal backrest angle is 120° as Figure 1. However, this angle will cause head flexion and neck pain. This is because drivers need to see through the windshield within their eye level. Thus, the optimal seat back angle needed is suggested to be at 100˚ (10˚ with respect to the vertical axis).

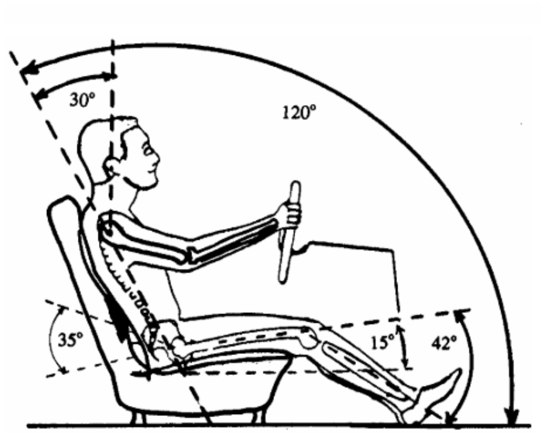


Figure 1 Backrest of 120˚ will cause abnormal 30˚ head flexion (Ferrari and Croft 2001)

Majid et al. (2013) investigated the driver's optimal seat adjustment using a rigid-body model. The model analyzed various seat-pan and backrest inclinations. The result proposed that the optimal adjustment for the car seat is 10˚ for seat inclination angle and 0˚ degree to 5˚ for the seat-pan inclination (Majid et al. 2013).

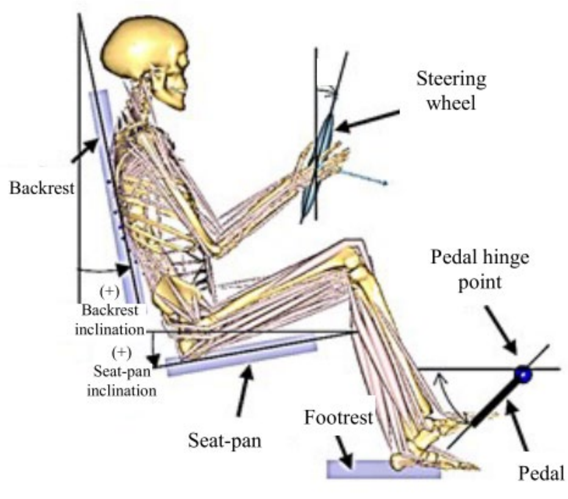


Figure 2 Model of a seated human by Majid et al. (2013)

Li, Zhang and Lv (2015) conducted a study to determine the effects of backrest inclination and vibration frequency on muscle activity using a full-body musculoskeletal system of a seated person in an adjustable car seat. Ten subjects were exposed to whole-body vibration with different backrest inclination angle of 5˚ to 30˚ with increment of 5˚. In this study, muscle oxygenation was measured using near-infrared spectroscopy. This study concludes that vibration frequency significantly influenced the muscle activity of the lumbar area. In addition, they suggested that a small backward of the backrest's inclination angle (approximately 10˚) may lessen the driver's muscle fatigue.

Figure 3 Experimental setup by Li et al. (2015)

In summary, the best seat inclination angle for a driver during driving is 10˚ as concluded in previous research. Therefore, in this research, the seat inclination angle is set to 10˚ throughout the experiment.

2.3 ELECTROMYOGRAPHY (EMG)

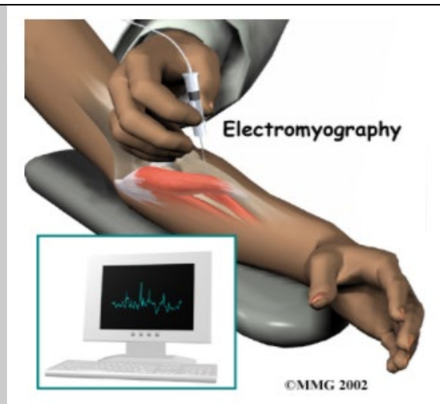
EMG stands for electromyography which is a sensor that measure the response of electrical activity of the muscle in response to the muscle's nerve stimulation (Elamvazuthi et al. 2015). EMG is an experimental practice for generating, recording, and analyzing myoelectric signals. The physiological difference in the condition of muscle fiber membranes produces myoelectric signals (Konard 2012). EMG is widely used especially in medical research, rehabilitation, ergonomics and sports science.

2.3.1 EMG SENSORS

There are two types of EMG muscle sensors available in the market namely intramuscular EMG and surface EMG (Kiswanto et al. 2018). Intramuscular EMG is also called an invasive electrode which uses a needle to penetrate the skin. This type of EMG is not preferable to be used by researchers because only certified personnel is able to perform this, test and this type of EMG will make the subject feel uncomfortable.

On the other hand, surface EMG uses a non-invasive electrode for measuring muscle electrical activity on the skin's surface of the subject. Surface EMG signal measures electrical activities. Two electrodes or more are needed for this type of EMG to measure the difference in potential (voltage) between them. The electrode is affordable and can be placed in any muscle for any purposes (Al-Mulla et al. 2011).

Table 1 Different types of EMG sensors

INTRAMUSCULAR EMG	SURFACE EMG
	

2.3.2 SURFACE ELECTRODE PLACEMENT

The position and orientation of EMG electrodes plays vital role for accuracy and repeatability of EMG signal amplitude(Toro et al. 2019). In order to avoid strength and quality reduction of EMG signal, the electrode needs to be placed in the middle of the muscle and parallel with the muscle fiber's orientation (Technologies 2015). As mentioned earlier, non-invasive EMG needs two electrodes. Thus, the placement of these electrodes should be based on the Surface EMG for Non-Invasive Assessment of Muscles (SENIAM) standard. This standard aims to standardize the placement procedure of EMG sensors for 27 different muscles, the processing of the EMG signals and the modeling of EMG signals (Stegeman and Hermens 2007)(Toro et al. 2019). Figure 1 shows the orientation of the electrode pair in ratio to muscle fiber direction as suggested by SENIAM (Konard 2012). In this work, the targeted muscle was trapezius muscle and the location of EMG electrode is marked with green circle in Figure 4. Another electrode besides two electrodes is needed for measuring muscle electrical activity. This electrode is called a reference electrode and is placed at the unaffected muscle area such as the bony area and joint but close to the targeted measuring muscle. In this research, the trapezius muscle is selected as the targeted muscle.

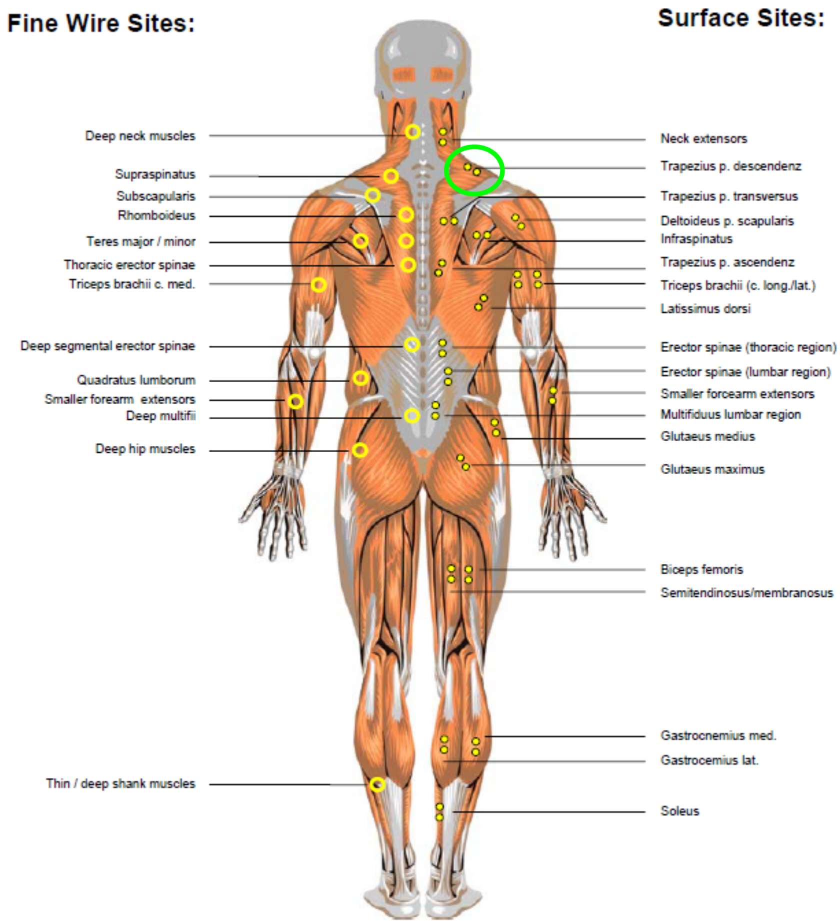


Figure 4 Anatomical position of invasive (fine wire sites) and non-invasive (surface sites) of EMG electrodes as suggested by SENIAM (Konard 2012)

2.3.3 EMG FREQUENCY SAMPLING AND FILTERING

For the EMG signal, the sampling frequency value must be greater than or equal to 1000 Hz to avoid the loss of the signal (Konard 2012). In addition, Chang (2012) also suggested that the sampling frequency of surface EMG measurement should be more than 1,000Hz.

The EMG signal needs to be filtered in order to remove unwanted components or features of the signal and remove noise. A suitable filter configuration will boost the visibility of a faulty signal greatly. (Tengku Zawawi et al. 2018). Table 2 shows the list of the sampling frequency, type of filter and bandwidth frequencies from previous research. From the table, most of the previous research used a sampling frequency of 1,000Hz - 2,000Hz.

Table 2 List of sampling frequency, type of filter and bandwidth used in previous research

References	**Sampling Frequency (Hz)**	**Type of Filter**	**Bandwidth (Hz)**
Karthick, Ghosh, & Ramakrishnan, (2018)	10,000	Band pass Notch	10-400 50Hz
Venugopal, Navaneethakrishna, & Ramakrishnan, (2014)	10,000	Band pass Notch	20-400 50

Khairul Amri Kamarudin et al., (2018) and Khairuddin et al.,(2021)	1,200	Band pass	5-500
Papakostas, Kanal, Abujelala, Tsiakas, & Makedon, (2019)	1,926	Median filtering technique	Not mention
Lejun Wang et al., (2018)	1,000	Forth order band pass	5-500
Lin Wang, Wang, & Jiang, (2017)	1,000	Notch	50
Menotti et al., (2015)	1,000	Band pass	10 - 400
(Balasubramanian & Adalarasu, (2007)	1,000	6th Order Band pass	15 - 500
Katsis, Ntouvas, Bafas, & Fotiadis, (2004)	800	Band pass	100 - 200

In the work, the frequency sampling was set to 1000Hz as suggested by previous research. Then, the raw EMG signal recorded was filtered using band pass filter with a range of 20-500Hz to remove noise at the high-end cut-off and motion artefacts at the low-end cut-off.

2.4 SELECTION OF MUSCLE

Balasubramanian & Adalarasu, (2007) conducted an experiment to analyze muscle activity changes in the shoulder and neck muscles while gaming in an automobile simulator. According to their study, the muscles in the upper part of the body, including the trapezius muscle, are the predominant muscles to be studied while driving. The EMG electrodes were placed at the right splenius capitius (neck area), right trapezius and right medial deltoid. There were two groups of participants: non-professional drivers and professional drivers. Figure 5 shows the experimental setup of their study. According to their findings, both groups showed a statistically significant (p<0.05) change in all muscle activity during a brief (15 min) gaming session. (Balasubramanian and Adalarasu 2007).

Figure 5 Experimental setup by Balasubramaniam et al. (2007)

Another study by Hostens et al. (2005) measured the trapezius and deltoid muscles of the driver during different driving conditions. Work that is physically boring (monotonous) or repetitive is linked to an increase in the low back, shoulder, and neck pain. The result of the study stated that for 1 hour of driving, signs of fatigue were present in both muscles (Hostens and Ramon 2005). Lee et al. (2017) stated that the trapezius muscle has a high amount of muscular activation when in stress conditions (Boon-Leng, Dae-Seok, and Boon-Giin 2016). Their study concluded that muscle activity is one of the reliable indicators to differentiate emotion (relaxation, fatigue, stress).

Trapezius muscle has also been studied in office syndrome where a worker needs to sit in one position for a long duration of time. Pratummas et al. (2022) conducted an experiment to classify the muscle activity of a worker in non-fatigue and fatigue conditions. Another research by Mork et al. (2007) aimed to find the effect of arm posture and movement of the trapezius muscle for computer workers who work a full day (Mork and Westgaard 2007).

Based on previous studies about muscle fatigue of the driver during driving and office worker stated earlier, the trapezius muscle is the type of muscle normally studied by the research. Thus, in this study, the trapezius muscle was selected as the tested muscle for the driver in the experiment.

2.5 MUSCLE FATIGUE

Monitoring muscle fatigue is essential in all fields because it can prevent injury in a long term. An EMG sensor measures electrical activity over the surface of the skin and is widely used to study muscle fatigue. Muscle fatigue is defined as a condition where the muscle is unable to produce the force needed (Karthick et al. 2018).

Muscle fatigue can occur due to the contraction of muscle. There are two types of muscle contraction namely isotonic contraction and isometric contraction (Chang, Liu, and Wu 2012). The isometric contraction is also known as dynamic contraction. During isometric contraction (dynamic contraction), muscle contract and relax rhythmically in order to maintain the same force (Tanvi Khurana & Suman Singh 2017). However, for isometric contraction, muscle remains contract in same state for a long period to maintains the same position. Even with low contraction, isometric contraction will result in muscle fatigue in a long period of time(Bhardwaj, Parameswaran, and Balasubramanian 2018).

During prolonged driving, the driver needs to maintain the position for controlling the steering wheel and pedal and to keep looking at the road and surrounding environment. Therefore, to maintain the position, the muscle need to maintain the neck, upper and lower limb position and stabilize the trunk (Lecocq et al. 2020)(Jung et al. 2021). This situation is an example of isometric (static) contraction where the posture is maintained without any movement. The same situation (isometric contraction) is explained by other researchers for isometric contraction during prolonged sitting in office work (Jia 2020)(Kett, Milani, and Sichting 2021).

In previous research, the median frequency (MDF) and mean frequency (MNF), based on the Fourier Transform of EMG signals have been used for muscle fatigue assessment (Hostens and Ramon 2005)(Sonmezocak and Kurt 2021). Nowadays there are many other advanced features of EMG signal can be computed to analyze muscle fatigue for example Normalized Spectral Moment, Spectral Entropy(Karthick et al. 2018)(Wang et al. 2017), multifractal detrended moving average algorithm (Marri and Swaminathan 2016) and Spectral Flux, Zero Crossing Rate, Willson Amplitude (Papakostas et al. 2019). However, in this work, only MDF and MNF were used to analyze muscle fatigue because MDF and MNF value were easier to compute, served as fundamental of muscle fatigue index and still widely used in the research(Yousif et al. 2019)(Ramos et al. 2020)(Zhao et al. 2022).

When muscle fatigue occurs, the blood flow to the muscle decreases because the muscles contract intensely, reducing the blood flow and thus the availability of oxygen. Otherwise, the muscle is working so hard that there is not enough oxygen to keep up with the demand (Ani et al. 2017). Energy reserves (sugar and phosphorus) are depleted, lactic acid and carbon dioxide levels rise, and muscle tissue becomes acidic (Lal and Craig 2001). As a result, the conduction velocity of the motor action potential on the muscle membrane slows down(Toro et al. 2019). Thus, the power spectrum of the EMG signals recorded from the muscle shifts towards lower frequencies when the muscles are in a fatigue condition. Consequently, both MDF and MNF values in non-fatigue conditions are higher than those obtained in fatigue conditions (Venugopal et al. 2014).

The regression coefficient of the MDF or MNF slope towards lower frequencies can be used as a non-invasive fatigue index for the investigated muscle. Moreover, the slope coefficient of the linear regression analysis of MDF is an important index of muscle fatigue (Candotti et al. 2009). The negative amplitude of the slope coefficient shows a higher level of fatigue (Chang et al. 2012).

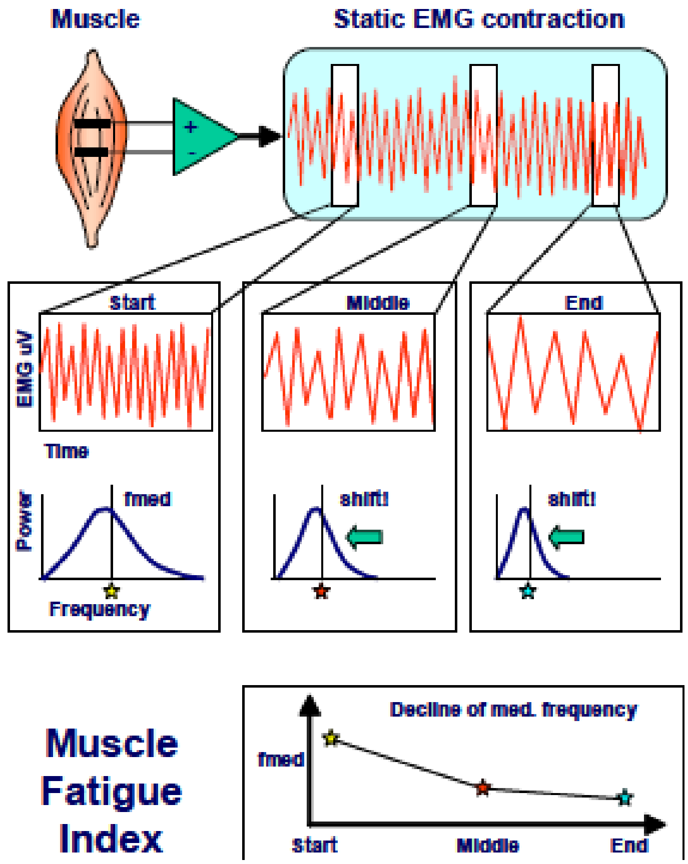


Figure 6 Illustration of frequency shifting to low frequency during fatigue by Konard et al. (2012)

2.6 CLASSIFICATION OF MUSCLE FATIGUE

As mentioned before, muscle fatigue detection is important in many fields. One of the common fields to detect muscle fatigue is sports science. The early detection of muscle fatigue is important to evaluate the performance of athletes and to avoid any injury. Papakostas, Kanal, Abujelala, Tsiakas, & Makedon (2019) studied physical fatigue due to muscle exhaustion based on objective EMG measurement and identify the presence of physical fatigue based on subjective user-report (Papakostas et al. 2019). In this research, the subjects were asked to hold and move the end-effectors of the robotic arm while performing exercises. During this process, the robotic arm would provide resistive forces to the subject. The EMG data were recorded from the deltoid and triceps of the subject. When the subject felt fatigued, they would inform the researcher who would mark the time point. The classification of muscle fatigue was done using machine learning algorithm.

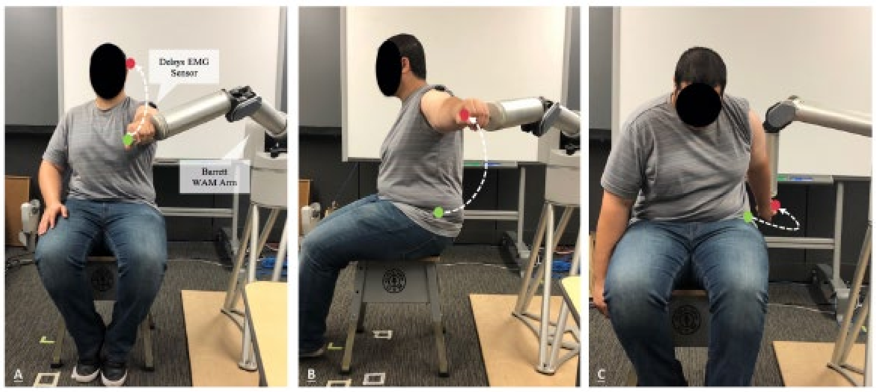


Figure 7 Experimental setup by Papakostas et al. (2019)

Karthick et al. (2018) successfully classified non-fatigue and fatigue conditions of 52 healthy subjects using Machine Learning algorithm by performing biceps curl exercise continuously. The dataset used in the machine learning training and testing was MDF and MNF (Karthick et al. 2018). The most accurate classifier was the Support Vector Machine (SVM) with 91% accuracy. Venugopal et al. (2014) conducted research to differentiate the EMG signals recorded from the biceps brachii of 50 subjects in non-fatigue and fatigue conditions. The EMG signal demonstrated a spectral shift towards the low-frequency area using MDF and MNF from the fatigued muscle. The most accurate classifier was the k-Nearest Neighbor (kNN) classifier with 93% accuracy (Venugopal et al. 2014).

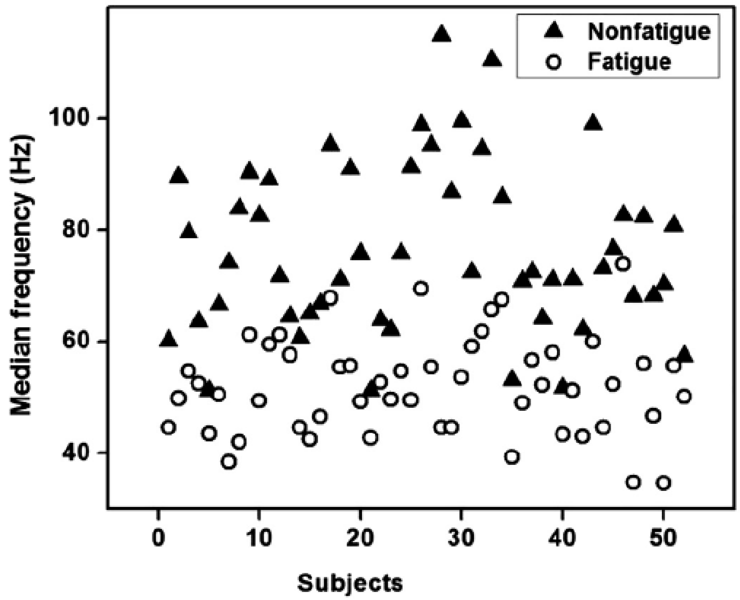


Figure 8 Classification of median frequency of non-fatigue and fatigue conditions by Karthick et al. (2018)

Marri et al. (2015) employed machine learning algorithm to classify muscle fatigue of 26 healthy subjects. The EMG signals were recorded from the biceps brachii (Marri and Swaminathan 2016). The EMG signal were preprocessed and segmented into six sections. The first section was selected as the non-fatigue condition while the sixth section was selected as the fatigue condition. The EMG signal were successfully classified using Logistic Regression and kNN classifiers.

Based on the literature review, the research on muscle fatigue classification mainly focuses on the area of rehabilitation, sports science, human-computer interaction and medical research. However, less research has been done in the field of driving. Classification of muscle fatigue during driving is also important as driving fatigue will lead to accidents and loss of life.

2.7 MATHEMATICAL MODEL OF DRIVER FATIGUE

A mathematical model can be described as a representation of how actual equipment or object will behave in mathematical equations. Ani, Kamat and Husin (2017) performed a study to develop mathematical model of psychophysical factors for driver's fatigue, that can predict the relationship between the process input parameters and output response. The EMG signals were recorded from the trapezius and biceps of the subject during driving. The mathematical model was developed to find the relationship between the input process parameters (exposure time, type of road, and gender) and muscle fatigue. The modeling process used Response Surface Methodology (RSM) and Design Expert 8.0.6 software. The mathematical model developed was successfully validated by computing residual errors (Ani et al. 2017).

Figure 9 Experimental setup by Ani et al. (2017)

Fu et al. (2016) developed a mathematical model using Hidden Markov Model (HMM). Twelve professional bus drivers participated in this research and they needed to drive real highway in 3.5 hours. This research recorded EMG, Electroencephalogram (EEG), respiration signals and also contextual information, such as sleep quality, driving conditions and circadian rhythm of the driver to estimate the fatigue of the driver (Fu et al. 2016).

Wang et al. (2016) established a new method using non-contact sensors to develop a mathematical model to detect driver fatigue. The sensors used were EMG and electrocardiogram (ECG) sensors located inside the cushion of the driver's seat. Twelve subjects were selected and they were requested to continuously drive for two hours using a driving simulator, as shown in Figure 10. The result showed that the model was successfully developed with the model accuracy of 91% using ten-fold cross validation and state validation techniques (Wang et al. 2017).

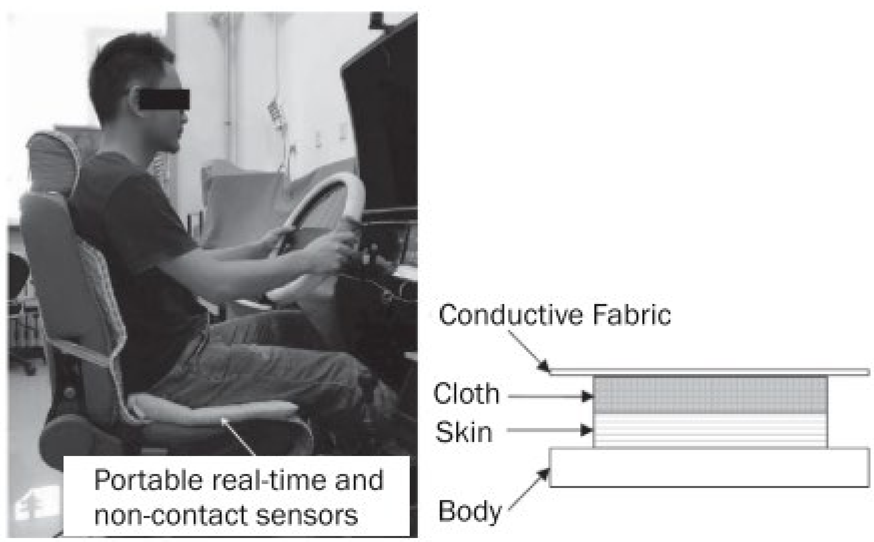


Figure 10 Driving simulator and portable non-contact sensors by Wang et al. (2017)

In summary, mathematical model of muscle fatigue during driving have been investigated previously. However, to date, no research has been undertaken to develop a mathematical model based on the physical factors of body mass index (BMI), age, and years of driving (YOD) to determine muscle fatigue during driving.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 INTRODUCTION

To achieve the stated objectives of this research, non-invasive EMG was used to measure the muscle activity of the driver during driving. The sensor used in this study is a BITalino biosignal acquisition board. This board is inexpensive and multipurpose hardware designed to build any projects by anyone using physiological sensors (Da Silva et al. 2014). A BITalino board consists of seven sensors which are electromyography (EMG), electrocardiography (ECG), electrodermal activity (EDA), electroencephalography (EEG), accelerometer (ACC), push button (BTN) and light sensor (LUX). As compared to other low-cost EMG sensors, BITalino is easier to use because the board is embedded with processors. Hence, users can collect data easily without any programming needed. Moreover, the data are transmitted via Bluetooth directly to the computer. Therefore, it can eliminate the noise coming from the wiring. In addition, an MPU6050 accelerometer and gyroscope sensor were used to set the driver's seat inclination angle.

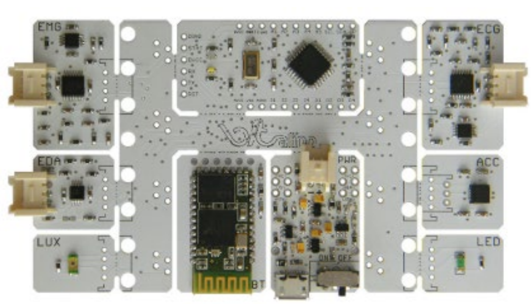


Figure 11 BITalino Biosignal Acquisition board

For the detection of muscle fatigue, six machine learning algorithms were used in this research. The recorded EMG data during the experiment was filtered and the values for mean frequency (MNF) and median frequency (MDF) were extracted during non-fatigue and fatigue conditions based on the subject's report. Lastly, the relationship between driver's information (BMI, age and years of driving) and muscle fatigue was identified and later, mathematical model was developed and validated.

3.2 PARTICIPANTS

Ten healthy subjects (five males, five females; age: 30.8±5.77 years old; height: 164.4±6.06 cm, mass: 64.2±12.70 kg) with no record of sleep-related problems volunteered in this experiment. The age of the subjects is between 20 and 40 years old. All volunteers should have at least 2 years of driving experience. The subjects were prohibited from drinking coffee, tea and other energy drink. The nature of the study and the procedure of the experiment were fully explained to the subjects. The study was approved by the Ethics Committee of the International Islamic University Malaysia (ID No: IREC 2020-069) and a written consent form was obtained from all subjects before the onset of the experimental procedures.

Before starting the experiment, all subjects needed to fill up the consent form (Appendix I) and provide their information in the questionnaire to evaluate the initial condition of the driver. The information asked were gender, age, height, weight and years of driving. After getting all the information, Body Mass Index (BMI) was calculated by dividing the weight (in kilograms) by square of the height (in meters) of the subject (Golmohammadishouraki 2022). The subjects also need to provide their years of active driving. Table 3 provides the summary of the subjects' information.

Table 3 Information about the subjects participated in the experiment

Subject No	Gender	Age (year)	Height (cm)	Weight (kg)	BMI	Years of Driving
001	Female	27	152	43	18.61	9
002	Female	30	168	56	19.84	12
003	Male	23	171	60	20.51	6
004	Male	35	166	70	25.40	18
005	Female	35	158	62	24.84	12
006	Female	32	165	82	30.12	7
007	Female	22	160	49	19.14	4
008	Male	30	170	65	22.49	12
009	Male	38	170	85	29.41	20
010	Male	27	167	65	23.31	12
Mean		**30.8**	**164.4**	**64.2**	**23.37**	**11.4**
Standard Deviation		**5.77**	**6.06**	**12.70**	**4.09**	**5.34**

3.3 EXPERIMENTAL PROCEDURES

3.3.1 ACCELEROMETER SENSOR READING

In this work, MPU 6050 is the sensor used to measure the driver's seat inclination angle. MPU6050 is an integrated board embedded with 3-axis accelerometer and 3-axis gyroscope (Albaghdadi and Ali 2019). MPU6050 selected because the size is small, easy to use, precise and affordable (Al-Hussein et al. 2021). This sensor is widely use nowadays for many applications such as to stabilize the position of Unmanned Arial Vehicle (UAV), self-balancing robot, detecting elderly fall and detecting human posture(Albaghdadi and Ali 2019)(Jian 2017)(Zhang et al. 2020).

In the previous research, Zhang et al., (2020) developed a wearable system to monitor fatigue for neck bending experiment. The MPU6050 is used to detect the tilt angle of the subject's head (Zhang et al. 2020). Albagdadi et al., (2019) used MPU6050 to stabilize the Unmanned Arial Vehicle by measuring and optimizing the roll, yaw and pitch angle of the UAV. According to them, the angles assumed to be zero degrees when the UAV is on the ground and once the UAV takes off, the angle measured change over time accordingly (Albaghdadi and Ali 2019). Al-Hussein et al., conduct research to investigate the social and cultural factor affecting Malaysia's driver. The MPU6050 is used to measure the steering angle of the car during driving and the result stated that male drivers drive aggressively as compared to female drivers during steering maneuvers.

Tlili et al., studied the bad posture in real time in order to prevent spinal pains. According to them, remain prolonged slouching during working or playing with tablet and phone is the common cause of the back pain. The MPU6050 is used to measure the inclination angle of the object in different axes (Tlili et al. 2021). Figure 12 below shows the angle measurement by MPU6050 according to X, Y and Z axis.

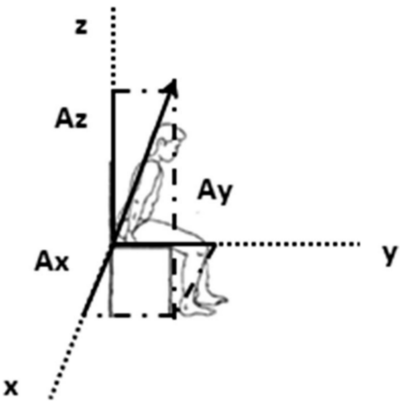


Figure 12 The angle measured by MPU6050 for detecting bad posture of the subject (Tlili et al. 2021)

The MPU6050 accelerometer and gyroscope sensor is able to measure the pitch, roll, and yaw angles of the sensor. The direction and orientation of pitch, roll and yaw are shown in the Figure 13.

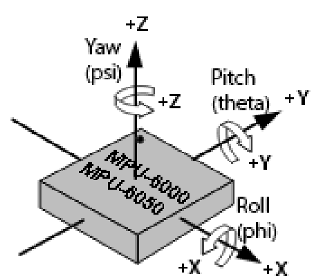
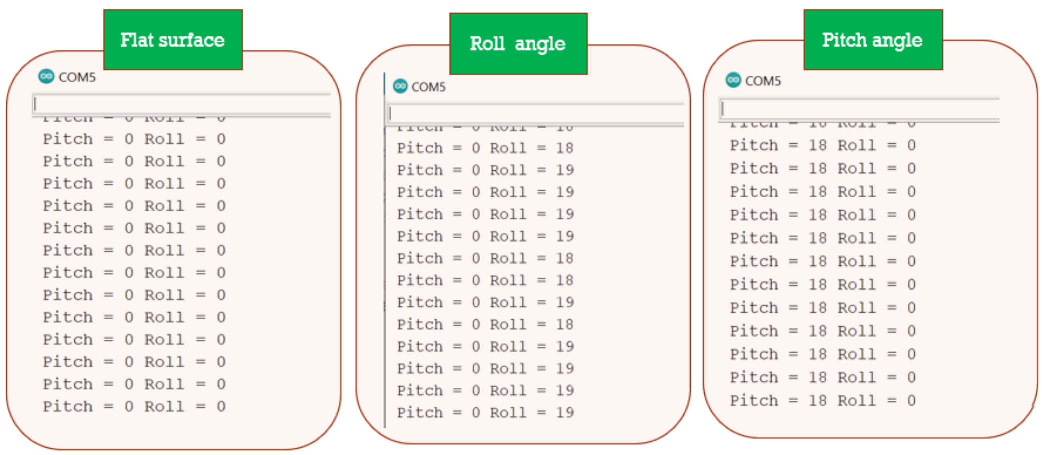


Figure 13 Pitch, roll and yaw of MPU6050 sensor

For this study, the seat angle was set to 10° with respect to y-axis which is suggested as discussed in the literature review chapter as the optimum angle for the driver's comfort (Ferrari and Croft 2001) (Majid et al. 2013)(Li et al. 2015). Initially, the MPU6050 need to be calibrated before starting the experiment. The sensor was placed on a flat surface and the angle was calibrated to zero degree for roll angle. Figure 14 (a) shows the result for the pitch and roll angles of MPU6050 on a flat surface. It can be seen that on flat surface the pitch and roll angles are zero. When the MPU6050 is tilted in the roll angle (Figure 14(b)), the value of the roll angle changes respectively. Meanwhile, when the MPU6050 is tilted in the pitch angle (Figure 14(c)), the value of the pitch angle also changes. In this research, only roll angle was measured as it represents the seat inclination angle of the driver's seat.



(a) (b) (c)

Figure 14 Measurement of pitch and roll angles of MPU6050

Before starting the experiment, the MPU6050 sensor was placed at the upper part of the driver's seat, as shown in the Figure 15 and the set inclination angle was set to zero. After that, the seat was adjusted to 10° with respect to y-axis as shown in the Figure 16. The seat angle was set for every subject and it was done once only before starting the experiment. The seat inclination angle reading was shown in Arduino IDE using serial monitor.

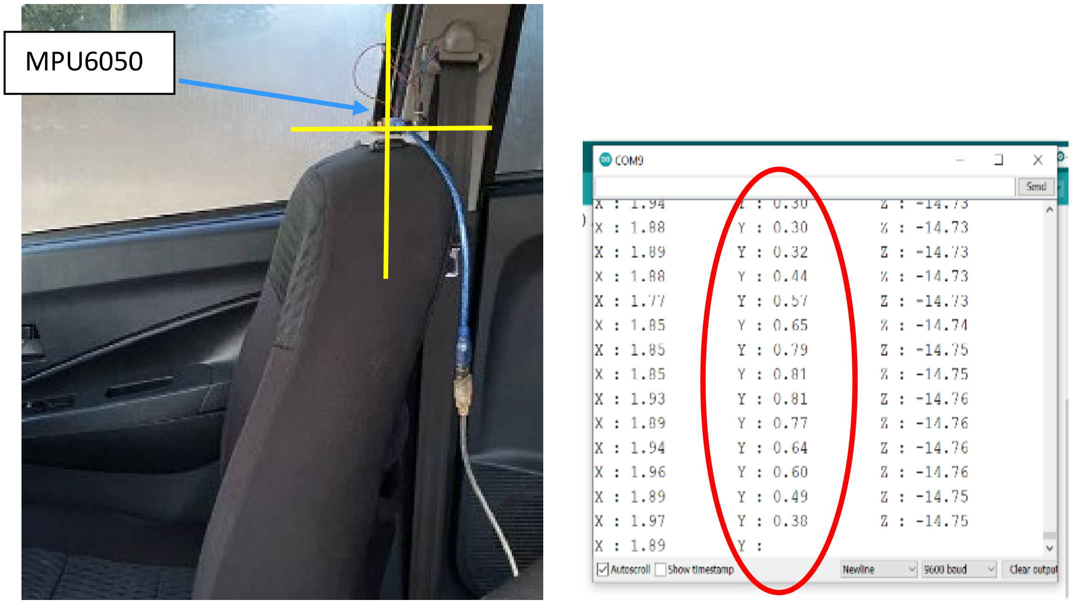


Figure 15 The 0° setting of the seat inclination angle

Figure 16 The 10° setting for the seat inclination angle

This sensor was connected to Arduino Uno as the microcontroller. The circuit of the MPU6050 accelerometer sensor was constructed in Figure 17. The SCL and SDA pin of accelerometer sensor were connected to the SCL and SDA pins of Arduino respectively. In addition, the VCC and GND of accelerometer were also connected to the 5V and GND pins of Arduino respectively.

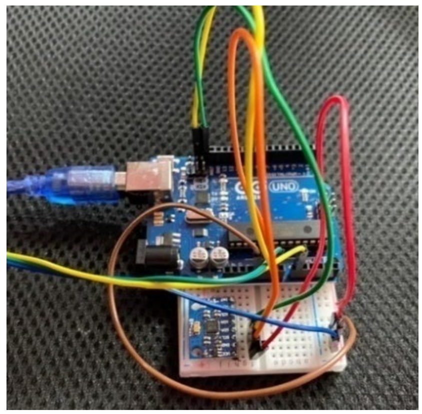


Figure 17 Circuit for Arduino and MPU6050 sensor

3.3.2 EMG SENSOR READING

The EMG sensor used in this study is the BITalino biosignal acquisition board with a sampling rate of 1,000Hz. Before the EMG electrodes were placed over the muscle, the skin surface needs to be cleaned to remove dried skin and dirt by applying alcohol swab (Isopropyl alcohol, approximately 70%). The electrodes were placed on the left trapezius muscle of the driver and the location and configuration of the electrodes conformed with the SENIAM recommendation. The reference electrode was placed on the bony surface of the C7 vertebra as shown in Figure 18.

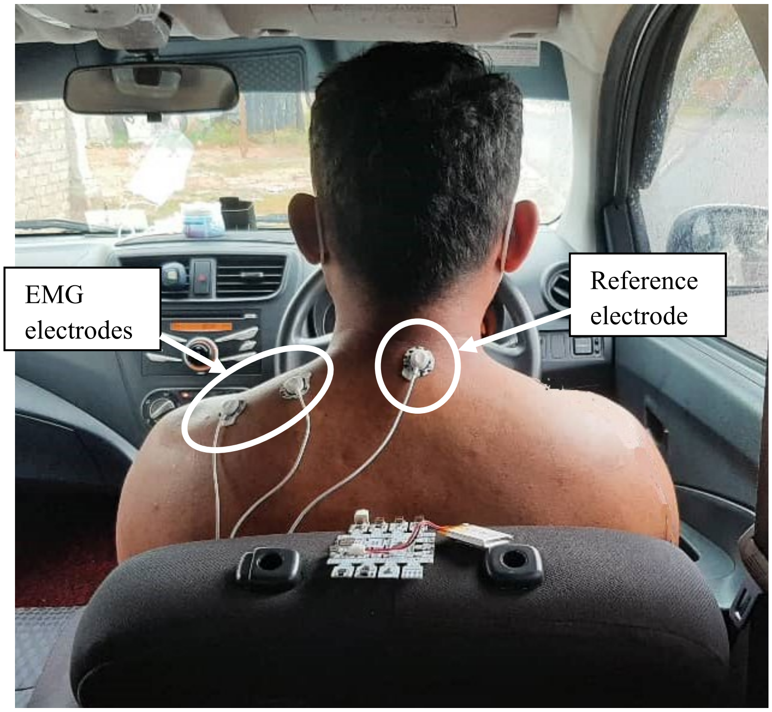


Figure 18 The location of EMG electrodes during experiment

All subjects were suggested to minimize the movement and refrain to use any in-car devices to reduce artefacts during EMG data reading (Fu et al. 2016) and to avoid false EMG reading due to huge movement of the hand. As discussed in Chapter 2, the driver will experience muscle fatigue even in low levels of muscle contraction (isometric contraction) during prolonged driving because the muscle were forced to maintain the same position in longer time. Controlling steering wheel and pedal, observing surrounding and road, needs the driver to maintain the neck position, stabilize the trunk and balancing the lower and upper limb which will result in muscle fatigue (Lecocq et al. 2020)(Tanvi Khurana & Suman Singh 2017).

The EMG data read by the BITalino biosignal acquisition board were transferred to a computer via Bluetooth. The data can be observed using OpenSignal software and can be downloaded to text files. The data were then processed using MATLAB software.

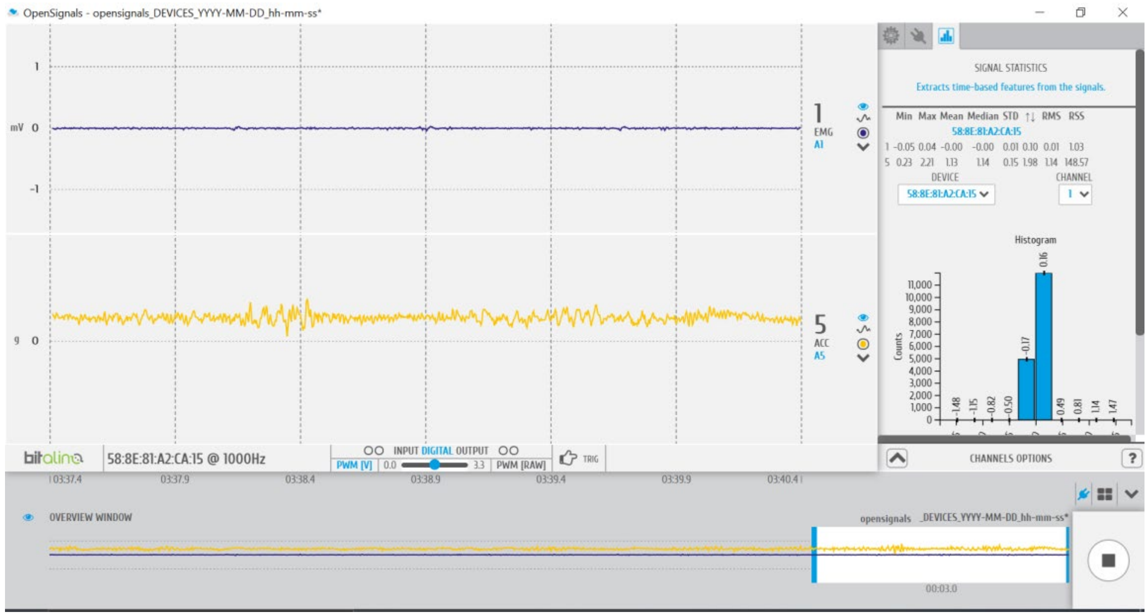
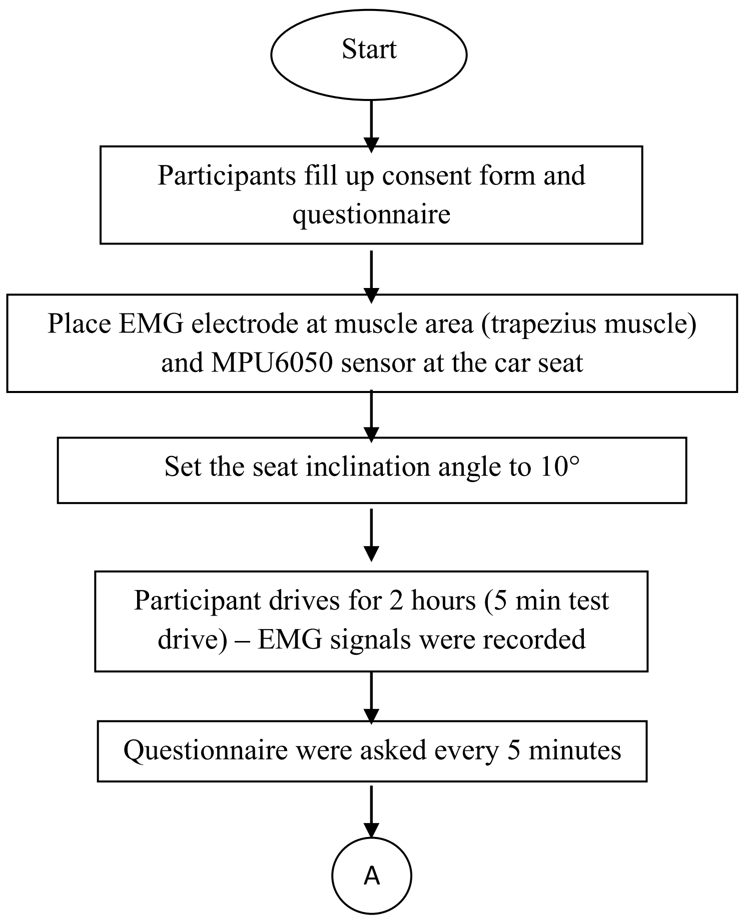


Figure 19 OpenSignal software interface

The subjects needed to drive for 2 hours using the same highway route on the East Coast Expressway Phase 2, Malaysia. The driving duration was chosen for 2 hours based on the pilot study by El Falou et al., (2003) where the subjects reported experiencing muscle pain after two hours in sitting position (El Falou et al. 2003). In addition, Baker et al., (2018) also concluded that discomfort increases after 90 to 120 minutes during prolonged sitting (Baker et al. 2018). The highway route is a monotonous environment with the straight feature but also with some slanted ramps, unexpected downhill and bumpy features. The same car model, Perodua Axia with automatic transmission was used as the test vehicle. All the subjects needed to maintain a driving speed of 90km/h. Before starting the experiment, the subjects were given a 5 min test drive for them to familiarize themselves with the car and the road.

When driving, the researcher verbally asked questions from a questionnaire (Appendix II) every 5 minutes if they feel drowsy, sleepy or experience any muscle pain. Wang et al., (2017) asked a 10 minutes interval questionnaire to the subject during the simulation driving to minimize the fluctuation and difference between subject (Wang et al. 2017). In this work, it is assumed that 5 minutes interval is more accurate to detect muscle fatigue perceived by the subjects. The muscle fatigue perceived by the driver is known as a subjective measure and the time of muscle fatigue occurred was recorded. Five EMG signals before and another five EMG signal after the time of the subjective muscle fatigue were extracted and were considered as non-fatigue and fatigue conditions, respectively. The flowchart of this research is provided in Figure 20.



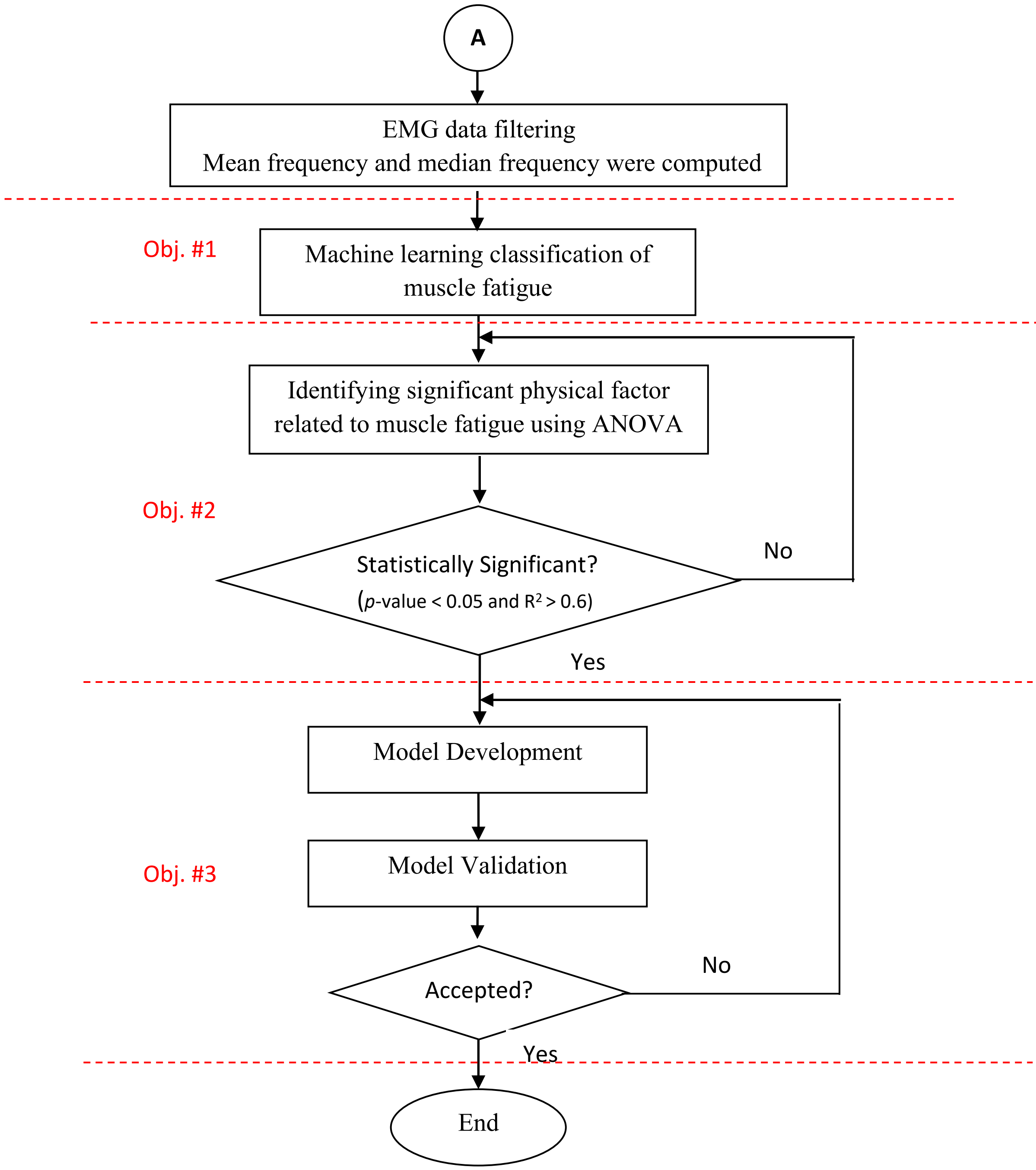


Figure 20 Overall flowchart of the experiment

3.4 DATA PROCESSING

The next step is to process the collected data from the EMG sensor. The data were filtered to remove noise and motion artefacts using a fourth-order Butterworth band pass filter with a range of 20 - 500 Hz to remove noise at the high-end cut-off and motion artefacts at the low-end cut-off. According to Nyquist's theory, the sampling frequency must be twice as high as the maximum signal frequency (Chang et al. 2012). The filtered EMG data of the trapezius muscle were then further processed to calculate MNF and MDF using the sliding window technique. An EMG signal consists of hidden useful information but the signal itself is very complex and consists of noise. Thus, the sliding window technique as suggested by Thongpanja et al. (2013) is the best technique to eliminate noise and inference while extracting important features. Figure 21 shows the concept of the sliding window technique.

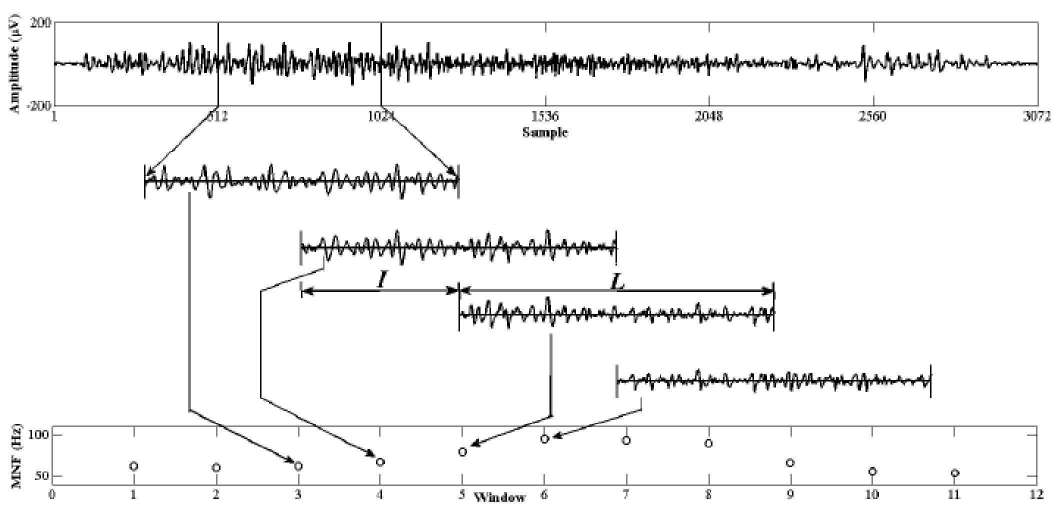


Figure 21 Sliding window concept with (L) is the window size and (I) is the increment of the window by Thongpanja et al. (2013)

A total of n^{th} window segment of MNF and MDF is calculated by:

$$n = \left(\frac{N-L}{I}\right) + 1 \tag{1}$$

where N is the total number of EMG signal recorded throughout the experiment, L is the window size and I is the increment. In this work, a window size of 250 samples and an increments of 125 samples, as suggested by Thongpanja et al. (2013) as the optimum value for window size and increment (Sirinee Thongpanja et al. 2013). For every window, MDF and MNF value were calculated. The MDF is the frequency value at which the EMG spectrum is divided into two region equally (Qassim et al. 2022). The equation for MDF is as follow:

$$MDF = \sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2}\sum_{j=1}^{M} P_j \tag{2}$$

The MNF which is defined as the sum of the product of the EMG power spectrum and frequency, then divided by the total sum of the power spectrum. MNF is also defined as the average frequency value. The equation for MNF is as follow:

$$MNF = \frac{\sum_{j=1}^{M} f_j P_j}{\sum_{j=1}^{M} P_j} \tag{3}$$

where P_j is the EMG power spectrum at frequency bin *j* and M is the length of frequency bin. In this work, the length of frequency bin was calculated using sliding window. Thus, in every window divided by sliding window technique, the MDF and MNF value were calculated. As stated in the literature review, MNF and MDF are normally used as features to identify muscle fatigue. In addition, the slope of linear regression for MNF and

MDF is computed using regression analysis is also used as the index of muscle fatigue (Chang et al. 2012).

3.5 CLASSIFICATION

Based on the time of the subjective muscle fatigue for every subject, the MNF and MDF were extracted before and after the fatigue time. The reason of extracting five signal before and another five data after subjective fatigue is because of great variability between self-report and actual measurement between the subjects and situations based on subject endurance limit and mental state (Marri and Swaminathan 2016; Papakostas et al. 2019). In addition, according to Sahayadhas et al. (2013), subjective measures and physiological measures do not exactly occur at the same time (Sahayadhas, Sundaraj, and Murugappan 2013). The values before the time of subjective muscle fatigue were considered non-fatigue conditions, whereas the values after the time of subjective muscle fatigue were considered fatigue conditions. Hence, for every subject, ten values consisting of five values for non-fatigue data and another five values for fatigue data were collected.

According to Venugopal et al., (2014), the dataset need to be normalized so that the differences in data range between the subjects will be eliminated (Venugopal et al. 2014). Normalizing is done by dividing the EMG signal with the Maximum Voluntary Contraction (MVC) of each subject and multiply with 100 to make the normalized value to percentage. Buchanan et al., (2004) suggested that the MVC value for particular subject is the maximum EMG value recorded during the subject's experimental procedure (Buchanan et al. 2004). In this work, both MDF and MNF were normalized and were used as the muscle fatigue features throughout this research.

The normalized MNF and MDF values were used as the features dataset for classification. The data were divided accordingly, where twenty-five percent were used as the test set, while the remaining data were used as the training set. Six machine learning classifiers (Logistic Regression, Support Vector Machine, Naïve Bayes, k-nearest Neighbours, Decision Tree and Random Forest) were used to classify non-fatigue and fatigue conditions.

Lastly, a ten-fold cross validation method was implemented to evaluate the performance (accuracy) of the classifiers. The acceptable accuracy range for the analysis of all classifiers are between 0.80±0.16 and 0.94±0.02 (Golmohammadishouraki 2022). The confusion matrix is computed to analyze the performance of the classifier (Bhardwaj, Parameswaran, et al. 2018). The confusion matrix is depicted in the Figure 22 below. There are four groups of prediction result which are True Positive (TP), True Negative (TN), False Positive (FP) and False Negatives (FN). From the confusion matrix, the equivalent performance metrics were evaluated (Narudin et al. 2016). The accuracy which is the measure of model correctly predict the output and the equation is as follow:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Prediction} \tag{4}$$

Precision is the proportion of the true positive value with positively predicted as positive by the classifier and the equation is as follow:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{5}$$

The Recall is the proportion of the true positive output that is positively predicted by the classifier and the equation is as follow:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{6}$$

The Specificity which is the proportion of true negative output that is negatively predicted by the classifier and the equation is as follow:

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \tag{7}$$

Lastly the F1 score which is the harmonic mean between precision and recall and good for imbalanced datasets. The equation is as below:

$$F1\ score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall} \tag{8}$$

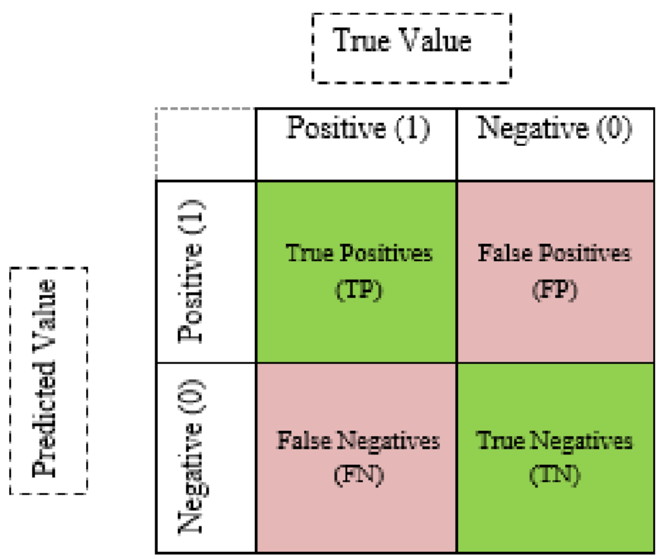


Figure 22 Confusion Matrix

3.6 REGRESSION

For the development of the mathematical model, the MDF and the slope of MDF were used in the regression technique. Based on the MDF value for each subject, the slope coefficient of linear regression was computed. The slope coefficient is also known as the muscle fatigue index and is used to signify the trend of muscle fatigue (De Santana et al. 2014). The data of MDF and slope coefficient of MDF were used as the output of the regression. The inputs of the regression are the subjects' BMI, age and YOD. Each input variable model was analyzed using Analysis of Variance (ANOVA). The ANOVA analysis will calculate the significance F which is defined as the probability of null hypothesis proposed will not be rejected. The value of significance F should be less or equal than 0.05 for the better performance (Wang et al. 2019).

In this work, the null hypothesis (H_0) was set: There is no statistically significant difference between each physical factors (BMI, age and YOD) while for alternative hypothesis (H_a) as follow: There is statistically significant difference between each physical factors. The statistical significance's threshold (α) was set to 5% which indicate the null hypothesis (H_0) is rejected when *p*-value is less and equal to 0.05 and thus alternative hypothesis (H_a) is accepted. On the contrary, when p-value is bigger than 0.05, the null hypothesis (H_0) is accepted (Davidović, Pešić, and Antić 2018). The model of input variables with a *p*-value less than 0.05 indicates that the model is significant to the output and is acceptable.

Another test that is normally used for evaluating a model is the coefficient of determination (R squared, R^2). The range of R^2 is between 0 to 1 where value of R^2=1 shows better prediction (Das C, Lucia MS et al. 2017). The R^2 value must be higher than or equal to 0.6 for the model to be accepted (De Santana et al. 2014). From the result of ANOVA and R^2, significant input variable was selected as the mathematical model of the muscle fatigue of the driver. Figure 23 shows the block diagram of the mathematical model developed in this research. A mathematical model is a model used to represent the physical behavior of an actual system.

Figure 23 Block diagram of the mathematical model

3.7 VALIDATION OF MATHEMATICAL MODEL

Model Validation is an essential step in accepting a model. The definition of validation is a process of confirming that the model is an accurate representation of real world by comparing the prediction from the model with the value from the real world (Mayer and Butler 1993). In this research, validation was carried out by calculating the residual error. The residual error was calculated as the difference in the predicted value derived from the developed model and the actual data (Ani et al. 2017). The formula of the percentage of residual error is shown in Equation 1.

$$\%\ Residual\ Error = \frac{Predicted\ value - \ Actual\ value}{Predicted\ value} x\ 100 \qquad (9)$$

The residual error computed should be less than 10% for the model to be considered suitable for its intended use.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 DATA PRE-PROCESSING

The EMG signals recorded from the trapezius muscle of the driver were transferred to MATLAB software for pre-processing. The original EMG signal was filtered with the fourth-order band pass filter with a range of 20-500 Hz to remove noise and motion artefacts. The raw EMG signal is shown in Figure 24 and the filtered EMG signal is shown in Figure 25.

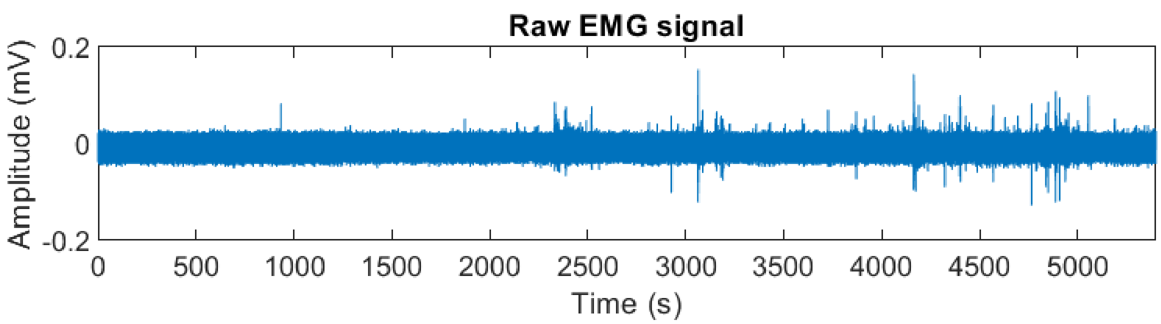


Figure 24 Raw EMG signal for a representative subject.

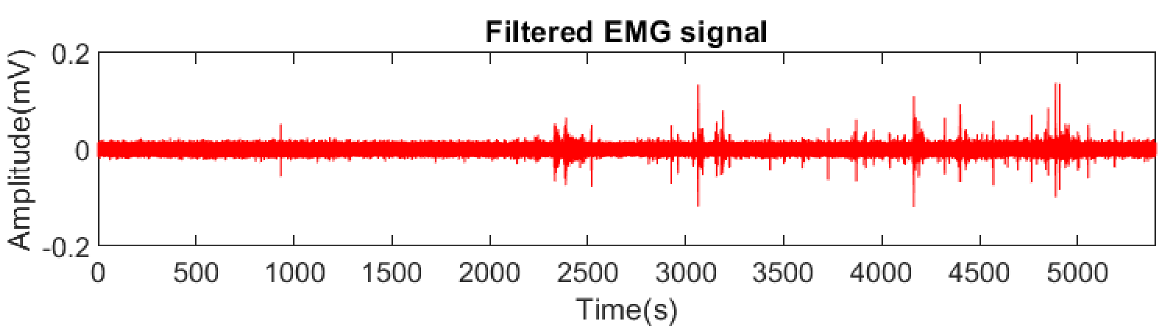


Figure 25 Filtered EMG signal for a representative subject.

Figure 24 and 25 above show the EMG signal for the representative subject. According to the subjective muscle fatigue which had been recorded by verbally questioning the subject during experiment, the time that the subject felt fatigue was in the 45th minutes of driving which is after 2,700 second. From the graph, it can be seen that after 2,700 second, the amplitude of EMG starting to increase and fluctuate. This is in line with the literature that stated that during fatigue, the amplitude of EMG signal will increase (Baker et al. 2018) (Toro et al. 2019).

Next, the Median Frequency (MDF) (Figure 26) and Mean Frequency (MNF) (Figure 27) were computed using the sliding window technique as explained in Chapter 3. From these graphs, after 21,599 window which is equivalent to 2,700 second, the MDF and MNF starting to fluctuate and decreasing. The decreasing of MDF and MNF indicate that the subject experienced fatigue due to the reduction in the propagation velocity of the muscle's action potential (Karthick et al. 2018).

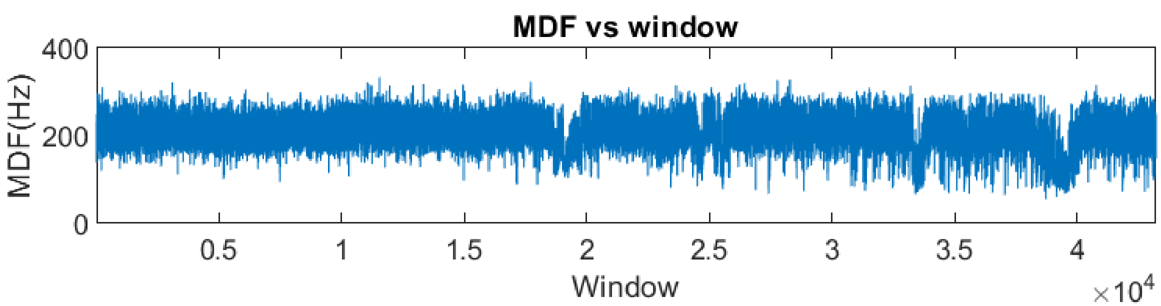


Figure 26 Graph of MDF for representative subject**.**

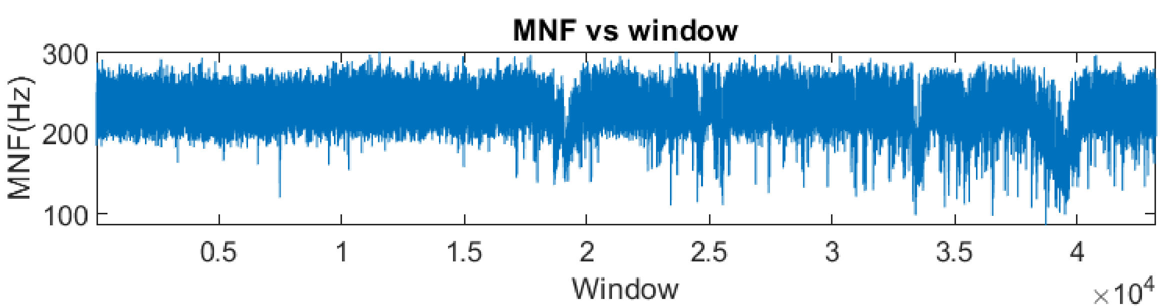


Figure 27 Graph of MNF for representative subject.

Based on the subjective fatigue and fatigue measured by EMG signal shown in the graphs, it can be concluded that both occurred around the same time. However, the exact time of fatigue occurrence of subjective do not fully coincide with physiological measure (EMG signal)(Sahayadhas et al. 2013). Thus, in this work, the subjective fatigue will be the baseline time of subject's fatigue and five values of MDF and MNF value were extracted before and after subjective fatigue for every subject. The MDF and MNF values before the time of subjective fatigue were considered non-fatigue conditions, whereas the values after the time of subjective fatigue were considered as fatigue condition. A total of 50 non-fatigue and fatigue datasets obtained from the 10 subjects and later will be used as the dataset in machine learning classification. All the data were normalized to eliminate differences in the subjects' EMG signal value range (Venugopal et al. 2014).

4.2 CLASSIFICATION OF MUSCLE FATIGUE

The normalized MDF and MNF computed were plotted in Figures 28 and 29 below. Based on the graphs, the normalized MDF and MNF values were higher in non-fatigue conditions which in agreement with literature discussed in Chapter 2 (Sonmezocak and Kurt 2021). The results indicate that the MDF and MNF could be used for the study of muscle fatigue.

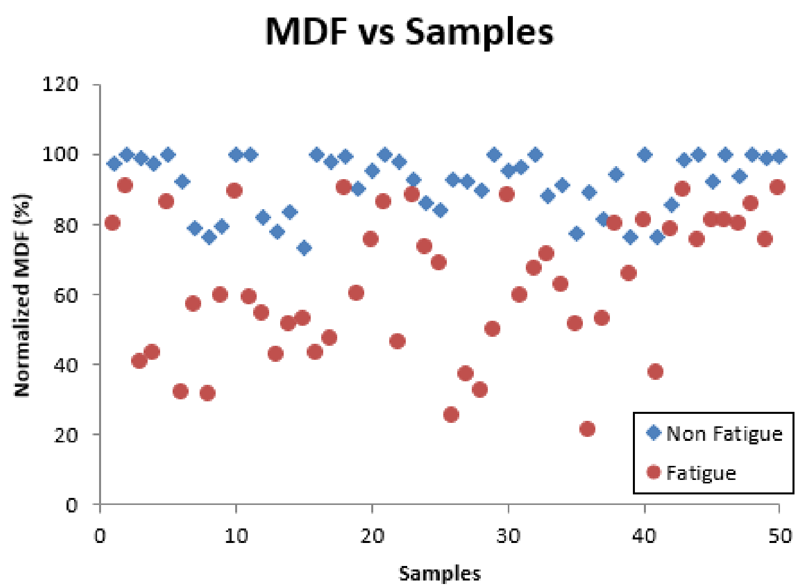


Figure 28 MDF for all subjects during non-fatigue and fatigue conditions

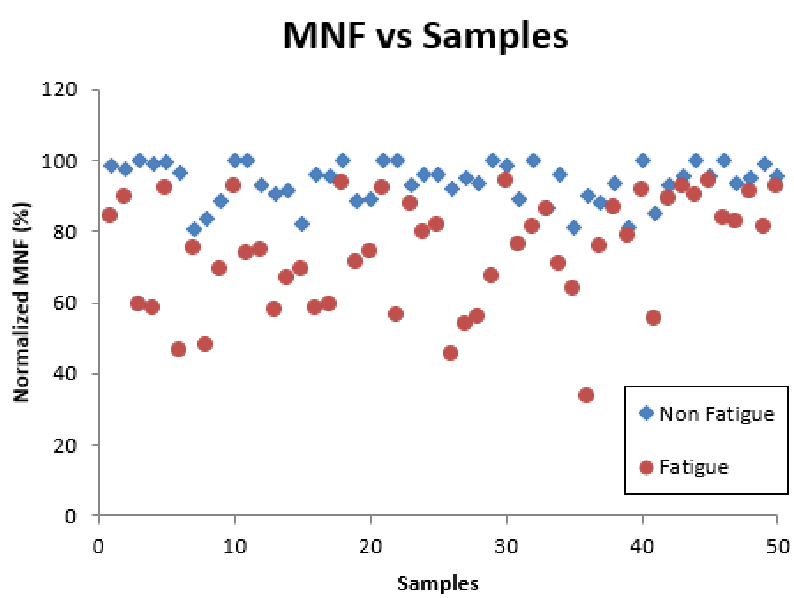


Figure 29 MNF for all subjects during non-fatigue and fatigue conditions

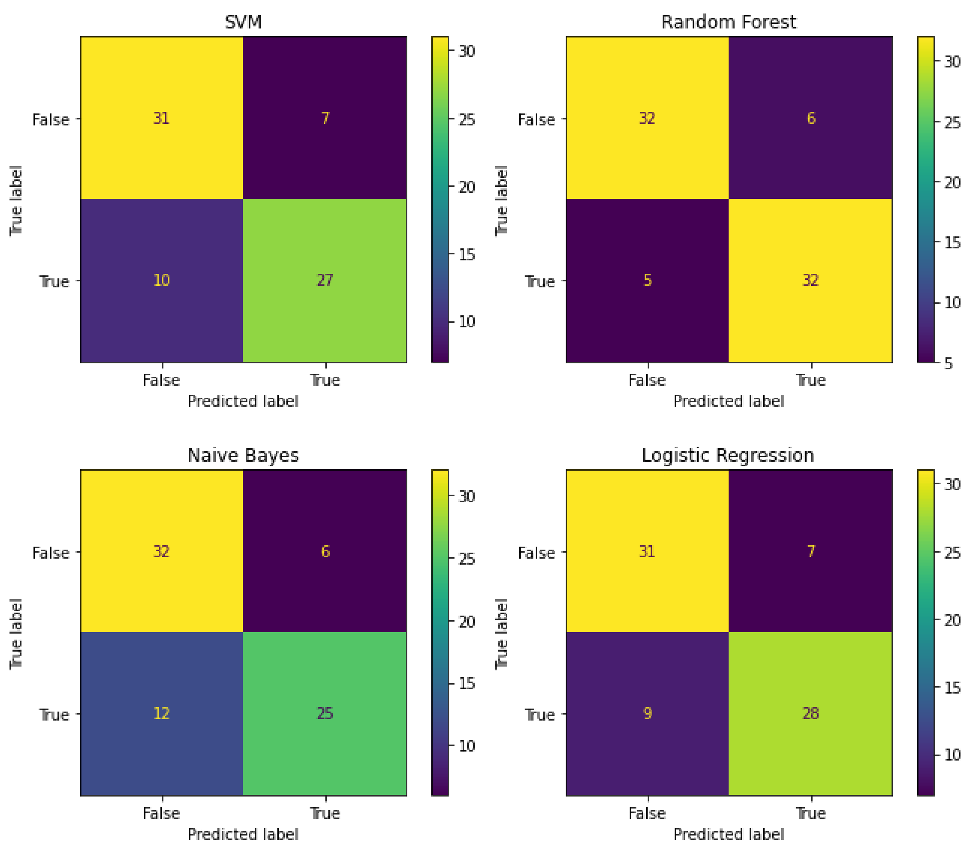
Next, the normalized MDF and MNF datasets were used as the features of the Machine Learning model. In this work, six Machine Learning models, namely Logistic Regression, Support Vector Machine, Naïve Bayes, k-nearest Neighbors, Decision Tree and Random Forest were used for classification.

Lastly, the performance of the classification of non-fatigue and fatigue conditions was evaluated using ten-fold cross-validation. The results of 10-fold Cross Validation are summarized in Table 4. The best validation accuracy for the normalized MDF dataset was obtained using the Random Forest classifier with 81.96%. On the other hand, when only using the normalized MNF dataset, the best accuracy was obtained by the Logistic Regression classifier with 77.68%. Lastly, when both the normalized MDF and MNF were used as the features in the Machine Learning model, the Random Forest classifier was the most accurate classifier, improving the accuracy to 85.00%.

Table 4 Cross-validation accuracy results for MDF and MNF

Classifier	MDF Accuracy (%)	MNF Accuracy (%)	MDF and MNF Accuracy (%)
SVM	80.36	76.25	77.86
Random Forest	81.96	73.57	85.00
Naïve Bayes	79.29	76.25	76.43
Logistic Regression	79.11	77.68	79.11
k-Nearest Neighbors	79.46	73.21	81.25
Decision Tree	79.46	74.64	83.75

From the classification accuracy results, it can be concluded that using more features yields more accurate classification in most of the classifier. Combining more than one features will improve performance accuracy in muscle fatigue classification (Pratummas and Khemapatpapan 2021) (Yousif et al. 2019). For example, the accuracy of the Random Forest classifier when using MDF dataset was only 81.96%, while using the MNF dataset only produced the rate of 73.57%. However, when both the MDF and MNF were used as the dataset of the model, the accuracy improved to 85.00%. In this situation, the accuracy also improved for the Logistic Regression, k-nearest Neighbor and Decision Tree classifiers. Figure 30 below shows the confusion matrix for all classifier.



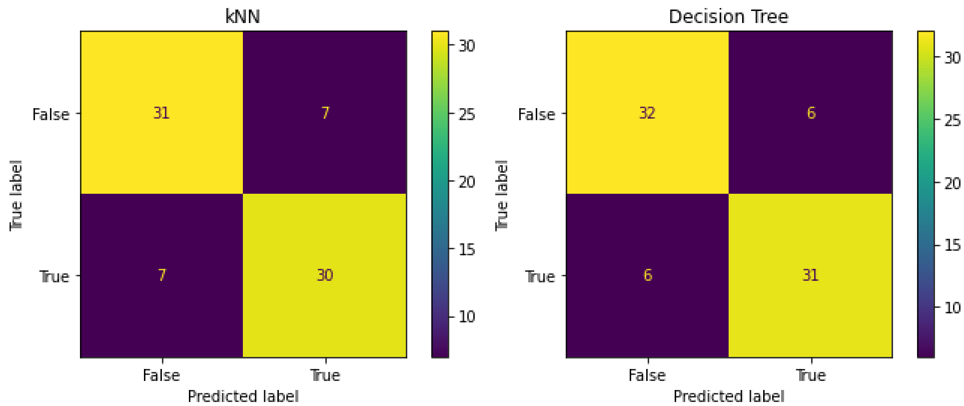


Figure 30 Confusion Matrix for all classifier

Based on the result from the confusion matrix, the performance matrices were computed and summarized in the Table 5 below. For the accuracy matrix, Random Forest produced highest true fatigued and true non-fatigued prediction over entire prediction as compared to another classifier with 0.85. The Precision which the proportion of correctly predicted fatigued over entire positive prediction is highest for Random Forest and Decision Tree. As for Recall, the highest value was produced by Random Forest with 0.86. The recall represents the proportion of the subject was correctly predicted fatigue by the classifier. Specificity depicted how good the classifier classifies non fatigue condition. Random Forest, Naïve Bayes and Decision Tree produced the highest specificity among another classifier. Lastly, Random Forest performed the best for F1 score where the precision and recall is considered regarding the true positive prediction.

Table 5 Summary of Confusion Matrix Measure

Classifier	Accuracy	Precision	Recall	Specificity	F1 score
SVM	0.77	0.79	0.73	0.82	0.76
Random Forest	0.85	0.84	0.86	0.84	0.85
Naïve Bayes	0.76	0.81	0.68	0.84	0.74
Logistic Regression	0.79	0.80	0.76	0.82	0.76
k-Nearest Neighbors	0.81	0.81	0.81	0.81	0.81
Decision Tree	0.84	0.84	0.84	0.84	0.84

In a nutshell, the most accurate form of muscle fatigue classification for prolonged driving was obtained by the Random Forest classifier using both the normalized MDF and MNF values of EMG signals. Random forest classifier is a very well-known and powerful classifier because of the stability and robustness of the data, which features only slight variation. This classifier is constructed using multiple distinct decision trees and the final decision is predicted by most of the trees. Each decision tree is trained with different subsets of the training data using random sample from the original training set (Karthick et al. 2018).

The second highest accuracy in this work was obtained using the Decision Tree classifier, which yielded 83.75%. This classifier performs well with an enormous volume of information, while unrelated features do not influence its results. However, the drawback is over-fitting as it is sensitive to information(Pratummas and Khemapatpapan 2021). This is because when the result will extremely change to huge degree when the information changes.

Lastly, the *k*-Nearest Neighbours classifier produced an accuracy rate of 81.25%, making it the third-best classifier. With training data, the *k*-Nearest Neighbours algorithm sets a group of *k* objects closest to the test object. It then assigns a class to the test object based on the neighbours. The three main stages of the *k*-Nearest Neighbours algorithm are initializing dataset and *k*-Nearest Neighbours, computing the distance between neighbors, and classifying the test data based on the majority of the neighbouring class data (Venugopal et al. 2014). The value of *k* was iterated and set as five in this study based on the highest classification accuracy obtained when tested with MDF classification. The result is shown in Table 6 below where *k*=5 produces the highest classification accuracy. The usage of *k*=5 also been used in the research by Marri et al. (2016) in classifying muscle fatigue (Marri and Swaminathan 2016).

Table 6 Selection of *k* value for *k*NN classifier

	***k*-Nearest Neighbours classifier**				
	k=3	*k*=4	*k*=5	*k*=6	*k*=7
Classifier Accuracy (%)	79.29	75.54	79.46	74.46	73.04

For the selection of *k*-value for k-fold cross validation, the same method applied as the selection of *k* value for *k*NN classifier. The accuracy result for ten-fold cross validation is highest as compared to other value. Table 7 below summarized the result of different *k*-fold cross validation tested. The ten-fold cross validation was also used by previous researcher to classify muscle fatigue (Zhao et al. 2022) (Pratummas and Khemapatpapan 2021)(Zhang et al. 2020).

Table 7 Selection of *k* value for *k*-fold cross validation

	***k*-fold cross validation value**			
	k=3	*k*=5	*k*=10	*k*=15
Classifier Accuracy (%)	73.33	77.33	79.46	78.67

4.3 REGRESSION OF MUSCLE FATIGUE

Based on the literature discussed in Chapter 2, muscle fatigue can be evaluated using MDF and slope coefficient from linear regression of MDF (S. Thongpanja et al. 2013)(Ostojić et al. 2018). Thus, by using the same normalized MDF from the experiment conducted before, further analysis was done to study the effect of physical factors of the driver (body mass index (BMI), age and years of driving (YOD)) on the MDF and slope coefficient. As mentioned before, the slope coefficient of MDF represents the rate of muscle fatigue occurrence for the driver.

The slope coefficient of MDF for all subjects was obtained using linear regression. Figure 31 shows the result of the linear regression of MDF for subject number one. Based on the result from the graph, it was confirmed that muscle fatigue happened because the slope coefficient value is negative which represents the decrease of power in the muscle.

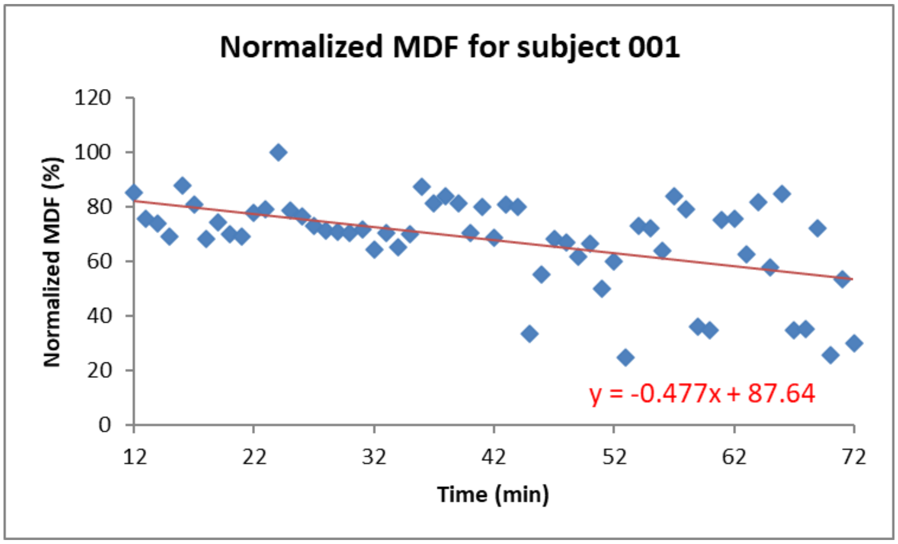
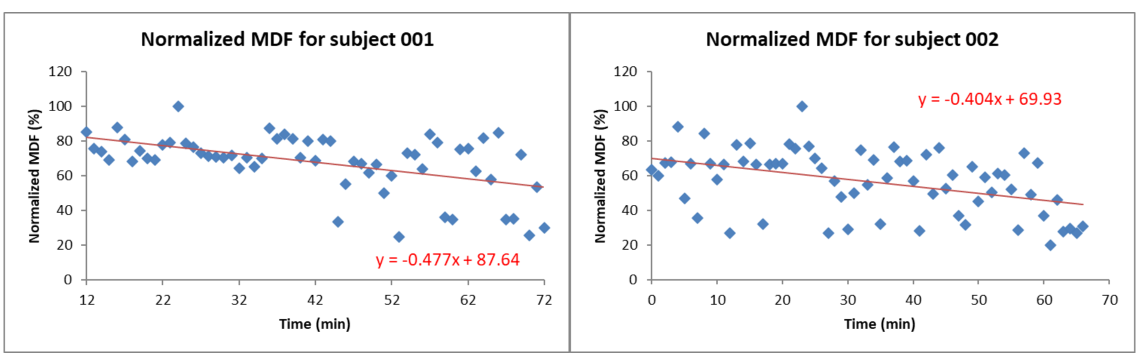


Figure 31 Normalized MDF linear regression for subject number one (female)

The result of the slope coefficient for all subjects is shown in Figure 32 and summarized in Table 8 below. Based on the result, the slope coefficient for all subjects was negative, indicating that all subjects experienced muscle fatigue during prolonged driving.



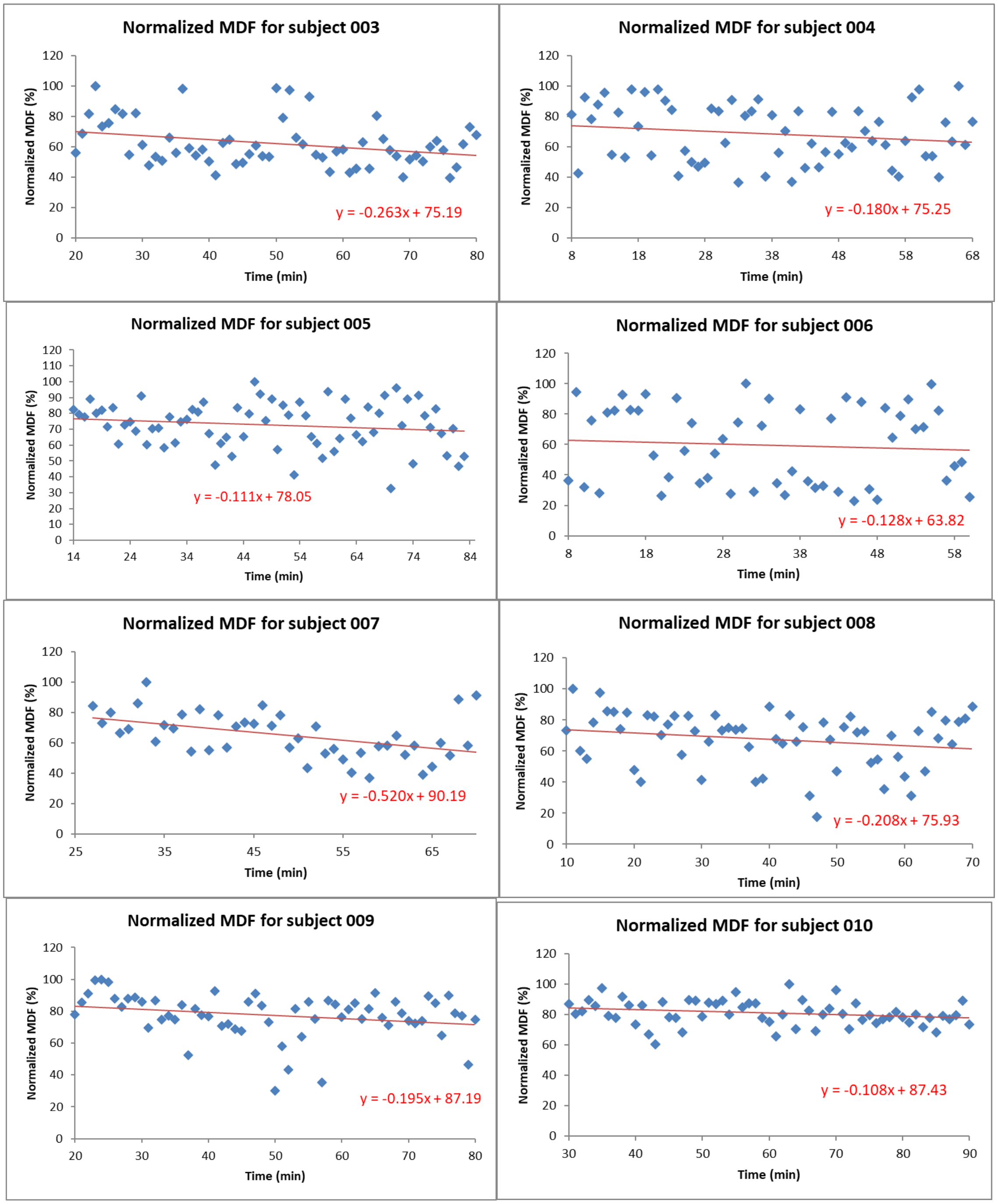


Figure 32 Regression result for all subjects

Table 8 Summary of slope coefficient results for all subjects

Subject No	Gender	Age	BMI	YOD	Slope Coefficient
001	Female	27	18.61	9	-0.477
002	Female	30	19.84	12	-0.404
003	Male	23	20.51	6	-0.263
004	Male	35	25.40	18	-0.180
005	Female	35	24.84	12	-0.111
006	Female	32	30.12	7	-0.128
007	Female	22	19.14	4	-0.520
008	Male	30	22.49	12	-0.208
009	Male	38	29.41	20	-0.195
010	Male	27	23.31	12	-0.108

Next, the slope coefficient for every subject was analyzed, and the result shows the difference in slope coefficients between genders. Figure 26 shows the slope coefficient result from subject number one which is a female subject. Meanwhile, Figure 33 shows the linear regression result for a male subject. The slope coefficient for the male subject is less negative compared to the female subject. According to Chang et al. (2012), the more negative value of the slope coefficient shows that the person has high muscle fatigue conditions (Chang et al. 2012). Thus, the result concludes that the female subject tends to experience faster muscle fatigue as compared to the male subject.

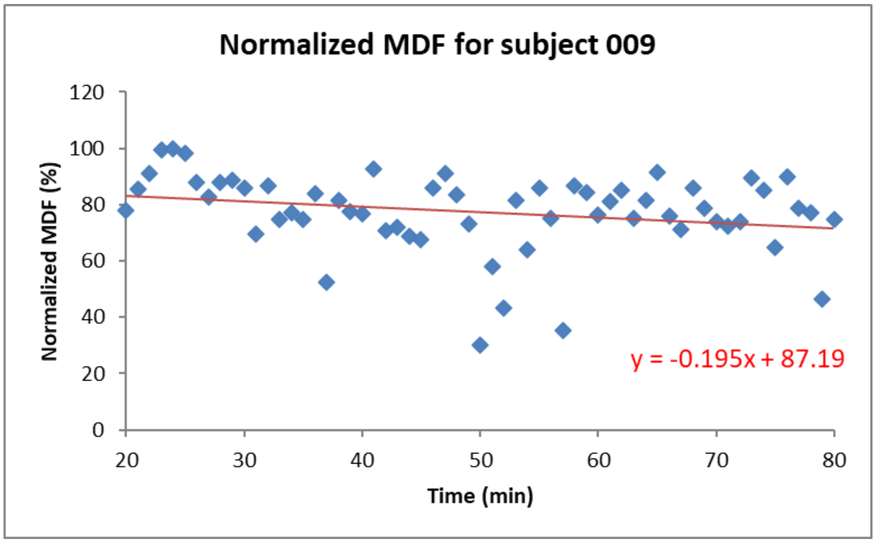


Figure 33 Normalized MDF linear regression for subject number nine (male)

Figure 34 summarizes the result of the slope coefficient between genders. Most of the male subjects produce a lower negative value of slope coefficient as compared to female subjects which in line with the finding from previous research (Chang et al. 2012)(Carneiro et al. 2010). From the graph, it is concluded that the majority of female subjects produce a lower negative value of slope coefficient, which indicates that female subjects experienced muscle fatigue faster. Although many studies have been done in comparing muscle fatigue between gender, other factors such as BMI, age and YOD should not be neglected.

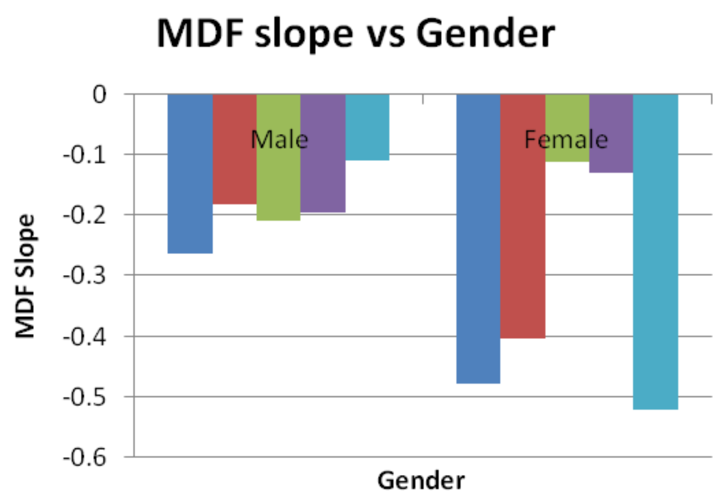


Figure 34 Difference of slope coefficients between genders

The second objective of this research is to identify the physical factors of the driver related to muscle fatigue. The physical factors selected are BMI, age, and YOD. Before the experiment was conducted, the subjects needed to provide these three parameters through the questionnaire. By using Analysis of Variance (ANOVA), the significant physical factor was analyzed and determined by polynomial regression. The model with a *p*-value less than 0.05 and a coefficient of determination (R^2) greater than 0.6 will be selected as the significant model for the driver during prolonged driving. For the regression analysis, eight subject's data will be used for ANOVA analysis while another two subject's data will be used for the validation step randomly. The following subsection will discuss the results of regression analysis using MDF and the slope coefficient of MDF.

4.3.1 REGRESSION ANALYSIS OF MDF

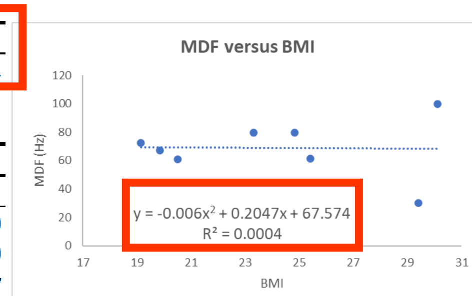
From the normalized MDF computed in the classification step before, the same values were used in this step which is regression analysis to develop a mathematical model. The normalized MDF values were considered as the output of the regression while BMI, age and YOD served as the input parameters. Each input parameter was analyzed using ANOVA using second-order polynomial and third-order polynomial regression.

As mention before, in order for the physical factor to be accepted and suitable for intended use, the ANOVA analysis result for *p*-value (Significance F) should be less or equal than 0.05 and the coefficient of determination R^2 should be higher than 0.6. Figure 35 shows the result of ANOVA and R^2 value of second-order polynomial regression for every physical factor using MDF. Meanwhile, Figure 36 shows the result of ANOVA and R^2 value of third-order polynomial for every physical factor using MDF.

ANOVA analysis for MDF vs BMI

	df	*SS*	*MS*	*Significance F*
Regression	2	1.0672	0.5336	0.9991
Residual	5	2843.3590	568.6718	
Total	7	2844.4262		

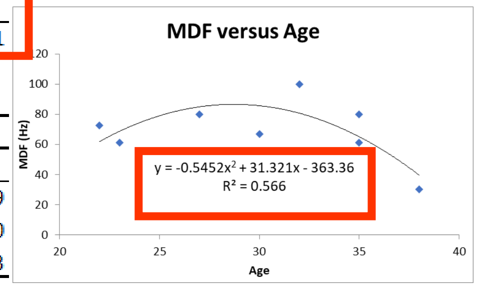
	Coefficients	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	67.5744	0.8760	-990.2202	1125.3690
BMI	0.2047	0.9954	-87.4587	87.8680
BMI^2	-0.0060	0.9934	-1.7828	1.7707



ANOVA analysis for MDF vs AGE

	df	SS	MS	Significance F
Regression	2	1610.0002	805.0001	0.1241
Residual	5	1234.4260	246.8852	
Total	7	2844.4262		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-363.3559	0.1269	-873.9656	147.2539
AGE	31.3207	0.0717	-4.0166	66.6580
AGE^2	-0.5452	0.0650	-1.1396	0.0493



ANOVA analysis for MDF vs YOD

	df	SS	MS	Significance F
Regression	2	1889.1478	944.5739	0.0654
Residual	5	955.2784	191.0557	
Total	7	2844.4262		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	48.1222	0.1171	-17.2793	113.5237
YOD	7.2856	0.1864	-4.9490	19.5201
YOD^2	-0.3950	0.0972	-0.8936	0.1035

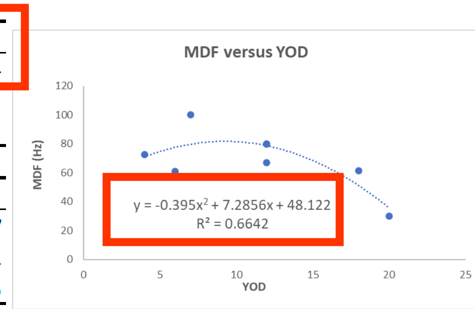
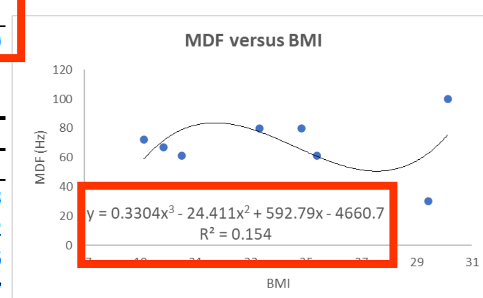


Figure 35 Second-order polynomial analysis for MDF

ANOVA analysis for MDF vs BMI

	df	SS	MS	Significance F
Regression	3	438.0191	146.0064	0.8629
Residual	4	2406.4071	601.6018	
Total	7	2844.4262		

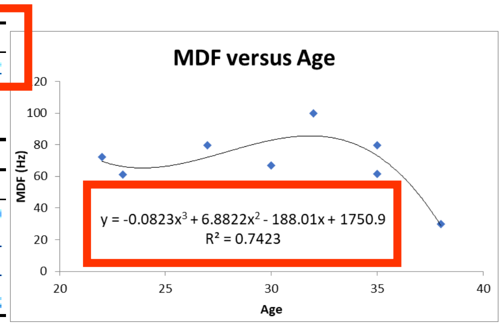
	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-4660.6662	0.4494	-20109.1722	10787.8398
BMI	592.7936	0.4425	-1340.2080	2525.7952
BMI^2	-24.4115	0.4421	-103.9446	55.1216
BMI^3	0.3304	0.4421	-0.7459	1.4067



ANOVA analysis for MDF vs AGE

	df	SS	MS	Significance F
Regression	3	2111.5024	703.8341	0.1132
Residual	4	732.9238	183.2309	
Total	7	2844.4262		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	1750.9155	0.2460	-1828.9909	5330.8219
AGE	-188.0083	0.2307	-557.5589	181.5424
AGE^2	6.8822	0.2004	-5.5949	19.3594
AGE ^3	-0.0823	0.1734	-0.2204	0.0558



ANOVA analysis for MDF vs YOD

	df	*SS*	*MS*	*Significance F*
Regression	3	1894.1891	631.3964	0.1843
Residual	4	950.2371	237.5593	
Total	7	2844.4262		

	Coefficients	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	59.1090	0.5039	-164.6161	282.8341
YOD	3.5747	0.8974	-68.6703	75.8197
YOD^2	-0.0433	0.9866	-6.7733	6.6866
YOD^3	-0.0098	0.8912	-0.1964	0.1768

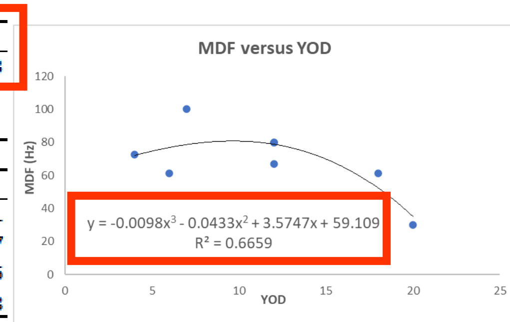


Figure 36 Third-order polynomial analysis for MDF

Table 9 summarizes the ANOVA result of MDF. From the results, it can be concluded that none of the physical factors have a statistically significant relation with MDF. This is because the *p*-value for all input parameters exceeds 0.05 and indicating that no sufficient evidence at the 95% confidence level that a significant linear relationship exist between dependent variable (MDF) and independent variable (BMI, age and YOD). Even though the value of R^2 for the input parameters age (3rd order polynomial) and YOD (2nd order polynomial and 3rd order polynomial) are greater than 0.6, the models were not accepted as the *p*-value is greater than 0.05.

Table 9 ANOVA of MDF and the physical factors for muscle fatigue.

	2nd Order Polynomial			**3rd Order Polynomial**		
Physical Factor	**R^2**	***p*-value**	**MSE**	**R^2**	***p*-value**	**MSE**
BMI	0.0004	0.9991	568.67	0.1540	0.8629	601.60
Age	0.5661	0.1241	246.89	0.7423	0.1132	183.23
YOD	0.6642	0.0653	191.06	0.6659	0.1843	237.56

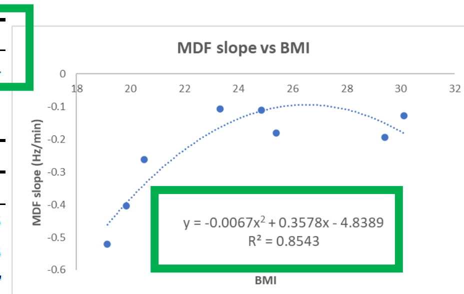
4.3.2 REGRESSION ANALYSIS OF THE SLOPE COEFFICIENT OF MDF

As mentioned earlier, the slope coefficient represents important features in muscle fatigue assessment. The slope coefficient indicates how fast muscle fatigue will happen and it can represent a person's endurance to muscle fatigue (Ostojić et al. 2018). The input parameters to analyze were BMI, age and YOD of the driver. The slope coefficients of MDF for all subjects were used as the output response of the polynomial regression of second-order and third-order polynomial. Each physical factors were analyzed using ANOVA. Figure 37 shows the result of ANOVA and R^2 value of second-order polynomial regression for every physical factor using MDF slope coefficient. Meanwhile, Figure 38 shows the result of ANOVA and R^2 value of third-order polynomial for every physical factor using MDF slope coefficient.

ANOVA analysis for MDF vs BMI

	df	*SS*	*MS*	*Significance F*
Regression	2	0.1350	0.0675	0.0081
Residual	5	0.0230	0.0046	
Total	7	0.1580		

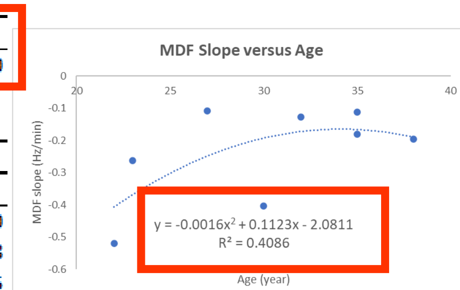
	Coefficients	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-4.8389	0.0091	-7.8495	-1.8283
BMI	0.3578	0.0142	0.1083	0.6073
BMI^2	-0.0067	0.0187	-0.0118	-0.0017



ANOVA analysis for MDF vs AGE

	df	*SS*	*MS*	*Significance F*
Regression	2	0.0646	0.0323	0.2690
Residual	5	0.0935	0.0187	
Total	7	0.1580		

	Coefficients	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-2.0811	0.2825	-6.5242	2.3620
AGE	0.1123	0.3908	-0.1951	0.4198
AGE^2	-0.0016	0.4505	-0.0068	0.0035



ANOVA analysis for MDF vs YOD

	df	SS	MS	Significance F
Regression	2	0.0548	0.0274	0.3453
Residual	5	0.1033	0.0207	
Total	7	0.1580		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-0.6309	0.0628	-1.3109	0.0492
YOD	0.0655	0.2431	-0.0617	0.1927
YOD^2	-0.0022	0.3165	-0.0074	0.0029

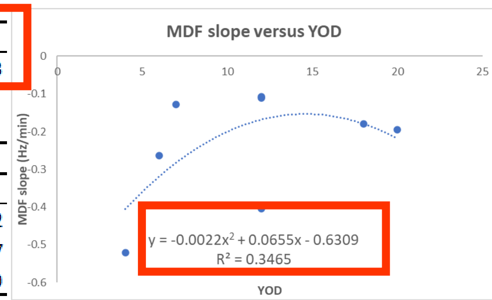
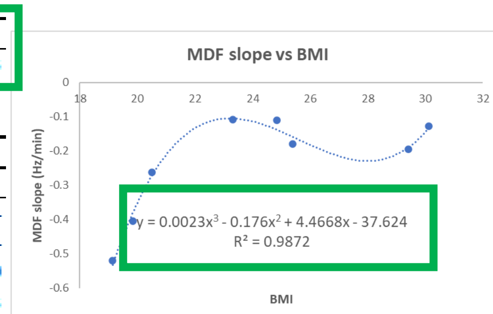


Figure 37 Second-order polynomial analysis for MDF slope coefficient

ANOVA analysis for MDF vs BMI

	df	SS	MS	Significance F
Regression	3	0.1560	0.0520	0.0003
Residual	4	0.0020	0.0005	
Total	7	0.1580		

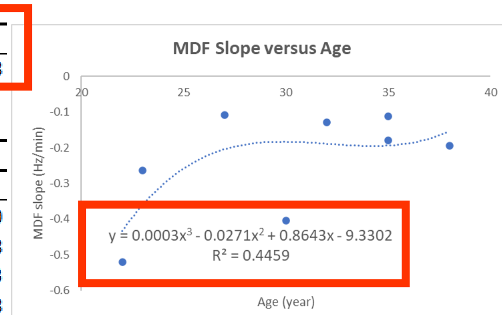
	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-37.6242	0.0018	-51.7911	-23.4574
BMI	4.4668	0.0022	2.6941	6.2394
BMI^2	-0.1760	0.0026	-0.2489	-0.1030
BMI^3	0.0023	0.0030	0.0013	0.0033



ANOVA analysis for MDF vs AGE

	df	SS	MS	Significance F
Regression	3	0.0705	0.0235	0.4548
Residual	4	0.0876	0.0219	
Total	7	0.1580		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-9.3302	0.5442	-48.4613	29.8010
AGE	0.8643	0.5844	-3.1751	4.9038
AGE^2	-0.0271	0.6104	-0.1635	0.1093
AGE^3	0.0003	0.6312	-0.0012	0.0018



ANOVA analysis for MDF vs YOD

	df	SS	MS	Significance F
Regression	3	0.0846	0.0282	0.3353
Residual	4	0.0734	0.0184	
Total	7	0.1580		

	Coefficients	P-value	Lower 95%	Upper 95%
Intercept	-1.4764	0.1055	-3.4431	0.4904
YOD	0.3510	0.1997	-0.2840	0.9861
YOD^2	-0.0293	0.2410	-0.0885	0.0299
YOD^3	0.0008	0.2713	-0.0009	0.0024

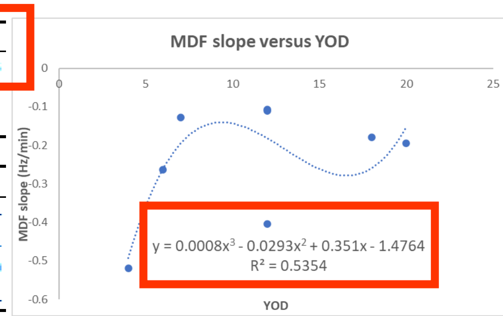


Figure 38 Third-order polynomial analysis for MDF slope coefficient

Table 10 summarizes the ANOVA result using MDF slope coefficients. From the results, age and YOD were found not statistically related to the slope coefficient of MDF because the *p*-value is greater than 0.05. On the other hand, the *p*-value for BMI is less than 0.05 and the R^2 value is greater than 0.6 for both second-order and third-order polynomial. This concludes that there is a relationship between BMI and the rate of muscle fatigue using MDF.

For the second-order polynomial regression graph for the BMI versus MDF slope, the R^2 value is 0.85 which indicates that 85% of dependent variable can be explained by the independent variable. Meanwhile, the significance *F*, which is also known as the *p*-value is less than 0.05 which indicate that 95% confidence that there is a significant linear relationship between independent variable (BMI) and dependent variable (MDF slope). The *p*-value for individual input variables (BMI and BMI^2) is also less than 0.05 which explains that the quadratic coefficient is significant. For the third-order polynomial regression of BMI versus MDF slope, the R^2 value and *p*-value also show a good fit of data (*p*-value = 0.001 and R^2=0.92). Based on the ANOVA result for individual input variables (BMI, BMI^2, BMI^3), it is shows that the third-order coefficient is significant because the *p*-value is less than 0.05.

Table 10 MDF slope ANOVA analysis of the physical factors for muscle fatigue

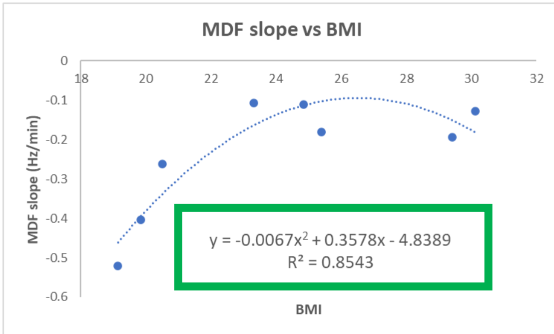
	2^{nd} Order Polynomial			**3^{rd} Order Polynomial**		
Physical Factor	**R^2**	***p*-value**	**MSE**	**R^2**	***p*-value**	**MSE**
BMI	0.85	0.0081	0.0046	0.92	0.0011	0.0029
Age	0.41	0.2690	0.0187	0.45	0.4548	0.0218
YOD	0.35	0.3453	0.0207	0.54	0.3353	0.0184

In conclusion, only BMI is found to be a statistically significant physical factor that affect muscle fatigue of the driver during prolonged driving. Therefore, the mathematical model of the second-order and third-order polynomial were successfully develop using ANOVA analysis where both the *p*-value was less than 0.05 and the R^2 value is higher than 0.6. In addition, based on the regression result, it is shown that the higher the BMI value, the faster muscle fatigue occurrence to the driver during prolonged driving.

4.4 VALIDATION OF THE MATHEMATICAL MODEL

The last step in developing a mathematical model is the validation of the model. Model validation refers to the process of confirming that the model is accurate representation of the real world from the perspective of its intended use. This is done by comparing other actual data from the experiment conducted with the data predicted by the model developed (Ani et al. 2017). The model is considered validated and suitable for its intended use when the residual error is less than 10%. This step is important to examine the model's accuracy in the real world.

Based on the ANOVA result, both second order and third-order polynomial models of BMI were statistically significant and the mathematical model was successfully developed as shown in Figure 39 below.



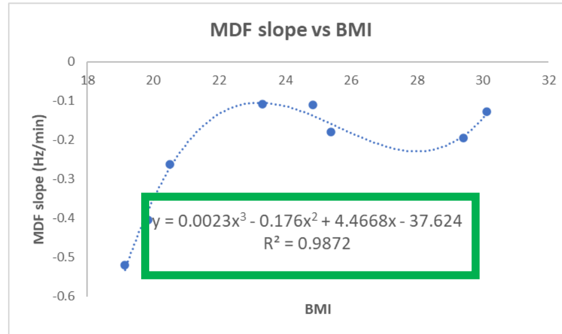


Figure 39 The mathematical models developed based on ANOVA analysis

Next, validation test needs to be done in order to allow the model to be accepted. As mention before, eight subject's data were chosen randomly for mathematical model development step while another two subject's data were used as the validation step. The residual errors were calculated based on Equation 6. Table 11 and Table 12 show the results of model validation and the residual error for second-order and third-order polynomial respectively.

Table 11 Validation result of the mathematical model developed using second-order polynomial

BMI	Prediction	Actual	Residual	Error (%)
18.61	-0.516	-0.477	0.039	7.55
22.49	-0.204	-0.208	0.004	1.96

Table 12 Validation result of the mathematical model developed using third-order polynomial

BMI	Prediction	Actual	Residual	Error (%)
18.61	-0.671	-0.477	0.194	28.91
22.49	-0.107	-0.208	0.101	94.39

Based on the results, the residual error for the second-order polynomial model is less than 10% and it can be concluded that the model developed was successfully validated whereby the model is suitable for its intended purpose. However, the third-order polynomial model is not successfully validated because the residual error is greater than 10%. In a conclusion, for the physical factors of the driver, only BMI is related to muscle fatigue. Furthermore, the mathematical model of the driver during prolonged driving can be developed using second-order polynomial regression and the equation is as follows:

$$Rate\ of\ Muscle\ Fatigue = -4.8390 + 0.3578BMI - 0.00675BMI^2 \qquad (10)$$

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

Due to the increasing use of road transportation and accident rates in Malaysia, this study focused on the classification of non-fatigue and fatigue conditions and predicting the rate of muscle fatigue of the drivers during prolonged driving using Electromyography (EMG) signal. Three main steps were taken to achieve the goals. The first goal is to classify non-fatigue and fatigue condition of the driver during prolonged driving. Secondly, to identify significant physical factor (Body Mass Index (BMI), age and year of driving (YOD)). Lastly is to develop and validate the mathematical model with respect to the driver's physical factor.

The EMG signals from the trapezius muscle were recorded and the mean frequency (MNF) and median frequency (MDF) were computed. For muscle fatigue classification, the MNF and MDF dataset were trained and tested using six machine learning models: Logistic Regression, Support Vector Machine, Naïve Bayes, k-nearest Neighbors, Decision Tree and Random Forest classifiers. The result shows that both MNF and MDF values were lower in fatigue conditions compared to non-fatigue conditions. In addition, by combining MNF and MDF data to the classifier dataset, the accuracy was improved as compared to a single dataset. Yousif et al. (2019) also stated that combining features is advisable in order to obtain more information and avoid losing information (Yousif et al. 2019). Pratummas et al. (2022) also combined features for the dataset of machine learning classification in their research (Pratummas and Khemapatpapan 2022). In this work, the non-fatigue and fatigue condition was successfully classified and the Random Forest classifier produced 85% of accuracy by using MNF and MDF as the dataset.

The second objective of this work is to identify physical factors related to muscle fatigue. The physical factors are body mass index (BMI), age and years of driving (YOD). The most significant factor was determined using ANOVA analysis. The MDF was further analyzed to find the slope coefficient of the linear regression for MDF. The slope of MDF is normally used as a fatigue index. From the slope coefficient results, all subjects produced a negative value of the slope coefficient, indicating that all subjects experienced muscle fatigue during prolonged driving. Based on the result of the slope coefficient, female subjects tend to feel muscle fatigue faster than male subjects. By using the MDF value and the slope coefficient of MDF, the relation between physical factor (BMI, age and YOD) was studied. Based on ANOVA analysis, there is no statistically significant relation between MDF and all of the physical factors. The models were not accepted as the p-value>0.05 thus accepting the null hypothesis mentioned before which was: There is no statistically significant difference between each physical factors (BMI, age and YOD).

As for the regression analysis results using the slope coefficient of MDF, it is shown that only BMI is significantly related to muscle fatigue. This is due to the p-value for both second-order and third-order polynomial of BMI analyzed by ANOVA produce a value less than 0.05 which rejected the null hypothesis (H_0) and accepted the alternative hypothesis (H_a) which was: There is statistically significant difference between physical factors (BMI). The R^2 value for both second-order and third-order polynomial of BMI produced values higher than 0.6 which indicate the good fit of data. Based on the regression trend, it is concluded that the higher the BMI, the faster muscle fatigue occurrence to the driver during prolonged driving. Other physical factors which were age and YOD was not statistically significance as the p-value is greater than 0.05 thus accepting the null hypothesis (H_0). Moreover, the R^2 value calculated for age and YOD was less than 0.6 indicating that the model does not fit the actual data. Based on the results, age and YOD do not affect muscle fatigue rate of occurrence.

For the last objective of this research, two sets of mathematical models of second-order polynomial and third-order polynomial using BMI were successfully developed using ANOVA. The final step of development of mathematical model is to validate the model. In this work, the models were validated by calculating the residual error of predicted value and actual value (Ani et al. 2017). Based on the result, the residual error calculated was less than 10% indicating the model was accurately able to predict the muscle fatigue occurrence of the driver during prolonged driving with respect to the BMI value.

From the result of this study, the following conclusions were achieved:

- The amplitude of EMG signal was increasing with the increasing of muscle fatigue.
- Median frequency and Mean Frequency value of EMG signal were decreasing during muscle fatigue condition.
- Random Forest classifier successfully classify non-fatigue and fatigue condition using EMG signal with accuracy of 85%.
- Using more features as the input dataset in Machine Learning classifier improved the classifier's performance analysis.
- The negative value of the slope coefficient of MDF indicate that the subjects experienced fatigued during prolonged driving
- The slope coefficient value for female subjects were more negative value compared to male thus concluded that female subject tends to experience fatigue faster than male subjects.
- Only BMI affect the muscle fatigue of the driver during prolong driving. The regression trend shows that the higher the BMI value, the higher the rate of muscle fatigue.
- The mathematical model of second-order polynomial was successfully developed and validated using ANOVA analysis with respect to BMI and muscle fatigue.

5.2 RECOMMENDATION

Further research is suggested to be conducted on classifying and predicting muscle fatigue during prolonged driving. For the classification of muscle fatigue, it is suggested to include other psychophysical signals like electrocardiogram (ECG), electromyogram (EMG), electrooculogram (EoG) and electroencephalogram (EEG) signals to improve the classification accuracy. Moreover, it is hoped that more studies on muscle fatigue classification are carried out in the field of driving. To date, muscle fatigue classification mainly focuses on the area of rehabilitation, sport science, human-computer interaction and medical research.

For mathematical model development, it is suggested that the sample size should be bigger for better accuracy and effectiveness. In addition, a broader range of age, YOD and BMI is suggested for the population of the subject. In this research, the physical factors of the driver are only limited to BMI, age and YOD. There are other factors affecting muscle fatigue of the driver such as the type of road, driving environment, health condition, driving time and type of car need to be studied. In this study only trapezius muscle was under study. For future, it is suggested to study other muscle in the body and examine which muscle is the most fatigue.

The outcomes of this work forms important guidelines that can be used when studying driver’s muscle fatigue to reduce fatigue, avoid musculoskeletal disorders and prevent accidents. Indirectly, the number of lives lost due to road accidents can be reduced.

REFERENCES

Al-Hussein, Ward Ahmed, Miss Laiha Mat Kiah, Lip Yee Por, and Bilal Bahaa Zaidan. 2021. “Investigating the Effect of Social and Cultural Factors on Drivers in Malaysia: A Naturalistic Driving Study.” *International Journal of Environmental Research and Public Health* 18(22). doi: 10.3390/ijerph182211740.

Al-Mulla, Mohamed R., Francisco Sepulveda, and Martin Colley. 2011. “An Autonomous Wearable System for Predicting and Detecting Localised Muscle Fatigue.” *Sensors* 11(2):1542–57. doi: 10.3390/s110201542.

Albaghdadi, Ahmed, and Abduladhem Ali. 2019. “An Optimized Complementary Filter For An Inertial Measurement Unit Contain MPU6050 Sensor.” *Iraqi Journal for Electrical and Electronic Engineering* 15(2):71–77. doi: 10.37917/ijeee.15.2.8.

Ani, Mohammad Firdaus, Seri Rahayu Binti Kamat, and Kalthom Husin. 2017. “A Study of Psychophysical Factor for Driver Fatigue Using Mathematical Model.” *Journal of Mechanical Engineering* SI 3(2):109–22.

Baker, Richelle, Pieter Coenen, Erin Howie, Ann Williamson, and Leon Straker. 2018. “The Short Term Musculoskeletal and Cognitive Effects of Prolonged Sitting during Office Computer Work.” *International Journal of Environmental Research and Public Health* 15(8). doi: 10.3390/ijerph15081678.

Balasubramanian, Venkatesh, and K. Adalarasu. 2007. “EMG-Based Analysis of Change in Muscle Activity during Simulated Driving.” *Journal of Bodywork and Movement Therapies* 11(2):151–58. doi: 10.1016/j.jbmt.2006.12.005.

Bhardwaj, Rahul, Priya Natrajan, and Venkatesh Balasubramanian. 2018. “Study to Determine the Effectiveness of Deep Learning Classifiers for ECG Based Driver Fatigue Classification.” *2018 13th International Conference on Industrial and Information Systems, ICIIS 2018 - Proceedings* (978):98–102. doi:

10.1109/ICIINFS.2018.8721391.

Bhardwaj, Rahul, Swathy Parameswaran, and Venkatesh Balasubramanian. 2018. "Comparison of Driver Fatigue Trend on Simulator and On-Road Driving Based on EMG Correlation." *2018 13th International Conference on Industrial and Information Systems, ICIIS 2018 - Proceedings* 2(978):94–97. doi: 10.1109/ICIINFS.2018.8721431.

Boon-Leng, Lee, Lee Dae-Seok, and Lee Boon-Giin. 2016. "Mobile-Based Wearable-Type of Driver Fatigue Detection by GSR and EMG." *IEEE Region 10 Annual International Conference, Proceedings/TENCON* 2016-Janua(December). doi: 10.1109/TENCON.2015.7372932.

Buchanan, Thomas S., David G. Lloyd, Kurt Manal, and Thor F. Besier. 2004. "Neuromusculoskeletal Modeling: Estimation of Muscle Forces and Joint Moments and Movements from Measurements of Neural Command." *Journal of Applied Biomechanics* 20(4):367–95.

Candotti, C. T., J. F. Loss, M. La Torre, M. O. Melo, L. D. Araújo, and V. V. Marcks. 2009. "Use of Electromyography to Assess Pain in the Upper Trapezius and Lower Back Muscles within a Fatigue Protocol." *Revista Brasileira de Fisioterapia* 13(2):144–51. doi: 10.1590/S1413-35552009005000018.

Carneiro, J. G., E. M. Gonçalves, T. V. Camata, J. M. Altimari, M. V. Machado, A. R. Batista, G. Guerra, A. C. Moraes, and Leandro Ricardo Altimari. 2010. "Influence of Gender on the EMG Signal of the Quadriceps Femoris Muscles and Performance in High-Intensity Short-Term Exercise." *Electromyography and Clinical Neurophysiology* 50(7–8):326–32.

Chang, Kang Ming, Shin Hong Liu, and Xuan Han Wu. 2012. "A Wireless SEMG Recording System and Its Application to Muscle Fatigue Detection." *Sensors* 12(1):489–99. doi: 10.3390/s120100489.

Davidović, Jelica, Dalibor Pešić, and Boris Antić. 2018. “Professional Drivers’ Fatigue as a Problem of the Modern Era.” *Transportation Research Part F: Traffic Psychology and Behaviour* 55:199–209. doi: 10.1016/j.trf.2018.03.010.

Elamvazuthi, I., Zulika Zulkifli, Zulfiqar Ali, M. K. A. Ahame. Khan, S. Parasuraman, M. Balaji, and M. Chandrasekaran. 2015. “Development of Electromyography Signal Signature for Forearm Muscle.” *Procedia Computer Science* 76(Iris):229–34. doi: 10.1016/j.procs.2015.12.347.

El Falou, Wassim, Jacques Duchêne, Michel Grabisch, David Hewson, Yves Langeron, and Frédéric Lino. 2003. “Evaluation of Driver Discomfort during Long-Duration Car Driving.” *Applied Ergonomics* 34(3):249–55. doi: 10.1016/S0003-6870(03)00011-5.

Ferrari, R., and A. C. Croft. 2001. “Sitting Biomechanics, Part Ii: Optimal Car Driver’s Seat and Optimal Driver’s Spinal Model [1] (Multiple Letters).” *Journal of Manipulative and Physiological Therapeutics* 24(2):140–41. doi: 10.1067/mmt.2001.112550.

Freitas, Inês Flores Mendes de. 2008. “Fatigue Detection in EMG Signals.” 2008:1–82.

Fu, Rongrong, Hong Wang, and Wenbo Zhao. 2016. “Dynamic Driver Fatigue Detection Using Hidden Markov Model in Real Driving Condition.” *Expert Systems with Applications* 63:397–411. doi: 10.1016/j.eswa.2016.06.042.

Golmohammadishouraki, Mahdokht. 2022. “Prediction of Fatigue in Lower Extremity Using EMG Sensor and Machine Learning Affairs in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science.”

Hawa Harith, Siti, and Norashikin Mahmud. 2018. “Technical Determinant of Road Accident: A Systematic Review.” *International Journal of Engineering & Technology* 7(3.36):34. doi: 10.14419/ijet.v7i3.36.29074.

Hostens, I., and H. Ramon. 2005. “Assessment of Muscle Fatigue in Low Level Monotonous Task Performance during Car Driving.” *Journal of Electromyography and Kinesiology* 15(3):266–74. doi: 10.1016/j.jelekin.2004.08.002.

Jia, Bochen. 2020. “The Application of EMG-Based Methods in Evaluating the Impact of Prolonged Sitting on People’s Health.” *Sedentary Behaviour-A Contemporary View, Intech*.

Jian, Huang. 2017. “Design of Angle Detection System Based on MPU6050.” 73(Icemc):7–9. doi: 10.2991/icemc-17.2017.2.

Jung, Kyoung Sim, Jin Hwa Jung, Tae Sung In, and Hwi Young Cho. 2021. “Effects of Prolonged Sitting with Slumped Posture on Trunk Muscular Fatigue in Adolescents with and without Chronic Lower Back Pain.” *Medicina (Lithuania)* 57(1):1–8. doi: 10.3390/medicina57010003.

Kamat, Seri Rahayu, Mohammad Firdaus Ani, Minoru Fukumi Fukumi, and Nor Azila Noh. 2020. “A Critical Review on Driver Fatigue Detection and Monitoring System.” *International Journal of Road Safety* 1(2):53–58.

Karthick, P. A., Diptasree Maitra Ghosh, and S. Ramakrishnan. 2018. “Surface Electromyography Based Muscle Fatigue Detection Using High-Resolution Time-Frequency Methods and Machine Learning Algorithms.” *Computer Methods and Programs in Biomedicine* 154:45–56. doi: 10.1016/j.cmpb.2017.10.024.

Katsis, C. D., N. E. Ntouvas, C. G. Bafas, and D. I. Fotiadis. 2004. “Assessment of Muscle Fatigue during Driving Using Surface EMG.” *Proceedings of the IASTED International Conference on Biomedical Engineering* (January 2004):259–62.

Kett, Alexander R., Thomas L. Milani, and Freddy Sichting. 2021. “Sitting for Too Long, Moving Too Little: Regular Muscle Contractions Can Reduce Muscle Stiffness During Prolonged Periods of Chair-Sitting.” *Frontiers in Sports and Active Living* 3(November):1–9. doi: 10.3389/fspor.2021.760533.

Khairuddin, Ismail Mohd, Shahrul Naim Sidek, Anwar P. P. Abdu. Majeed, Mohd Azraai Mohd Razman, Asmarani Ahmad Puzi, and Hazlina Md Yusof. 2021. "The Classification of Movement Intention through Machine Learning Models: The Identification of Significant Time-Domain EMG Features." *PeerJ Computer Science* 7:1–15. doi: 10.7717/PEERJ-CS.379.

Khairul Amri Kamarudin, Mohd, Noorjima Abd Wahab, Roslan Umar, Ahmad Shakir Mohd Saudi, Muhammad Hafiz Md Saad, Nik Rozaireen Nik Rosdi1, Sarah Alisa Abdul Razak, Muhamad Murtadha Merzuki, Abdul Salam Abdullah, Siti Amirah, and Asyraff Mohd Ridzuan. 2018. "Road Traffic Accident in Malaysia: Trends, Selected Underlying, Determinants and Status Intervention." *International Journal of Engineering & Technology* 7(4.34):112. doi: 10.14419/ijet.v7i4.34.23839.

Kiswanto, Gandjar, Muhammad Fathin Juzar, Adjeng Ayu Setiani, Dody Rakhmat Ramadhan, Ferdiansyah Zhul-, and Rachmad Muhammad Suryantoro. 2018. "Muscle Contraction Sensor Filtering and Calibration for Virtual Manufacturing Development." *International Journal of Engineering and Technology (UAE), 7(4.16 Special Issue 16), 13-17* 7:13–17.

Konard, Peter. 2012. "The ABC of EMG." *Noraxon: Scottsdale* (April):1–60.

Lal, Saroj K. L., and Ashley Craig. 2001. "A Critical Review of the Psychophysiology of Driver Fatigue." *Biological Psychology* 55(3):173–94. doi: 10.1016/S0301-0511(00)00085-5.

Lecocq, Mathieu, Pascaline Lantoine, Clément Bougard, Jean Marc Allègre, Laurent Bauvineau, Christophe Bourdin, Tanguy Marqueste, and Erick Dousset. 2020. "Neuromuscular Fatigue Profiles Depends on Seat Feature during Long Duration Driving on a Static Simulator." *Applied Ergonomics* 87(Umr 7287). doi: 10.1016/j.apergo.2020.103118.

Li, Wenhao, Ming Zhang, Guomin Lv, Qingyu Han, Yuanjin Gao, Yan Wang, Qitao Tan, Manyu Zhang, Yixun Zhang, and Zengyong Li. 2015. “Biomechanical Response of the Musculoskeletal System to Whole Body Vibration Using a Seated Driver Model.” *International Journal of Industrial Ergonomics* 45:91–97. doi: 10.1016/j.ergon.2014.12.006.

Liu, Chi, and Yi Qiu. 2021. “Mechanism Associated with the Effect of Backrest Inclination on Biodynamic Responses of the Human Body Sitting on a Rigid Seat Exposed to Vertical Vibration.” *Journal of Sound and Vibration* 510(June):116299. doi: 10.1016/j.jsv.2021.116299.

Mahat, Norpah, Nurdiyana Jamil, and Siti Sarah Raseli. 2020. “Analysing Road Accident Triggers in Malaysia By Using Analytical Hierarchy Process.” *GADING Journal of Science and Technology* 3(2):118–25.

Majid, Noor Aliah binti Abdul, Mohd Fareez Edzuan Abdullah, Mohd Syahmi Jamaludin, Mitsuo Notomi, and John Rasmussen. 2013. “Musculoskeletal Analysis of Driving Fatigue: The Influence of Seat Adjustments.” *Advanced Engineering Forum* 10(December):373–78. doi: 10.4028/www.scientific.net/aef.10.373.

Marri, Kiran, and Ramakrishnan Swaminathan. 2016. “Classification of Muscle Fatigue Using Surface Electromyography Signals and Multifractals.” *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2015* 669–74. doi: 10.1109/FSKD.2015.7382022.

Mayer, D. G., and D. G. Butler. 1993. “Statistical Validation.” *Ecological Modelling* 68(1–2):21–32. doi: 10.1016/0304-3800(93)90105-2.

Menotti, Federica, Luciana Labanca, Luca Laudani, Arrigo Giombini, Fabio Pigozzi, and Andrea Macaluso. 2015. “Activation of Neck and Low-Back Muscles Is Reduced with the Use of a Neck Balance System Together with a Lumbar Support in Urban Drivers.” *PLoS ONE* 10(10):10–18. doi: 10.1371/journal.pone.0141031.

Mork, Paul Jarle, and Rolf H. Westgaard. 2007. "The Influence of Body Posture, Arm Movement, and Work Stress on Trapezius Activity during Computer Work." *European Journal of Applied Physiology* 101(4):445–56. doi: 10.1007/s00421-007-0518-4.

MOTM, MINISTRY OF Transport Malaysia. 2022. "Statistik Pengangkutan Malaysia 2021."

Narudin, Fairuz Amalina, Ali Feizollah, Nor Badrul Anuar, and Abdullah Gani. 2016. "Evaluation of Machine Learning Classifiers for Mobile Malware Detection." *Soft Computing* 20(1):343–57. doi: 10.1007/s00500-014-1511-6.

Ostojić, S., S. Peharec, V. Srhoj-Egekher, and M. Cifrek. 2018. "Differentiating Patients with Radiculopathy from Chronic Low Back Pain Patients by Single Surface Emg Parameter." *Automatika* 59(3–4):400–407. doi: 10.1080/00051144.2018.1553669.

Papakostas, Michalis, Varun Kanal, Maher Abujelala, Konstantinos Tsiakas, and Fillia Makedon. 2019. "Physical Fatigue Detection through EMG Wearables and Subjective User Reports - A Machine Learning Approach towards Adaptive Rehabilitation." *ACM International Conference Proceeding Series* (June):475–81. doi: 10.1145/3316782.3322772.

Pratummas, Parama, and Chaiyaporn Khemapatpapan. 2021. "Static Fatigue Detection in Office Syndrome Using SEMG and Machine Learning." Pp. 64–69 in *2021 International Conference on Intelligent Cybernetics Technology and Applications, ICICyTA 2021*. Institute of Electrical and Electronics Engineers Inc.

Pratummas, Parama, and Chaiyaporn Khemapatpapan. 2022. "Static Fatigue Detection in Office Syndrome Using SEMG and Machine Learning." *2021 International Conference on Intelligent Cybernetics Technology and Applications, ICICyTA 2021 (2021) 64-69* (December 2021):64–69. doi: 10.1109/icicyta53712.2021.9689169.

Qassim, Hassan M., Wan Zuha Wan Hasan, Hafiz R. Ramli, Hazreen Haizi Harith, Liyana Najwa Inche Mat, and Luthffi Idzhar Ismail. 2022. “Proposed Fatigue Index for the Objective Detection of Muscle Fatigue Using Surface Electromyography and a Double-Step Binary Classifier.” *Sensors (Basel, Switzerland)* 22(5). doi: 10.3390/s22051900.

Ramos, G., J. R. Vaz, G. V. Mendonça, P. Pezarat-Correia, J. Rodrigues, M. Alfaras, H. Gamboa, and Liang Zou. 2020. “Fatigue Evaluation through Machine Learning and a Global Fatigue Descriptor.” *Journal of Healthcare Engineering* 2020. doi: 10.1155/2020/6484129.

Sahayadhas, Arun, Kenneth Sundaraj, and Murugappan Murugappan. 2013. “Drowsiness Detection during Different Times of Day Using Multiple Features.” *Australasian Physical and Engineering Sciences in Medicine* 36(2):243–50. doi: 10.1007/s13246-013-0200-6.

Sanjaya, Kadek Heri, Soomin Lee, and Tetsuo Katsuura. 2016. “Review on the Application of Physiological and Biomechanical Measurement Methods in Driving Fatigue Detection.” *Journal of Mechatronics, Electrical Power, and Vehicular Technology* 7(1):35–48. doi: 10.14203/j.mev.2016.v7.35-48.

De Santana, Ligia Moreira, Paulo Robertocarvalho Do Nascimento, Thais De Sousa Lima, Ana Carolina Tocilo Lopes, Amanda Costa Araujo, Fábio Mícolis De Azevedo, and Rúben De Faria Negrão Filho. 2014. “Electromyographic Analysis of the Vertebral Extensor Muscles during the Biering-Sorensen Test.” *Motriz. Revista de Educacao Fisica* 20(1):112–19. doi: 10.1590/S1980-65742014000100017.

Shariff, S. S. R., F. N. M. Nusa, H. A. Hanizan, and M. N. Taib. 2022. “An Analysis of Accidents Involving Public Transport along Highway.” *IOP Conference Series: Earth and Environmental Science* 1022(1). doi: 10.1088/1755-1315/1022/1/012021.

Silva, Hugo Plácido, José Guerreiro, André Lourenço, Ana Fred, and Raúl Martins. 2014. "BITalino: A Novel Hardware Framework for Physiological Computing." *PhyCS 2014 - Proceedings of the International Conference on Physiological Computing Systems* 246–53. doi: 10.5220/0004727802460253.

Sonmezocak, Temel, and Serkan Kurt. 2021. "Machine Learning and Regression Analysis for Diagnosis of Bruxism by Using EMG Signals of Jaw Muscles." *Biomedical Signal Processing and Control* 69(June):102905. doi: 10.1016/j.bspc.2021.102905.

Stegeman, D., and H. Hermens. 2007. "Standards for Surface Electromyography: The European Project Surface EMG for Non-Invasive Assessment of Muscles (SENIAM)." *Línea). Disponible En: Http://Www. Med. ...* (January):108–12.

Tanvi Khurana & Suman Singh. 2017. "Understanding Static Muscular Contractions and Bodily Movements." *International Journal of Applied and Natural Sciences (IJANS)* 6(4):91–96.

Technologies, Advancer. 2015. "Myoware Datasheet." 1–8.

Tengku Zawawi, T. N. S., A. R. Abdullah, M. H. Jopri, T. Sutikno, N. M. Saad, and R. Sudirman. 2018. "A Review of Electromyography Signal Analysis Techniques for Musculoskeletal Disorders." *Indonesian Journal of Electrical Engineering and Computer Science* 11(3):1136–46. doi: 10.11591/ijeecs.v11.i3.pp1136-1146.

Thongpanja, S., A. Phinyomark, P. Phukpattaranont, and C. Limsakul. 2013. "Mean and Median Frequency of EMG Signal to Determine Muscle Force Based on Time Dependent Power Spectrum." *Elektronika Ir Elektrotechnika* 19(3):51–56. doi: 10.5755/j01.eee.19.3.3697.

Thongpanja, Sirinee, Angkoon Phinyomark, Franck Quaine, Yann Laurillau, Booncharoen Wongkittisuksa, Chusak Limsakul, and Pornchai Phukpattaranont. 2013. "Effects of Window Size and Contraction Types on the Stationarity of Biceps Brachii Muscle EMG Signals." *I-CREATe 2013 - International Convention on Rehabilitation Engineering and Assistive Technology, in Conjunction with SENDEX 2013* (July 2014).

Tlili, Ferdews, Rim Haddad, Ridha Bouallegue, and Neila Mezghani. 2021. "A Real-Time Posture Monitoring System Towards Bad Posture Detection." *Wireless Personal Communications* 120(2):1207–27. doi: 10.1007/s11277-021-08511-2.

Toro, Sergio Fuentes Del, Silvia Santos-Cuadros, Ester Olmeda, Carolina Álvarez-Caldas, Vicente Díaz, and José Luís San Román. 2019. "Is the Use of a Low-Cost SEMG Sensor Valid to Measure Muscle Fatigue?" *Sensors (Basel, Switzerland)* 19(14). doi: 10.3390/s19143204.

Venugopal, G., M. Navaneethakrishna, and S. Ramakrishnan. 2014. "Extraction and Analysis of Multiple Time Window Features Associated with Muscle Fatigue Conditions Using SEMG Signals." *Expert Systems with Applications* 41(6):2652–59. doi: 10.1016/j.eswa.2013.11.009.

Wan Husin, Wan Zakiyatussariroh, Adlina Sofia Afdzal, Nur Lisa Hashim Azmi, and Siti Auni Taqiah Sheikh Hamadi. 2021. "Box-Jenkins and State Space Model in Forecasting Malaysia Road Accident Cases." *Journal of Physics: Conference Series* 2084(1). doi: 10.1088/1742-6596/2084/1/012005.

Wang, Fan, Hong Chen, Cai Hua Zhu, Si Rui Nan, and Yan Li. 2019. "Estimating Driving Fatigue at a Plateau Area with Frequent and Rapid Altitude Change." *Sensors (Switzerland)* 19(22). doi: 10.3390/s19224982.

Wang, Lejun, Yuting Wang, Aidi Ma, Guoqiang Ma, Yu Ye, Ruijie Li, and Tianfeng Lu. 2018. "A Comparative Study of EMG Indices in Muscle Fatigue Evaluation Based on Grey Relational Analysis during All-Out Cycling Exercise." *BioMed Research International* 2018. doi: 10.1155/2018/9341215.

Wang, Lin, Hong Wang, and Xin Jiang. 2017. "A New Method to Detect Driver Fatigue Based on Emg and Ecg Collected by Portable Non-Contact Sensors." *Promet - Traffic - Traffico* 29(5):479–88. doi: 10.7307/ptt.v29i5.2244.

Yousif, Hayder A., Ammar Zakaria, Norasmadi Abdul Rahim, Ahmad Faizal Bin Salleh, Mustafa Mahmood, Khudhur A. Alfarhan, Latifah Munirah Kamarudin, Syed Muhammad Mamduh, Ali Majid Hasan, and Moaid K. Hussain. 2019. "Assessment of Muscles Fatigue Based on Surface EMG Signals Using Machine Learning and Statistical Approaches: A Review." *IOP Conference Series: Materials Science and Engineering* 705(1). doi: 10.1088/1757-899X/705/1/012010.

Zaleha, S. H., Nur Haliza Abdul Wahab, Norafida Ithnin, Johana Ahmad, Noor Hidayah Zakaria, Chinonso Okereke, and A. K. Nurain. Huda. 2021. "Microsleep Accident Prevention for SMART Vehicle via Image Processing Integrated with Artificial Intelligent." *Journal of Physics: Conference Series* 2129(1):2–7. doi: 10.1088/1742-6596/2129/1/012082.

Zhang, Weizhe, Shuohan Wang, Nan Bao, and Wenbin Li. 2020. "A Wearable Cervical Fatigue Monitoring System Based On Multi-Sensor Data." 2(1):1–11. doi: 10.25236/FMSR.2020.020101.

Zhao, Kai, Jian Guo, Shuxiang Guo, and Qiang Fu. 2022. "Design of Fatigue Grade Classification System Based on Human Lower Limb Surface EMG Signal." *2022 IEEE International Conference on Mechatronics and Automation, ICMA 2022* 1015–20. doi: 10.1109/ICMA54519.2022.9855927.

APPENDIX I – CONSENT FORM

Department of Mechatronics Engineering
Kulliyah of Engineering
International Islamic University Malaysia
53100, Gombak
Selangor

Version 1, 20 May 2020

Subject Identification Number for this trial: __________________

CONSENT FORM

MATHEMATICAL MODELING FOR THE ERGONOMIC ANALYSIS OF DRIVER DURING PROLONGED DRIVING

Investigators: Dr. Nur Liyana Azmi, Dr. Khairul Affendy Md Nor, Dr Nor Hidayati Diyana Nordin, Noor Azlyn Ab Ghafar

Please initial the boxes

1 I confirm that I have read and understand the Participant Information Sheet dated 20/05/2020 for the above study and have had the opportunity to ask questions.

2 I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, and without my medical care, education or legal rights being affected.

3 I agree to my anonymised data.

4 I agree to take part in the above study.

5 I would like to be provided with a summary report of our findings at the end of the study, at my request

________________________________ ______________

Name of the subject Date

Signature

________________________________ ______________

Name of the person taking consent Date

Signature

APPENDIX II – QUESTIONNAIRE FORM

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

Garden of Knowledge and Virtue

LEADING THE WAY

KHALĪFAH • AMĀNAH • IQRA' • RAHMATAN LIL-ĀLAMĪN

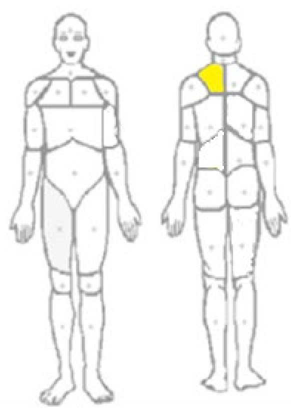
(Company No. 101067-P)

Questionnaire

Subject ID				**Age**	
Weight	kg	**Height**	cm	**Gender**	F / M
Years of Driving					

Instruction: Please tick √ if the driver feels drowsy, sleepy or experiences any muscle pain during experiment is conducted.

Location of EMG sensor



Minutes	Fatigue
0 (start)	
5	
10	
15	
20	
25	
30	
35	
40	
45	
50	
55	
60	

Minutes	Fatigue
65	
70	
75	
80	
85	
90	
95	
100	
105	
110	
115	
120	
125 (end)	

APPENDIX III – MPU6050 PROGRAMMING CODE

```
#include "Wire.h"
 #include <MPU6050_light.h>
 MPU6050 mpu(Wire);
 unsigned long timer = 0;
 void setup() {
   Serial.begin(9600);
   Wire.begin();
 byte status = mpu.begin();
   Serial.print(F("MPU6050 status: "));
   Serial.println(status);
   while (status != 0) { } // stop everything if could not connect to MPU6050
 Serial.println(F("Calculating offsets, do not move MPU6050"));
   delay(1000);
   mpu.calcOffsets(); // gyro and accelero
   Serial.println("Done!\n");
 }
 void loop() {
   mpu.update();
 if ((millis() - timer) > 10) { // print data every 10ms
     Serial.print("X : ");
     Serial.print(mpu.getAngleX());
     Serial.print("\tY : ");
     Serial.print(mpu.getAngleY());
     Serial.print("\tZ : ");
     Serial.println(mpu.getAngleZ());
     timer = millis();
   }
 }
```

APPENDIX IV – CLASSIFICATION CODE

Random Forest Classification

Importing the libraries

```
[ ] import numpy as np
    import matplotlib.pyplot as plt
    import pandas as pd
```

Importing the dataset

```
[ ] dataset = pd.read_csv('COMMIN.csv')
    X = dataset.iloc[:, :-1].values
    y = dataset.iloc[:, -1].values
```

```
[ ] print(X)
```

Splitting the dataset into the Training set and Test set

```
[ ] from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)
```

Feature Scaling

```
[ ] from sklearn.preprocessing import StandardScaler
    sc = StandardScaler()
    X_train = sc.fit_transform(X_train)
    X_test = sc.transform(X_test)
```

Training the Random Forest Classification model on the Training set

```
[ ] from sklearn.ensemble import RandomForestClassifier
    classifier = RandomForestClassifier(n_estimators = 10, criterion = 'entropy', random_state = 0)
    classifier.fit(X_train, y_train)
```

Making the Confusion Matrix

```
[ ] from sklearn.metrics import confusion_matrix, accuracy_score, plot_confusion_matrix
    y_pred = classifier.predict(X_test)
    cm = confusion_matrix(y_test, y_pred)
    print(cm)
    accuracy_score(y_test, y_pred)
```

```
[ ] from sklearn.model_selection import cross_val_score
    accuracies = cross_val_score(estimator = classifier, X = X_train, y = y_train, cv = 10)
    print("Accuracy: {:.2f} %".format(accuracies.mean()*100))
    print("Standard Deviation: {:.2f} %".format(accuracies.std()*100))
```