DETECTING MENTAL FATIGUE IN ONLINE LEARNERS USING EEG-BASED EMOTIONAL METRICS AND ENSEMBLE LEARNING

BY

MD FARHAD HOSSAIN

A dissertation submitted in fulfillment of the requirement for the degree of Master of Computing (Computer Science and Information Technology).

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

DECEMBER 2024

ABSTRACT

Mental fatigue is a brain state that causes energy levels to decline due to engagement in a prolonged period of cognitive tasks, such as online learning. It results attention deficits and poor academic performance by reducing focus and engagement during online learning sessions. It is essential to address mental fatigue early because it leads to permanent cognitive deficits and emotional impairments. Existing subjective assessments like surveys, interviews and self-report questionnaires are time-consuming and often reported as bias, since humans cannot accurately measure their cognitive performance. Although there are numerous machine learning (ML) approaches for mental fatigue detection using electroencephalogram (EEG) signals, there is a lack of study on mental fatigue detection among online learners. Therefore, this study aims to use EEG-derived emotional states metrics for mental fatigue detection by identifying existing EEG-based emotional metrics and mental fatigue detection techniques through literature review, then developing and evaluating machine learning models. Thus, this study employs Emotiv Performance Metrics (EPM) including engagement, excitement, interest, stress and relaxation. The mental fatigue detection consists of EPM metrics collection using an Emotiv Insight headset, and ML model development including Logistic Regression (LR), Support Vector Machine (SVM), Multilayer Perceptron (MLP) and Ensemble Learning model. A total of 10 students participated in the data collection session. However, data from only 8 participants were included in the mental fatigue detection. The two participants were excluded based on the Chalder fatigue Scale (CFS). One participant pre-CFS score is 4, which defines that he is already in a fatigue state. Other participants pre-CFS and post-CFS score was less than 4, that indicates the participants did not fatigue during the learning session. The findings of descriptive analysis, correlation analysis and statical analysis show the significance of EPM between the fatigue and non-fatigue sessions. The mental fatigue detection models show reliable and consistent performance. Notably, the Ensemble model exhibits the most consistency and reliability in the subject-dependent analysis with average accuracy 0.91, F1-score 0.91, ROC AUC score 0.96 and cross validation 0.87, where subject-independent analysis with accuracy 0.75, F1-score 0.76, ROC AUC 0.83 and cross-validation mean 0.75. This evident that the subject-dependent models are more consistent than the subject-independent model. The findings are also evident that EPM emotional states are crucial features in the mental fatigue detection for online learners. Furthermore, the findings of this research lead to proposes a real-time mental fatigue intervention framework for online learners, where mental fatigue detection is achieved using EPM emotional metrics and Ensemble learning. Future studies need to investigate the influence factors of emotional states and mental fatigue to evaluate the mental fatigue intervention framework. A comprehensive investigation is required for subject-independent and subject-dependent based mental fatigue detection and intervention techniques.

ملخص البحث

الإرهاق العقلي هو حالة دماغية تسبب انخفاض مستويات الطاقة بسبب الانخراط في فترة طويلة من المهام الإدراكية، مثل التعلم عبر الإنترنت. ويسبب ذلك ضعف في الانتباه والأداء الأكاديمي نتيجة قلة التركيز والمشاركة أثناء جلسات التعلم عبر الإنترنت. ومن الضروري معالجة الإرهاق العقلي في وقت مبكر لأنه يؤدي إلى عجز إدراكي دائم واضطرابات عاطفية. التقييمات الذاتية الحالية، مثل الاستطلاعات والمقابلات والاستبيانات الذاتية، تستغرق وقتًا طويلاً وغالبًا ما يتم الإبلاغ عنها كتحيز، لأن البشر لا يمكنهم قياس أدائهم الإدراكي بدقة. على الرغم من وجود العديد من الأساليب التي تعتمد على التعلم الآلي (ML) للكشف عن الإجهاد العقلي باستخدام إشارات تخطيط كهربية الدماغ(EEG) ، إلا أن هناك نقصًا في الدراسات المتعلقة بالكشف عن الإرهاق العقلي بين المتعلمين عبر الإنترنت. لذلك، تهدف هذه الدراسة إلى استخدام مقاييس الحالات العاطفية المستمدة من EEG للكشف عن الإرهاق العقلي من خلال تحديد المقاييس العاطفية الحالية المستندة إلى EEG وتقنيات الكشف عن الإجهاد العقلي من خلال مراجعة الأدبيات، ثم تطوير وتقييم نماذج تعلم الآلة. لذلك، تعتمد هذه الدراسة على مقاييس الأداء من Emotiv (EPM)التي تشمل المشاركة، والإثارة، والاهتمام، والضغط، والاسترخاء. تتكون عملية الكشف عن الإرهاق العقلي من جمع مقاييس EPM باستخدام سماعة رأسEmotiv Insight، وتطوير نماذج تعلم آلي تشمل الانحدار اللوجستي(LR) ، وآلة دعم المتجهات(SVM) ، والشبكات العصبية متعددة الطبقات(MLP) ، ونموذج التعلم التجميعي(Ensemble Learning) . شاركت مجموعة من 10 طلاب في جلسة جمع البيانات، ولكن تم تضمين بيانات 8 مشاركين فقط في عملية الكشف عن الإرهاق العقلي. حيث تم استبعاد مشاركين اثنين بناءً على مقياس تشالدر للإجهاد(CFS) . حيث أظهر أحد المشاركين درجة CFS قبل الجلسة تساوي 4، مما يشير إلى أنه كان في حالة إجهاد بالفعل. بينما كانت درجات CFS قبل وبعد الجلسة لبقية المشاركين أقل من 4، مما يشير إلى أنهم لم يتعرضوا للإجهاد أثناء جلسة التعلم. أظهرت نتائج التحليل الوصفي، التحليل الارتباطي، والتحليل الإحصائي أهمية مقاييس EPM بين الجلسات المتعبة وغير المتعبة. وأظهرت نماذج الكشف عن الإرهاق العقلي أداءً موثوقًا ومتسقًا. بشكل ملحوظ، أظهر نموذج التعلم التجميعي (Ensemble) الاتساق والموثوقية الأكبر في التحليل المعتمد على الموضوع بمتوسط دقة 0.91، ودرجة F1 تساوي 0.91، ودرجة ROC AUC تساوي 0.96، وتحقق تقاطع متوسط 0.87. بينما في التحليل المستقل عن الموضوع، حقق النموذج دقة تبلغ 0.75، ودرجة F1 تساوي 0.76، ودرجة ROC AUC تساوي 0.83، ومتوسط التحقق المتقاطع 0.75. وهذا يدل بوضوح على أن النماذج المعتمدة على الموضوع أكثر اتساقًا من النماذج المستقلة عن الموضوع. كما أظهرت النتائج أن الحالات العاطفية وفقًا لمقاييس EPM هي سمات أساسية للكشف عن الإرهاق العقلي بين المتعلمين عبر الإنترنت. بالإضافة إلى ذلك، تقود نتائج هذه الدراسة إلى اقتراح إطار تدخل في الوقت الفعلي للإرهاق العقلي للمتعلمين عبر الإنترنت، حيث يتم الكشف عن الإجهاد العقلي باستخدام مقاييس EPM العاطفية والتعلم التجميعي. تحتاج الدراسات المستقبلية إلى التحقيق في العوامل المؤثرة على الحالات العاطفية وإرهاق العقل لتقييم إطار التدخل المقترح. وهناك حاجة إلى تحقيق شامل للكشف عن إرهاق العقل وتقنيات التدخل المستندة إلى التحليل المعتمد على الموضوع والمستقل عن الموضوع.

ASST. PROF. DR. ZAINAB SENAN ATTAR BASHI
Assistant Professor
Department of Computer Science
Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

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 Computing (Computer Science and Information Technology).

………………………………

Assoc. Prof. Ts. Dr. Hamwira Sakti
Bin Yaacob
Supervisor

………………………………

Prof. Dr. Faizah Binti Idrus
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 Computing (Computer Science and Information Technology).

………………………………

Ts. Dr. Dini Oktarina Dwi
Handayani
Internal Examiner

………………………………

Assoc. Prof. Ts. Dr. Mohd Ibrahim
Bin Shapiai
External Examiner

This dissertation was submitted to the Department of Computer Science and is accepted as a fulfilment of the requirement for the degree of Master of Computing (Computer Science and Information Technology).

....

Head, Department of Computer Science

This dissertation was submitted to the Kulliyyah of Information and Communication Technology and is accepted as a fulfillment of the requirement for the degree of Master of Computing (Computer Science and Information Technology).

....

Kulliyyah of Information and Communication Technology

DECLARATION

I hereby declare that this dissertation 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.

Md Farhad Hossain

23/12/2024

Signature…………………………… Date……………………

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

DETECTING MENTAL FATIGUE IN ONLINE LEARNERS USING EEG-BASED EMOTIONAL METRICS AND ENSEMBLE LEARNING

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

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

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

Affirmed by Md Farhad Hossain

……………………………… 23/12/2024 ……………………..

Signature Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

DETECTING MENTAL FATIGUE IN ONLINE LEARNERS USING EEG-BASED EMOTIONAL METRICS AND ENSEMBLE LEARNING

I declare that the copyright holder of this thesis/dissertation is International Islamic University Malaysia.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

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

Affirmed by Md Farhad Hossain

	23/12/2024
……………………………….	…………………..
Signature	Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

DETECTING MENTAL FATIGUE IN ONLINE LEARNERS USING EEG-BASED EMOTIONAL METRICS AND ENSEMBLE LEARNING

I declare that the copyright holder of this thesis/dissertation is Name of the Student.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

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

Affirmed by Md Farhad Hossain

23/12/2024

……………………………… Signature

………………….. Date

This thesis is dedicated to my respected parents, beloved wife, caring younger sisters, my professors and schools teachers for laying the foundation of what I turned out to be

ACKNOWLEDGEMENTS

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

I am most indebted to by supervisor, Assoc. Prof. Ts. Dr. Hamwira Sakti Bin Yaacob, whose enduring disposition, kindness, promptitude, thoroughness and friendship have facilitated the successful completion of my work. I put on record and appreciate his detailed comments, useful suggestions and inspiring queries which have considerably improvedthis thesis. His brilliant grasp of the aim and content of this work led to his insightful comments, suggestions and queries which helped me a great deal. Despite his commitments, he took time to listen and attend to me whenever requested. The moral support he 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, Prof. Dr. Faizah Idrus, whose support and cooperation contributed to the outcome of this work.

Lastly, my gratitude goes to my respected parents, my beloved wife, lovely children and caring younger sisters; for their prayers, understanding and endurance while away.

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAR	Automatic Artifact Removal
ADHD	Attention-Deficit/Hyperactivity Disorder
ANN	Artificial Neural Network
BCI	Brain-Computer Interface (BCI)
CFQ	Chalder Fatigue Questionnaire
CFS	Chalder Fatigue Scale
CNN	Convolutional Neural Network
COVID 19	Coronavirus Disease 2019
CSP	Common Spatial Pattern
DE	Differential Entropy
DFT	Discrete Fourier Transformation
DNN	Deep Neural Network
DQN	Deep Q-Network
EC	Eyes Closed
ECG	Electrocardiogram
EEG	Electroencephalogram
EMA	Exponential Moving Average
EMD	Empirical Mode Decomposition
EMG	Electromyography
EO	Eyes Open
EOG	Electrooculogram
EPM	Emotiv Performance Metrics
ESTCNN	EEG-Based Spatial-Temporal Convolutional Neural Network
FAA	Frontal Alpha Asymmetry
FAT	Fatigue
FFT	Fast Fourier Transform

FIR	Finite Impulse Response
FLDA	Fisher Linear Discriminant Analysis
FL-RSEFNN	Functional-Link Recurrent Self-Evolving Fuzzy Neural Network
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional near-infrared spectroscopy
FWNN	Fuzzy Wavelet Neural Network
GD	Gradient Descent
GP	Gaussian Process
ICA	Independent Component Analysis
IIR	Infinite Impulse Response
IIUM	International Islamic University Malaysia
IRA	Independent Residual Analysis
IREC	IIUM Research Ethics Committee
KNN	K-Nearest Neighbor
KPLS	Kernel Partial Least Squares Regression
LDA	Linear Discriminant Analysis
LE	Logarithmic Energy
LightFD	Lightweight Flow Detection
LORETA	Low-Resolution Brain Electromagnetic Tomography
LR	Logistic Regression
LSTM	Long Short-Term Memory
MF	Mental Fatigue
MFB-CNN	Multiple-Feature-Branch Convolutional Neural Network
ML	Machine Learning
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
NB	Naive Bayes
NF	Neurofeedback
NIRS	Near Infrared Spectroscopy

Non-FAT	Non-Fatigue
NU	Negative-Unlabeled
ONFIN	Online Neural Fuzzy Inference Network
PET	Positron Emission Tomography
PP	Power Percentage
PPO	Proximal Policy Optimization
PSD	Power Spectral Density
PSO-HELM	Particle Swarm Optimization-Holomorphic Embedding Load-flow Method
QDA	Quadratic Discriminant Analysis
ROC AUC	Receiver Operating Characteristic Area Under The Curve
RSEFNN	Recurrent Self-Evolving Fuzzy Neural Network
SCP	Slow Cortical Potentials
SDG	Sustainable Development Goal
SG	Savitzky Golay
SMR	Sensory Motor Rhythm
SOBI	Second-Order Blind Identification
SONFIN	Online Self-Constructing Neural Fuzzy Inference Network
SPECT	Single Photon Emission Computed Tomography
SSP	Signal-Space Projection
STFT	Short-Time Fourier Transform
SURE	Stein's Unbiased Risk Estimate
SVM	Support Vector Machines
SVM-P	Support Vector Machines Probabilistic
SVM-RBF	Support Vector Machines Radial Basis Function
SVR	Support Vector Regression
TBR	Theta/Beta Ratio
TRFN	TSK-Type Recurrent Fuzzy Network

LIST OF APPENDICES

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF STUDY

The use of technology become a central feature of education prior to the COVID-19 epidemic. The increasing acceptance of online learning platforms has resulted in a fundamental shift of the educational landscape. Despite this, there are several particular challenges associated with online learning. According to the self-directed nature of online learning, it might be especially difficult for an individual to demonstrate high levels of motivation, time management abilities, and self-discipline (Moralista et al., 2022; Ngien & Hogan, 2022). Lengthy online learning sessions generally require extended screen time, which can result in reduced attention spans and mental fatigue (Salim et al., 2022). Several research works (Y. Chen & Qin, 2023; Maloney et al., 2023) have explored the complex field of mental fatigue in an effort to pinpoint its symptoms as well as comprehend the numerous risk factors that are connected to it.

Mental Fatigue occurs when a person is involved in an activity for an extended amount of time. It's similar to physical tiredness, but it affects your mind rather than your body and brain's energy levels drop (Rudroff et al., 2020; Van Der Linden & Eling, 2006). It is tough to pay attention to a task when someone is mentally fatigue. Mental fatigue can result in significant concerns such as worry, tension, and physical ailments (Maciej Serda et al., 2000; Marcora et al., 2009; van der Linden et al., 2003). The impact of COVID-19 is huge on human mental conditions. Long term of work from home, study from home, and quarantine makes everyone very easily stressful (Jin et al., 2022; Rudroff et al., 2020). Researchers (Y. Chen & Qin, 2023) found a negative relationship between academic performance and fatigue level, indicating that higher levels of fatigue were linked to worse academic achievement. Mental fatigue Addressing is one of the most important issues to improve cognitive performance. Mental fatigue can lead to cognitive impairments (Van Der Linden & Eling, 2006).

Over the last few decades, there has been a huge growth in brain signal analysis to understand the affective state of various cognitive and mental states due to improvements in high-throughput technology. A Brain-Computer Interface (BCI) is a cutting-edge technology in which researchers conducted a series of experiments to demonstrate that direct human-machine communication is possible. The electroencephalogram (EEG) is a common, non-invasive approach for monitoring and evaluating the brain's condition (Thakor & Sherman, 2013). EEG recording systems become more widely available as low-cost wearable gadgets, a wide range of non-therapeutic uses has arisen.

EEG signals, a non-invasive brain wave, are used by BCI. The EEG power is dynamically and non-linearly linked to human cognition. BCI is a hardware and software framework for monitoring machines and other communication devices for a specific job. As an example, recent research has used wireless EEG devices for brain-computer interfaces and neurofeedback applications (Jap et al., 2009). Additionally, EEG has been used by numerous research groups to improve road safety by detecting and quantifying driver drowsiness (Kar et al., 2010). Beyond the evaluation of drowsiness and fatigue, the recognition of various mental states based on EEG data, as well as the research of cognitive performance, are important strides forward. The use of EEG-based BCI to identify and measure mental fatigue has shown to be reliable (Gao et al., 2019; Lee et al., 2020a; Y. T. Liu, Lin, Wu, Hsieh, et al., 2016; Myrden & Chau, 2017; Roy et al., 2014; H. Zeng et al., 2019).

The human brain is extremely flexible to process and learn new information and forming new neural pathways in the process. This is referred as neuroplasticity (Demarin et al., 2014). Usually, EEG based BCI uses scalp-mounted sensors to record electrical impulses from the brain and then uses for analyzing different mental states. Researchers demonstrated that using different techniques including machine learning (ML) techniques with EEG signals, emotional states can be detected (Hassan et al., 2023; Jafari et al., 2023; Nandini et al., 2023; Sidharth et al., 2023). Machine learning algorithms can find complex pattern in EEG signal. However, mental fatigue has impact on different emotional states evident in (Maloney et al., 2023; Schlichta et al., 2022; D. Wang et al., 2022). Emotional regulation is impaired by mental fatigue, while emotional reactivity is unaffected (Grillon et al., 2015). This suggests a strong relation between mental fatigue and emotional states.

A variety of emotions, such as attention, focus, engagement, excitement, interest, relaxation, stress, sad, happy and many more. These emotions can have a major effect on the learning outcomes. Emotional metrics are measurable metrics or indicators that are used to evaluate or quantify particular features of emotional states. Emotiv performance metrics are different emotional state metrics which are engagement, excitement, interest, stress and relaxation (Emotiv, 2023). These metrics showed reliable in many research including for different mental condition analysis (Asif et al., 2023; Faruk et al., 2021; Holman & Adebesin, 2019; Paranthaman et al., 2021; Santoyo-Mora et al., 2022; Strmiska & Koudelkova, 2018).

EEG signals are converted into useful commands by a BCI, which makes it possible to determine emotional states and mental fatigue. The goal of this research is to analyze different emotional metrics produced from EEG signals in order to construct a Machine Learning-based model for mental fatigue detection. The approach aims to optimize online learning settings by offering a mental fatigue detection model in online learners utilizing emotional metrics. Additionally, this study offers future research by proposing a framework for mental fatigue intervention for online learners.

1.2 PROBLEM STATEMENT

It is essential to early address mental fatigue because it causes cognitive deficits (Kamal et al., 2021; Moralista et al., 2022; Ngien & Hogan, 2022; Salim et al., 2022; Van Cutsem et al., 2022; Van Der Linden & Eling, 2006). Traditional self-report questionnaire techniques are subjective measure of mental fatigue which are often reported as bias because people are not able to judge their cognitive ability (Kunasegaran et al., 2023; Schmidt et al., 2009; Tempelaar et al., 2020), in this reason an objective measure is required with the subjective measure (Tempelaar et al., 2020). Recent studies show complex relationship between mental fatigue and emotion (Lewczuk et al., 2022). Although numerous methods exist for detecting mental fatigue using EEG based BCI (Hossan et al., 2017; Myrden & Chau, 2017; H. Zeng et al., 2019), but there is a lack of investigation on the detection of mental fatigue using EEG based emotional metrics.

1.3 RESEARCH OBJECTIVES

This research purpose to detect mental fatigue using EEG based emotional metrics through machine learning technique by following the objectives:

RO1: To explore existing EEG emotional metrics and mental fatigue detection techniques.

RO2: To develop machine learning model for mental fatigue detection using EEG-based emotional metrics among online learners.

RO3: To evaluate the mental fatigue detection model.

1.4 RESEARCH QUESTIONS

This research works on the following questions:

RQ1: What is the existing EEG derived emotional metrics and mental fatigue detection techniques?

RQ2: How to develop a robust ML model for mental fatigue detection using EEG-derived emotional metrics?

RQ3: How to evaluate the mental fatigue detection model?

1.5 RESEARCH FRAMEWORK

The research framework shown in Table 1 is the result of combining the research objectives and questions. There are both descriptive and quantitative aspects of this research work. The comprehensive literature review is performed and described in chapter two, which refers as descriptive research sections. The literature review helps to identify the existing EEG based mental fatigue detection techniques, EEG-based emotional states metrics and its available application programming interface (API) to use as feature for mental fatigue detection. The literature review also used to investigate neurofeedback as mental fatigue intervention and neurofeedback modality for mental fatigue intervention. The quantitative study helps to design, develop and evaluate the mental fatigue detection model using EEG-based emotional metrics those described in chapter three, four, and five. A data collection process during an online learning session, then data preparation performed for descriptive analysis, correlation analysis, statical testing and develop mental fatigue detection model. The mental fatigue detection model evaluation includes the ML model evaluation metrics: (i) accuracy, (ii) F1-score, (iii) receiver operating characteristic area under the curve (ROC AUC) score and (iv) cross-validation. A real-time mental fatigue intervention framework is proposed as the future study focus in this research. In this purpose a comprehensive literature review is performed to understand the neurofeedback for mental fatigue intervention and neurofeedback modality that has relationship with mental fatigue for online learners.

Table 1 Research Framework

Research Objectives	Research Questions	Methods	Outcomes
RO1: To explore existing EEG emotional metrics and mental fatigue detection techniques.	RQ1: What is the existing EEG derived emotional metrics and mental fatigue detection techniques?	Literature review	Identified existing mental fatigue detection techniques, available emotional metrics API. Publication on literature review (LR) (F. Hossain & Yaacob, 2022; Yaacob et al., 2023) and a proposal (F. Hossain & Yaacob, 2023b)
RO2: To develop machine learning model for mental fatigue detection using EEG-based emotional metrics among online learners.	RQ2: How to develop a robust ML model for mental fatigue detection using EEG-derived emotional metrics?	Data collection, descriptive analysis, correlation analysis, statical testing, ML model development	A ML model for mental fatigue detection using EEG-based emotional states metrics
RO3: To evaluate the mental fatigue detection model.	RQ3: How to evaluate the mental fatigue detection model?	ML model evaluation metrics including accuracy, F1-score, ROC AUC score and cross-validation	Evaluated mental fatigue detection model.

1.6 RESEARCH SIGNIFICANCE

The significance of this research is extensive, since it is in accordance with the Malaysian Science, Technology, Innovation and Economy (MySTIE) Framework, Islamic ethical standards, and the Sustainable Development Goal (SDG). To begin with, this research supports the third SDG of supporting Good Health and Well-Being. In addition, maintaining mental capabilities and encouraging intellectual health are fundamental principles in Islamic ethics. Furthermore, the findings of this study coincide with the Neurotechnology-Education section of the 10-10 MySTIE framework for Malaysia.

This research shows that EEG based emotional metrics can be used as feature for mental fatigue detection. This identifies association between emotions and mental fatigue. This research also evident EPM are the reliable emotional metrics for mental fatigue detection. The changes on emotional states are the measurement of mental fatigue. The research finds 15 minutes as a crucial mental state changes for online learner. This suggests a duration of getting fatigue for an online learner. This research will assist in empowering mental health to guard against long-term cognitive deficiencies by enhancing mental fatigue detection strategy. The early detection of mental fatigue may increase cognitive power as well as improvement on emotional responses which is crucial for an individual. This research also evident a significance on both subject-dependent and subject-independent mental fatigue detection using emotional metrics. This research leads the way for further research and development of real-time personalized or generalized mental fatigue detection and intervention model. While there is a mental fatigue detection, an intervention strategy may apply to reduce fatigue and improve academic performance for online learners. Early detection of mental fatigue for online learners may assist in improving cognitive performance.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

The purpose of the literature review is to comprehend the current research controversies and approaches regarding mental fatigue detection. Therefore, this literature review includes review on mental fatigue assessment, mental fatigue detection, Emotiv Performance Metrics, neurofeedback signal generation, and effect of light and temperature on mental fatigue.

2.2 SELF-REPORT MENTAL FATIGUE ASSESSMENT

There are many ways to measure mental fatigue states, including self-reporting and observing other people's behavior. The applied self-reporting tools outlined in Table 2 were used to determine the target labels of mental fatigue in accordance with the participants' feelings, attitudes, and/or opinions.

The Chalder Fatigue Questionnaire (CFQ) was introduced by (Chalder et al., 1993). The severity of fatigue in 444 people with multiple sclerosis was evaluated using the CFQ in (Chilcot et al., 2015). Then, (Cella & Chalder, 2010) used CFQ to assess the chronic fatigue syndrome among 361 participants. A CFQ with by modal scoring is used to measure fatigue among 1887 participants in (Jing et al., 2016). Mental fatigue was measured by (Talukdar et al., 2020b) while participants were performing motor imagery BCI task. The CFQ questionnaire has a scoring range of 0 to 3, and the total score is determined by summing the ratings for each component. The total score can vary from 0-33. (Chilcot et al., 2015) scoring using the following labels: less than normal (0), up to but not beyond usual (1), above average (2), and significantly above average (3). High scores suggest severe fatigue. To assess the validity of one, two, and bi-factor models of fatigue, the researchers utilized confirmatory factor analysis (CFA) with weighted least-squares with mean and variance adjustment estimates. In the (Cella & Chalder, 2010) study, the response options "less than usual" and "no more than usual" were rated as 0, while "more than usual" and "much more than usual" were rated as 1. To examine the outcomes, this study performed statistical analyses such as distribution plots and skewness values, the Shapiro-Wilk test, Cronbach's alpha coefficient, principal component analysis (PCA), and receiver-operating characteristic (ROC). Although, CFQ is one of the popular self-assessment tools but there are some limitations reported. Since, it relies on self-reported data, it can be prone to subjectivity. The scoring may be influenced by individual mood or cultural factors. Besides this, the two-factor structure (physical and mental fatigue) is not consistently supported, because the high correlations between factors suggesting limited distinctiveness (Cella & Chalder, 2010). The two different scoring methods (bimodal and Likert) may vary across different studies by complicating comparisons and potentially missing nuances in severity. Therefore, CFQ is a widely used tool for assessing fatigue because of its simplicity, and broad applicability across various populations and conditions

Table 2 Mental Fatigue Assessment Tools

Assessment	Applications	Number of Items	Subjects	Scoring	Evaluation	References
Chalder Fatigue	Chronic fatigue syndrome (CFS)	11	361	A score ranging 0-3	Statistical Analysis	(Cella & Chalder, 2010)

Questionnaire (CFQ)	Motor imagery BCI	11	11	Five-point Likert scale ranges from 1 to 5	Statistical Analysis	(Talukdar et al., 2020b)
	Healthy Participants fatigue	11	1887	Bimodal score where 1= “Yes”, and 0 = “No”	Statistical Analysis	(Jing et al., 2016)
	Severity in multiple sclerosis	11	444	A score ranging 0-3	Confirmatory Factor Analysis (CFA)	(Chilcot et al., 2015)
Extended Positive and Negative Affect Schedule (PANAS-X)	Impact in Parkinson’s disease	60	100	Five-point Likert scale ranges from 1 to 5	Statistical Analysis	(Schiehser et al., 2013)
Positive and Negative Affect Schedule (PANAS)	Effect on auditory temporal order judgments	20	41	Five-point Likert scale ranges from 1 to 5	Statistical Analysis	(Simon et al., 2020)
Karolinska Sleepiness Scale (KSS)	Comparing two versions of the KSS for drowsiness	1	12	Nine-point scale ranges from 1 to 9	Statistical Analysis	(Miley et al., 2016)
	Driver’s Passive fatigue	1	29	Nine-point scale ranges from 1 to 9	Statistical Analysis	(Foong et al., 2019)
Visual Analogue Scale - Fatigue (VAS-F)	Mental fatigue on task for motor imagery BCI	18	11	Five-point Likert scale ranges from 1 to 5	Statistical Analysis	(Talukdar et al., 2020b)

While (Simon et al., 2020) looked into PANAS to test the short-term cognitive fatigue effect on auditory temporal order judgments, (Schiehser et al., 2013) employed PANAS-X to measure the impact of fatigue on people with Parkinson's disease. A five-point Likert scale with a range of 1 (very slightly or not at all) to five (extremely or very much) was utilized for rating in both study questionnaires. The outcomes of the (Schiehser et al., 2013) study was evaluated using mean, standard deviation, and P-value, whereas (Simon et al., 2020) study employed PCA, mean, and P-value. Besides its widespread use it has a significant drawback as it is relying on self-reported data similar to CFQ. Additionally, it focuses on broad emotional states and may not capture the nuances of specific emotions. In this reason, need to use extended version PANAS-X, but it can increase the burden on participants. Although it is adaptable for different time frames, but its reliability may vary depending on the context or population studied. These drawbacks suggest that PANAS should be used cautiously and supplemented with other methods for a comprehensive assessment.

The Karolinska Sleepiness Scale (KSS) is another technique that was used in (Miley et al., 2016) to compare two KSS versions for drowsiness and in (Foong et al., 2019) to measure passive fatigue. Nine points were assigned to each response on a scale from 1 (extremely alert), 2 (very alert), 3 (alert), 4 (rather alert), and 5 (neither alert nor sleepy), 6 (some signs of sleepiness), 7 (sleepy, but not making any effort to stay awake), 8 (sleepy, making some effort to stay awake), and 9 (very sleepy, making a great effort to stay awake and fighting sleep). (Miley et al., 2016) used statistical analyses such as the mean, standard deviation, and P-value to evaluate the results, while (Foong et al., 2019) performed this by employing the McNemar test, Cohen's unweighted Kappa, mean, and P-value. Since KSS is a self-reported assessment tool it has similar limitations of CFQ and PANAS. However, KSS may not always align with objective measures of sleepiness (Gillberg et al., 1994; Miley et al., 2016). KSS does not describe understanding long-term sleepiness patterns or trends, only a snapshot of sleepiness at a specific moment. It relies on a single item that reduces its sensitivity variations in sleeping levels. In this reason, KSS is often paired with objective assessments for a comprehensive evaluation.

Another study (Talukdar et al., 2020b) employed the Visual Analogue Scale - fatigue (VAS-F) to assess mental fatigue and looked at the adaptive feature extraction in EEG-based motor imagery BCI. The degree of fatigue is determined using a subjective scale with a value from 1 to 5 that falls between two extremes: 1 being the least fatigued and 5 being the most. To express their level of exhaustion, subjects made a selection along a scale. The statistical analysis used in this study to assess the outcomes included mean, correlation coefficient, standard deviation, and P-value. As a self-reported tool, it has similar limitations like CFQ, PANAS and KSS, that can be influenced by factors like individual's mood, and personal interpretation of the scale. Additionally, VAS-F measures overall fatigue without distinguishing physical or mental components. This limits ability to capture the full complexity of fatigue.

2.3 MENTAL FATIGUE DETECTION THROUGH BCI

There is a systematic literature is conducted on the mental fatigue detection (Yaacob et al., 2023). The literature review done by following PRISMA (Moher et al., 2009) framework and data publications were collected from Scopus, Institute of Electrical and Electronics Engineers (IEEE) Explore, PubMed and Web of Science (WOS). The selection of databases was included because of their comprehensive coverage of interdisciplinary research. The databases chosen to ensure high-quality and relevant studies. The data collection conducted up to 13^{th} September 2022. The keyword for data extraction was Boolean search expression including "electroencephalogram and BCI", and "fatigue classification and Brain-Computer Interface". The literature search consisted of identification, screening and inclusion that depicted in the PRISMA flowchart in Figure 1. The selection of 39 articles from the identification step of 562 studies reflects the rigorous screening process. These articles offering diverse perspectives on mental fatigue detection.

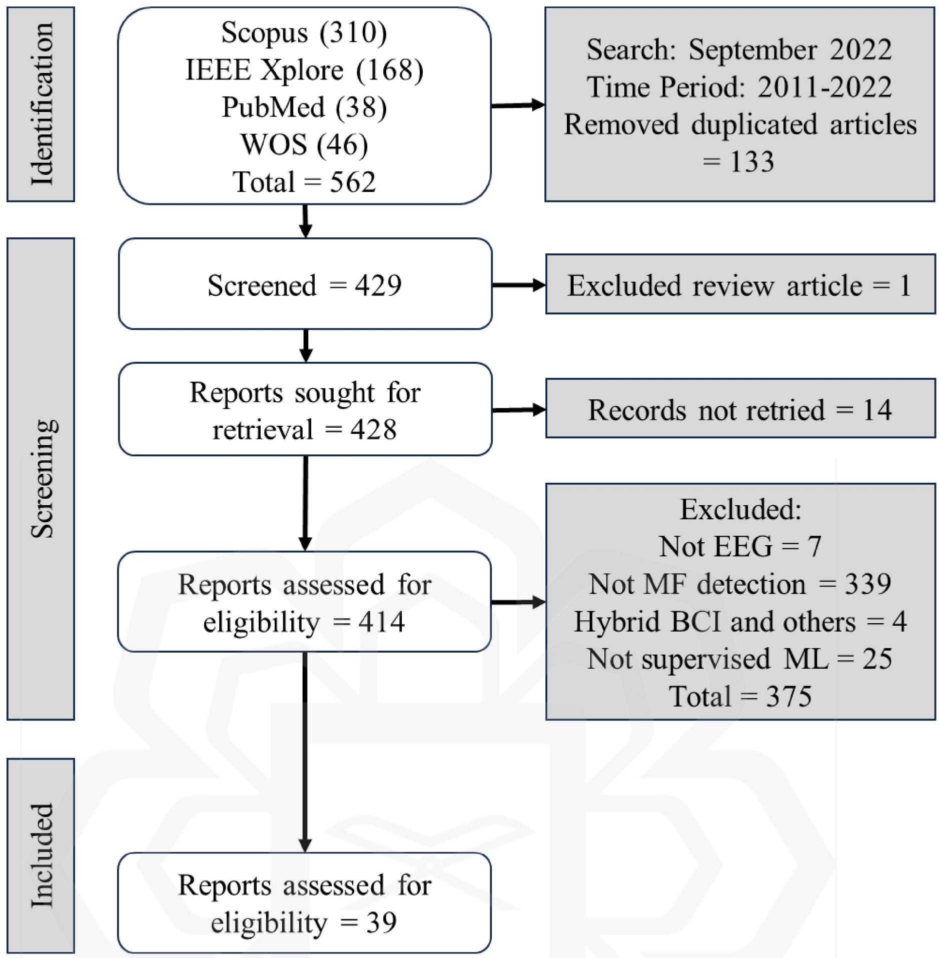


Figure 1 PRISMA Framework for Mental Fatigue Detection Review

This review provides a guideline for the implementation of mental fatigue detection through EEG based BCI. EEG signals are non-invasive brain waves. Five EEG signal frequency bands are used for analysis: delta (0.3–4 Hz), theta (0.8–13 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (> 30 Hz) (Subha et al., 2008). Measuring mental fatigue involves tracking changes in the EEG spectrum. The parietal and central areas of the brain were seen to be more active when higher frequency activity decreased and alpha and theta band activity increased (Craig et al., 2012; Klimesch, 1999; Lal & Craig, 2002; Paus et al., 1997; Tanaka et al., 2012). The steps for mental fatigue detection have been addressed in this review: signal acquisition, pre-processing, feature extraction, and classification/pattern recognition. These four steps are known as the Basic Functions of BCI that allows users to interact without the involvement of muscles or peripheral nerves (Wolpaw et al., 2002). Therefore, Table 3 summarizes the details of EEG-based mental fatigue detection.

2.3.1 Signal Acquisition

The analysis of brain activity can be done using a variety of invasive, non-invasive, and semi-invasive signals (Mudgal et al., 2020). The technique of collecting and analyzing brain signals (EEG, fMRI, fNIRS, PET, and other) is quite complicated because of their high sensitivity nature (Fernández Lucas et al., 2010; Lippé et al., 2009; Schinkel et al., 2009; Takahashi, 2013). However, it has been shown that using existing brain signals is a reliable way to study brain activity. SPECT (BRUYANT, 2002; ELHENDY et al., 2002) and fMRI (Perer A. Rinck, 2014) are used to track changes in blood flow, whereas PET (Bailey et al., 2005) and NIRS (Cui et al., 2011; Ferrari & Quaresima, 2012) are used to track changes in metabolic activity. The electrical activity of the brain is directly observed by EEG (Lesser, 1986; Niedermeyer & da Silva, 2005) . Table 3 presents EEG devices used in EEG signal collection for mental fatigue detection.

Table 3 Mental Fatigue Detection

Mental fatigue Domain	Signal Acquisition	Pre-processing	Feature Extraction	Classification / Pattern	References
Driver's drowsiness estimation	Subjects: 37 Device: Scan SynAmps2	Bandpass filter	Convolutional Neural	Deep Q-Network (DQN)	(Ming et al., 2021)

			Network (CNN)		
Student Fatigue Detection	Subjects: 11 Device: g.USBamp	Notch filter	FFT, Daubechies wavelet	SVM	(Peng et al., 2021)
Driver Fatigue State Classification	Subjects: 8 Device: Neuroscan	Bandpass filter	Differential Entropy (DE)	KNN, SVM, DNN	(Hwang et al., 2021)
Driver's Drowsiness Detection	Subjects: 27 Device: Scan SynAmps2	Bandpass filter, Automatic Artifact Removal (AAR)	Attention Metric, PSD	CNN	(Paulo et al., 2021)
Driver fatigue detection	Subjects: 11 Device: NeuroScan	FIR, AAR	STFT, PSD	LSTM	(Kuang & Qu, 2021)
Driver fatigue detection	Subjects: 6 Device: BrainAmp	Wavelet, Gaussian	STFT, PSD, EMD	SVM, KNN, PSO-HELM	(C. Zeng et al., 2022)
Student's fatigue detection	Subjects: 8 Device: Six-electrodes	Butterworth filter	FFT, PSD, Hjorth	CNN	(K. Chen et al., 2021)
Driver fatigue detection	Subjects: 23 Device: NeuroScan	Butterworth filter	Logarithmic Energy (LE)	Bayesian Linear Regression, Gaussian Process (GP)	(Tabejamaat & Mohammadzade, 2022)
Driver Mental States Classification	Subjects: 10 Device: gUSBamp	Butterworth bandpass filter, ICA	CSP	LightFD	(H. Zeng et al., 2019)
Detection of Changes in Mental State	Subjects: 11 Device: B-Alert X24	FIR Bandpass filter, ICA	FFT, PSD	SVM, LDA, NB	(Myrden & Chau, 2017)

Assessment of Mental Fatigue	Subjects: 6 Device: NeuroScan NuAmps	IIR Bandpass filter	FFT, PSD	FL-RSEFNN, SVR, SONFIN and TRFN	(Y. T. Liu, Lin, Wu, Hsieh, et al., 2016)
Predicting Driving Fatigue	Subjects: 20 Device: NeuroScan NuAmps	IIR Bandpass filter	FFT, PSD	RSEFNN, SVR, ONFIN, FWNN, TRFN	(Y. T. Liu, Lin, Wu, Chuang, et al., 2016)
Decoding of Pilots' Mental States	Subjects: 7 Device: BrainAmp	Butterworth bandpass filter, ICA	FFT, PSD	MFB-CNN	(Lee et al., 2020a)
Driver Vigilance	Subjects: 10 Device: Homemade wireless	Wavelet	FFT, PSD	SVM	(X. Zhang et al., 2017)
Detection of mental fatigue	Subjects: 20 Device: BrainAmpTM	Butterworth bandpass filter, SOBI	CSP	FLDA	(Roy et al., 2014)
Driving fatigue prediction	Subjects: 10 Device: Ag / AgCl active electrode	IIR Bandpass filter, ICA	FFT, PSD	RSEFNN, GD	(Y. T. Liu, Wu, et al., 2016)
Driver Fatigue Evaluation	Subjects: 8 Device: Neuroscan	Butterworth bandpass filter	Spatio–temporal feature	ESTCNN	(Gao et al., 2019)
Mental Fatigue Detection	Subjects: 7 Device: Neurosky Mindwave	SavitzkyGolay (SG) filter	Power Percentage (PP)	LDA, QDA, SVM-P, SVM-RBF	(Hendrawan et al., 2018)
Mental Fatigue Estimation	Subjects: 18 Device: Neuroscan SymAmp2	FIR Bandpass filter	PSD, Welch's periodogram method	MLR	(Tian et al., 2018)
Real-time fatigue level detection	Subjects: 15 Device: U-Wake	Butterworth bandpass filter	FFT, Hamming window, power spectrum array	LR	(Ko et al., 2015)

Mental fatigue for motor-imagery BCI	Subjects: 11 Device: Biosemi	EAWICA, Low pass	CSP	LDA	(Talukdar et al., 2020a)
Tracking mental fatigue	Subjects: 11 Device: Biosemi	EAWICA, Low pass	ADCSP	KPLS	(Talukdar et al., 2020b)
Quantifying passive fatigue	Subjects: 29 Device: Muse	Butterworth bandpass filter, Notch	FFT, PSD	Negative-Unlabeled (NU)	(Foong et al., 2019)
Driver's fatigue detection	Subjects: 109 Device: BCI2000	Butterworth bandpass filter	FFT, PSD	Wave ratio (slow to fast) / threshold	(Hossan et al., 2017)
Bus Drivers Fatigue Measurement	Subjects: 20 Device: Neurosky Mindwave	ThinkGear	ThinkGear	eSense	(C. L. Chen et al., 2017)
Detection of drowsiness	Subjects: 20 Device: Ag/AgCl dry electrode el120	Butterworth bandpass filter, Notch	Alpha, Theta	Threshold	(Pathak & Jayanthy, 2017)

The initial phase in BCI architecture is EEG signal acquisition. The placement of the electrodes and the frequency bands are the two key factors that control the acquisition of the EEG signal. Electrodes must be positioned over or inside the brain in order to record brain impulses. Electrodes produce brain waves. The EEG data were collected using the 10-20 electrode placement method (Homan, 2015; Myrden & Chau, 2017; H. Zeng et al., 2019). Different environmental variables were created for different studies on the generation of feedback signals (Table 20 & Table 21) and the detection of mental fatigue (Table 3). EEG signal acquisition involves different EEG headset. This headset has its own efficiency and drawback. The popular EEG headset in mental fatigue detection is depicted in Figure 2. The most commonly used device is NeuroScan, because of its reliability, signal quality, and broad application in research. BrainAmp and gUSBamp are popular but cater to slightly different needs like scalability and portability. Muse is favored in consumer and neurofeedback research because of its accessibility and ease of use. Then, Biosemi and U-Wake are also specialized but less commonly used, may be due to cost or the specific research needs they address. Other devices, such Neurosky, Ag/AgCl active electrode, homemade wireless, B-Alert X24, and six-electrodes, were used less frequently. The use of different EEG devices in different studies may influenced by several factors including cost, ease of use, portability, signal quality, and specific research needs.

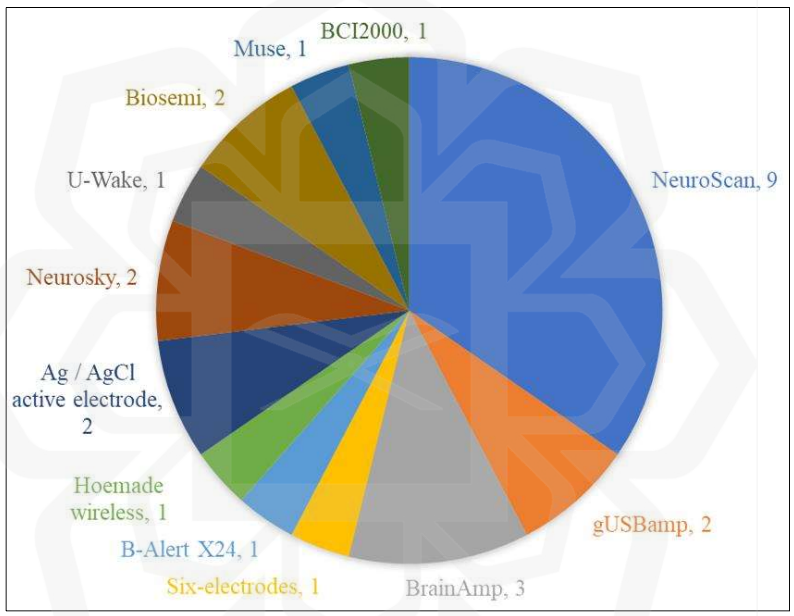


Figure 2 Popular Neuroheadset Headsets for EEG Signal Acquisition

2.3.2 Signal Pre-Processing

EEG signals are vulnerable to artefacts, which are noise. The sources of the artefacts are the amplifier, power line, Electrocardiogram (ECG), Electrooculogram (EOG), and Electromyography (EMG) as well as weak electrode contacts with the scalp (Bi et al., 2013;

Lakshmi et al., 2014; Subha et al., 2008). Artefacts must be removed in order to move forward since they make it difficult to identify mental states. The identified signal pre-processing techniques are presented in Table 3. Figure 3 depicted the widely used pre-processing technique identified by this research.

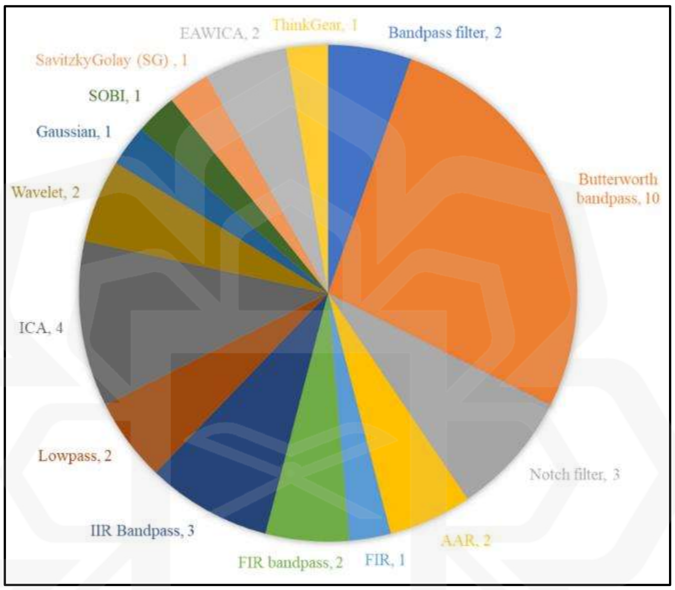


Figure 3 EEG Pre-Processing Techniques

The popularity of EEG processing techniques is often driven by the ability to address specific challenges in EEG signal quality, artifact removal, and feature extraction. The Butterworth filter is the most widely used due to its smooth frequency response and non-oscillatory behavior. It is effective in removing noise while preserving signal integrity. This makes it especially useful for bandpass filtering. Therefore, fixed frequency bands may not well adapt in dynamic EEG signals which vary across individuals or states. Butterworth filter can introduce phase shifts in the signal, and affecting time-sensitive analysis. However, the passband frequency response is made as flat as possible with the help of this filter. Simple solutions for electrical grounding noise through digital filters includes bandpass filters, low pass filters, and high pass filters (Suto & Oniga, 2018). Lowpass filter is used to remove high-frequency noise from EEG signal data such as electromagnetic interference.

Frequency-specific artefacts can be eliminated using a notch filter, finite impulse response (FIR), or infinite impulse response (IIR). IIR filters are commonly used for bandpass filtering because of efficiency in the computational resources. It is effective for separating desired frequencies from noise or artifacts, which can introduce phase distortion. Big data quantities are successfully handled using Independent Component Analysis (ICA) to determine good calculation efficiency (Korats et al., 2013; Lakshmi et al., 2014). ICA is powerful blind source separation technique. ICA is computationally expensive and can be slow in real-time applications. It is very effective in removing artifacts from EEG signals like eye blinks or muscle activity. It separates the observed mixed signals into statistically independent components. This is crucial technique for cleaning EEG data for further analysis.

The wavelet transformations can be used to decompose the time series into a collection of components with various frequency bands (C. Zhang et al., 2014). The wavelet coefficients are thresholder to decrease noise using Stein's unbiased risk estimate (SURE) threshold (Geetha & Geethalakshmi, 2011). The most precise approximation of states is provided by the Kalman filter, which is used to combine data from multiple sensors while noise is present (Aznan & Yang, 2013). To reject noise and suppress outside disturbances, the Signal-Space Projection (SSP) technique can be applied (Uusitalo & Ilmoniemi, 1997). A noisy signal with a broad frequency range is frequently smoothed using the Savitzky-Golay algorithm. The eye blink and muscle activity artefacts can be eliminated with EAWICA. EAWICA is more computationally intensive than standard ICA because of the additional entropy-based weighting. The exponential moving average (EMA) filter is an infinite impulse response (IIR) discrete low-pass filter that exponentially discounts past data to give more weight to current data(Binti Mustaffa et al., 2018; Durantin et al., 2014). IIR filters can be unstable especially when it implemented in a digital system. EEG signal preprocessing involves trade-offs between artifact removal and signal preservation, that can impact downstream analysis. Effective EEG signal preprocessing requires expertise in selecting and configuring suitable methods based on study objectives.

2.3.3 Feature Extraction

The pattern of brain activity is represented by features. The feature extraction process requires the signals to be formatted in a way that classification algorithms can fully characterize the hypothesis being represented. It is challenging to extract features from EEG signals because they are dynamic, volatile, non-linear, and non-stationary. Table 3 provide an overview of the approaches used in this review. The popular feature extraction techniques are visually presented in Figure 4.

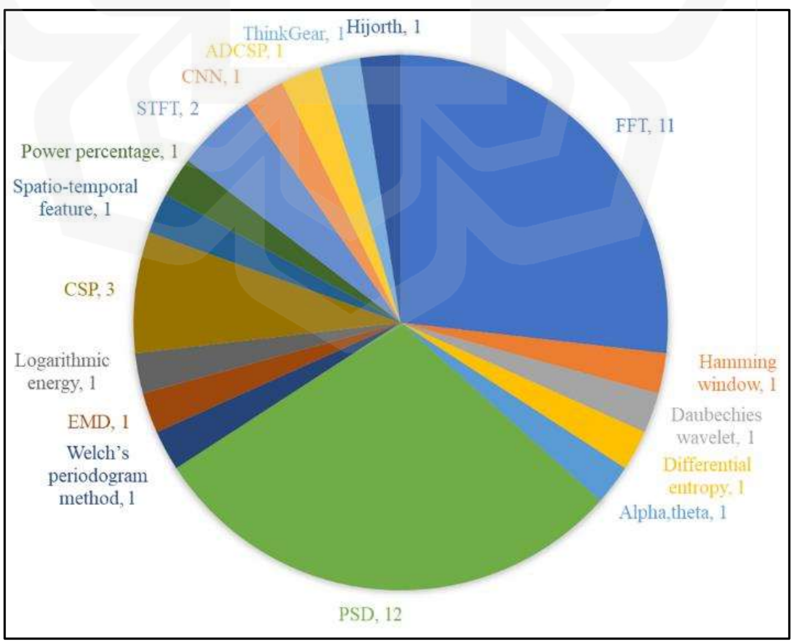


Figure 4 Feature Extraction Techniques

The most used feature extraction technique is Fast Fourier Transform (FFT). The signals are converted from the time domain to the frequency domain in order to retrieve their features. FFT is often referred to as the Discrete Fourier Transformation (DFT). It either does spectrum analysis or calculates the power spectral density (PSD) (Al-Fahoum & Al-Fraihat, 2014; Kołodziej et al., 2012). Then, Alpha, Theta, and Alpha Ratio properties can be derived as EEG bands using FFT (Ansari et al., 2019). PSD and FFT provide limited temporal resolution, it is a challenge for studying dynamic brain processes that change rapidly over time. EEG sensorimotor rhythms (SMR, 12–15 Hz) are another EEG feature that can be used in conjunction with neurofeedback to enhance cognitive performance, particularly in the areas of memory and attention (Campos da Paz et al., 2018). The average surprise of a random variable is represented by the differential entropy (DE) in continuous probability distributions. Spatial and temporal properties can be obtained using independent residual analysis (IRA) and common spatial patterns (CSP) (Q. Zhao & Zhang, 2007). CSP works well in a small dataset or in limited electrodes, so it might struggle in scalability when handling a large dataset or a large number of electrodes. Then, beamformer (Wittevrongel & Van Hulle, 2017), welch's (L. Zhao & He, 2013), and wavelet transforms (Akin, 2002) techniques are also used for feature extraction. STFT and Wavelet transforms need significant computational power for real-time processing. This is a limit of their use in resource-constrained environments or in an application that require low-latency processing.

2.3.4 Classification / Pattern Recognition

The classification or pattern recognition or threshold analysis of EEG data helps to identify the different mental conditions. Table 3 demonstrated that the state of the art in terms of classification algorithms is extremely diverse because there is no classification approach that dominates among the EEG-based mental fatigue detection. However, the appropriate classification technique should be selected based on the particular features of the dataset.

The are model from different domain like statistics, ML and deep learning. SVM is the most popular model applied in mental fatigue detection with different kernels. SVM is effective in handling high-dimensional data and capable of separating complex, non-linear patterns by using different kernels. The deep learning algorithms like CNN, SONFIN, DNN and other are used to detect mental fatigue fatigued. Besides these the statical models like LR and threshold analysis also used to identify mental fatigued. The EEG based mental fatigue model are depicted in Figure 5.

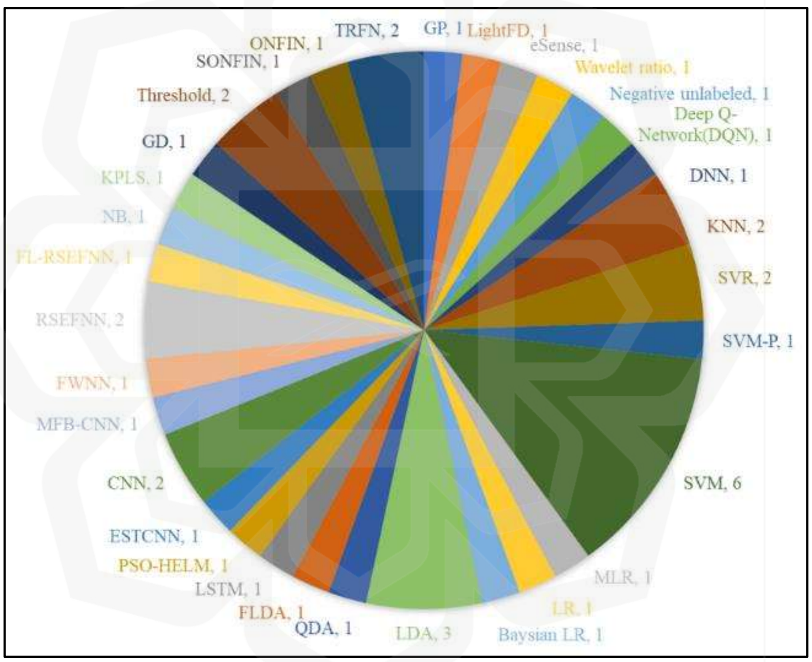


Figure 5 EEG Based Mental Fatigue Detection Model

Two states of mental exhaustion are used in the majority of the works. In order to illustrate the change from the normal condition to the fatigued state, several works take the potential of intermediary stages into consideration (Lee et al., 2020b; H. Zeng et al., 2019).

SVM classifiers with practical applications sort input items into several categories. It is capable of performing various linear and non-linear classifications by simply changing its kernel function. When used to address function approximation and regression estimation issues, SVM is referred to as support vector regression (SVR) (Y. T. Liu, Lin, Wu, Chuang, et al., 2016; Y. T. Liu, Lin, Wu, Hsieh, et al., 2016). In other words, it is useful for continuous mental fatigue estimation rather than binary classification and handles non-linear relationships between features and fatigue levels. SVM struggles in large datasets because it requires high computational cost during training. Additionally, it requires careful tuning of hyperparameters such as kernel and regularization. A convolutional neural network (CNN) based deep learning system with temporal-spatial EEG filters utilized to distinguish four mental states (Y. T. Liu, Lin, Wu, Chuang, et al., 2016). (Hossan et al., 2017) utilized the ratio of slow waves to fast waves to identify drowsiness. CNN automatically detects spatial patterns in EEG signals. CNN variants like MFB-CNN and ESTCNN are designed to handle effectively EEG signal-specific features. However, it requires high-performance GPU for training and large amount of labeled data to prevent underfitting.

K-Nearest Neighbors (KNN) (Murphy et al., 2009) is easy to implement and does not require extensive training. It works well for small datasets and quickly identifies patterns based on proximity. KNN requires storing entire dataset and may be slow during prediction in using large datasets. Performance depends on scaling and selection of features. Deep Q-Network (DQN), (Ming et al., 2021) automatically learns complex features from EEG signals without any manual feature engineering. It performs good with large datasets and diverse feature spaces, but large dataset may not always available. In this reason it is prone to overfitting with small or imbalanced datasets. Long Short-Term Memory (LSTM) (Matic et al., 2019) specifically designed for sequential data model. It is ideal for capturing temporal patterns in EEG signals. LSTM retains information over long periods which is critical for fatigue detection. LSTM models need significant time to train and the model can overfit on small EEG datasets.

Recursive Self-Evolving Fuzzy Neural Network (RSEFNN) (Y. T. Liu, Lin, Wu, Chuang, et al., 2016) is suitable for identifying dynamic EEG patters because of its complex structure and parameters during training. It incorporates uncertainty handling that is crucial for noisy EEG data. The self-evolving structure of RSEFNN requires significant computational resources. Linear Discriminant Analysis (LDA) (Talukdar et al., 2020a) simplifies high-dimensional EEG signals into a lower-dimensional space by maximizing class separability. LDA is suitable for real-time applications because it is fast and computationally light. However, LDA may significantly impacted by noise and outliers in the EEG signals. Since LDA assumes linear boundaries between classes that might not identify complex EEG patterns.

There are many more algorithms available based on needs for classification or pattern recognition like Artificial Neural Network (ANN) (Grossi et al., 2017; Ventouras et al., 2005), Naive Bayes (Katkar & Kulkarni, 2013; Mehmood et al., 2017), LightGBM (H. Zeng et al., 2019), eSense (Srimaharaj et al., 2021), Multiple Linear Regression (MLR) (Tian et al., 2018) and many more listed in Table 3 and Table 20. Literature review did not identify any specific techniques that has a certain advantage, but the best classification result can be achieved by applying the appropriate classification model for the given sample of data and research goal. The widespread application of distinct classification algorithms is the current development trend.

2.4 EMOTIV PERFORMANCE METRICS (EPM) FOR MENTAL STATE DETECTION

There are many studies available for EEG based emotional state detection indexed by (Al-Nafjan et al., 2017; Suhaimi et al., 2020; Zangeneh Soroush et al., 2017). There are many EEG device (Soufineyestani et al., 2020) available for EEG signal acquisition for different mental condition detection including emotional state. But No device provided API or system to use emotional metrics except Emotiv. Numerous research employed the Emotiv headset and its six-performance metrics: engagement, excitement, focus, stress, relaxation, and interest. The Emotiv Performance Metrics (EPM) showed reliable results in the detection of mental disorders. Some of these investigations are summarized in Table 4. Most research assessed all the 6 EPMs to determine the level of mental disorder. This entails the decoding of spoken language from the human brain straight into a digital display (Faruk et al., 2021). This research used Emotiv Epoc X for data collection from three participants. The headset needs to maintain contact quality of 98% and it is necessary for EPM readings. This research reported missing EPM data because of poor contact quality. This research used statistical analysis included correlations between performance metrics and histogram analysis of experimental datasets and external datasets. This study reported 69% and 62% improved accuracy respectively for the Naïve Bayes and Linera Regression classifiers than their first research attempt. These levels are not optimal for clinical applications. The findings suggest room for improvement in EPM analysis methods.

There is a study (Santoyo-Mora et al., 2022) merely examines 3 EPMs (focus, interest, & engagement) to identify a person's mental state. This study utilized Emotiv Epoc X to collect data from total of 147 participants. The simple reaction time (SRT) and psychophysics tests of Two-Alternative Forced Choice (2AFC) used for the cognitive process evaluation including visual attention, decision-making and information processing speed. The statistical analysis includes mean, median, standard deviation, z-value, and p-value calculations on response time and alternative-forced choice on recovered patients, EPM is utilized to evaluate the long-term effects of COVID-19. This research suggested a follow-up for COVID-19 patients because there is a significant reduction in the performance cognitive process abilities. There is a report of 42.42 percent reduction in the mild-moderate patients where 46.15 percent in the severe-critical patients for information processing speed. There was a 46.15 percent less performance on visual sensitivity for decision making task. The uneven distribution among groups limits the generalizability of findings to broader populations. Participants were excluded if less than 80% signal quality, else can cause bias results by removing data from the potentially affected individuals.

A study analyzed the reliability of 6 EPM metrics acquired using Emotiv Epoc X from 14 participants (Paranthaman et al., 2021). Statistical tools like mean, standard deviation, and p-value (t-test) used to evaluate the dependability of EPM in a virtual reality game. There was a significance between the actual player's experience and EPM. The adhoc linear model directly estimates player's emotional states from raw EEG. This research also reported various personalized brain activity maps to shows association between emotions and brain activity. Wearing the VR headset alongside the Emotiv headset introduced additional challenges such as device interference that affected EPM data accuracy. There was a poor EEG signal quality that resulted in missing EPM values during the experiment. The computed EPM did not align well with player ratings (PR) that indicating discrepancies in the EPM and PR assessment.

Table 4 Emotiv Performance Metrics (EPM) in Mental State Detection

Applications	Emotiv Headsets	Subjects	EPM	References
Decode human speech directly	Epoc X	3	Focus, interest, excitement, engagement, stress, relaxation	(Faruk et al., 2021)
COVID-19 Long-Term Effects	Epoc X	147	Focus, interest, & engagement	(Santoyo-Mora et al., 2022)
Reliability of performance metrics	Epoc X	14	Focus, interest, excitement, engagement, stress, relaxation	(Paranthaman et al., 2021)
Emotional responses to the visual patterns of urban street	Epoc X	26	Focus, interest, excitement, engagement, stress, relaxation	(Z. Zhang et al., 2021)
Compare level of executive functions	Insight	60	Focus, interest, excitement, engagement, stress, relaxation	(Asif et al., 2023)
Determine the presence of the emotion and EPM	Epoc X	10	Focus, interest, excitement, engagement, stress, relaxation	(Holman & Adebesin, 2019)
Analysis of Performance Metrics	Epoc X	3	Focus, interest, excitement, engagement, stress, relaxation	(Strmiska & Koudelkova, 2018)

Emotional response was analyzed using EPM while participants were experiencing virtual reality environment of various urban street scenes (Z. Zhang et al., 2021). The data collection involved Emotiv Epoc X from 26 participants. The examination of EPMs using statistical methods, such as the mean, SD, correlation coefficients among metrics, and linear regression with standardized coefficients as statistical analysis. This research found that less fragmented but high diverse scenes stimulated positive emotional response. The device reliability is affected by factors such as discomfort of wearing the headset. The EEG data quality compromised due to poor electrode contact particularly with hair that introduce noise into the EEG measurements. The study evident inconsistencies between EPM measure and subjective measure of emotional responses. This concerns the validity of EPM results with participants subjective experiences.

Another study examined the EPM and compared the levels of executive function in gamified and non-gamified activities for 60 participants using statistical methods such as the Kolmogorov-Smirnov test, sample Wilcoxon signed test, t-test, mean, standard deviation, and p-value (Asif et al., 2023). Emotiv Insight used to collect data when participants were performing train making test (gamified task) and scrabble (non-gamified task) for around 1-2 minutes. The study reported that there was no significance in the stress, excitement, relaxation, focus and focus, but interest metrics showed significance for the participants who performed both cognitive tasks. This study reported that the Emotiv headset does not fit for all participants properly and may results in reduced connectivity and signal quality. The headset requires sensors to be placed directly on the skin and the use of conductive gel or saline solution to enhance EEG signal quality. EPM are computed in general percentages that may oversimplify complex brain activities and emotional states. The EPM did not show significances between gamified and non-gamified tasks except for "interest", and this raising concerns about EPM sensitivity to task variations.

Through monitoring brain activity for 10 participants in (Holman & Adebesin, 2019), the EPM can also be used to assess whether anger is present or absent. This research evaluated the overall user experience in information system by analyzing EPM and emotion of anger from raw EEG. The data collection involved Emotiv Epoc X while participants reading online news article for 20 minutes. This research fount the Emotiv headset as reliable device on the collection of raw EEG data and EPM for evaluation user experience objectively. The Emotiv headset was uncomfortable for the participants as well as not perfect fit for all participants. During the experiment session there was electrical interference and causes loss of EPM data. Besides this long and think hair may causes noise or less contact between skin and electrodes. There were fluctuations in the raw EEG data when participants involved in task. Some of the EPM metrics like stress, engagement, and relaxation lacked consistency. The presence of anger emotional response was identified during the task session.

Another research examined the six EPM to show the efficiency of BCI (Strmiska & Koudelkova, 2018). The data collection performed using Emotiv Epoc X headset from three participants while participants are performing cognitive task including math exam, rehabilitation and football match. This research performed statistical analysis using the Emotiv pro default data visualization system. Similar to other study, this research also reported electrical interference, contact quality and discomfortable fell on the use of headset. Then, stress and relaxation metrics showed delayed responses in some cases that may cause potential limitations in real-time monitoring accuracy. This research identified that the Emotiv EPOC and the system takes some time on preparation and connection, but Emotiv shows reliable result on mental condition detection while person doing math exam, rehabilitation exercise and watching football match.

In summary, EPM is widely used in different mental condition analysis. There are different Emotiv device available and shows reliability in the use of EPM on different settings. The analysis of EPM includes different statical methods and machine learning algorithms to detect mental conditions. Therefore, further robust analysis is required on EPM in different settings including mental fatigue detection.

2.5 SUMMARY

It is evident that researchers across a range of disciplines utilize a variety of scales and questionnaires to evaluate and quantify fatigue in a range of populations and medical situations. The reviewed questionnaires, including the CFQ, PANAS, PANAS-X, KSS, and VAS-F, that provide arbitrary measurements of the level of fatigue. As statistical methods, researchers analyze the results of fatigue assessments using confirmatory factor analysis, principal component analysis, mean, standard deviation, and P-value. This research used a bimodal scoring based CFQ self-assessment tool, because it is one of the widely used and established questionaries across various populations and conditions that offers subjective measures of both physical and mental fatigue. This means the CFQ measure general fatigue of an individuals. CFQ suitable for practicing large-scale studies and routine clinical assessments because of its time efficiency. The CFQ may work well as a pre- and post-assessment tool to measure fatigue level before and after learning sessions to monitor changes over time.

This review showed that signal acquisition, pre-processing, feature extraction and classification are the procedure for EEG based mental fatigue detection in different settings. This literature review also identified various methods for each procedure. Table 3 Showed there are many EEG signal acquisition device available. Then variety of signal processing and feature extraction methods are used, where Butterworth bandpass filter, ICA and notch are the poplar signal pre-processing technique despite their limitations. FFT and PSD are dominating as feature extraction technique because of the reliability in capturing frequency domain features. There are many ML models used as classification or detection model. Therefore, the neural network-based algorithms, MLP, LR and SVM with its different kernel are the popular methods and showed reliable results on the mental fatigue detection that earlier presented in Table 3. However, there are some limitations such as scalability, computational demands, and interpretability challenges persist. This suggest that it is very important to choose appropriate pre-processing, feature extraction and classification techniques for mental fatigue detection. Emerging classification or detection technique like ensemble learning may address the drawbacks by integrating multiple algorithms for improved accuracy and robustness. This research employed LR, SVM and MLP as the base models for the Ensemble Learning technique. The model performance evaluation includes the accuracy, F1-score, ROC AUC score and cross-validation. The mental fatigue model developed for both subject-dependent and subject-independent analysis.

The review identified that EEG-based EPMs are used to examine mental health and cognitive capabilities in different settings. Although, Emotiv headset has some limitations like electrode contact quality with skin in different scalp size or thick and long hair or discomfort while wearing headset, but this headset also provides good quality of EPM data in different settings. The limitation may overcome using different additional techniques like use of saline solution or conductive gen for better contact quality and consistent EPM values. Then, may close all the Bluetooth device and wifi connection to reduce connection interference between headset and the system. Researchers analyze data using statical methods like mean, standard deviation, t-test, p-value, and correlation coefficients to determine the reliability and efficiency of EPMs in various contexts. The experiments showed how effective and useful EPMs may be in comprehending and assessing human mental conditions. There is no study used EPM for mental fatigue detection. This study used five EPM includes stress, relaxation, engagement, excitement and interest to detect mental fatigue. The five EPM are the emotional metrics can get from the Emotiv API. A descriptive analysis and ML algorithms used to assess the EPM for mental fatigue detection. The descriptive analysis includes mean, median, standard deviation, 25th percentiles and 75th percentiles. Then a correlation analysis and Welch's T-test performed to understand the association and significance between the fatigue and non-fatigue session.

CHAPTER THREE

METHODOLOGY

3.1 INTRODUCTION

The research methodology includes systematic literature review, experimental design, data collection, mental fatigue detection, and evaluation of the mental fatigue detection model. After that a description about the collected data includes participants details, dataset details and data labelling. The research methodology illustrated in the Figure 6.

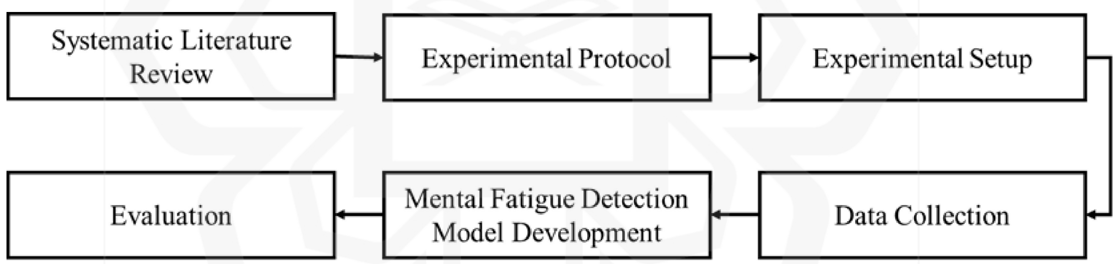


Figure 6 Research Methodology

3.2 SYSTEMATIC LITERATURE REVIEW

A literature has performed to understand the mental fatigue detection using EEG signals and identify the factors related to mental fatigue to design the intervention framework. The literature review includes neurofeedback signal generation, mental fatigue detection, mental fatigue assessment, Emotiv performance metrics and effect of light and temperature on mental fatigue. This research aims to use available emotional metrics and Emotiv provide six performance metrics which are measure of six emotional states know as emotional metrics. The findings from literature review illustrated in the previous chapter two as entitled literature review.

3.3 EXPERIMENTAL PROTOCOL

Figure 7 shows the experimental procedure. The subject initially receives a briefing about the entire session. The participant fills out the pre-task CFQ questionnaire form, the demographic information form, and the online consent form before signing it. Then, the room temperature, screen brightness and contrast were setup by following participants preferred level. The room temperate pre-setup at 25°C and the monitor brightness and contrast pre-setup at 50 percent for adjusting by participants, but nobody changes these settings. They preferred the presetting level of temperature and monitor contrast.

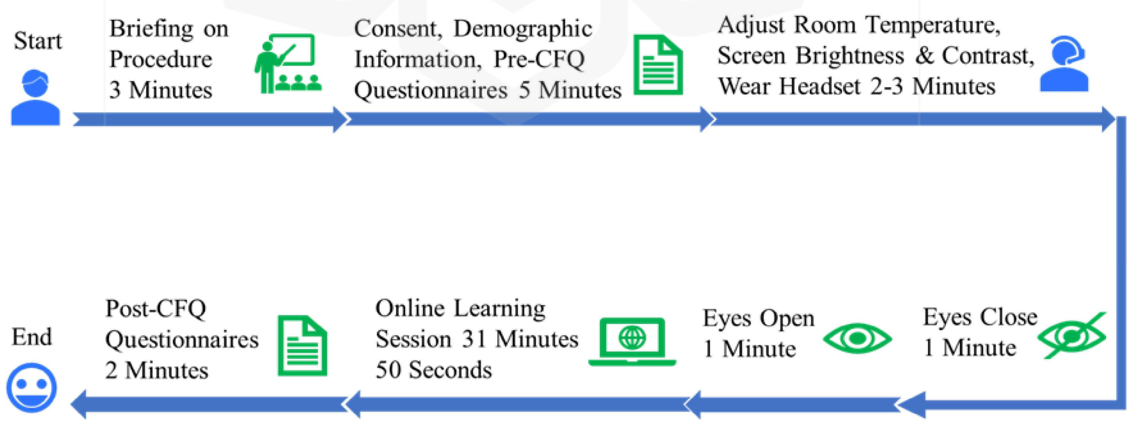


Figure 7 Experimental Protocol

Participants wear Emotiv Insight headset. Participants spends one minute with their eyes closed and another one minutes with their eyes open to get them ready for the online learning session. Then they saw a tutorial video on the display. The system collects EPM data while participants watching video. Participants completed a post-CFQ question form after the session.

3.4 EXPERIMENTAL SETUP

Figure 8 describe the experimental setup. The subject needs to wear Emotiv Insight. That is connected to laptop/desktop through Bluetooth dongle. The headset needs to maintain contact quality to avoid null or missing EPM data because of poor contact quality. The electrode placement is affected by the hair which cause noise into the EEG measurements. So, the headset needs to direct contact on the skin. Alcoholic pad was used to clean skin pointed for electrode placement. Then saline water used to maintain good contact between skin and electrodes.

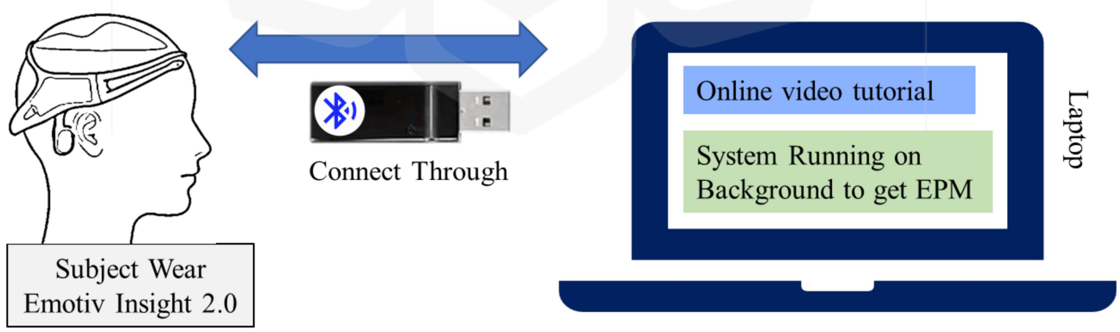


Figure 8 Experimental Setup

A long-wired USB cable was used to connect Bluetooth dongle to avoid device interference. The dongle was placed as near as possible to the Emotiv headset. To eliminate the Emotiv headset connection interference, the Wi-Fi connection was turned off. The internet connection was wired. Then there was no other Bluetooth device. The room setup was soundproof and temperature setup to 25 in air conditioner (preferred by participants). The screen brightness and contrast setup at the preferred level by the participants. Then, the tutorial video is presented. At the background the system is running to collect EPM data.

3.5 DATA COLLECTION

The data collection is one of challenging and complicated steps in the brain signal analysis. The process includes selection of participants, hardware, software selection based on research goal.

3.5.1 Ethical Consideration

Several ethical issues have been taken into account in this study to guarantee the protection of participants' rights, safety, and wellbeing. Each participant in the study has to be given their informed consent after being fully informed about the study's procedures, risks, rewards, and confidentiality policies. Participants are free to leave the study at any moment and without incurring any fees. The non-invasiveness and comfort of the EEG data gathering procedures for participants has been ensured. This study takes into account individual requirements and preferences while also respecting cultural diversity. The integrity and transparency of the research were upheld at every stage. Thus, get

International Islamic University Malaysia (IIUM), Kulliyyah of Information and Communication Technology (KICT) approval for conducting data collection survey before starting the project that is presented in Appendix A. The rules and principles for ethics were adhered to, as established by the ethics committee. Furthermore, IIUM RESEARCH ETHICS COMMITTEE (IREC) also approved (shown in Appendix B) the consent form (shown in Appendix C) and personal information form (shown in Appendix D) which contains questionnaire for selection of subjects. The data collection and associated protocols were reviewed and approved by IREC, IIUM. The IREC registration number is IREC 2023-047.

3.5.2 Participants

Number of subject selections done by following existing study that is illustrated in Table 4. There is total of 26 study with total 473 subjects counted. The highest number of subjects in a particular study was 109 and lowest number of subjects was 6 in two study. The most frequent number of subjects are 11 (5 research), 20 (4 research), 10 (3 research), 8 (3 research), 7 (2 research), 6 (2 research) and other research the numbers of subjects are unique. There is a huge variance in the sample size used in the existing research work, but no study reported any standard for a use of any specific number sample size. The bigger sample size may claim more better and reliable outcomes of a research work.

It is beneficial to provide detailed demographic information about the participants includes age range, gender distribution, educational background, and other relevant characteristics of the participants. This can help to demonstrate the sample reflects on the diversity of online learners. Furthermore, the recruitment process by considering inclusion and exclusion criteria is decided based on the demographic information. Addressing potential limitations of certain demographics and the effect on the generalization of the results that shows robustness to the study.

3.5.3 Neuroheadset

Table 2 and Table 4 showed that there are variety of EEG signal acquisition devices used. This study used Emotiv Insight 2.0 to collect the EPM. The device and the integrated electrodes placement are represented in Figure 9 (image source: https://www.emotiv.com/products/insight) and Figure 10.

Figure 9 Emotiv Insight 2.0 Neuroheadset

Emotiv Insight has 5 channels including AF3, AF4, T7, T8, and Pz. Then, Common Mode Sense (CMS) and Driven Right Leg (DRL) are 2 references on left mastoid process. The sensor material is Hydrophilic semi-dry polymer. The Emotiv Insight is wireless and can be connected to Bluetooth. It supports Bluetooth 5. However, the EEG signal sampling rate in a single electrode is 128 Hz per second. Frequency response is 0.5-43 Hz with digital notch filter at 50Hz and 60 Hz. This device comes with built in digital 5th order Sinc filter. Internal Lithium Polymer battery of 480mAh is used that can run for 20 hours in a single charge. Besides Emotiv Insight need a desktop or laptop to see video.

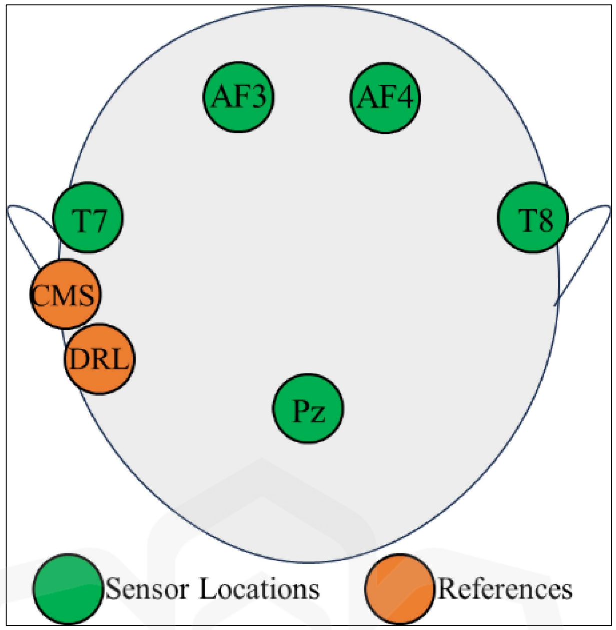


Figure 10 Emotiv Insight's Electrode Position

3.5.4 Video for Stimuli

A recorded video is selected as online learning environment. The video is long for 31 minutes 50 seconds. The video is on a topic of "Requirements Modeling in Software Development" domain. This video is chosen because this video already used for online learning among IIUM computer science students. The video is relevance to online learning and can occur fatigue because of cognitive demands like prolonged concentration on the learning of complex information. The video does not contain any music. The video sound, brightness and contrast level were setup by the participants preferred level. The video is stored and play from a laptop while the user view video from a different monitor. The video presentation monitor was a squared shape monitor and connected to a laptop. The video was integrated in a power point slide to present the participants.

3.6 MENTAL FATIGUE DETECTION MODEL DEVELOPMENT

Mental fatigue detection will be using EPM metrics. A self-reported chalder fatigue scale is used to identify a person's fatigue experience. Each dataset will be divided in training and testing sub dataset in a proportion of 80 percent and 20 percent with randomized selection process. There are ML algorithms including Logistic Regression (LR), Support Vector Machine (SVM) with polynomial kernel, Multilayer Perceptron (MLP) as base models for ensemble learning model to detect mental fatigued. The choice of base models lies in leveraging the unique strengths of their nature. The LR, SVM and MLP fall under distinct categories accordingly statistical model, machine learning, and neural networks. This ensures a diverse foundation for the ensemble learning model.

LR is excellent at identifying linear relationships in the data and gives a baseline that is easily interpretable. It may reliable for comprehending the EPM features associated with mental fatigue with its simplicity and statistical basis. SVM can better tackle the complex structures found in EPM by modeling with polynomial kernel. it is highly resistant to overfitting with high-dimensional data fields. MLP as a neural network model able to capture non-linearities and complex associations within the data by leveraging its multi-layer structure for better feature extraction. As discussed in earlier section 2.3.4 that every ML model has drawback. Similarly, LR has limitations of linear dependency, then SVM experiencing scalability issue on large dataset and MLP has limited interpretability compared to simpler models like LR or SVM.

The limitation can be mitigated using an Ensemble learning technique where LR, SVM and MLP are the base model. The robustness of SVM, the flexibility of MLP, and the statistical integrity of LR work in conjunction to improve the overall MF detection performance. This technique ensures better accurate, generalize and adaptable MF detection model.

3.6.1 Emotiv Performance Metrics (EPM)

There are six EPM emotional metrics available. Emotiv eliminate the focus metrics and added the attention metrics that was not available during data collection process. This research used stress, engagement, interest, excitement, and relaxation.

Stress: A measure of comfort with a task, stress is described as having low to moderate levels that improve productivity and greater levels that have detrimental effects on health and wellbeing.

Engagement: Engagement is characterized by heightened physiological arousal and beta waves, alertness, and conscious attention to task-relevant stimuli. The higher the effort, attention, and focus, the higher the output score the detection reports.

Interest: The level of attraction or aversion to a stimuli, activity, or environment is referred to as interest. Low interest ratings show a strong disdain for the assignment, high interest scores show a strong love for the task, and mid-range levels show that you are indifferent about the activity.

Excitement: Excitement entails pleasant physiological arousal. The output score for the detection increases as physiological arousal increases. The excitement detection is calibrated to produce output scores that capture short-changes in enthusiasm across time intervals as little as a few seconds.

Relaxation: Relaxation test measures one’s capacity to unwind and reset after periods of intense concentration.

The EPM scores ranged from 0 to 1. For instance, if the value of the engagement metrics is 0.1, the person is not engaged, whereas the value 1.0 indicates strong involvement on the part of the person (Emotiv, 2023, 2024a).

3.6.2 Mental Fatigue Detection

The EEG signal acquisition, pre-processing, feature extraction and classification model development are the primary part of EEG based mental fatigue detection. These steps are same as EEG based emotional state detection (Al-Nafjan et al., 2017; M. F. Hossain et al., 2021; Suhaimi et al., 2020). Then, Emotiv emotional metrics are including these four steps but the total process or techniques are not revealed by Emotiv. Emotiv process the EEG signal to produce the EPM which are emotional metrics. Therefore, if an EEG headset device provide the emotional metrics through an API like Emotiv, then a mental fatigue detection technique pipeline can be depicted as in Figure 11. The EEG processing is handled by Emotiv. The mental fatigue detection pipeline includes collection EPM and development of mental fatigue detection model.

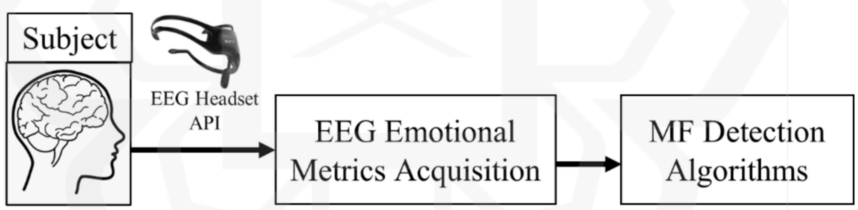


Figure 11 EEG-Derived Emotional Metrics-Based Mental Fatigue Detection Pipeline

Table 3 and Table 20 presented many Algorithms for different mental conditions detection including Mental fatigue detection. Every algorithm has its own efficiency, and limitations. However, the algorithms are shown reliable results in the detection of mental conditions. This research chooses Logistic Regression, Support Vector Machine, Multilayer Perceptron and Ensemble Learning algorithms to detect mental fatigue using Emotiv metrics. It is important to note that the mathematical term of the algorithms did not include because this research did not contribute on this. This research used the model from scikit-learn which is a free ML library for python programming language. However, important link and publications related to algorithms are included in the respective section.

3.6.2.1 Logistic Regression (LR)

A well-performing, easily comprehensible model that provides insight into the connections between features and the goal variable is logistic regression. In particular, for datasets with a moderate number of features and samples, logistic regression is computationally efficient. The LR introduced by (Cramer, 2002). This LR development based on (Pedregosa et al., 2011) guidelines. A key technique for binary classification is logistic regression, which offers comprehensible probabilities and a distinct boundary for the decision depending on parameters that have been learned. It uses a logistic (sigmoid) function to model the likelihood of the positive class as illustrated in Figure 12. By maximizing a cost function using labeled training instances, parameters are learned. Through feature engineering or interactions, this linear model could indicate non-linear relationships.

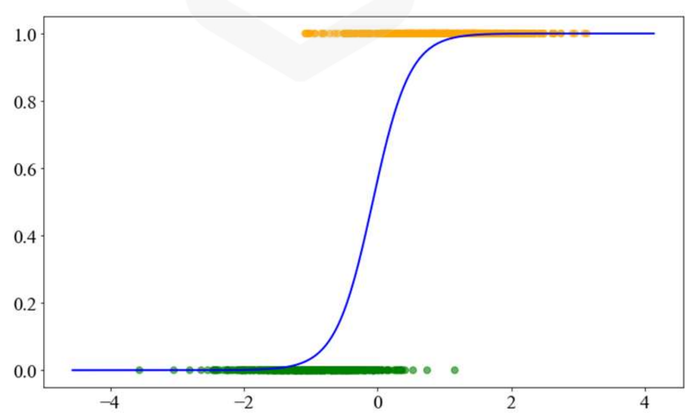


Figure 12 Logistic Function Curve

The LR parameters settings (scikit-learn library) in default with a "random_state" parameter to make the model outcomes consistent for every single run. A parameter "penalty" used for regularization which handle both the sparse and dense input. Then a default "solver" parameter utilized which support the "penalty" parameter to encounter the optimization problem. There is "max_iter" used to maximize the number of iterations for optimization. A floating "c" parameter used as inverse of regularization strength. To added an intercept in the model the default "fit_intercept" used where a "tol" parameter also used as tolerance as stopping criteria.

3.6.2.2 Support Vector Machine (SVM)

Complex patterns in the data can be captured with SVM. Support Vector Machines (SVM) can be used to address a wide variety of classification problems, including binary problem solving. SVM is an algorithm that is robust and efficient for datasets that are separable both linearly and non-linearly. Support Vector Machine (SVM) was introduced by Vladimir Vapnik and his colleagues (Boser et al., 1992; Cortes et al., 1995; Vapnik, 1997). The SVM implementation guidelines are from (Pedregosa et al., 2011). SVMs are supervised learning models that can be applied to both regression and classification problems. Figure 13 is a representation of the SVM with polynomial kernel margins and decision boundary.

The best hyperplane that optimizes the margin between the nearest data points of various classes is found by SVM. Support vectors are the closest data points. If the data is not linearly separable in its original feature space, SVM may translate data into a higher-dimensional space using a polynomial kernel function where the classes become separable. Then, within this higher-dimensional space, the SVM determines a hyperplane. The hyperplane defines the decision boundary that SVM creates. After being trained on labeled data SVM can predict their class. The side of the hyperplane on which new data points fall determines their classification.

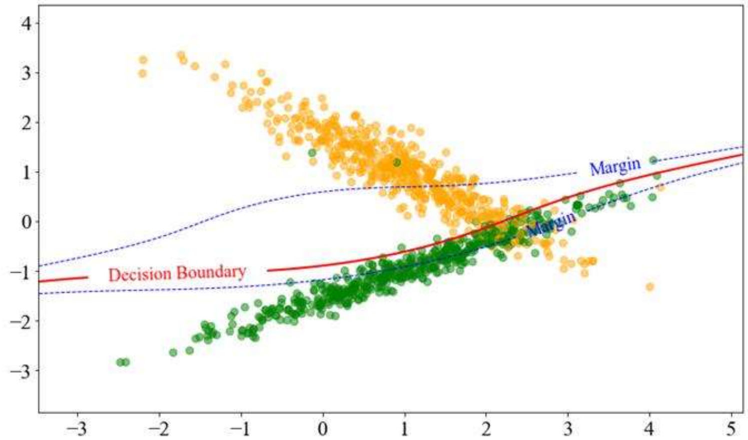


Figure 13 Support Vector Machine with Polynomial Kernel Decision Boundary & Margin

A support vector classifier (SVC) with polynomial kernel and degree of three is initialized for classification. Then the “c” and “tol” used the model with default settings as the inverse regularization strength and tolerance. A “gamma” parameter indicates impact of a single training example. The “shrinking” ensure the shrinking heuristic in the optimization of the model. Then, “coef0” defines the comparative influence of high degree and low degree polynomials. The “random_state” used to ensure the model’s reproducibility. The probability estimation enables by using the “probability”.

3.6.2.3 Multilayer Perceptron (MLP)

An artificial neural network (ANN) known as a Multilayer Perceptron (MLP) is a structure made up of several feedforward layers of nodes, or neurons. One of the most popular and adaptable deep learning architectures, it may be used to a variety of tasks, including regression, classification, and even unsupervised learning. Over several decades, numerous researchers have contributed to this MLP evolution. However, the term “perceptron” introduced by Frank Rosenblatt (Rosenblatt, 1958).

Multilayer Perceptron (MLP) use both feedforward and backpropagation. MLP processes input data and produces predictions or outputs by using feedforward. The input data is passed forward through the neural network layer by layer during training and inference. At first, every layer transforms the input linearly, then adds a non-linear activation function. This procedure is repeated until the data reaches the output layer, where the final predictions or probability estimates are produced. Following the feedforward pass, the neural network’s weights and biases are updated using the backpropagation algorithm in accordance with the differences between the true and predicted outputs. Following the feedforward pass, the neural network’s weights and biases are updated using the backpropagation algorithm in accordance with the differences between the true and predicted outputs. By calculating a loss function’s gradients according to the model’s parameters, backpropagation modifies the parameters to minimize the loss. Until convergence or a maximum number of iterations is achieved, the error is propagated backward through the network while the parameters are updated iteratively. Figure 14 illustrated MLP with single hidden layer.

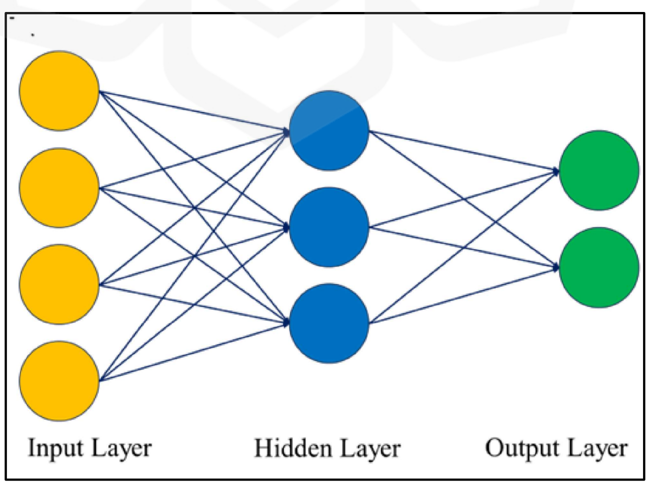


Figure 14 Multilayer Perceptron with a Hidden Layer

The development of MLP model includes scikit-learn guidelines (Pedregosa et al., 2011). The architecture of the neural networks is defined by the "hidden_layer_sizes" parameter that determines the number of neurons in a single hidden layer. The "activation" parameter used for nonlinearity in the neural networks which helps to learn more complex associations in data. The model's weight optimizations encounter by using the "solver" parameter. The reproducibility of the model ensures by the "random_state" parameter. The "alpha" is another parameter that used as the strength of the regularization term which prevent the overfitting issue by reducing the weights size. The "max_iter" indicates the maximum number of iterations for training of the neural network. The "learning_rate" is the steps size to determine the updated amounts of weights during the training. The initial leaning rate defines by the "learning_rate_init" parameter.

3.6.2.4 Ensemble Learning Model

Ensemble learning is a robust computational learning technique that combines the predictions of several different independent models to improve predictive performance. Ensemble approaches are widely used (Dong et al., 2020; Polikar, 2012) and provide substantial improvements over standalone models, particularly when dealing with complex relationships or noisy data. Ensemble learning may efficiently solve problems like overfitting, variance reduction, and improving generalization skills by taking advantage of the diversity across several models. A scikit-learn (Pedregosa et al., 2011) guidelines used to develop Ensemble learning model.

Ensemble learning model with soft voting classifier depicted in Figure 15. Ensemble learning makes use of the independence and variety among models to mitigate bias and variance. Several tactics (bagging, boosting, and stacking) can be employed for ensemble methods by considering the nature of the problem. An ensemble learning model implemented combined earlier designed LR, SVM, and MLP algorithms.

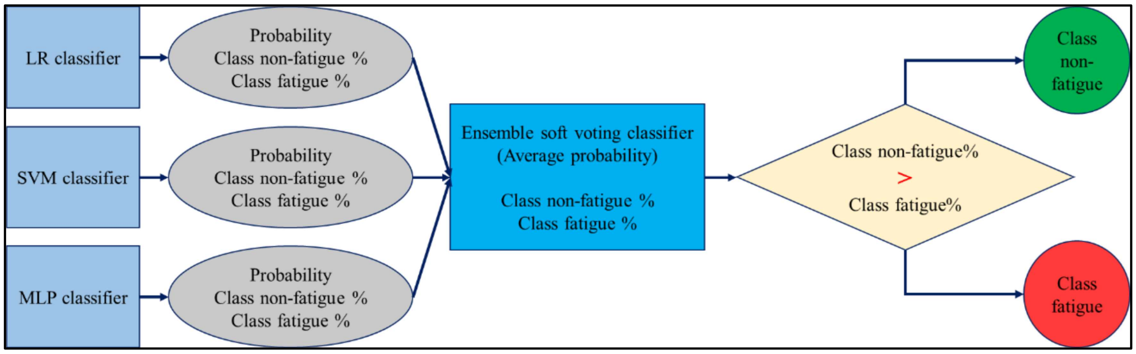


Figure 15 Ensemble Learning Model with Soft-Voting Classifier

In scikit-learn, the VotingClassifier is a meta-estimator (stacking). It combines multiple distinct classifiers and makes predictions by the majority vote. In the training set each base classifier (SVM, LR, and MLP) is trained separately which are the "estimators" of the model. The 'soft' voting method used to select the class with the highest average probability by averaging the class probabilities (scores) from each classifier. The "weights" parameter used to combine the predictions from the base models. In this case, the performance of each model is weighted proportionally to its ability. The number of required parallel jobs to ensure the fit is ensured by the "n_jobs". The "flatten_transform" used to affect the shape of the output transform where it returns as the matrix shape. The verbosity level defines by the "verbose" parameter to set the logging level of the model.

3.7 EVALUATION

Finding the Emotiv Performance Metrics' (EPM) effectiveness to detect mental fatigue is one of the evaluation's key components. EPM can identify a variety of mental states, according to a previous evaluation of the literature, but there is no research to back up its ability to identify mental fatigue. One of the purposes of this study is to assess EPM's ability to detect mental fatigue. At first it is important to measure individuals caseness of mental fatigue Using a self-reported assessment tool. Then statical analysis to see the difference between fatigued and non-fatigued session. After that correlation analysis to sees the correlation among EMP and mental fatigue, statical test to identify significance variable between fatigued and non-fatigued session and then ML model performance evaluation metrices to see the model performance.

3.7.1 Self-Report Mental Fatigue Assessment

To measure mental fatigue, a self-report CFQ questionnaire will be employed. The original CFQ (Chalder et al., 1993) included a total of 14 questions, while the updated version (Cho et al., 2007; Jackson, 2015; Jing et al., 2016) has 11 questions that are thought to be more dependable and better suited for model construction. The CFS can be scored using both a bimodal score and a Likert score, which depicted in Table 5. When scoring bimodal, "Yes" equals 1 and "No" equals 0, respectively.

Table 5 Chalder Fatigue Scale (CFS)

Questions	Scoring					
	Likert Scoring				Bimodal	
	Less than usual	No more than usual	More than usual	Much more than usual	No	Yes
1. Do you have problems with tiredness?	0	1	2	3	0	1
2. Do you need to rest more?	0	1	2	3	0	1
3. Do you feel sleepy or drowsy?	0	1	2	3	0	1

4. Do you have problems starting things?	0	1	2	3	0	1
5. Do you lack energy?	0	1	2	3	0	1
6. Do you have less strength in your muscles?	0	1	2	3	0	1
7. Do you feel weak?	0	1	2	3	0	1
8. Do you have difficulties concentrating?	0	1	2	3	0	1
9. Do you make slips of the tongue when speaking?	0	1	2	3	0	1
10. Do you find it more difficult to find the right word?	0	1	2	3	0	1
11. How is your memory?	0	1	2	3	0	1

The Bimodal score applied as the CFS to assess mental fatigue. The pre and post chalder fatigue questionnaires depicted accordingly in Appendix E and Appendix F. By contrasting with the CFS score, the mental fatigue identification is confirmed by measuring EPM. The CFS score 4 or more is “CASENESS” of mental fatigue (Chalder et al., 1993; Jackson, 2015; Loge et al., 1998).

3.7.2 Descriptive Statistics

The mean, median, standard deviation, and the 25th and 75th percentiles calculated for each EPM in order to detect differences in the data between fatigued and non-fatigued sessions. The mean is the average of a EPM in a session. So, this helps to identify the difference between two sessions. Therefore, there is challenge to depend on only mean because mean is sensitive to outliers or extreme maximum or minimum value. To solve this issue median is calculated, which is the middle point of data. If mean near to median then the data is likely to be balanced around the center, few outliers this suggest a symmetrically distributed data. That can be very useful for both sessions.

Standard deviation (SD) is calculated to see the data variability because SD is the square root of the variance. Low SD indicated the data points close to mean; higher SD value indicated wide ranges of data. Furthermore, the 25th and 75th percentiles are included to support SD and better observation of data distributions. 25th percentile refers the values which fall below 25% of the data and 75th marks the value that fall under 75% of the data. This can help to identify the data central ranges for each EPM in both sessions.

These statistical indicators provide insightful information on the data's distribution, central tendency, and variability that facilitating a comprehension of the differences in features between the two groups. A better understanding about the pattern or behavior of EPM that separate fatigued from non-fatigued sessions by looking at the statistics summary.

3.7.3 Correlation Analysis

Correlation analysis helps to understand baseline association between EPM and mental fatigue status. The linear relationship between two variables is measured in this study using the Pearson correlation coefficient (Pearson, 1895). It has a range of -1 to 1, where 1 represents a perfect linear association between the variables, -1 represents a perfect negative linear association, and 0 represents no linear association at all ("Pearson's Correlation Coefficient," 2008). The correlation presented for fatigued session, then non-fatigued session to identify the dynamics of EPM in the respective session. So, this will help to compare the both session EPM behavior. The two sessions were then combined, and the relationship between fatigue and EPM was presented. This helped to discover the association between EPM and mental fatigue for both individuals and generalized observations. This will help to understand the behavior or relationship of each EPM metrics against mental fatigue status.

3.7.4 Statical Testing

The significance of differences between fatigued and non-fatigued sessions is assessed using Welch's t-test, which improves the analysis's understanding and consistency. Student's t-test is adapted into Welch's t-test (WELCH, 1947). The t-statistic (t) is computed with the means, variances, and degrees of freedom using the Welch-Satterthwaite equation (Satterthwaite, 1946; WELCH, 1947). Next, using the t-distribution with degrees of freedom, find the p-value related to the t-statistic. The null hypothesis can be rejected if the p-value is smaller than the significance level (0.05), indicating significant difference between the means of the two groups. If p-value exceeds the significance level, the null hypothesis cannot be rejected which states there is no significant difference.

3.7.5 Mental Fatigue Detection Model Evaluation

ML model performance evaluation is the basis for building effective and reliable ML application. This helps decision-making, improves interpretability, and identifies overfitting and underfitting in machine learning applications.

3.7.5.1 Machine Learning Model Performance Metrics

There are several model performance metrics including accuracy (Flach, 2019), precision, recall, F1- score, and ROC ACU (Basha & Rajput, 2019) score can be used to determine model performance. A tabular representation of the model's predictions compared to the actual values is called a confusion matrix. A typical confusion matrix presented in Table 6. Confusion Matrix consist of True Negative (TN), False Positives (FP), False Negative (FN) and True Positive (TP). It is expected that a higher TP value will indicate that the model accurately predicted positive instances. Similarly, a greater number of TN is preferable since it shows that the model accurately predicted negative instances. On the other hand, a small FP value is expected since these indicate cases that were incorrectly classified as positive (Type I error). Similar to FP a small FN value will be expected, since it indicates instances that were incorrectly classified as negative (Type II error).

Table 6 Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	TN	FP
Actual Positive	FN	TP

Accuracy is most common metrics to evaluate a model performance (Flach, 2019; Yin et al., 2019). It calculates the percentage of correctly classified instances relative to all instances. A high (near to 1) accuracy value indicates a model with accurate predictions. The formula to calculate for accuracy is following:

$$accuracy = (TP+TN) / (TP+TN+FP+FN)$$

Precision is another metrics that finds the percentage of true positive prediction among all positive predictions (Basha & Rajput, 2019). It measures how well the model can recognize positive instances; in other words, it measures the ability to avoid false positives of a ML model. The formula is following:

$$\text{precision} = TP / (TP+FP)$$

Recall (sensitivity) calculates the percentage of accurate positive predictions among all actual positive instances (Basha & Rajput, 2019). High recall helpful while there is a low precision because it finds most true positive cases even though it is prone to false positives. The formula for recall is following:

$$\text{recall} = TP / (TP+FN)$$

F1-score is another important evaluation metric that establishes a balance between recall and precision (Basha & Rajput, 2019). It measures the model's overall performance. Better model performance can be witnessed by a higher F1-score (closer to 1), which reflects a well-balanced trade-off between precision and recall. The F1-score formula is following:

$$\text{F1-score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{recall})$$

Specificity is the proportion of actual negative instances that the model correctly classifies as negative (Basha & Rajput, 2019; Flach, 2019). A low specificity score implies possible problems with false positive predictions, whereas a high score indicates successful negative categorization. The formula is following:

$$\text{specificity} = TN / (TN+FP)$$

Receiver Operating Characteristic Area Under the Curve (ROC AUC) is calculated by plotting the true positive rate (recall) against the false positive rate (1-specificity) for a variety of threshold values (Basha & Rajput, 2019; Flach, 2019). Higher ROC AUC (near to 1) values demonstrate better discrimination between positive and negative instances. ROC AUC value 0.5 indicates random classification; means there is no discrimination between positive and negative instances. Figure 16 is the representation of a ROC AUC curve.

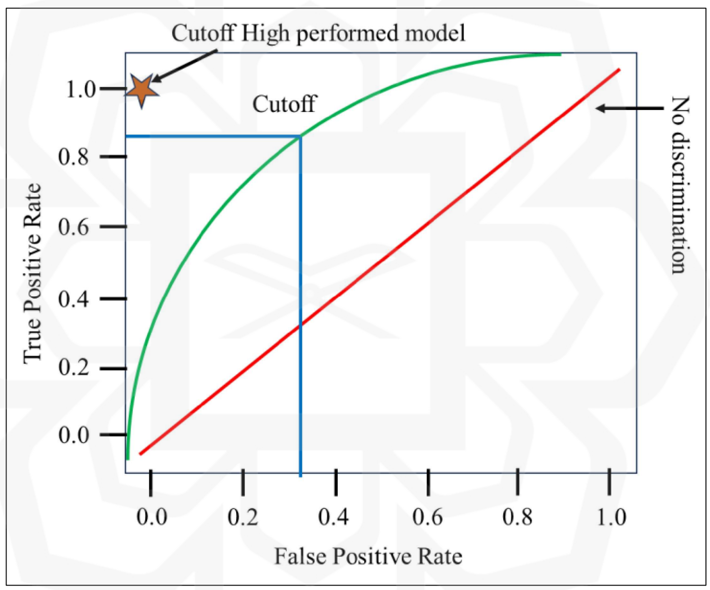


Figure 16 ROC AUC Curve of a Classification Model

3.7.5.2 Cross-Validation

Cross-validation is a vital technique to evaluate the performance of predictive model and ensure model generalization ability to new data (Browne, 2000). The dataset is divided into several subsets, the model is trained on some of these subsets, and its performance is evaluated on the remaining subsets. While there are various approaches to perform cross validation, this study uses k-fold cross-validation, specifically with five folds. The dataset is split into five equal-sized subsets. The model is trained and evaluated five times. Each time using a distinct subset as the validation set and the remaining subsets as the training set. After calculating cross-validation score for five folds, then mean, SD, training score and validation score is calculated.

Cross-validation mean score describes the overall performance of the model across all folds (Berrar, 2019; Browne, 2000). The SD measure the variability or spread the model's performance across different folds. A higher (closer to 1) mean vale and a lower (closer to 0) SD value defines better model performance with consistency.

Training Score describe the model performance on the same dataset that used to train model (Berrar, 2019; Browne, 2000). A higher training score indicate better performance on the respective training data. In contrast, validation score defines the model performance on a different data that is not used to train model. The model's prediction is evaluated on the validation data to identify model's generalization ability to new data. In this reason, validation score is important to identify overfitting or underfitting of model. If a model training score is high but validation score is significantly low the model will be overfitted model. The model will be identified as underfitting, if both training and validation scores are low. The model will be well generalized performed model if both training and validation scores are high consistent, means validation score closer to training score.

3.8 Data

This section describes a collection of unlabeled Emotiv performance metrics (EPM) data details which is recorded from 10 participants, under online video learning stimuli. The data includes five EPMs: stress, engagement, interest, relaxation, and excitement. A self-reported fatigue rating known as a chalder fatigue questionnaire that was obtained at the beginning of the session and the end of the session. Participants wore the Emotiv headset while participating in online learning. Figure 17 shows a data collection session. The dataset will contribute to the detection of mental fatigue detection for online participants.

Figure 17 A Participant in the Data Collection Session

3.8.1 Participants Details

This research aims to perform mental fatigue detection with different subjects. The subjects are from International Islamic University Malaysia (IIUM) students. There are 10 subjects participated in the experiment. A volunteer participation adds shared in social media like

IIUM Facebook page, different WhatsApp and Telegram group. Initially, total 13 individuals interested to participate in the data collection. Three individuals did not participate in the data collection session. Total 10 individuals participated for the EPM data collection. All the participants were given an identification number to protect their identity. The subject inclusion-exclusion criteria are sickness, smoking, and drug addiction. If the participants are sick and/or smoking and/or drug addicted and/or any psychological disorder accounted as exclusion criteria for subject selection. The demographic information collection includes age, job, medicine and sleeping hours which are depicted in Table 7.

Table 7 Demographic Information

Information criteria		Number of participants
Age group	20-25	7
	25-30	2
	30-35	1
Marital status	Single	9
	Married	1
Gender	Male	10
	Female	0
Body Mass Index (BMI)	Underweight (<18.5)	0
	Normal (18.5-25)	6
	Overweight (25-30)	3
	Moderately obese (30-35)	0
	Severe Obese (>35)	1
Education	Doctor of Philosophy (PhD)	2
	Master	1
	Bachelor	7
Housemate	Alone	4
	Friends	2
	Family	4
Job	Full-time	0

	Part-time	0
	Unemployed	10
Health Condition	History of Covid 19	5
	Physical disorder	0
	Psychological disorder	0
	Current sickness	0
	Currently medication	1
Addiction	Smoking	0
	Drug addiction	0
Preferred study mode	Face-to-Face	10
	Online learning	0
Fatigue during learning	Face-to-Face (Frequently)	1
	Face-to-Face (Sometimes)	7
	Face-to-Face (No)	2
	Online learning (Frequently)	3
	Online learning (Sometimes)	7
Feeling loneliness	No	3
	Sometimes	4
	Frequently	3

The Subjects participate on their flexible time. After participation selection they were guided for data collection session. The instruction includes: (i) take enough sleep, (ii) cannot drink Alcoholic drink, (iii) cannot drink coffee or tea, (iv) wash hair, and (v) cannot use hair oil or gel. All the participants preferred 50% of brightness and contrast level.

3.8.2 Screening Based on Chalder Fatigue Scale (CFS)

The data screening process used for sample inclusion and exclusion for mental fatigue detection. There are two participants data is excluded from the analysis. Subject_1001 starts the session with fatigued caseness that is defined from his CFQ outcomes. He has gotten pre-CFS score of 4 which is caseness of fatigue according to the CFS scale. This indicates he is in fatigue states already. So, this need to be excluded, because need to have both brain states (fatigue and non-fatigue). Another participant's (Subject_1006) data is excluded because his pre-CFS score is 0 and post-CFS score is 3 which defines he was not fatigued during the learning session. Since only non-fatigue brain state, so excluded from the further analysis. This two participants inclusion may cause of data imbalance since their fatigue status pattern are different than other 8 participants. This will lead to a person-specific bias rather than generalization model for fatigue detection. Also, there is a high chance to be overfitting issue. Since, their fatigue status is static (means both persons have only one fatigue state), this may lead to lower mental fatigue detection model performance. Finally, total 8 person data is included for mental fatigue detection. The CFS presented in Table 8.

Table 8 Participants CFS Score

Partcipants_ID	pre_CFS(out_of_11)	post_CFS(out_of_11)
Subject_1001	4	10
Subject_1002	2	6
Subject_1003	1	9
Subject_1004	0	5
Subject_1005	0	7
Subject_1006	0	3
Subject_1007	0	4
Subject_1008	3	7
Subject_1009	0	6
Subject_1010	0	6

3.8.3 Dataset Details

As earlier mentioned, five Emotiv performance metrics (EPM), demographic information and a Chalder Fatigue Questionnaire (CFQ) data were recorded. The EPM dataset saved in Comma-separated values (CSV) file format. Self-reported CFQ is scaled in binary form 1 and 0. The CFQ and demographic data were recorded using Google Form. After that the data is recorded in a CSV file. Table 9 is a sample of the unlabeled EPM data. The EPM's measurement from EEG signals is done by Emotiv but the details process is not described by the Emotiv. Emotiv headset connected through Emotiv launcher and generates performance metrics from Emotiv devices at 0.1 Hz (Emotiv, 2024). This indicates that the performance data are updated or transmitted once every ten seconds. As Emotiv describes each EPM is in a decimal number range between 0 and 1. If there is poor EEG signal quality then the EPM value is null and excluded from the dataset. There was no null value, since careful experimental setup has performed as describe earlier in Section 3.4.

Table 9 Sample of EPM Data

engagement	excitement	stress	relaxation	interest	average	timestamp
0.637614	0.431304	0.276254	0.26414	0.56341	0.434544	1.7E+09
0.534821	0.253394	0.340384	0.376263	0.521004	0.405173	1.7E+09
0.429153	0.162997	0.337936	0.366186	0.488786	0.357012	1.7E+09
0.503934	0.391049	0.434824	0.440165	0.484864	0.450967	1.7E+09
0.451405	0.252189	0.304995	0.481921	0.514529	0.401008	1.7E+09
0.396943	0.29764	0.319735	0.450011	0.479825	0.388831	1.7E+09
0.488003	0.573425	0.297895	0.41388	0.514818	0.457604	1.7E+09
0.465568	0.52055	0.36467	0.297609	0.463998	0.422479	1.7E+09
0.618003	0.561818	0.37487	0.412575	0.539995	0.501452	1.7E+09
0.901969	0.307723	0.272013	0.130901	0.572449	0.437011	1.7E+09

3.8.4 Data Labelling

Tutorial video used to induce fatigue. The data can be labelled in different time interval. Each participant's EPM data was divided into two different categories in the data labeling process: non-fatigued (0) and fatigued (1). Subject-dependent analysis refers to each participants individual analysis. Subject-independent analysis refers all participant data combine and make a generalize dataset to analyze. Brain signals are dynamics person to person. It is important to have both subject-dependent and Subject-independent analysis to show the dynamics among participants brain states. This will help to identify the best detection model for mental fatigue. In particular, each participant's first 15-minute phase was classified as "non-fatigued", whereas their second 15-minute segment was classified as "fatigued". Then merged the data from every participant to create a general dataset for subject-independent analysis. This research chooses to label first 15 minutes 50 seconds data as non-fatigue (0) and second 16 minutes as fatigued (1) because participants self-reported to fall in fatigue around this time. So, there is a significant difference expected in this time period. These time frame provides a reliable result in the detection of mental fatigue that described in the following chapters.

Each subject-dependent dataset and subject-independent dataset were divided into training and testing sub-dataset in a proportion of 80 percent and 20 percent with randomized selection methods. The number of observations is described in Table 10.

Table 10 Number of Observations in a Dataset

Dataset	Total observation	Training (80%)	Testing (20%)
Subject-dependent	191	152	39
Subject-independent	1528	1222	306

CHAPTER FOUR

RESULT ANALYSIS

4.1 INTRODUCTION

This section represents the result and discussion on the findings. Subject-independent and subject-dependent analysis is employed to look at EPM for mental fatigue detection. Start with descriptive statistics for central tendency discrepancies to highlight the differences between the fatigued and non-fatigued segments. Then, these variations will be visually represented using data visualization. Correlation analysis explores the links between variables as well as shows difference between fatigue and non-fatigue session. Statistical testing carefully evaluates significance difference between both sessions. Finally, using available features, machine learning algorithms are used to create mental fatigue classification model.

4.2 SUBJECT-DEPENDENT ANALYSIS

This section represents descriptive analysis, correlation analysis, statical testing findings and mental fatigue detection findings for subject-dependent mental fatigue analysis. Subject-dependent models help to leverage individual EPM pattern and mental fatigue status. Subject-dependent models are able to exhibit higher accuracy because personalized model can capture individual variability. This may enable a highly personalized applications like personalized neurofeedback.

4.2.1 Descriptive Analysis

Descriptive statistics provides a concise description of the significant features of the data, such as mean, standard deviation (SD), and median, that enables the quick identification of potential differences between the fatigued and non-fatigued session. The descriptive statistics depicted in Table 11 for all subjects.

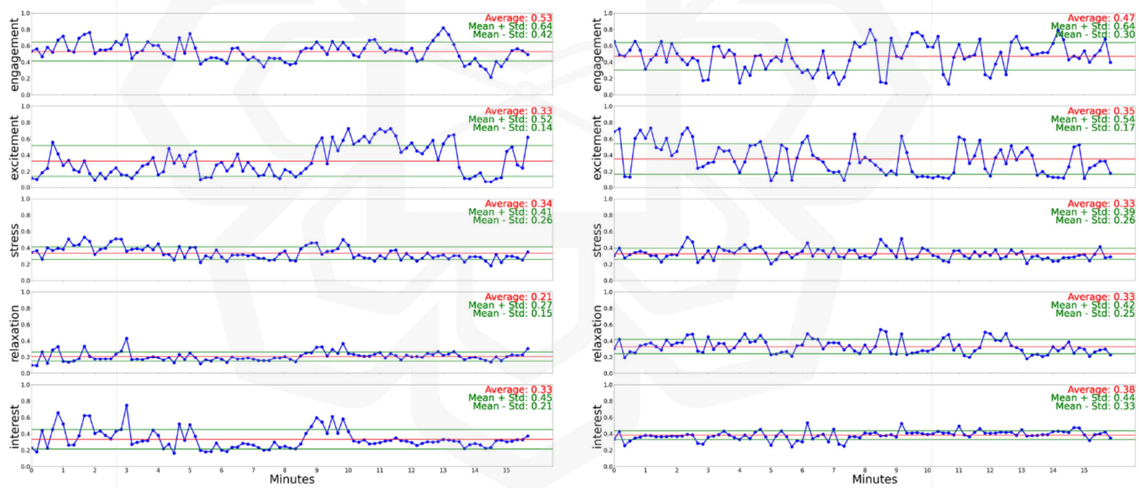
Table 11 Subject-Dependent Center Tendency Fatigue (FAT) vs Non-Fatigue (Non-FAT)

Participants_ID	Center Tendency	Engagement		Excitement		Stress		Relaxation		Interest	
		non-FAT	FAT	non-FAT	FAT	non-FAT	FAT	non-FAT	FAT	non-FAT	FAT
Subject_1002	Mean	0.53	0.47	0.33	0.35	0.34	0.33	0.21	0.33	0.33	0.38
	SD	0.11	0.17	0.19	0.19	0.08	0.07	0.06	0.09	0.12	0.05
	Median	0.54	0.48	0.28	0.32	0.32	0.31	0.2	0.32	0.3	0.39
	25%	0.45	0.38	0.18	0.19	0.28	0.28	0.17	0.26	0.26	0.36
	75%	0.58	0.59	0.49	0.5	0.39	0.37	0.24	0.38	0.38	0.42
Subject_1003	Mean	0.57	0.51	0.34	0.49	0.39	0.37	0.47	0.4	0.67	0.42
	SD	0.14	0.16	0.11	0.14	0.09	0.06	0.2	0.1	0.2	0.05
	Median	0.61	0.51	0.32	0.51	0.38	0.36	0.43	0.4	0.69	0.42
	25%	0.47	0.41	0.26	0.37	0.33	0.32	0.3	0.32	0.51	0.4
	75%	0.68	0.64	0.4	0.6	0.44	0.41	0.69	0.48	0.84	0.44
Subject_1004	Mean	0.54	0.38	0.42	0.26	0.38	0.3	0.43	0.26	0.54	0.43
	SD	0.11	0.09	0.18	0.11	0.07	0.04	0.12	0.08	0.1	0.05
	Median	0.53	0.39	0.41	0.25	0.37	0.29	0.42	0.25	0.53	0.42
	25%	0.45	0.33	0.25	0.16	0.32	0.27	0.33	0.2	0.47	0.4
	75%	0.63	0.45	0.56	0.31	0.42	0.32	0.49	0.31	0.59	0.47
Subject_1005	Mean	0.57	0.33	0.41	0.31	0.31	0.26	0.31	0.38	0.45	0.37
	SD	0.13	0.13	0.16	0.18	0.05	0.06	0.12	0.16	0.12	0.09
	Median	0.59	0.35	0.43	0.31	0.31	0.26	0.27	0.35	0.43	0.39

	25%	0.47	0.19	0.3	0.12	0.28	0.22	0.23	0.22	0.37	0.29
	75%	0.66	0.41	0.53	0.46	0.34	0.3	0.35	0.52	0.47	0.45
Subject_ 1007	Mean	0.41	0.39	0.36	0.42	0.29	0.29	0.26	0.24	0.43	0.47
	SD	0.14	0.14	0.16	0.17	0.07	0.04	0.06	0.05	0.05	0.04
	Median	0.41	0.36	0.32	0.39	0.28	0.29	0.25	0.23	0.44	0.48
	25%	0.31	0.29	0.23	0.28	0.24	0.26	0.2	0.21	0.42	0.47
	75%	0.49	0.49	0.46	0.58	0.35	0.32	0.31	0.26	0.46	0.48
Subject_ 1008	Mean	0.47	0.36	0.51	0.32	0.38	0.28	0.4	0.25	0.54	0.39
	SD	0.1	0.1	0.12	0.16	0.09	0.05	0.11	0.09	0.09	0.06
	Median	0.46	0.37	0.51	0.31	0.36	0.28	0.39	0.23	0.52	0.39
	25%	0.39	0.28	0.42	0.19	0.32	0.25	0.32	0.18	0.48	0.35
	75%	0.52	0.43	0.61	0.44	0.42	0.31	0.47	0.29	0.6	0.43
Subject_ 1009	mean	0.59	0.46	0.5	0.32	0.44	0.37	0.51	0.37	0.61	0.49
	SD	0.11	0.1	0.11	0.12	0.07	0.06	0.12	0.09	0.08	0.09
	median	0.58	0.47	0.52	0.31	0.43	0.37	0.52	0.37	0.61	0.5
	25%	0.52	0.4	0.42	0.24	0.39	0.33	0.42	0.31	0.55	0.46
	75%	0.65	0.54	0.59	0.38	0.47	0.41	0.57	0.43	0.66	0.56
Subject_ 1010	mean	0.58	0.47	0.51	0.34	0.38	0.33	0.37	0.32	0.49	0.45
	SD	0.11	0.09	0.17	0.11	0.06	0.07	0.11	0.09	0.04	0.04
	median	0.55	0.48	0.55	0.33	0.39	0.33	0.35	0.32	0.48	0.46
	25%	0.51	0.4	0.42	0.26	0.33	0.29	0.29	0.25	0.46	0.44
	75%	0.64	0.53	0.65	0.42	0.42	0.37	0.41	0.39	0.51	0.47

Several significant changes between the fatigued and non-fatigued sessions for all participants can be seen in the Table 11. Subject_1002 experience a minor decline in engagement mean value during fatigue, but an increase in relaxation and interest. There is a stability in the excitement and stress levels in both sessions. The median values are relatively same as mean value and shows similar trends for all the metrics in both sessions. The engagement shows a decrease median value while increase in excitement, relaxation and interest. There is a stability in the stress levels. Changes on the variability of metrices can see by the SD. The standard deviation for interest is higher, but lower in engagement and relaxation in non-fatigued sessions compared to fatigued session.

There is a consistent variability in excitement and stress for both sessions. Therefore, the variability is high for both session in engagement, excitement and interest but low in stress and relaxation. Furthermore, there is a comparable distribution of all metrices between fatigued and non-fatigued session is noticed by the 25th percentile and the 75th percentile. Engagement, and relaxation experience wide spread, then less spread in stress and interest during fatigued session, but excitement remain same spread for both sessions. Interestingly the higher quartiles exhibit minor fluctuations, especially for interest, indicating a possible reduction in sustained attention during fatigue, even if the median values for the majority of metrics stay relatively stable across both states. The difference between fatigued and non-fatigued is shown in Figure 18. The line chart shows a significant difference between two mental condition session.

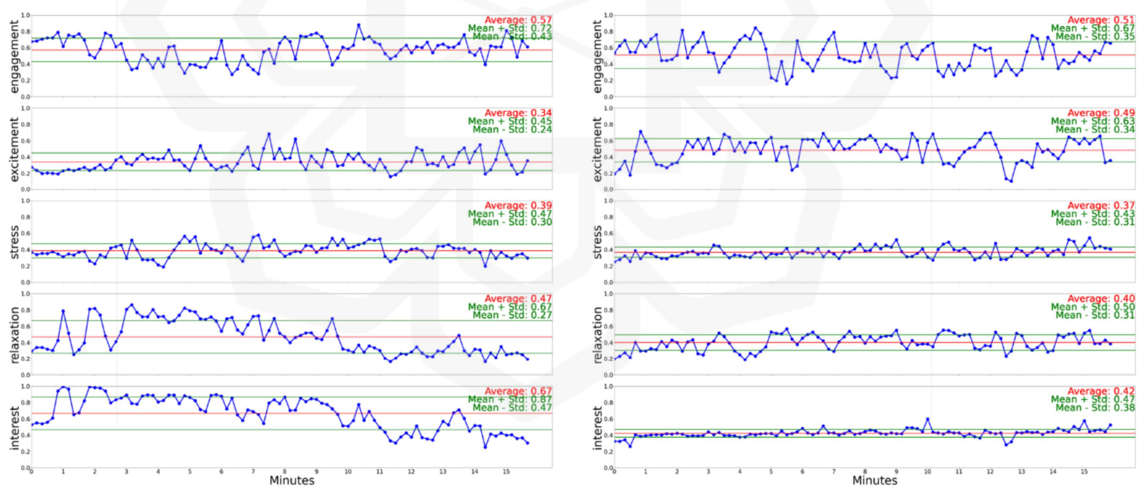


Non-Fatigued Fatigued

Figure 18 Subject_1002 Non-Fatigue vs Fatigue

Table 11 shows that Subject 1003 exhibits a decline in the mean value of engagement, relaxation, and interest but increases in excitement during a fatigued session compared to non-fatigued session. Stress levels are similar during the two sessions. Similar patterns across measures in both sessions are indicated by median values, which are shown by mean values patterns. Interestingly during fatigued session, the median and mean value are identically same for engagement, relaxation and interest.

Then, standard deviations show larger variability for all metrices in both session except stress and interest during fatigued session. When comparing a class of fatigued to one of non-fatigued students, the variability in stress is about constant, but it is greater in engagement and excitement and lower in relaxation and interest. Furthermore, the distributions of the two sessions are different, as the 25th and 75th percentiles show. The distribution ranges are higher in engagement, excitement and lower in other metrices for the fatigued session. Interest, relaxation shows significant different distributions. Figure 19 graphically represents these contrasts between fatigued and non-fatigued states.



Non-Fatigued Fatigued

Figure 19 Subject_1003 Non-Fatigue vs Fatigue

Table 11 and Figure 20 show that Subject 1004's non-fatigued and fatigued periods differ noticeably from one another. It is interesting to point out that during fatigued session, all metrics' mean and median values decreased. It is also noticeable for all metrics' mean and median values were nearly identical in the respective session. The standard deviations indicate less variability for each metrics in the fatigued session compared to non-fatigued session. Similar to this, the data spread revealed a lower range in fatigued session is indicated by the difference between the 25th and 75th percentiles.

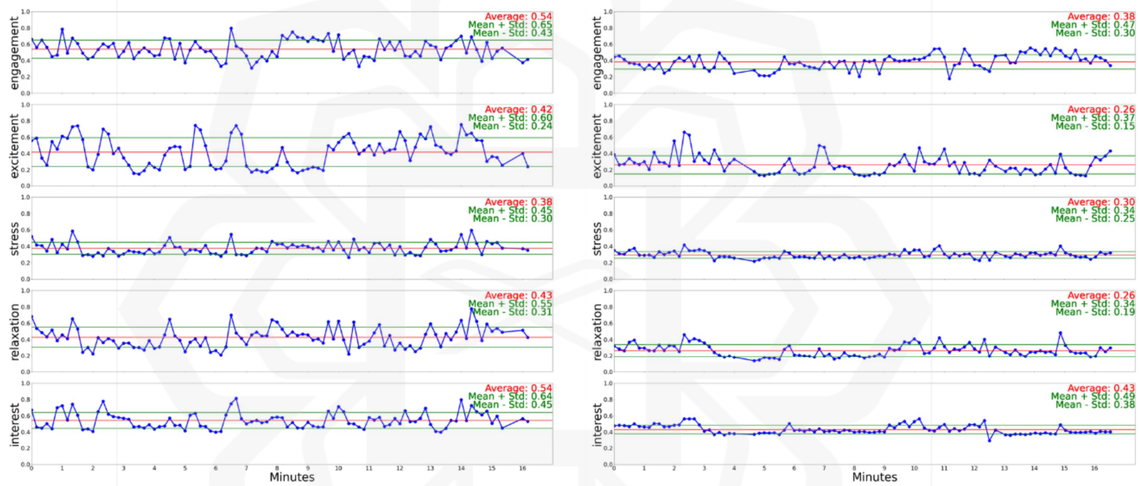
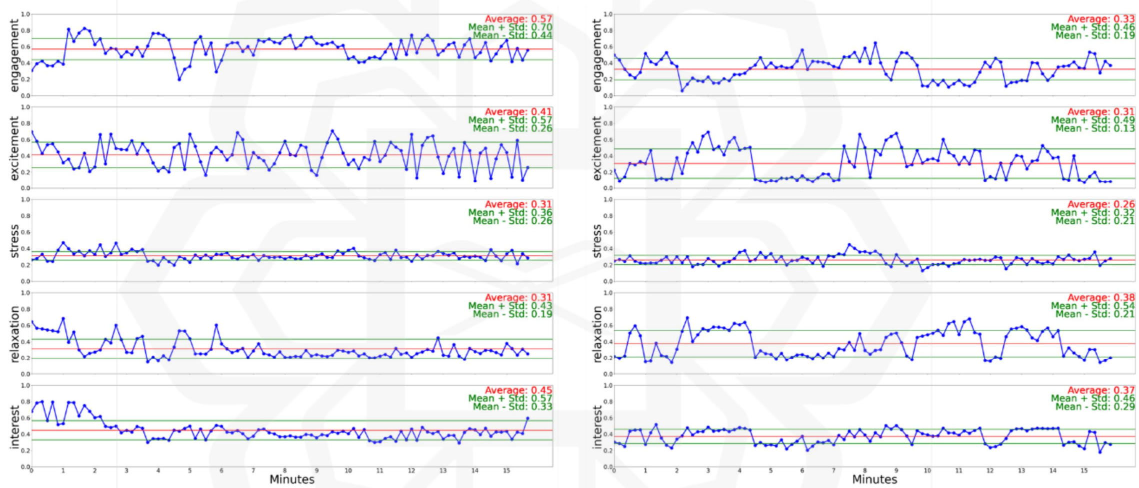


Figure 20 Subject_1004 Non-Fatigue vs Fatigue

Subject 1005's emotional metrics exhibit trends that are very similar to those of Subject 1004's. It can be observed from Table 11 that during the fatigued session, the mean and median values of every metrices decreased and these values for every metrices were nearly identical in its respective session except relaxation metric. Relaxation metric's shows a slightly increase in fatigued session. The SD is high and identical for engagement.

Excitement, relaxation and interest metrices data variability is almost same and high for both session but stress variability is low in both sessions.

The data destructions are in higher ranges (75th percentile - 25th percentile) among all metrices in fatigued session compared to non-fatigued session. Another important aspect is that the data distribution shows minimal in stress but high in interest, engagement, excitement and relaxation for both sessions. Figure 21 is the representation of Subject_1005 data for both sessions.

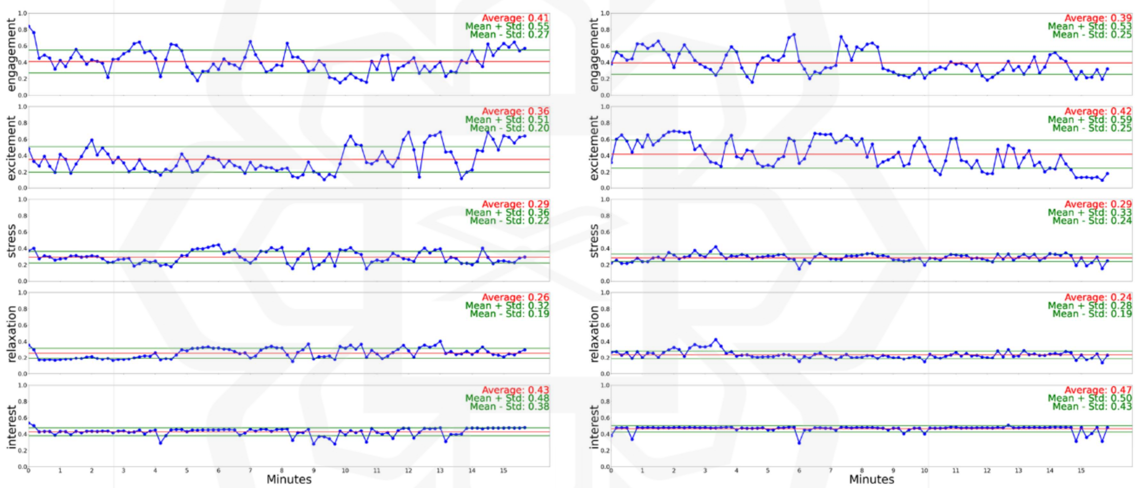


Non-Fatigued Fatigued

Figure 21 Subject_1005 Non-Fatigue vs Fatigue

Table 11 describes that Subject_1007 exhibits a decline in the mean and median values of engagement, and relaxation, but increases in excitement and interest during a fatigued session compared to non-fatigued session. Stress mean values are identical and median values are almost identical for both sessions. The mean and median values are almost same for all emotional metrices in their respective session.

The SD is higher in engagement and excitement while lower in stress, relaxation and interest. The variability is identical in engagement, but slightly lower in other metrices in fatigued class. Fatigued session showed high data spread range for engagement and excitement but lower spread in stress, relaxation and interest. Engagement and excitement showed high and different distribution for both sessions. On the other hands stress, relaxation showed high distribution in non-fatigued class but minimal in fatigued session. Interest shows low and different data spread in both sessions. Figure 22 graphically illustrates the contrasts between fatigued and non-fatigued states.

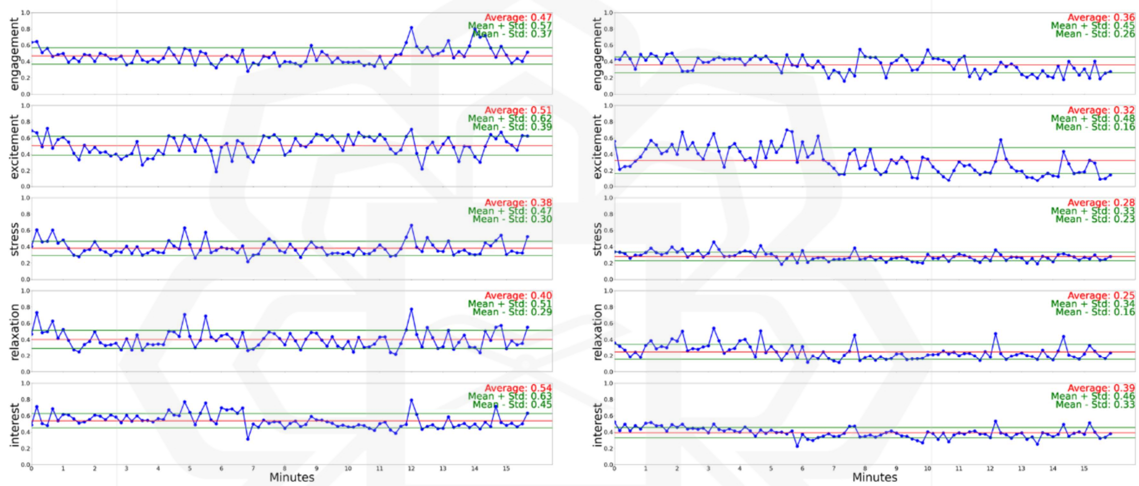


Non-Fatigued Fatigued

Figure 22 Subject_1007 Non-Fatigue vs Fatigue

Table 11 demonstrates significant differences between non-fatigued and fatigued states for Subject_1008. When compared to the non-fatigued session, the mean and median values of every metric demonstrate a considerable reduction during the fatigued session. Although SD highlight less variability differences for stress, interest and relaxation, but engagement variability is identical and excitement is higher in fatigued session.

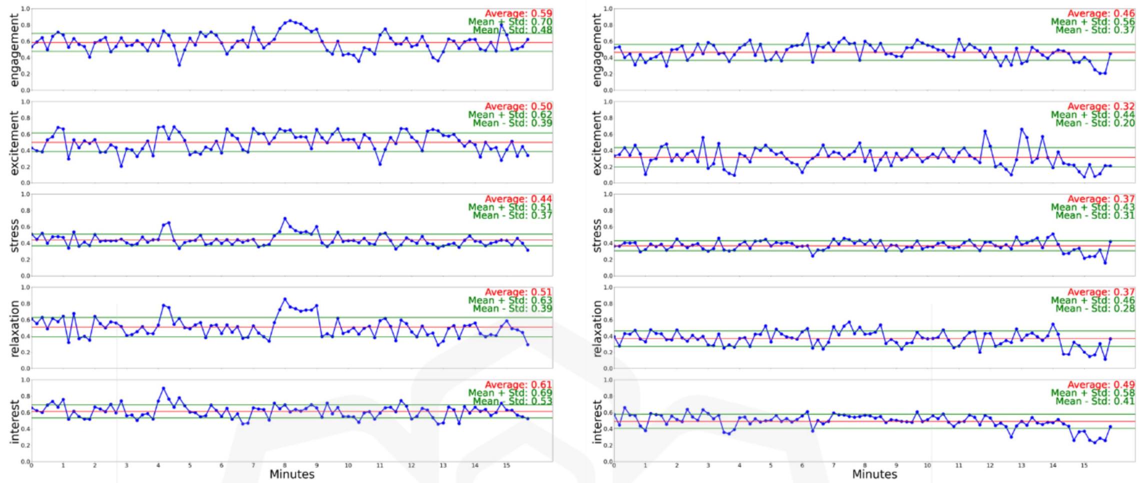
The excitement and engagement distribution ranges increase during the fatigued session. The stress shows consistent decrease in the fatigue session. Similarly, the interest and relaxation show a consistent trend the fatigue session. The 25th and 75th percentiles are greater in non-fatigued session than fatigued session. These results highlight the emotional states distinction, as shown graphically in Figure 23. The difference between the two sessions is significant.



Non-Fatigued Fatigued

Figure 23 Subject_1008 Non-Fatigue vs Fatigue

Comparable to Subject_1008, the mean and median values of Subject_1009 exhibit a similar pattern. It's interesting to note that SD indicates almost same variability for all metrics in both sessions. The 25th and 75th percentiles are lower in fatigued sessions for all metrics, similar to the mean and mean values. However, the data spread range is lower for interest, relaxation, and excitement, but larger for excitement when the participants are fatigued and same for excitement in both sessions. Figure 24 shows how two sessions differ from one another.

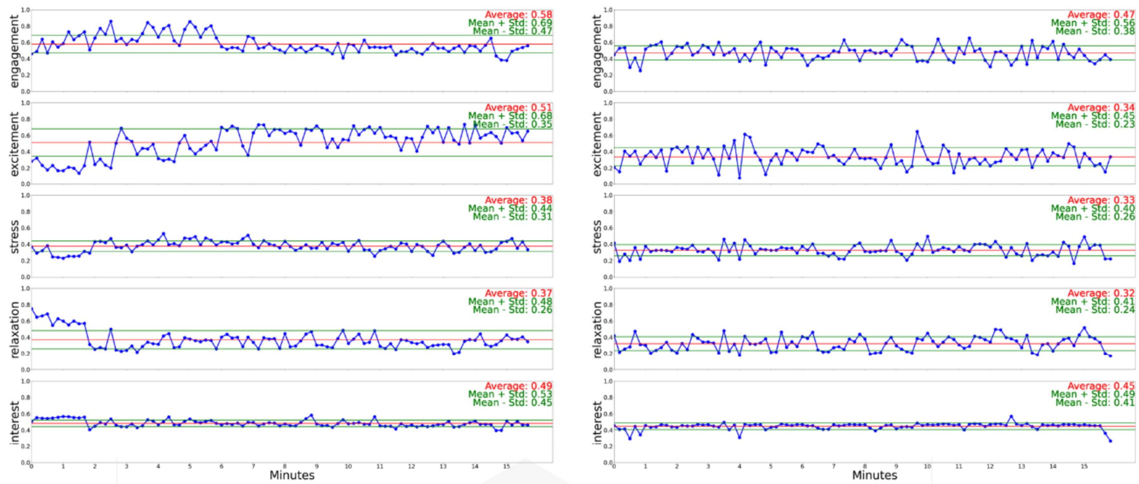


Non-Fatigued   Fatigued

Figure 24 Subject_1009 Non-Fatigue vs Fatigue

The mean and median value of all EPM in the fatigued class is decreased for Subject_1010. The SD shows a decreasing in excitement, engagement, and relaxation, a minor rise in stress, but the same in interest. The variability changes for excitement, engagement, relaxation and stress, this suggest that there is a significance in the fatigue and non-fatigue session.

The 25th and 75th percentiles also show a difference between the both sessions. All the EPM shows a decrease in the fatigue session compared to non-fatigue session. Similar, trend also noticed in the 75th percentiles. The difference between 75th percentiles and 25th percentiles shows that the data ranges is lower levels in excitement, stress, and interest but higher levels in relaxation; but engagement is the same range. The spread is different in the both sessions. The differences between the fatigue and non-fatigue can be visually observed in the Figure 25.



Non-Fatigued Fatigued

Figure 25 Subject_1010 Non-Fatigue vs Fatigue

There is significant difference between fatigue and non-fatigue labelled data. EPM shows different pattern for each participant. The mean, and median, are almost similar with a low SD for the respective individuals. The low SD defines the low variability in the data. It is interesting to note that the mean, median and SD although different, but there is no common pattern for any EPM metrics. The 75th percentiles and 25th percentiles also identified the difference between the fatigue and non-fatigue labelled data, but there is no common pattern in the EPM among the all participants. The EPM behavior are unique based on personal.

4.2.2 Correlation Analysis

There are significant differences between the non-fatigued and fatigued phases, according to the correlation analysis for Subject_1002 which depicted in Figure 26. Moderate to significant positive connections have been found between engagement and other metrices in the non-fatigued state, including excitement (0.42), stress (0.56), relaxation (0.52), and interest (0.63). Besides these there are moderate to high correlation can be found in interest-stress, excitement-relaxation and interest-relaxation. In contrast, fatigued session did not show any high correlations between matrices, only a moderate positive correlation in stress-relaxation (0.57) and negative correlation in engagement-relaxation (-0.54). The fatigued and non-fatigued integrated data shows less correlation among metrices except relaxation-interest with moderate association. However, the association between fatigue status and other metrices are low, only relaxion showed moderate association.

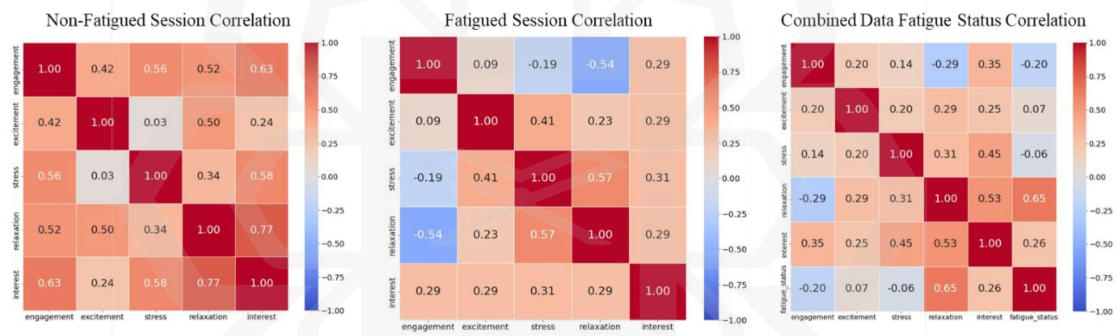


Figure 26 Subject_1002 Comparative Correlation Analysis

Figure 27 demonstrates that during non-fatigued sessions, Subject_1003 metrics have low correlation; only interest-relaxation exhibits a high correlation, while engagement and relaxation have a moderate association. Metric associations are stronger in the fatigued session than in the non-fatigued session, but the interest-relaxation relationship is lower. The combined data indicates a negative correlation in engagement and relaxation and a moderately positive correlation for fatigue status with excitement (0.49). There find a high negative association (-0.65) between interest and fatigue status.

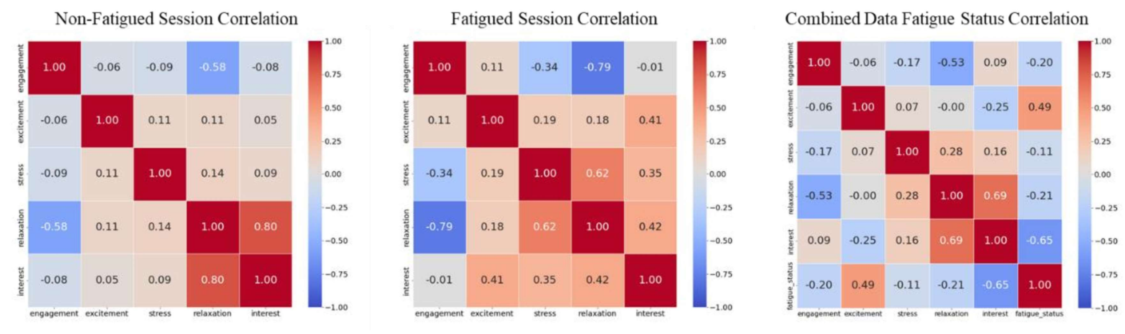


Figure 27 Subject_1003 Comparative Correlation Analysis

Figure 28 shows the difference between the fatigued and non-fatigued sessions is evident in the association between the metrics for Subject_1004. Only the stress-relaxation correlation (0.89) is high during the non-fatigued session; the majority of other correlations are at a low to moderate level. On the other hand, fatigued session has a strong association in stress-relaxation and interest-relaxation. Other correlations differ from the non-fatigued session in terms of correlation value; however, they are at a low to moderate level. In both sessions, there is no negative association. When combining the two sessions' amounts of data, all metrics have moderate to strong negative correlations with fatigue status. Other than these, all of the metrics have moderate to high association.

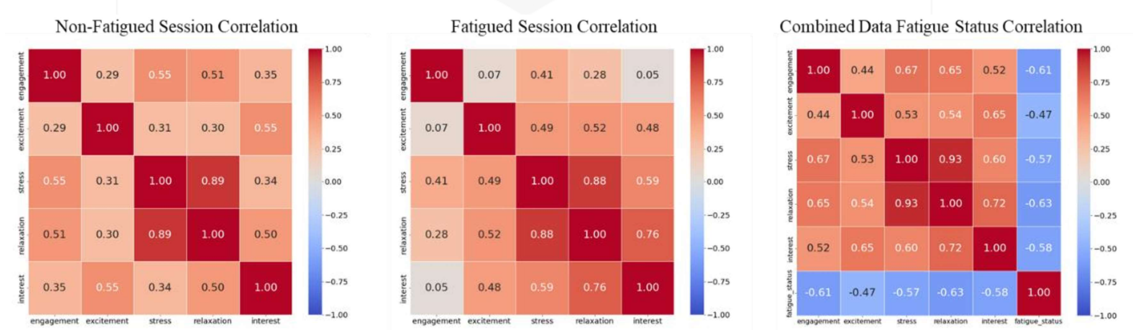


Figure 28 Subject_1004 Comparative Correlation Analysis

Subject_1005, there is not a significant association between the metrics during the non-fatigued session by Figure 29. Only engagement-relaxation (-55) and interest-relaxation (58) exhibit a moderate link among the metrices, with the majority of correlations between them being low. Hence, fatigued session exhibits noticeably strong relationships between the relevant metrics. Interestingly, there is no correlation between stress and relaxation. The combined data from both sessions shows a negative association (0.22) between fatigue status and all metrics except relaxation. Only engagement shows a strong association (-0.69) with fatigue status. With the exception of engagement-relaxation (-0.61) and interest-relaxation (0.54), the correlation between the metrics is minimal.

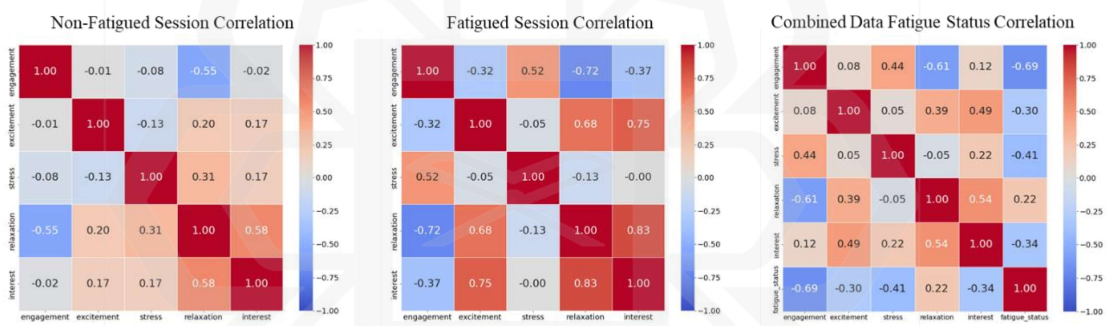


Figure 29 Subject_1005 Comparative Correlation Analysis

For both sessions Subject_1007 exhibits a distinct association between the metrics. The only relaxation-stress association that is high during non-fatigued sessions is 0.73; during fatigued sessions the correlation is 0.66. On the other hand, during fatigued sessions, interest and stress have a higher association (0.66) than in non-fatigued sessions. Less correlation exists between fatigue status and metrices as well as between metrices in the aggregated data from the two sessions. Fatigue has a positive and negative relationship with other metrices. Figure 30 displays the associations for Subject_1007.

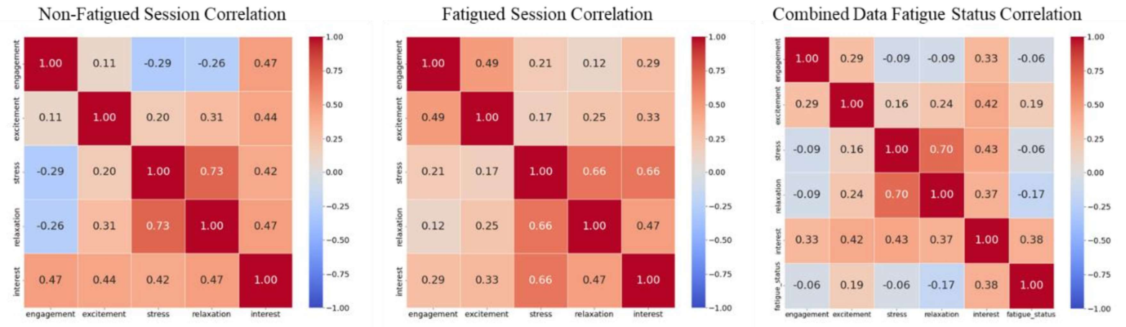


Figure 30 Subject_1007 Comparative Correlation Analysis

Figure 31 shows similarity to other participants by comparing the metrics of the fatigued and non-fatigued sessions, Subject_1008 likewise displayed a clear association between the metrics. The association between stress and relaxation is strong in both sessions, but it is lower in the fatigued session (0.88) than in the non-fatigued session (0.94). Both sessions had strong but lower in non-fatigued session for interest-stress and interest-relaxation correlations. While engagement and other metrics correlations declined in fatigued sessions, there is an increase in other correlations in. There is no negative correlation in either session. However, when the two sessions are combined, there is a negative and moderate to high association between fatigue and metrics.

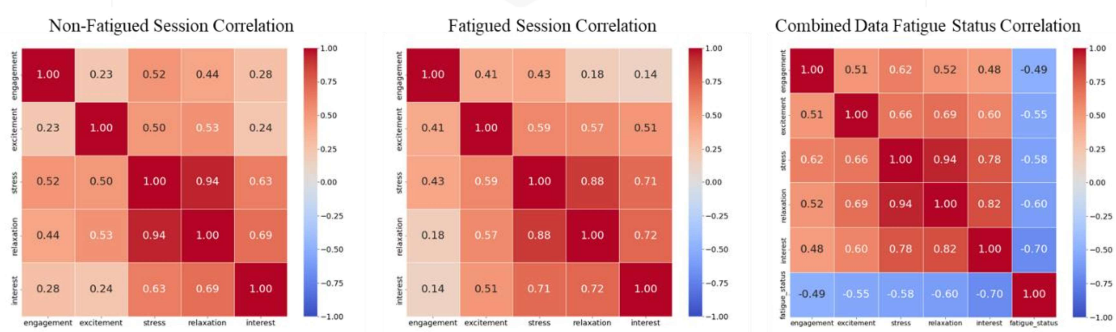


Figure 31 Subject_1008 Comparative Correlation Analysis

The correlation between metrices for Subject_1009 in Figure 32 clearly illustrates the variations between fatigued and non-fatigued sessions. For both sessions, there is a very weak association between excitement and engagement. Engagement exhibited strong correlations with relaxation (0.63) and stress (0.65), but weak correlations with interest (0.36) in session that is not fatigued. High relationships have been shown excitement with stress, relaxation, and interest in the fatigued session compared to non-fatigued session. In non-fatigued sessions, there is a strong association between stress and relaxation (0.87) and a low correlation between stress and interest (0.43) compared to fatigued session. The aggregated data reveals moderate to high correlations among metrices.

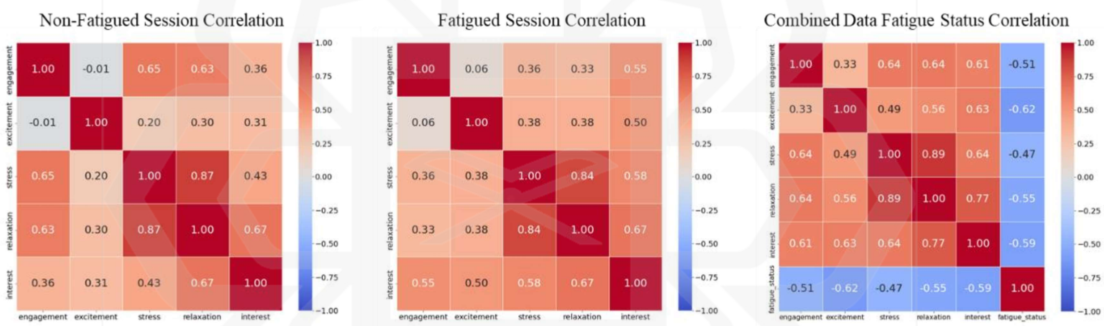


Figure 32 Subject_1009 Comparative Correlation Analysis

Subject_1010's fatigued and non-fatigued sessions can be distinguished using the correlation analysis shown in Figure 33. The majority of the correlations between metrics are weak (less than 0.50), while the interest-relaxation correlation is strong (0.75) during non-fatigued sessions and decreases in fatigued sessions (0.54). On the other hand, compared to non-fatigued sessions, the relaxation-stress correlation (0.69) is higher in fatigued sessions. With the exception of interest-relaxation correlation (0.68), there is a negative and low to moderate correlation between the metrics in the combined data from the two sessions. The fatigue status shows low to moderate correlations with the EPM. The EPMs are negatively correlated with the participants fatigue states. The highest correlation shown by excitement which is 0.53, and the lowest correlation shown by relaxation which is -0.23. The different correlations in the fatigue and non-fatigue session shows there are significance in the both sessions.

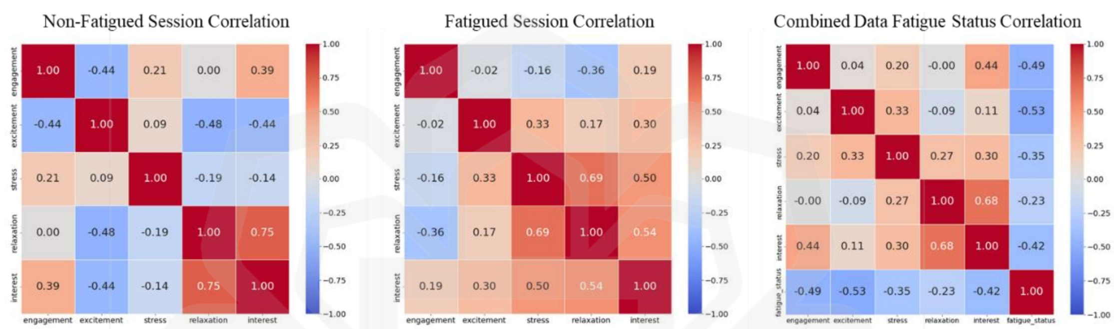


Figure 33 Subject_1010 Comparative Correlation Analysis

The correlation analysis clearly identifies the difference between the fatigue and non-fatigue data. The correlations show different pattern for the fatigue and non-fatigue session. The EPM has unique association with the fatigue status for all the individuals. Every single EPM shows different association for an individual.

4.2.3 Statical Testing

Table 12 displays the results of the Welch's t-test, which demonstrate that there are statistically significant differences for a number of variables between sessions with participants who are fatigued and those who are not. Subject_1002 shows significant differences in interest, relaxation, and engagement, but not in excitement and stress. While there is less significant difference in stress, Subject_1003 exhibits significant differences in engagement, excitement, relaxation, and interest. Subject_1004 exhibits high significance for all the EPM. Subject_1005 variability in the significance level of EPM as well where only relaxation has less significance compared to other metrics.

Table 12 Subject-Dependent Welch's_T-Test for Fatigue vs Non-Fatigue

Participants_ID	Participants_Welch's_T-test_P-value				
	Engagement	Excitement	Stress	Relaxation	Interest
Subject_1002	0.004742363	0.36881638	0.420056	2.6865E-23	0.000325
Subject_1003	0.004618094	5.7698E-13	0.114914	0.00311026	1.46E-20
Subject_1004	8.09559E-21	1.5139E-11	3.42E-17	9.9523E-22	1.52E-17
Subject_1005	8.05483E-28	2.3544E-05	2.43E-09	0.0020187	1.62E-06
Subject_1007	0.391715974	0.0078469	0.415584	0.01934068	7.66E-08
Subject_1008	4.08069E-13	3.0277E-16	8.03E-18	4.253E-20	3.32E-28
Subject_1009	8.81014E-14	1.0943E-21	5.27E-12	1.4698E-16	2.92E-19
Subject_1010	1.21902E-12	9.8421E-15	9.62E-07	0.00112291	2.24E-09

The excitement, relaxation, and interest differ significantly between the two states for Subject_1007. On the other hand, there is no noticeable variation in stress or engagement. Subject_1008 showed noticeable significance for all EPM with variability. Subject_1009 p-value is significantly less that evident changes of brain states as well.

Subject_1010 similarly show variations across all metrics between the two sessions, with significant differences observed in engagement, excitement, stress, relaxation, and interest. The EPM exhibits inconsistent significance among the participants that suggest the experience of the emotional states during the fatigue and non-fatigue session.

Statical analysis also, support the descriptive analysis and correlation analysis which identifies the significant difference between the fatigue and non-fatigue data. Almost all the EPMs shows p-value less than 0.05 which defines the significance of the both sessions. The p-value shows that there is importance of EPM indicators varies significantly throughout participants. The importance of EPM alters based on how each person feels about being tired. Although, the engagement, relaxation and interest are the most significance EPM, but there is less significance in excitement and stress compared to those three metrics. But the significance level is varied in every EPM for subject by subject, so cannot give an exclusion of any EPM for mental fatigue analysis in general. This variability in the significance of EPM evident that EPM are dynamics for a personal.

It can be summarized from descriptive analysis, correlation analysis and Statical testing that EPM are dynamics for person to person. There is no common pattern for EPM in general. The analysis also evident that the EPM are valuable feature for mental fatigue detection. The notable variability in the significance of EPM between fatigue and non-fatigue session evident EPM are the significant features to detect mental fatigue. In this reason all the EPM metrics will be used as feature for mental fatigue detection model.

4.2.4 Mental Fatigue Detection

This section illustrated the mental fatigue detection models performance including logistic regression, support vector machine, multilayer perceptron and ensemble learning model.

4.2.4.1 Logistic Regression

The Logistic Regression (LR) model performance for each participant is represented in Table 13, which demonstrates that mental fatigue detection performance ranges from modest to high. Then, Figure 34 represent the ROC AUC plots for each participant. With an accuracy of 0.67, the logistic regression model does reasonably well to detect fatigue for Subject 1002. Its F1-score is 0.65 due to its acceptable recall (0.60) and precision (0.71). With a ROC AUC of 0.75 and plots in Figure 33 evident the discrimination ability appears to be satisfactory. Considering a standard deviation of 0.06 and a mean accuracy of 0.78, cross-validation shows steady performance. The model fits the training data relatively well, according to the training score of 0.7829, and it generalizes to new data satisfactorily, according to the validation score of 0.6667.

Table 13 Subject-Dependent LR Model Performance Evaluation

Participants_ID	LR Model Performance Metrics								
	Accuracy	Precision	Recall	F1-Score	ROC AUC	Cross-Validation			
						Mean	SD	Training Score	Validation Score
Subject_1002	0.67	0.71	0.60	0.65	0.75	0.76	0.06	0.78	0.67
Subject_1003	0.82	0.74	1.00	0.85	0.94	0.87	0.08	0.86	0.82
Subject_1004	0.87	0.86	0.90	0.88	0.93	0.86	0.04	0.86	0.87
Subject_1005	0.92	0.95	0.90	0.92	0.98	0.83	0.07	0.83	0.92
Subject_1007	0.69	0.70	0.70	0.70	0.70	0.60	0.12	0.61	0.69
Subject_1008	0.79	0.77	0.85	0.81	0.89	0.82	0.06	0.82	0.79
Subject_1009	0.85	0.89	0.80	0.84	0.93	0.84	0.09	0.85	0.85
Subject_1010	0.87	0.89	0.85	0.87	0.94	0.86	0.05	0.86	0.87

The logistic regression model does reasonably well in detecting fatigue for Subject_1007, with an accuracy of 0.69. Recall and precision are balanced, contributing to its F1-score of 0.70. Having a ROC AUC of 0.70 that plotted in Figured 33, the model appears to have an acceptable level of discriminating ability, similar to Subject_1002. The model has moderate stability in cross-validation, with a mean accuracy of 0.61 and a standard deviation of 0.12. This variability suggests that performance varies slightly throughout data folds. The model fits fairly well, according to the training score of 0.61, and it generalizes to new data satisfactorily, according to the validation score of 0.69.

Then, the LR models also demonstrate strong performance in detecting fatigue for Subjects_1003, Subjects_1004, Subjects_1005, Subjects_1008, Subjects_1009, and Subjects_1010, with corresponding high accuracies of 0.82, 0.87, 0.92, 0.79, 0.85, and 0.87. These models have significantly higher F1-scores of 0.85, 0.88, 0.92, 0.81, 0.84, and 0.87, in accordance with high precision and recall values.

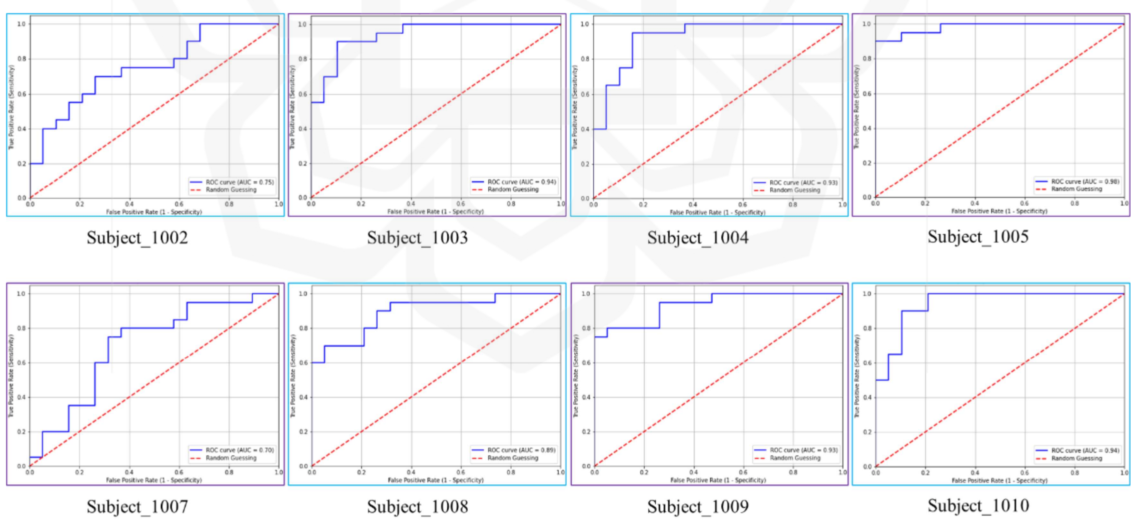


Figure 34 Subject-Dependent LR Models ROC AUC

The ROC AUC values for each of these individuals vary from 0.89 to 0.98 that visually represented in Figure 33, indicating good classifying ability. Considering high mean accuracies ranging from 0.82 to 0.87 and low standard deviations from cross-validation results show consistent performance and stability in the model's performance across different data folds. High scores for both training (ranges from 0.82 to 0.86) and validation (ranges from 0.79 to 0.92) indicate an excellent fit to the training set as well as strong generalization for new data. There is neither overfitting nor underfitting, based on the very small differences between the training and validation scores among these individuals.

4.2.4.2 Support Vector Machine

Support Vector Machine model's performance illustrated in Table 14 and Figure 35. Among all the participants the least classification accuracy for Subject_1009 is 0.77, which is good classification accuracy. The precision 0.72 and recall is 0.90 is lead to get high F1-score of 0.80. Additionally, excellent classification ability evident by ROC AUC score of 0.92 depicted in Figure 35. Then, the average (0.79) and low SD (0.07) scores of cross validations defines the consistent performance and stability in the model's performance across different data folds. The model fits good and there is no overfitting and underfitting determined by the relatively same Training (0.82) and validation (0.77) score. SVM shows excellent performance for mental fatigue detection among all other seven subjects which accuracy ranges from 0.82 to 0.95 that depicted in Table 14.

The lower (less than 80) precision among participants are 0.74 (Subject_1004), 0.76 (Subject_1007) and 0.77 (Subject_1010), but the recall is extremely high accordingly 1, 0.95 and 1. Then, Subject_1008 recall is 1 while precision is low about 0.65, and F1-score of 0.79. Other three participants precision and recall are above 0.80 that makes the F1-score high. The similar accuracy and F1-score suggest a consistent model.

Table 14 Subject-Dependent SVM Model Performance Evaluation

Participants_ID	SVM Model Performance Metrics								
	Accuracy	Precision	Recall	F1-Score	ROC AUC	Cross-Validation			
						Mean	SD	Training Score	Validation Score
Subject_1002	0.92	0.95	0.90	0.92	0.93	0.86	0.06	0.91	0.92
Subject_1003	0.90	0.83	1.00	0.91	0.95	0.85	0.04	0.90	0.90
Subject_1004	0.82	0.74	1.00	0.85	0.93	0.84	0.05	0.84	0.82
Subject_1005	0.95	1.00	0.90	0.95	0.97	0.82	0.04	0.82	0.95
Subject_1007	0.82	0.76	0.95	0.84	0.85	0.75	0.08	0.87	0.82
Subject_1008	0.82	1.00	0.65	0.79	0.93	0.80	0.07	0.80	0.82
Subject_1009	0.77	0.72	0.90	0.80	0.92	0.79	0.07	0.82	0.77
Subject_1010	0.85	0.77	1.00	0.87	0.98	0.82	0.05	0.86	0.85

Furthermore, Subject_1007, Subject_1004, Subject_1010, Subject_1003, Subject_1002 and Subject_1005 having high F1-score that ranges from 0.84 to 0.95. High ROC AUC score also ensures the classification ability of SVM that visually presented in Figure 34. The mean value of the cross validation is above 0.8 except Subject_1007 (0.75) and SD is low for all seven subjects. This cross-validation results show consistent performance and stability in the model's performance across different data folds. There is high and relatively same training score (range from 0.80 to 0.91) and validation score (range from 0.82 to 0.95) indicates that there is a significant good fit model with no overfitting and underfitting for individuals' detection model.

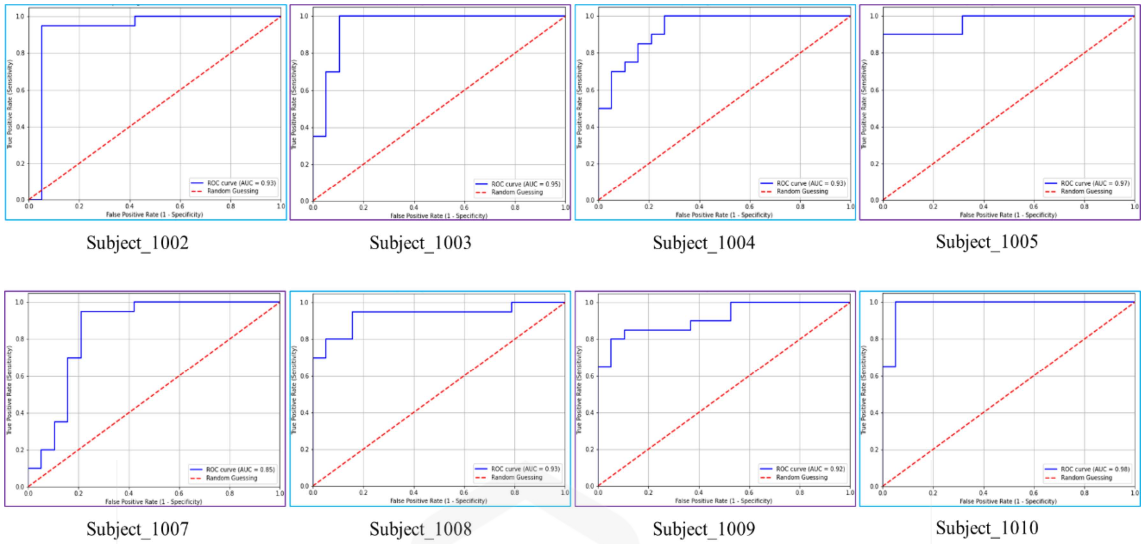


Figure 35 Subject-Dependent SVM Models ROC AUC

4.2.4.3 Multilayer Perceptron

Table 15 illustrates the feasible performance of the Multilayer Perceptron (MLP) model in detecting fatigue for Subject_1007, with an accuracy of 0.74. Its F1-score of 0.76 indicates an acceptable balanced recall (0.80) and precision (0.73). The ROC AUC plots in Figure 35 shows the classification ability visually. However, having a good ROC AUC of 0.85, the model appears to have adequate discriminating capacity.

The MLP model performs steadily in cross-validation, with a mean accuracy of 0.78 and a standard deviation of 0.06. This suggests that performance is consistent across data folds. The training score of 0.78 indicates that the model fits the training data quite well. Additionally, the validation score of 0.74 indicates that it generalizes to new data adequately.

Table 15 Subject-Dependent MLP Model Performance Evaluation

Participants_ID	MLP Model Performance Metrics								
	Accuracy	Precision	Recall	F1-Score	ROC AUC	Cross-Validation			
						Mean	SD	Training Score	Validation Score
Subject_1002	0.92	0.90	0.95	0.93	0.92	0.85	0.04	0.88	0.92
Subject_1003	0.90	0.83	1.00	0.91	0.98	0.86	0.05	0.91	0.90
Subject_1004	0.85	0.85	0.85	0.85	0.92	0.85	0.03	0.86	0.85
Subject_1005	0.85	0.82	0.90	0.86	0.97	0.84	0.05	0.86	0.85
Subject_1007	0.74	0.73	0.80	0.76	0.85	0.77	0.06	0.78	0.74
Subject_1008	0.87	0.83	0.95	0.88	0.92	0.83	0.02	0.87	0.87
Subject_1009	0.87	0.94	0.80	0.86	0.92	0.87	0.04	0.89	0.87
Subject_1010	0.85	0.89	0.80	0.84	0.98	0.86	0.04	0.86	0.85

The other 7 participants also get high accuracy on MLP classification model. The accuracy ranges from 0.85 to 0.92. The balanced and high precision (ranges from 0.82 to 94) and recall (range from 0.80 to 1) score, leads to high F1-score ranges from 0.84 to 0.93. The significant high ROC AUC scores (range from 0.92 to 0.98) define the significantly good performance of MLP. The ROC AUC is visually represented in Figure 36 that also support the good classification ability. The cross validation mean score range from 0.83 to 0.87 and very low SD also indicates the models are performed well. Relatively same and high training and validation scores show evident that all the respective models are good fit for mental fatigue detection and no evident on overfitting or underfitting model.

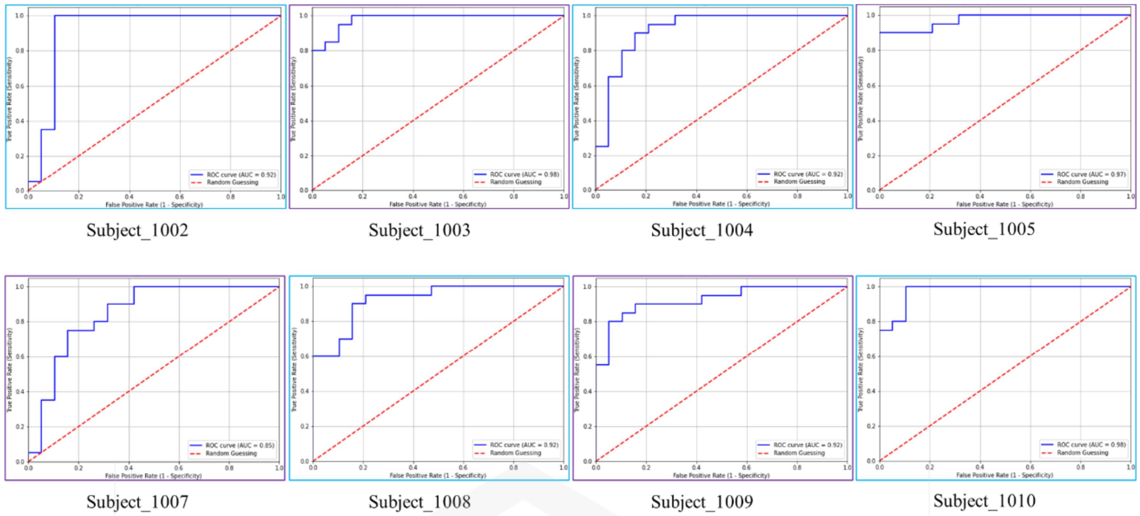


Figure 36 Subject-Dependent MLP Models ROC AUC

4.2.4.4 Ensemble Learning Model

Ensemble method consists the earlier described three based methods including LR, SVM, and SVM. The Ensemble method shows a better performance than its base models. Like the base model all the ML model evolution metrics are analyzed for model evaluation. The findings of the Ensemble detailed in Table 16 and Figure 37.

Table 16 evident that Ensemble model performance is significantly high for all the participant's model. The accuracy is ranges from 0.87 to 0.97 that defines excellent detection individual's models. Only two Ensemble modes (Subject_1007 and Subject_1005) show below 0.9 accuracy which is 0.87 for both models. But both of the models F1-score is 0.9. The precision and recall are high and relatively same on the respective models. This resulted a high F1-score very high ranges from 0.88 to 0.98.

Table 16 Subject-Dependent Ensemble Learning Model Performance Evaluation

Participants_ID	Ensemble Learning Model Performance Metrics								
	Accuracy	Precision	Recall	F1-Score	ROC AUC	Cross-Validation			
						Mean	SD	Training Score	Validation Score
Subject_1002	0.97	0.95	1.00	0.98	0.99	0.88	0.04	0.93	0.97
Subject_1003	0.90	0.86	0.95	0.90	0.98	0.91	0.04	0.95	0.90
Subject_1004	0.90	0.90	0.90	0.90	0.94	0.87	0.03	0.89	0.87
Subject_1005	0.87	0.86	0.90	0.88	0.97	0.82	0.05	0.89	0.87
Subject_1007	0.87	0.86	0.90	0.88	0.91	0.88	0.07	0.91	0.87
Subject_1008	0.90	0.86	0.95	0.90	0.93	0.86	0.04	0.89	0.90
Subject_1009	0.90	0.94	0.85	0.89	0.93	0.85	0.07	0.89	0.90
Subject_1010	0.95	0.95	0.95	0.95	1.00	0.91	0.06	0.94	0.95

A graphical representation of ensemble models depicted in Figure 37. The well performed models are evident by ROC AUC score for all the individual's model that are range from 0.93 to 1. Considering high cross validation mean and very low SD indicates the stable model performance. Every model relatively same and high training (range from 0.89 to 0.95) and testing (range from 0.87 to 0.97) score to it respective model's evident a good fit mental fatigue detection model with no overfitting or underfitting model.

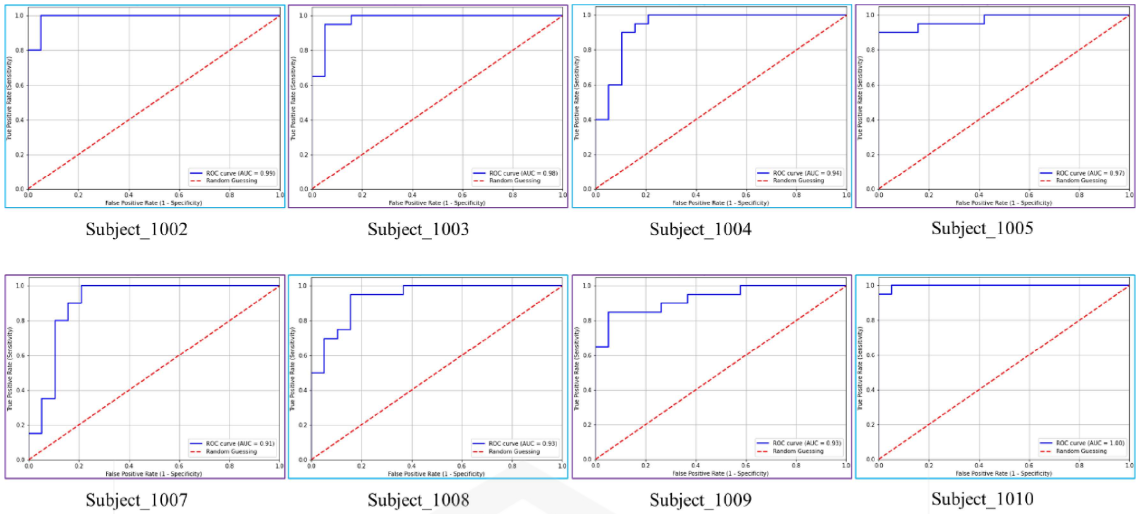


Figure 37 Subject-Dependent Ensemble Models ROC AUC

4.3 SUBJECT-INDEPENDENT ANALYSIS

This section represents subject-independent descriptive analysis, correlation analysis, statical testing and mental fatigue detection model. Subject-independent analysis offers generalizability across individuals by training and testing on variety of datasets. This approach is important for real-world applications because this model able to perform reliably for different users. Subject-independent models may perform slightly less well due to inter-subject variability. Therefore, subject-independent analysis is crucial for creating widely used application for general fatigue detection systems. This analysis did not include line graph to picturized the difference between fatigued and non-fatigued session because the data collection time is different for every participant. So, cannot see the changes of EPM over time. Subject-independent mental fatigue detection models and evaluation of models are same as subject-dependent model.

4.3.1 Descriptive Analysis

Table 17 describes that subject-independent or combined all participant's data shows differences between fatigued and non-fatigued class. There is a decline in the mean, SD and median values of all among all metrices, but almost same SD in engagement during a fatigued session compared to non-fatigued session. The variability is high for excitement, engagement and relaxation in both sessions.

The 25th percentile and 75th percentile value is high in non-fatigued session compared to fatigued session. The 25th in a non-fatigued session for engagement, excitement, stress, relaxation and interest are 0.44, 0.28, 0.3, 0.25 and 0.43 where in fatigued session respectively 0.33, 0.23, 0.27, 0.23 and 0.39. So, the 25th percentile data value is bigger in non-fatigued session and similar for 75th percentile. The data spread is high in both sessions for engagement, excitement, relaxation, but there is a decline in the distribution ranges among both sessions. Interest and stress data spread are high in non-fatigued session.

Table 17 Subject-Independent Center Tendency Fatigue (FAT) vs Non-Fatigue (Non-FAT)

Center Tendency	Engagement		Excitement		Stress		Relaxation		Interest	
	Non-FAT	FAT	Non-FAT	FAT	Non-FAT	FAT	Non-FAT	FAT	Non-FAT	FAT
Mean	0.53	0.42	0.42	0.35	0.36	0.31	0.37	0.32	0.51	0.43
SD	0.13	0.14	0.17	0.16	0.09	0.07	0.16	0.11	0.15	0.07
Median	0.53	0.42	0.43	0.33	0.36	0.31	0.34	0.3	0.48	0.43
25%	0.44	0.33	0.28	0.23	0.3	0.27	0.25	0.23	0.43	0.39
75%	0.63	0.52	0.56	0.47	0.41	0.36	0.46	0.4	0.59	0.48

4.3.2 Correlation Analysis

There are significant differences between the non-fatigued and fatigued class, according to the correlation analysis which depicted in Figure 38. Low to high positive connections have been found between two sessions. There is a high interest and relaxation correlation in non-fatigued session (0.79) which is declined to 0.43 at fatigued session. There is a moderate correlation for relaxation-stress is 0.58, but declined to 0.53. Other correlations are low in both sessions with different values. There is no negative correlation in non-fatigued session but can be found in fatigued session between relaxation and engagement (-0.13). The fatigued and non-fatigued combined data shows low and negative correlation between fatigue and metrices. However, there are moderate correlations between stress-relaxation and interest-relaxation. Other correlations are less than 0.5. The fatigue correlation with engagement, excitement, stress, relation and interest are -0.37, -0.21, -0.3, -0.18 and -0.33.

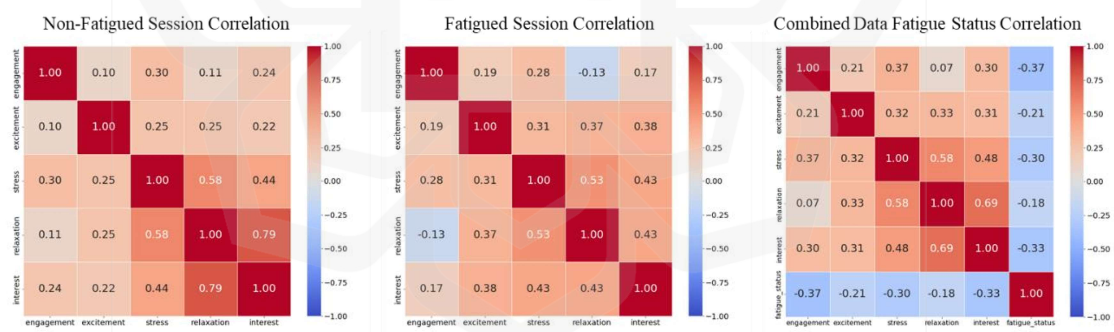


Figure 38 Subject-Independent Comparative Correlation Analysis

The correlation for fatigue and non-fatigue shows two different association, which identify there is enough evident to evident to measure fatigue and non-fatigue session. The combined dataset also shows there is an association between fatigue status. The EPM shows negative correlations. That suggest that if there is an increase in the level of fatigued status then then EPM scores will be low, if the individuals fatigued status low then the EPM sore will be high. This suggest that the EPM can be used as feature for mental fatigue detection. Statical testing presented next section to justify the EPM as feature for mental fatigue detection.

4.3.3 Statical Testing

Table 18 displays the results of the Welch's t-test, which demonstrate statistically significant differences for a number of variables between fatigued and non-fatigued sessions. There is a significant difference evident by the p-value between both session in EPM metrics. The engagement shows the most significance (p-value=3.29E-52) among all the EPM. Then interest also shows second most significance level (p-value=3.92E-39) for between fatigue and non-fatigue session. Stress also shows very high significance (p-value=2.69E-32). Similarly, the significance also very high for excitement (p-value=4.32E17) and relaxation (p-value=3.92E-13). The noticeable significance of EPM enhances the reliability of EPM as features that effectively differentiate between states. The significant differences exhibit a high accuracy model that can capture variations across fatigue status. This reduces risk of underfitting because EPM as the input data is relevant to the target mental fatigue classification.

Table 18 Subject-Independent Welch's_T-Test for Fatigue vs Non-Fatigue

EPM metrics	Welch's T-test_P-value
Engagement	3.29E-52
Excitement	4.32E-17

Stress	2.69E-32
Relaxation	3.92E-13
Interest	3.79E-39

4.3.4 Mental Fatigue Detection

This section presents the subject-independent mental fatigue detection model performance including LR, SVM, MLP and Ensemble learning model.

4.3.4.1 Logistic Regression

With a precision of 0.70, recall of 0.67, and accuracy of 0.68, the Logistic Regression model performed satisfactorily that depicted in Table 19. A satisfactory F1-score is 0.69 which is calculated by recall and precision in proportion. In Figure 38 is the graphical representation of the ROC AUC. The ROC AUC score of 0.74 shows the discriminating capacity of the LR model. The mean accuracy during cross-validation was 0.69, exhibiting a relatively stable performance with a standard deviation of 0.02. The detection model with moderate fit is considered to have training and validation scores of 0.69 and 0.68, respectively. There is no indication of an overfitting or underfitting model because the training and validation scores do not differ significantly.

4.3.4.2 Support Vector Machine

Table 19 shows similar to LR, the SVM model perform moderately to detect mental fatigue where the accuracy is at 0.68. The balance and moderate score of precision (0.66) and recall

(0.77) gives a F1-score of 0.72 which also indicates the satisfactory model performance. The ROC AUC score is 0.74 and Figure 38 shows the same significance of classification ability. The cross validation means 0.71 is considerably acceptable with a low standard deviation 0.03 which shoes the consistent performance in various data fold. The training score and validation score is relatively same and acceptable that indicates a moderate level of fit detection model.

4.3.4.3 Multilayer Perceptron

In Table 19, MLP shows a higher accuracy of 0.73 for mental fatigue detection than SVM and LR. The precision and recall are relatively high and balanced that calculated F1-score of 0.74 that defines better model performance. ROC AUC score of 80 is high and define the model efficiency as classification of mental fatigue that is visualized in Figure 39. Cross-validation mean of 0.74 and significantly low SD of 0.02 indicated stable model performance. The training (0.77) and validation (0.75) scores are close and considerably high that defines the model as moderately good fit. This also indicates there is no overfitting and underfitting issue in the model.

4.3.4.4 Ensemble Learning Model

The ensemble learning model perform well as described in Table 19. The accuracy is 0.75 with precision 0.76 and recall 0.75. The F1-score is 0.76 that defines a relatively good perform model. To support F1-score, the ROC AUC score is calculated at 0.83 that defines the model can classify fatigue.

Table 19 Subject-Independent Mental Fatigue Detection Model Performance Evaluation

Classification Model	Accuracy	Precision	Recall	F1-score	ROC AUC	Cross-Validation			
						Mean	SD	Training Score	Validation Score
LR	0.68	0.70	0.67	0.69	0.74	0.69	0.02	0.69	0.68
SVM	0.68	0.66	0.77	0.72	0.76	0.71	0.03	0.72	0.68
MLP	0.73	0.73	0.75	0.74	0.80	0.74	0.02	0.76	0.73
Ensemble	0.75	0.76	0.75	0.76	0.83	0.75	0.02	0.77	0.75

The discrimination ability of model graphically presented in Figure 39. The cross validation mean value of 0.75 and low SD of 0.02 is defined the consistency of the model performance. The training (0.77) and validation (0.75) defines a considerably good fit detection model where the model is not an overfitting or underfitting model.

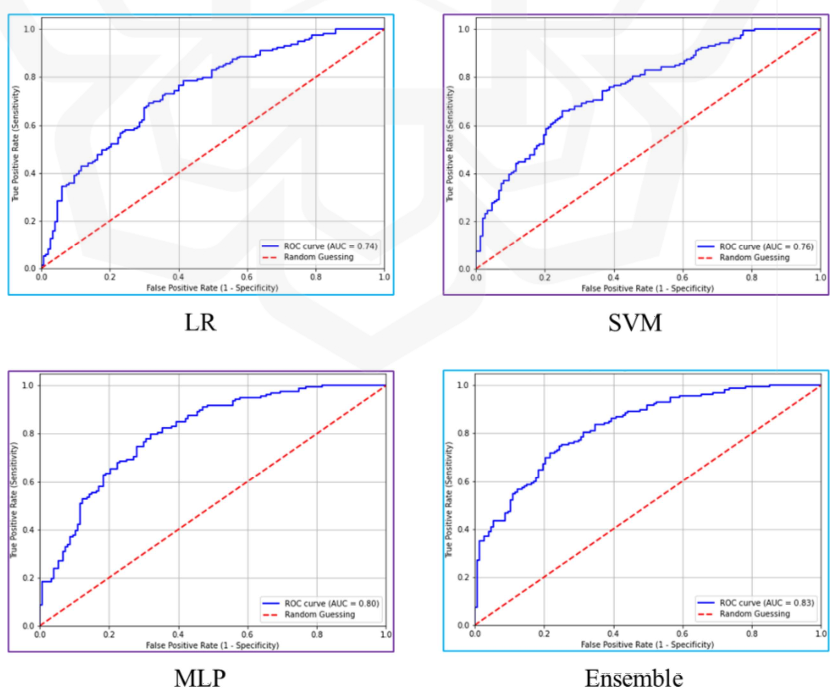


Figure 39 Subject-Independent Mental Fatigue Detection Models ROC AUC

4.4 DISCUSSION

This section maps the results to the relevant research questions. This research aims to develop a mental fatigue detection technique using emotional metrics. There is a literature review to identify existing mental fatigued detection and identify the available emotional metrics for different mental conditions including mental fatigue. After selection of participants a guided data collection protocol has been followed by getting the approval from IREC as earlier mentioned. Then a screening is performed to exclude data based on exclusion criteria. After data labelling a statical analysis performed to understand difference between fatigued and non-fatigued session. Then multiple algorithms used to detect mental fatigue. The evaluation of model includes analysis of several model evaluation metrics and cross validation.

4.4.1 Data

There was total 10 participants. The participants were good in health. All of the participants are full-time students and prefer face-to-face learning. As self-reported answer, the participants are feeling fatigue more during online learning session compared to face-to-face learning. There is total five person has the history of COVID-19 and one person is taking medicine. Two person data is excluded from analysis after observing there CFS score. Because one person starts with the caseness of fatigue and another person did not fall in fatigue. Among 8-person five-person pre-CFS score is 0 and post-CFS score are 5, 7, 4, 6 and 6. This suggest, they start without fatigue level but they experience different fatigue level in online learning session. Subject_1003 score updated from 1 to 9 the most significant among all the participants. Subject_1002 and Subject_1008 participants CFS score changes from 2 to 6. So, it is clearly seen that almost every participant's fatigued experience is in different level. The comparative description summarized in the Figure 40.

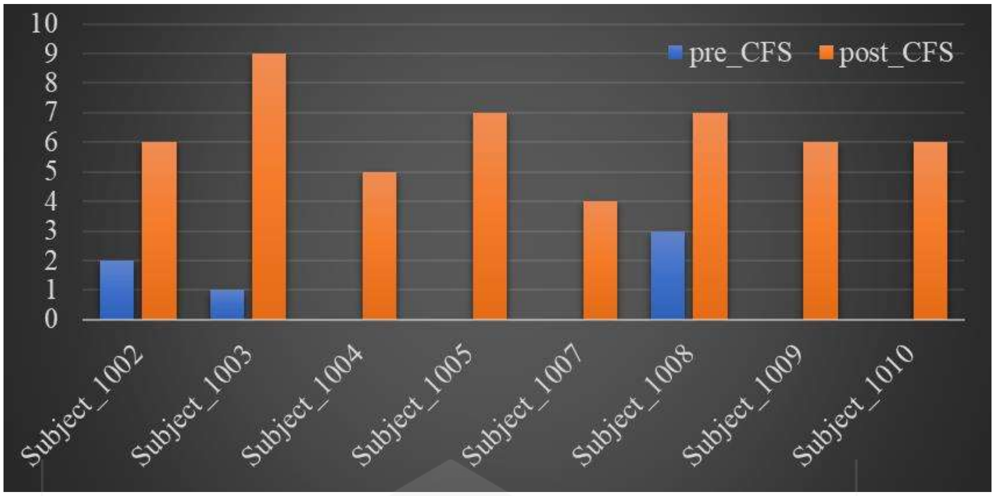


Figure 40 Pre-CFS Score vs Post-CFS Score

This research used five emotional EPM including: engagement, excitement, interest, stress and relaxation. The EPM data is collected in a CSV file format. The data divided into two sections with same ration. First half of the data which is 15 minutes 50 seconds, is labeled as non-fatigue and second half which is 16 minutes, is labelled as fatigue data. This portion is giving the equal chance for both sessions. The variability and different score of CFS shows the changes of mental states, which identify the persons was in fatigue states during the learning session. Then multiple algorithms used to detect mental fatigue. The evaluation of model includes analysis of model evaluation metrics and cross validation.

4.4.2 Mental Fatigue Detection

Mental fatigue detection involves the self-report CFS. The structure is described in section 3.7.1 and findings detailed in section 3.8.2 and Table 8. Based on the CFS score two participants were excluded from ML model development because one of them did not experience fatigue in the whole learning session and another participant was already in fatigue caseness level. Total eight participants data used to develop the mental fatigue detection model. The data set were labelled in 15 minutes 50 second- and 16-minutes interval where first period labeled as non-fatigue and second period labeled as fatigue. This labelling shows a significant difference between two sessions that evident by descriptive analysis, correlation analysis and statical testing. Descriptive analysis shows a significant difference in mean, median, standard deviation, and data spread between both sessions.

4.4.2.1 Subject-Dependent Mental Fatigue Detection Model

The subject-dependent findings show in Table 11 which written in section 4.2.1. The engagement for all participants shows a decrease in mean, median in the fatigued session. The mean and median values are almost same in the respective session; this suggests the consistency and difference in both sessions. SD values are higher in fatigued session for Subject_1002, Subject_1003, identical for Subject_1005, Subject_1007, and Subject_1008; then lower for Subject_1004, Subject_1009 and Subject_1010. So, the data variability pattern is not common among participants. The 25th percentiles and 75th percentiles with different values are higher in non-fatigued session compared to the fatigued session. Additionally, the data spread ranges (75th percentile – 25th percentile) are greater in fatigued session only Subject_1010 shows identical. These suggest that the data values for 25th percentile and 75th percentile will be higher value in non-fatigue but the data spread can be low in this session.

The excitement means and median values are relatively close that suggest the data are not influenced by the outliers for the respective session. Through the fatigued session the mean and median values are high for Subject_1002, Subject_1003, Subject_1007, but low for Subject_1004, Subject_1005, Subject_1008, Subject_1009 and Subject_1010; these shows different pattern than the engagement. Similarly, SD values are not in similar pattern the both session among all participants. The 25th percentile (except Subject_1007) and 75th percentiles values are significantly high in non-fatigued session among all the participants. It is note that the data spread ranges (75th percentile – 25th percentile) does not show similar pattern among all the participants.

The stress means and median values are high in non-fatigued session for all subjects except Subject_1007 which shows identical mean value and almost same median values in both sessions. There is a common patter like engagement and excitement in the relatively closed value of mean and median. So, the stress values are not influenced by outliers. The SD values indicate low data variability in fatigued session for all participants except Subject_1005 and Subject_1010. Similar to excitement, through the non-fatigued session, the 25th percentiles are high in non-fatigued session except for Subject_1007, but 75th percentiles are high among all participants. Total six participants show a lower data spread ranges (75th percentile – 25th percentiles) except Subject_1009 and Subject_1005.

It interesting to observed the relaxation mean, median and SD values. The values are low in fatigued session for six participants except Subject_1002 and 1007. There is a closed mean and median values on its respective session which indicates the relaxation is not affected by outliers. It is also interesting that there is different pattern than other EPM metrics in terms of 25th percentile and 75th percentiles. The 25th percentiles are high in non-fatigued session for Subject_1002, Subject_1003, and Subject_1007. Also, there are low 75th percentile values in non-fatigued session for Subject_1002 and Subject_1005. The data spread ranges (75th percentile – 25th percentile) are low in fatigued session except Subject_1002, Subject_1005 and Subject_1010. The mean and median values are different among the participants interest level. The mean and median values are higher in fatigued session for Subject_1002 and Subject_1007. Similar to other metrics the interest data are not influenced by outlier because the mean and median values are close. The data SD values indicate identical data variability in Subject_1010, higher for Subject_1009 and low for other six subject in the non-fatigued session. The 25th percentiles and 75th percentiles are high in value for Subject_1002, Subject_1007 in the fatigued session while other subjects experience lower value. It is interesting that only Subject_1005 shows high data spread range (75th percentile – 25th percentile) in the fatigued session.

In the subject-dependent analysis the EPM show dynamic behavior. There is no outlier influence among all the participants EPM data. The variability of EPM in both sessions are different and fall in different ranges (mean+SD and mean-SD) suggest that the participants are normally high in emotional metrics or EPM values in non-fatigued session compared to fatigued session. The data spread ranges calculated by the difference between 75^{th} and 25^{th} percentiles, also show the level of emotional metrics are low in the EPM value. These prove the Emotiv's claim (Emotiv, 2024) on their EPM performance are correct as example "0" means no engagement and "1" means high engagement.

The claim of difference between the two-time periods of non-fatigue and fatigue are evident by correlation analysis detailed in section 4.2.2. The finding for subject-dependent analysis shows that there are changes on association among metrices are significantly different in both sessions. The associations among participants are unique. Therefore, there are different correlation between fatigue and EPM metrices for each subject. Most of the EPM are shows negative correlation with fatigue status. The correlations are negation and moderate to high between fatigue and all metrices for Subject_1004, Subject_1008, Subject_1009 and subject_1010. Besides the negative association Subject_1007 shows a positive low association with fatigue status; relaxion shows a positive association with fatigue status for Subject_1005; Subject_1003s' fatigue experience is positive correlation with excitement and Subject_1002 observation shows a low to moderate positive association between EPM (excitement, interest and relaxation) and fatigue status. The correlation suggests there are difference between fatigue and non-fatigue session as well as all the EPM has dynamic relationship with fatigue status.

The significance of the both sessions further observed by a Welch's T-test p-value which detailed in section 4.2.3 and depicted in Table 12. Relaxation and interest show a significance (p-value < 0.05) for the both session among all participants. Engagement and excitement also show high significance between both session but excitement is less significance (p-value $=0.37$) for Subject_1002 and engagement is low significance for Subject_1007 (p-value $= 0.39$). Similarly, stress shows significant difference for six participants but Subject_1002 and Subject_1003 and Subject_1007 shows ess significance for both sessions. The Welch's T-test evident that there is a difference between the fatigued and non-fatigued session.

It can be summarized that the data labeling as non-fatigued and fatigued is significantly good to detect mental fatigue that is evident by descriptive analysis, correlation analysis and a statical analysis for the both sessions. It is important that the EPM are unique and dynamic among the participants. This lead to develop a subject-dependent ML model for mental fatigue detection. There are four different model is developed including LR, SVM, MLP and Ensemble, where LR, SVM, MLP are the base model of Ensemble learning model.

4.4.2.2 Subject-Independent Mental Fatigue Detection Model

As earlier evident that the emotions are dynamic and unique for all subjects and that lead to subject-dependent or personalized ML model development. Therefor it also important to have analysis subject-independent analysis to observed the ability of a general model for everyone to detect mental fatigue. The participant's fatigued and non-fatigued data combined and generates a new subject-independent dataset. The dataset shows a significant difference between non-fatigue and fatigue observations that detailed in section 4.3 and depicted in Table 17, and Table 18.

Table 17 defines that the mean and median values of all EPM metrics shows a decrease in the fatigued session data compared to non-fatigued session. The mean and median values are closed to each other for all EPM which indicates the EPM data are not influence by the outliers. The SD shows a large variability in both sessions, but non-fatigue data variability is slightly higher than the non-fatigue session among all EPM except engagement. Engagement shows a slightly increase variability in fatigued session compared to non-fatigued session. The 25th percentile and 75th percentiles values are greater in non-fatigued session among all EPM metrics. The data spread ranges (75th percentile –

25th percentile) are slightly higher for all EPM matrices in non-fatigue session. These descriptive analyses suggest that there is a common pattern in mean and median values which identified a decrease a EPM metrics mean and median are low in fatigue. The 25th percentiles and 75th percentiles also show a same pattern like mean and median.

The detailed correlation analysis described in section 4.3.2 and graphically presented in Figure 37. It is clearly seen that there is a difference between the non-fatigue and fatigued labelled data. The association among the metrices are different for both sessions. Then all the EPM metrics has shown a low and negative correlation with the fatigue status. These suggest that there is a difference between two sessions but not a strong difference exist. Table 18 is the representation of the Welch's T-test shows a significantly high difference for all metrices between fatigued and non-fatigued session evident by p-value (<0.05). Since there is a difference between fatigue and non-fatigue labelled data, a ML detection algorithm can be used to detect mental fatigue for a subject-independent analysis. The model structure describes in section 3.6.2.

After a comprehensive analysis including descriptive analysis, correlation analysis and statical testing decided to develop ML model. This important analysis evident that there is a fatigued and not fatigued data pattern available. In this purpose the LR, SVM, MLP and Ensemble model structure used to detect a generalized mental fatigue detection model. The choice of algorithms is from three different domain including LR is a statical analysis-based model, then SVM is widely used ML model and MLP is another popular deep learning model. Therefore, different model's comparative performance analysis is helpful to select the best model.

4.4.3 Mental Fatigue Detection Model Evaluation

The evaluation process includes the ML model evaluation metrices including accuracy, precision and recall to get F1-score, ROC AUC score and cross-validation. The detection model performance is detailed in section 4.2.4. The subject-dependent mental fatigue detection models are shows reliable. The comparative performance for all subjects is depicted in Figure 40 which included the accuracy, F1-score and ROC AUC score. The data details are presented in Table 13, Table 14, Table 15 and Table 16. It is clearly seen from Figure 41 that the accuracies of all LR models are moderate to high. Then evaluation metrics (accuracy, F1-score and ROC AUC) visualization breakdown to compare the mental fatigue detection model performance in APPENDIX G.

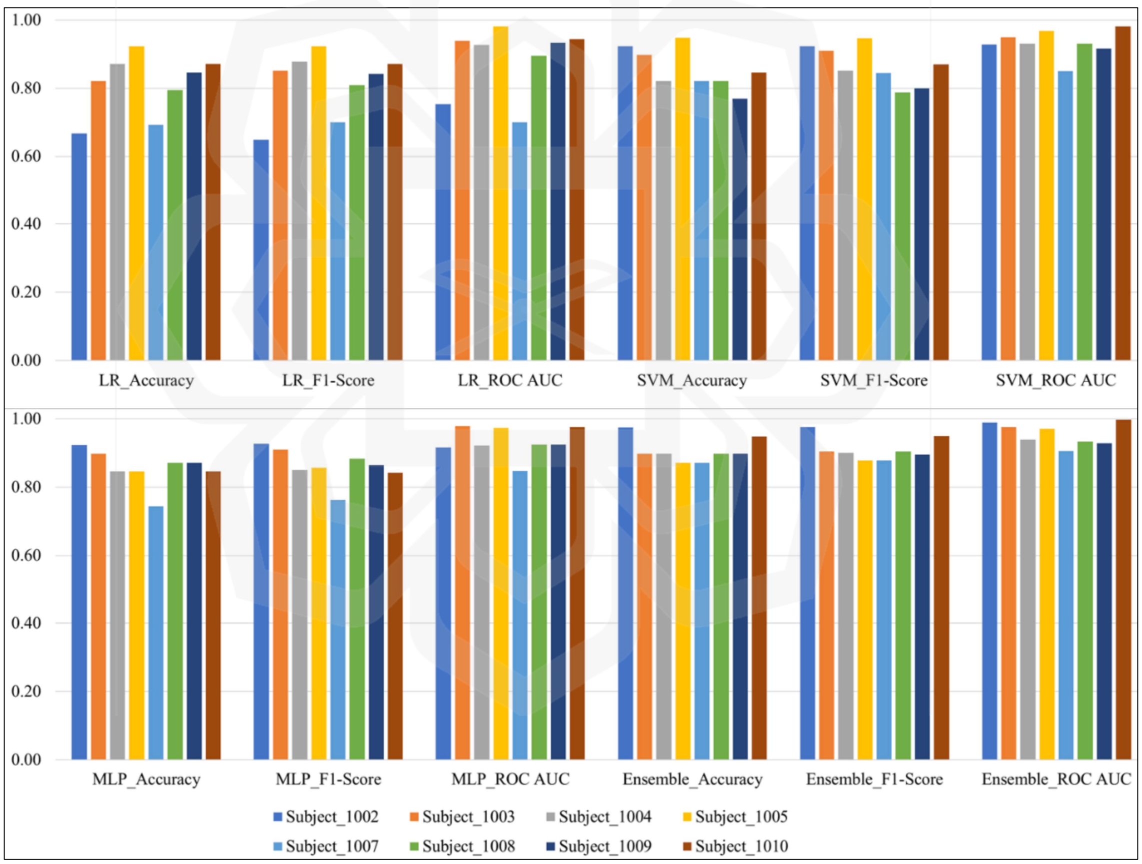


Figure 41 Subject-Dependent ML Models Performance Comparison

The moderate to high score of F1-score on the respective model that calculated from balanced precision and recall support the accuracy of moderate to high. The ROC AUC also shows the similar pattern. Table 13 evident a consistent performed LR model by presenting a high cross-validation mean score and low SD value among the respective LR models. Beside this Figure 41 also indicates a well fit and generalized LR model by presenting the validation and training scores which are high and relatively same on the respective LR model. So, there are no overfit or underfit model identified.

Figure 41 and Figure 42 indicates all the SVM, MLP and Ensemble models are consistent, reliable, well-fit and generalized model that evident by respective models high and consistent accuracy, F1-score (balanced high precision, and recall), ROC AUC score, training and validation score. It is clearly seen that the SVM perform better than LR; then MLP perform better that SVM and Ensemble model perform better than MLP. The subject-dependent Ensemble models show the best performance compare to its respective base LR, SVM and MLP models. Then cross-validation metrics (mean, trainin and validation score) visualization breakdown to compare the detection model performance in APPENDIX G.

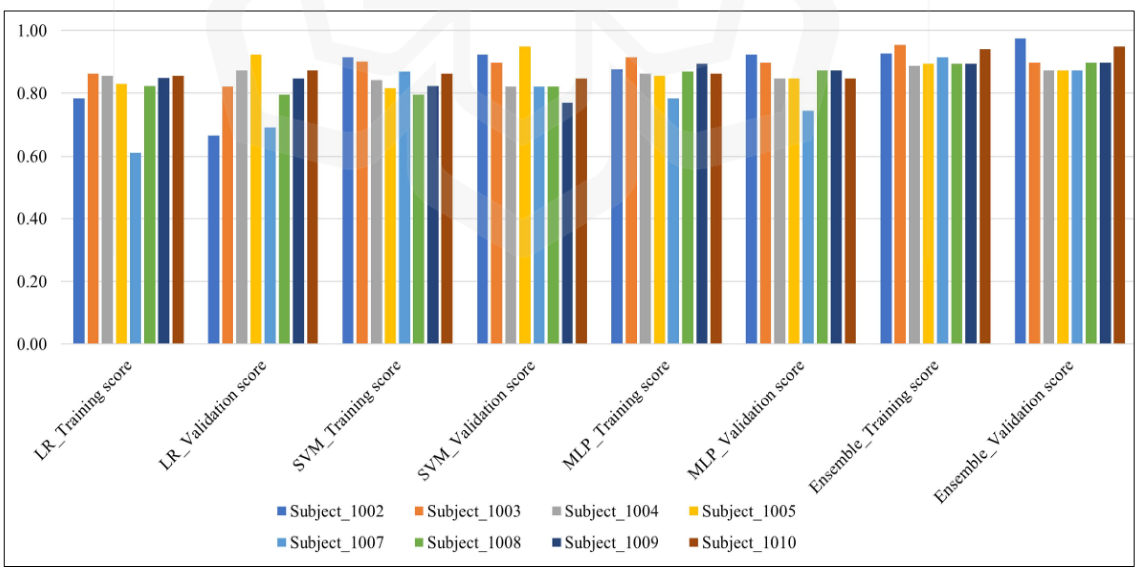


Figure 42 Subject-Dependent ML Models Cross-Validation Scores Comparison

The subject-independent LR, SVM, MLP and Ensemble models shows a moderate performance that illustrated in Table 19 under section 4.3.4.2 and at Figure 43. There is a moderate score of accuracy in the all four model performance. The balanced precision and recall (Table 19) provide a better F1-score for respective model that indicates better performance model. Additionally, the reliability of the four models is evident by ROC AUC score. The ML evaluation metrices shows gradually increase in the score of accuracy, F1-score and ROC AUC score, that defies the reliable and consistency of the model.

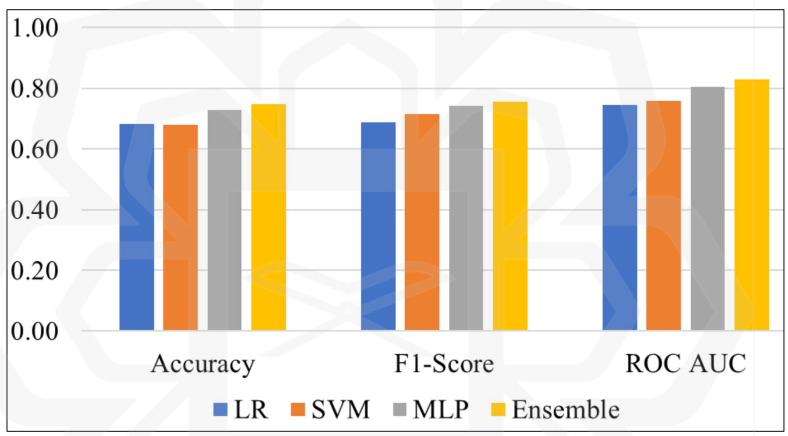


Figure 43 Subject-Independent ML Models Performance Metrics Comparison

The cross-validation mean score are moderately high and significantly low SD at Table 19 indicates reliable and consistent for all four-models performance. This supports the accuracy, F1-score and the ROC AUC score. Figure 44 present the training and validation score of cross-validation. It is clearly observed from Figure 43 and Figure 44 that Ensemble is the most well performed model than the MLP, then SVM and LR model. The model performance is moderate but shows a consistence performance. The training and validation scores are moderate and closed to each other in the respective model that indicates the model is fits well; hence there is no overfitted or underfitted model.

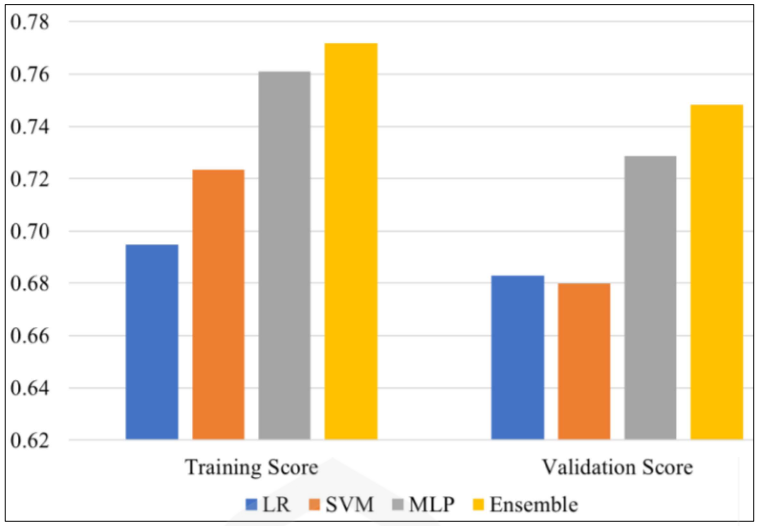


Figure 44 Subject-Independent ML Models Cross-Validation Scores Comparison

4.4.3.1 Mental Fatigue Detection Model: Subject-Dependent vs Subject-Independent

Subject-dependent models are well performed model for all of the models including LR, SVM, MLP and Ensemble. This comparison is evident by Figure 41, Figure 42, Figure 43 and Figure 44. Only two LR models is under 70 but same as subject-dependent model that assess by the accuracy, F1-score, ROC AUC score and cross-validation score. The personalized model shows Significantly better performance in SVM as well as for MLP and Ensemble method.

The performance of subject-dependent model is better than the subject-independent because of the unique behavior of an emotional metrices for the individuals. The descriptive analysis comparison between Table 11 and Table 17 shows different nature of emotional metrices. The mean of an emotional metrics for an individual is different than the mean of combined general data because different person shows different data ranges from 0 to 1. Then there is a different variability measure by SD, data spread ranges measure by difference of 75th percentiles and 25th percentiles. When the data combined and make a general dataset that gives different data pattern observation among all the EPM. This also can be evident by the correlation analysis because every person experiences different correlation among EPM metrices in both fatigue and non-fatigue session. The association between fatigue and EPM metrices also shows different pattern for the subject-dependent analysis which detailed in section 4.2.2. In contrast, the generalized data showed different relationship significance between fatigue and not fatigue session. Also, different association significance between EPM and fatigue statues that detailed in section 4.3.2. Furthermore, the statical shows that the significance level between two sessions is significantly different for personalized and subject-independent analysis.

Although, subject-independent models are well performed than the subject-independent model, but the subject-independent models performance shows acceptable and consistency evident by relatively same accuracy, F1-score, ROC AUC score and cross-validation score. The validation score where training and validation score closed to each other indicates not overfit or underfit model in both cases subject-independent and subject-dependent. A subject-dependent analysis and ML model for mental fatigue detection is the better choice than the subject-independent.

4.4.3.2 Ensemble Learning Model vs Base Model

Ensemble learning models require more computational power because it is more complex than its based model or compare to a standalone model. Therefore, this model should be considered to accurately find the complex relation between features like EEG-derived EPM metrics. This study shows that the Ensemble ML models are significantly better performed model than an individual ML model for mental fatigue detection that depicted in Figure 41, Figure 42, Figure 43 and Figure 44. The subject-independent performance level is in the sequence of Ensemble, MLP, SVM and LR that illustrated in Figure 43 and Figure 44. In the subject-independent analysis the winner is Ensemble method. Then in subject-dependent ML model, all the base model individually shows a reliable, consistent and well-fit model that shows the generalization behavior. Therefore, there are two subject-dependent LR model accuracy is less than 70 (67 and 69), one LR less than 80 (79), one SVM model less than 80 (0.77) and one MLP less than 80 (74). But there is no Ensemble accuracy fall under 80. The lowest Ensemble accuracy is 87 and higher accuracy is 0.97. The ranges of accuracy for Ensemble are 87 to 97 which is significantly high, where the accuracy ranges of LR is 67 to 87, accuracy range of SVM is 77 to 95 and accuracy ranges of MLP is 0.74 to 0.92. Therefore, Ensemble model shows similar findings to all base model in the training and validation score which are high and closed to each other that indicates the consistent and generalization performance with well-fit model.

Ensemble model shows the most reliable and consistent performance that the individual model in the both analysis whether it is subject-dependent or subject-independent.

CHAPTER FIVE

CONCLUSION

5.1 SUMMARY

This research aims to develop an Ensemble learning model for mental fatigue detection using EEG based emotional metrics. These research objectives are addressed though a systematic approach (detailed in Chapter 3) involving a comprehensive literature review, statical analysis, ML model development, and evaluation of the mental fatigue detection model for online learners. This study also included subject-independent and subject-dependent mental fatigue detection model analysis and development.

The study extensively reviewed existing ML based mental fatigue detection model to identify the mental fatigue detection pipeline, emotional metrices, self-reported mental fatigue assessment tool and mental fatigue detection model. It is clearly identified that mental fatigue detection pipeline involved signal acquisition, pre-processing, feature extraction and classification or detection model which detailed in section 2.3. ML based mental fatigue detection model applied in many contexts but it is rare for online participants detailed at Table 3. There are many EEG device available to collect the noise sensitive EEG signal with different sampling rates where Figure 2 showed NeuroScan, and BrainAmp are most popular headset. There are many complex pre-processing and feature extraction available. Butterworth bandpass filter and ICA are the most popular signal processing methods can be observed from Figure 3. Then Figure 4 presented the overview on feature extraction and shows that FFT, PSD and CSP are most popular feature extraction methods. Numerous ML model applied for me mental fatigue detection; among those SVM with different kernel methods, Neural Network or deep learning-based models are the most popular ML model depicted at Figure 5. A review on a mental fatigue assessment toll is included to identify the self-reported mental fatigue assessment and CFQ is used to assess the learning session for this research.

There are numerous studies on EEG based emotional states detection model available. This research review identified only Emotiv provide emotional metrics know as performance metrics through their API and those are used in different mental conditions analysis. The research also identified the existing statical methods for data analysis and ML model performance metrics for model evaluation.

Emotiv performance metrics (EPM) which are emotional metrics data utilized for mental fatigue detection. An ethical consideration is acknowledged to collect EEG emotional metrics that approved by IREC. A self-assessment tool CFQ is used to include-exclude the participants from analysis. Total two participants were excluded among 10 participants because one of them was already in fatigued level other participants did not experience fatigue according to CFS scale which detailed in section 3.8.2. To ensure data quality and reliable data labelling a comprehensive data analysis performed to observe the significance between fatigue and no-fatigue EPM data.

This analysis includes a descriptive analysis for center tendency (mean, median, SD, 25th percentile and 75th percentile) observation, correlation analysis to understand the behavior of each EPM metrics and a statical analysis to find the significance between the fatigue and non-fatigue session. These three comprehensive evaluations show consistent and reliable findings in the significant difference between both session and ensures the data quality to apply ML for mental fatigue detection. Three ML algorithms used to mental fatigue detection model includes LR, SVM with polynomial kernel, MLP and Ensemble model. The structure of these ML model described in the section 3.6 and finding detailed in chapter four. The model detection done for both subject-independent and subject-dependent. These selected ML model evaluations were performed using ML model performance metrics.

The ML model's performance metrics includes accuracy, precision, recall, F1-score, ROC AUC score and cross-validation. The details of these metrics described in section 3.7.5 and the findings detailed in chapter four. The LR shows moderate to high and consistent performance among the subject-dependent model. LR shows acceptable moderate level of performance in the subject-independent mental fatigue detection. SVM shows better performance than LR, similarly MLP shows better performance than SVM and ensemble model performance is the best among all the ML model for both subject-dependent and subject-independent analysis. All the metrics are relative same score shows which defines their consistency of mental fatigue detection. A comparative performance of the model picturized in Figure 41 (subject-dependent) and Figure 43 (subject-independent). All the ML model is well-fit model there is no overfit or under fit model identified and shows a generalization behavior on the respective model and this claim picturized in the Figure 42 and Figure 44 by comparing the training score and validation score. The comprehensive evaluation of ML model identified a successful detection of mental fatigue model. This research also find that an Ensemble ML model is better than the single models in both case subject-dependent and subject-independent mental fatigue detection. Although subject-independent shows acceptable and consistent performance on the Ensemble mental fatigue detection model, but further study is required to get more reliable result. A personalized analysis and Ensemble based mental fatigue detection model is suggested, especially when required to apply mental fatigue reduce strategy.

The findings have significant practical implications for enhancing the online learning experience by addressing mental fatigue using EPM and Ensemble learning model. This can apply in real-world educational setting to monitor remote students cognitive states in real-time. Future study needs to integrate this model into the existing learning management system to test the effectiveness on large and diverse populations.

This ML based mental fatigue detection model using emotional metrics leads to develop a brain-computer interface for mental fatigue intervention for online learners. This study shows that emotional metrics can be used to develop an Ensemble model for mental fatigue detection. This detection model development can be applied in real-time mental fatigue detection and then apply intervention when mental fatigue detected. The light, monitor brightness, and temperature high influence on mental fatigue as earlier evident in literature review. An intervention can be a sound as notification system, control of monitor brightness or control of room temperature or combination of multiple intervention modality that may reduce the mental fatigue.

5.2 LIMITATIONS

Efficient, reliable and comfortable hardware and API is desired in the research. The Emotiv device is reliable device for EEG data analysis which is evident by numerous studies. Therefore, the Emotiv Insight sensor tips head are not comfortable for the user. The participants feel uncomfortable while they wearing for a long time as reported by the participants. Therefore, different people have different head size. So, the headset needs to be flexible enough to adjust in the correct electrode placement. The Emotiv also shows a connection interruption where there is a message about wireless connection interference. This is a cause of data loose from the cognitive session. The Emotiv needs to improve the consistent connection for better performance.

Emotiv metrics are the measure of different emotional state. The calculation or process of EPM is not revealed by Emotiv. Additionally, there are not many studies available based on the EPM analysis. A known mechanism of emotional metrics can help to identify more insightful findings. A future study needs to be addressing the EEG based emotional metrics development. The Emotiv API provides six emotional metrics. This research cannot collect the Focus metrics which is one of the important metrics in this research. The API need to be well maintained to ensure the consistency and reliability. Furthermore, Emotiv provides only six metrics, but there are more emotional states in human brain. Future study needs to focus on getting data from different device as well as more emotional states to detect mental fatigue.

The EPM data sampling interval highlights another Emotiv limitations in this research. Rapid changes in emotional states are crucial for comprehending the delicate dynamics of mental fatigue in real-time, that may be captured by the high sampling rates than the current sampling rate of 1 Hz (one observation every 10 seconds) frequency. To overcome this constraint and allow for a more sophisticated assessment of mental fatigue states, future research could investigate with the high frequency sampling rate.

The data labelling is based on half period's segmentation, where first half is labelled as non-fatigue and second half as fatigue. Although this research shows 15 minutes is good enough for identifying brain states changes to fatigue states, however a comprehensive investigation is required with different time frame to identify the most optimal mental fatigue exhibits time. In this reason need have future study with high sampling rate of EPM to have enough observation especially in a shorter time frame compared to 15 minutes segmentations. To support this a video of the person facial expression can be recorded. So, the video analysis may help to identify fatigue status and these findings can be compared with the EEG data to label the fatigue and non-fatigue session.

This research includes all the supervised machine learning algorithms. Although this research shows that LR, SVM, MLP and Ensemble method perform reliably and consistent, but still require more analysis to find the generalize model mental fatigue detection. A comprehensive analysis with unsupervised and reinforcement learning ML approach may provide insightful and reliable mental fatigue detection model.

Another limitation of this research is lack of gender diversity because all the participants are male. This questions the generalizability of the findings that cannot account potential of EPM patterns for mental fatigue detection. Future studies should investigate more diverse participants to ensure the models effectiveness for mental fatigue detection.

This research identifies EPM as dynamics for every individual. Therefore, EEG-based emotional metrics engagement and excitement exhibits more significant variations than other metrics during fatigued sessions in the subject-independent analysis. Future studies need to investigate the mechanisms of driving these variations potentially by incorporating additional physiological or behavioral data. This may provide more comprehensive understanding the influence of EPMs that exhibits mental fatigue and enhance the interpretability of subject-independent ML models.

A subject-independent analysis needs to be further study because without robust and continuous study may have wrong impact. Every person's EEG signal and emotional states are dynamic in nature, so without proper investigation the reliability of the mental fatigue detection may lead to wrong impact on individuals brain waves. Although this study shows a reliable performance on the mental fatigue detection using emotional metrics, but further investigation with different approach is required to evident that emotional metrics can be used to detect mental fatigue.

5.3 FUTURE WORK

While a person becomes fatigued, an intervention method can be useful (Brandtner et al., 2022; Feldman & Dreher, 2012; Friese et al., 2012) to reduce fatigue. Intervention refers to Interfering with the result of anything, especially a condition or process, in order to avoid danger or enhance functionality (Sörensen et al., 2006). It can be described as intercession, treatment, assessment, breaking-in, preventative, mediation. Interventions are activities taken to alter a person's behavior, emotional state, or sentiments in order to bring about change. Interventions aim to address the underlying causes of mental conditions as well as their consequences. There are many different intervention strategies available that target different kinds of problems. Intervention strategies included Cognitive-behavioral, relaxation, social skills training, social support, mindfulness, meditation, psychoeducational, acceptance and commitment therapy, and resilience training (Axelsen et al., 2020; Friese et al., 2012; Sörensen et al., 2006; Zhu et al., 2020). Intervention like physical exercise, watching video (Friese et al., 2012), mental calculation task (WANG et al., 2017), changes of light brightness (L. Liu et al., 2005), temperature (Jin et al., 2022) and changes of monitor contrast (Kalra & Karar, 2022) can reduce fatigue. Different mental disorders can be treated using neurofeedback as an intervention technique (Enriquez-Geppert et al., 2019; Hariharan et al., 2018; Korfmacher et al., 2022; Loriette et al., 2021).

Neurofeedback (NF) is a form of biofeedback that uses cognitive strategies to improve healthy brain activity (Marzbani et al., 2016). Although there is controversy over neurofeedback's effectiveness, therefore research has shown that patients can benefit from the treatment (Enriquez-Geppert et al., 2019; Hariharan et al., 2018; Loriette et al., 2021; Noohi et al., 2017). NF is brain training that modifies brain signals to enhance cognitive function. NF strategy can be a design of sound or video or anything that alter brain signal to improve mental health (Kamiya & Francisco, 2011). The most common NF signal is the EEG signal. Usually, EEG based BCI uses scalp-mounted sensors to record electrical impulses from the brain and then uses sound or visual displays to provide feedback. NF as an intervention technique may reduce mental fatigue and improve daily performance.

There is a systematic literature review performed (F. Hossain & Yaacob, 2022) to understand the existing neurofeedback techniques for different mental conditions. The data collection performed in March 2022 and used google scholar which is not a database, it is a literature search engine in the internet to find article from various sources. Google Scholar provides extensive literature articles access. Google Scholar has a lack on the rigorous indexing standards of databases such as IEEE Xplore or PubMed or WoS. The search result may include non-peer-reviewed sources or duplicates that requires extra careful screening process. The search keywords were "mental fatigue intervention", "fatigue and intervention", "BCI and intervention", "eeg neurofeedback", "neurofeedback treatment", "BCI and neurofeedback" and "neurofeedback framework". There was total 240 articles from the search result and 19 articles were selected. A review is performed on the self-reported mental fatigue assessment to identify the existing tool for mental fatigue measurement. This helps to measure individual's fatigues level subjectively. The emotional states for mental conditions detection were reviewed and identified that Emotiv performance metrics are the emotional metrics that use in different mental conditions which is detailed in chapter two.

Neurofeedback is a strategy for intervention (Breteler et al., 2010; Lipp & Cohen Kadosh, 2020; Lubar, 1997). The step-by-step process for neurofeedback training to stimulate brain activity is outlined in this review. The procedures include signal acquisition, pre-processing, feature extraction, classification algorithm, and feedback signal generation.

The first four steps are signal acquisition, pre-processing, feature extraction and classification algorithm was identified as Basic functions of BCI (Wolpaw et al., 2002). These four steps are used to identify mental conditions. Furthermore, signal acquisition, pre-processing, feature extraction, and classification/pattern recognition were defined as the mental fatigue detection processes in the next section 2.3. So, these steps are detailed in the next step from the point of mental fatigue detection strategy. The summary of the different mental condition analysis depicted in Table 20.

Table 20 Mental Condition Detection

Brain Condition	Signal Acquisition	Pre-processing	Feature Extraction	Classification / Pattern	References
Improve brain activity performance	Signal: fMRI Subjects: 24 Device: MRI Scanner	EMA, Kalman filter	Fast Fourier Transform (FFT)	Threshold	(Baqapuri et al., 2021)
Attention training	Signal: EEG Subjects: 22 Device: Enobio	Low pass FIR, SSP	Spatial & temporal features	Logistic Regression	(Tuckute et al., 2019)
Alzheimer	Signal: EEG Subjects: 5 Device: Emotiv Epoc	Low-pass Butterworth	FFT	Proximal Policy Optimization (PPO)	(Ai et al., 2021)
Autism	Signal: EEG Subjects: 12 Device: BrainLink	Outlier filter	Threshold	eSense	(Mercado et al., 2018)
Improving Attention	Signal: EEG Subjects: 30 Device: Biosemi	Low-pass, High-pass Butterworth	Alpha ratio	Linear Discriminant Analysis (LDA)	(Arvaneh et al., 2019)

Attention diversion	Signal: EEG Subjects: 12 Device: g.USBamp	Butterworth Band-pass filter	Temporal & spectral, combination of both	Support vector machines (SVM)	(Aliakbaryhosseinabadi et al., 2020)
Effects of gender on the learning outcomes	Signal: EEG Subjects: 142 Device: Nexus-10	IIR Butterworth	SMR, Alpha, Theta	Threshold	(Wood & Kober, 2018)
Effortless awareness	Signal: EEG Subjects: 32 Device: Biosemi	IIR Butterworth Bandpass filter	Beamformer	Low-resolution brain electromagnetic tomography (LORETA)	(van Lutterveld et al., 2017)
Attention Deficit Hyperactivity Disorder (ADHD)	Signal: EEG Subjects: 114 Device: BrainquiryPET 4.0	Low pass filter	FFT	Threshold	(Krepel et al., 2020)
ADHD	Signal: EEG Subjects: 336 Device: Quikcap, Nuamps	Low pass filter	Wavelet transform	Threshold	(Arns et al., 2018)
Attention monitoring	Signal: EEG Subjects: 10 Device: OpenBCI	SURE threshold	FFT	Improved random forest-(IRF)	(B. Wang et al., 2021)
Chronic Tinnitus	Signal: EEG Subjects: 26 Device: BrainAmp	Butterworth Band-pass, ICA filter	FFT	Threshold	(Güntensperger et al., 2019)
Increase of Mindfulness	Signal: EEG Subjects: 45 Device: g.tec Amplifier	Notch, ICA, Bandpass filter	FFT	Threshold	(Navarro Gil et al., 2017)
ADHD	Signal: EEG Subjects: 112 Device: Biosemi	Notch, Bandpass filter	FFT	Threshold	(Janssen et al., 2020)

Table 21 provides a summary of neurofeedback signal generation. The two types of neurofeedback mentioned in this review are (i) EEG neurofeedback and (ii) Functional Magnetic Resonance Imaging (fMRI) neurofeedback. The popularity of EEG neurofeedback is mostly due to its accessibility, affordability, and effectiveness. Neurofeedback tries to improve a particular brain pattern to lessen a condition's symptoms. The learner actively participates in neurofeedback and electrical simulation, then continuously modifies strategies to shift his or her brain activity in the desired direction. Following the feature extraction from EEG signals, a feedback signal is produced that converts the features into a sensory input that is then presented to and processed by the learner. The effectiveness of neurofeedback training is influenced by the learner characteristics. When the targeted aspect of brain activity meets a predetermined threshold or condition, a feedback signal is sent out. It is challenging to design and create NF protocols for neurofeedback signal generation to solve the problem (Enriquez-Geppert et al., 2019; Van Doren et al., 2019).

Numerous NF protocols have been found. The majority of the stimuli combine visual and auditory components. Depending on the needs and research goal, various feedback modalities were developed in a distinct study. Theta/beta (Gloss et al., 2016; Saad et al., 2015) is the protocol that is used the most, followed by threshold, alpha, and other NF protocols. A resting EEG indication is the theta/beta ratio (TBR). By dividing the power of the slower frequency band (theta) by the power of the faster frequency band, the theta/beta power ratio was calculated independently for each electrode (beta). When comparing a signal's amplitude to a fixed or variable predefined value, a threshold is used. A cutting-edge therapy called alpha neurofeedback trains patients to increase their alpha strength to enhance their creative thinking and lessen depression. A successful implementation of the NF protocol requires the training design, EEG signal acquisition equipment, required feature extraction, establishment of a threshold, feedback mode, and characteristic or complexity of feedback signal.

Training design takes time and is very important (Mauro & Cermak, 2006). This stage requires consideration of pre-post training, the quantity and duration of sessions, and the type of study blinding (single, double, or triple). A factor is the EEG signal device's configuration, operation, and capabilities. The number of electrodes and their placement must also be considered because different electrode placements in neurofeedback have relatively distinct effects (Rogala et al., 2016). The basis for feature selection is typically evidence for a relationship between oscillation and cognition. Neurofeedback will function better when it depends on customized characteristics, determining the level of individualization for feature extraction is also critical (Moretti et al., 2004).

A particular threshold can be established using a variety of calculations based on variations from a resting condition. Thresholds are established for participants based on their z-score in relation to a normative sample rather than only examining their unique brain activity (Hernandez-Gonzalez et al., 2011). Feedback can be given visually, audibly, or in combination. Thus, practical considerations and participant characteristics are typically considered while selecting the feedback modality. The feedback signal displays the required brain waves' performance in comparison to a benchmark. The complex stimuli that are delivered have an impact on the properties of the feedback signal. It is challenging to design complex stimuli since the participants' reactions to them may be unpredictable. In clinical applications, complex stimuli like films are utilized. An ensemble of stimuli is preferable than single stimulus (Aliakbaryhosseinabadi et al., 2020; Wood & Kober, 2018).

Table 21 Neurofeedback Signal Generation

Stimuli Type	Stimuli (Feedback Modality)	NF For	Type Of NF	NF Protocol	References
Visual stimuli	VE-Shooter Game (movement speed)	Improve brain activity	fMRI NF	SAM	(Baqapuri et al., 2021)
Visual stimuli	Composite images (an overlay of male and female images)	Attention training	EEG NF	Sham NF	(Tuckute et al., 2019)
Visual & Audio stimuli	Immersive 3D virtual reality animal environment	Alzheimer	EEG NF	Reinforcement Learning (RL)	(Ai et al., 2021)

Audio stimuli	A mix of river, ocean waves, and background noise	Post-Traumatic Stress Disorder (PTSD)	EEG NF	Alpha/theta	(Noohi et al., 2017)
Visual stimuli	A BCI video game	Autism	EEG NF	Threshold	(Mercado et al., 2018)
Visual stimuli	P300-based speller	Improving attention	EEG P300 NF	Alpha ratio	(Arvaneh et al., 2019)
Visual & Audio stimuli	Visual cue, auditory cue (oddball task)	Attention diversion	EEG NF	Alpha, Theta	(Aliakbaryhosseinabadi et al., 2020)
Visual & Audio stimuli	Moving bar	Effects of gender on the learning outcomes	EEG NF	Threshold, SMR	(Wood & Kober, 2018)
Visual stimuli	Posterior Cingulate Cortex (PCC) activity	Effortless awareness	EEG NF	Gamma	(van Lutterveld et al., 2017)
Audio stimuli	Eyes open, Eyes closed, & auditory oddball task	ADHD	QEEG-informed NF	SMR, TBR, SCP	(Krepel et al., 2020)
A light & sound attenuated room	Eyes Open (EO), Eyes Closed (EC)	ADHD	EEG NF	TBR and alpha peak frequency (APF)	(Arns et al., 2018)
Visual and auditory	Three Serious Games	Attention monitoring	EEG NF	SAM	(B. Wang et al., 2021)
visual intervention	City car driving	Emotion and fatigue	EEG NF	Frontal Alpha Asymmetry (FAA)	(Sapta Pramana & Arlini Puspasari, 2020)
Audio stimuli	Eyes Open (EO), Eyes Closed (EC)	Chronic Tinnitus	EEG NF	Alpha/Delta	(Güntensperger et al., 2019)
Visual stimuli	count the change of color tones	Increase of Mindfulness	EEG NF	Alpha	(Navarro Gil et al., 2017)
Visual stimuli	Stop-signal task (SST)	ADHD	EEG NF	Theta/beta	(Janssen et al., 2020)
Visual and audio stimuli	Video (images) clip and sound	Improve NF performance	EEG NF	Beta/alpha threshold	(Sho'ouri et al., 2020; Shourie et al., 2018)
Visual	Cycle ergometer, depleting cognitive task	Improves cycling time	EEG NF	Alpha	(Mottola et al., 2021)

There are many factors related to mental and physical fatigue (Åkerstedt et al., 2004; Jang et al., 2021; Otani et al., 2017). This review focused on the effect of light and temperature. The data was extracted from google search in December 2022 and depicted in Table 22.

Table 22 Light and Temperature Effect on Mental Fatigue

Domain	Factors	Result	References
Video viewing	Light	medium level screen brightness caused minimum visual fatigue	(Kalra & Karar, 2022)
During Chemotherapy	Light	High level of fatigue associated with less light exposure	(L. Liu et al., 2005)
Performance and physiological arousal	Light	Less sleepy, more vital, and happier, if exposed to bright light	(Smolders & de Kort, 2014)
Traumatic Brain Injury	Light	High-intensity blue light reduced fatigue & daytime sleepiness	(Sinclair et al., 2014)
Effects of Display on visual fatigue	Light	Higher luminance contrast in dark mode, less visual fatigue	(Xie et al., 2021)
Multiple Sclerosis	Temperature	High temperature, high fatigue	(Bol et al., 2012)
Cycle ergometer exercise	Temperature	High temperature, high fatigue	(Gonzàlez-Alonso et al., 1999)
Office task	Temperature	Hogh temperature, high fatigue	(Tanabe et al., 2007)
Different Protection States of nurse	Temperature	High temperature and high protection high mental fatigue, low temperature and low protection had the least	(Jin et al., 2022)

Temperature and Lights are correlated to fatigue. Different lighting environment like display contrast/brightness, luminance (Kalra & Karar, 2022; Sinclair et al., 2014; Xie et al., 2021), and blue light intensity (L. Liu et al., 2005; Smolders & de Kort, 2014) was designed to analyze the mental fatigue. Three color temperatures were designed for fatigue assessment, where result shows a color temperature of 4500 K, fatigue was the lowest (Lu et al., 2021). A color temperature of 2700 K shows more fatigue, where a color temperature of 7500 K showed more productive but for prolonged working induce fatigue. Multiple sclerosis participants feel more fatigued with heat, but in UV light and day temperature there was no effect on Mental fatigue (Bol et al., 2012). In a same level of task performance, participants feel more fatigue in a hot environment than the thermal environment (Tanabe et al., 2007). Treatment with high-intensity blue light therapy reduced fatigue and daytime sleepiness during the treatment phase while controlling for age, gender, and baseline depression. Lower intensity yellow light therapy did not show these improvements (Sinclair et al., 2014). The effects of low, medium, and high brightness levels on visual fatigue were examined. Keeping the screen brightness at a medium setting resulted in the least amount of visual fatigue in a space that is well lighted by LED luminaries (Kalra & Karar, 2022).

This research shows that a ML model can detect mental fatigue for online learners. A review on neurofeedback signal generation shows that neurofeedback can improve cognitive performance by altering current brain pattern. There are many factors that has influence on mental fatigue for a participant including light, and temperature (detailed section 2.6). Although, there are various technique to detect mental fatigue through BCI (Hossan et al., 2017; Myrden & Chau, 2017; H. Zeng et al., 2019). BCI widely adopted for Neurofeedback in the treatment of various mental states, it is uncommon for mental fatigue intervention (Ai et al., 2021; Enriquez-Geppert et al., 2019; Noohi et al., 2017; Proost et al., 2022; Soylu et al., 2021). However, there is no real-time mental fatigue intervention model using brain-computer interface. Indeed, this research propose a real-time mental fatigue intervention framework by adapting EEG derived emotional metrics based mental fatigue detection method and a NF modality as a mental fatigue intervention method.

This research identified that EEG based emotional metrics can detect mental fatigue and presented a pipeline for mental fatigue detection in Figure 11. Then Section 2.5 identified neurofeedback pipeline for different mental conditions treatment including signal acquisition, pre-processing, feature extraction, classification model and feedback signal generation. This neurofeedback pipeline depicted at Figure 45. The NF pipeline can be divided into two main sub-working part. First part is the mental condition detection includes signal acquisition, pre-processing, feature extraction and detection or classification algorithms for the desired mental condition analysis. Next part is NF signal generation includes control of feedback modality. Neurofeedback modality can be auditory or visual (as example sounds, videos and images) based on real-time brainwave activity measured by EEG which describe in section 2.5 and depicted in Table 21.

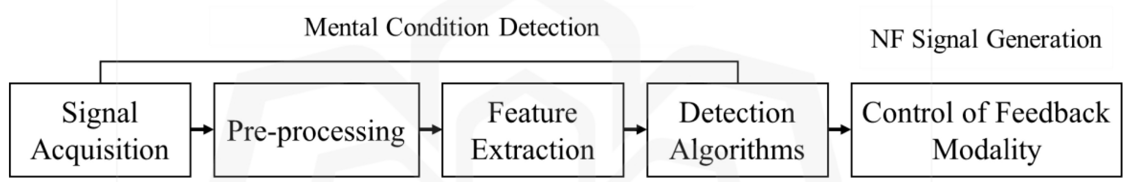


Figure 45 Neurofeedback Pipeline

Emotiv headset processed the EEG signal to get the EPM which are the emotional metrics. Since the EEG signal acquisition, preprocessing and feature extraction to get the EPM is done by Emotiv and deliver in 0.1 Hz means every 10 seconds there will be one instance of EPM metrices. By adapting EPM for mental detection from Figure 11 with neurofeedback pipeline at Figure 45 can get the real-time mental fatigue intervention pipeline which presented in Figure 46.

The intervention pipeline can be divided in two sections: (i) mental fatigue detection and (ii) intervention signal generation strategy. Figure 46 used the intervention instead neurofeedback because the purpose of the intervention to make the person active by reducing participants fatigue on the real-time online learning. So, there is no therapeutic modality like neurofeedback where need the participants need continuous follow-up neurofeedback session to improve attention or reducing fatigue.

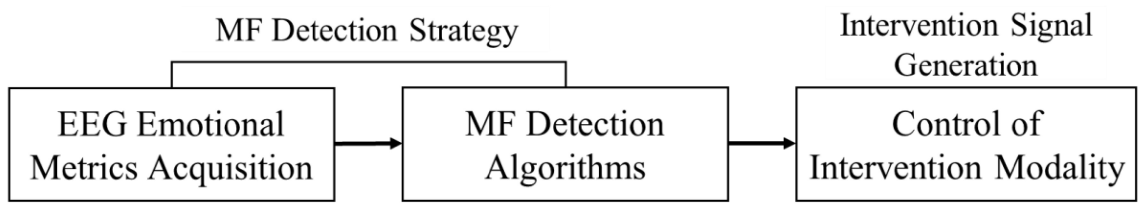


Figure 46 EEG-Derived Emotional Metrics Based Mental Fatigue Intervention Pipeline

Besides EEG derived emotional data, an EEG based Real-time mental fatigue intervention model pipeline can be done by including signal acquisition, pre-processing, feature extraction, mental fatigue detection technique (classification/pattern recognition/threshold) and control of intervention modality. Real-time mental fatigue intervention will acquire by the combination of EEG derived mental fatigue detection pipeline and neurofeedback pipeline. Figure 46 illustrated the mental fatigue intervention pipeline. This leads to design the general architecture real-time mental fatigue intervention model using brain-computer interface which includes EEG emotional metrics acquisition, Mental fatigue detection algorithm and intervention signal generation by controlling intervention modality. The architecture of the proposed EEG derived emotional metrics based real-time mental fatigue intervention model is depicted in Figure 47.

The model is designed as a closed loop for mental fatigue intervention. The system consists of two closed-loop functions: (i) mental fatigue detection and (ii) mental fatigue intervention. At first the model detects the mental fatigue. The continuous process of emotional metrics acquisition (EPM), and mental fatigue detection algorithm occur until the model detects the mental fatigue during the cognitive task period. If there is mental fatigue identified the model will activate the intervention strategy which can include an alert system or a control of intervention like monitor brightness or both. The mental fatigue detection and control of intervention modality occur until the detected mental fatigue reduced.

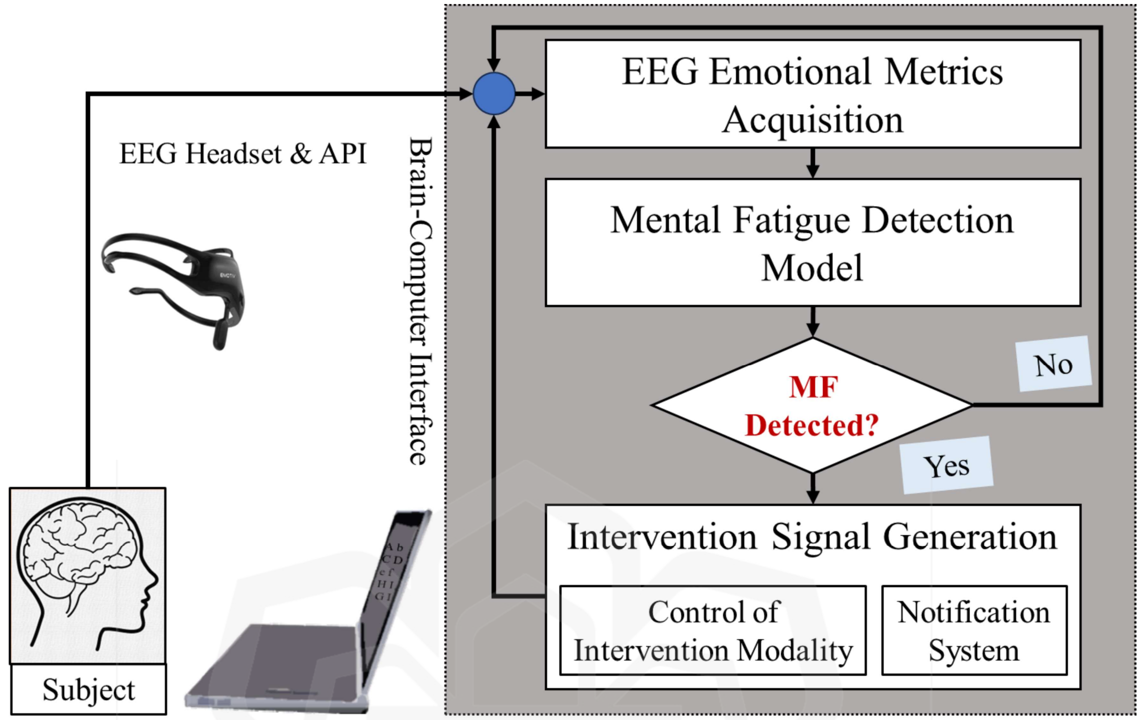


Figure 47 EEG-Derived Emotional Metrics Based Real-Time Mental Fatigue Intervention using Brain-Computer Interface

The intervention will be applied when a participants fall in fatigue during the learning session to make them active. In section 2.6 described that the temperature, and light has an influence on mental fatigue that detailed in Table 22. Online participants affected by monitor brightness and room temperature. So, an intervention can be done by the controlling of monitor brightness or control of room temperature or a both. These need to be further research to find the effectiveness of mental fatigue intervention framework with control of different intervention modality.

In conclusion, this research evident that EEG based emotional metrics can be used to detect mental fatigue using machine learning technique. Emotional metrics are unique and dynamic for an individual. The association between fatigue and emotional metrics are significance. The findings also suggest that, an Ensemble method rather than a single ML model can be a generalized model for detection of mental fatigue of an individual. The concept on Ensemble method with three different type of algorithms shows a reliable and consistent performance. A numerous number of participants mental fatigue detection required to evident the generalization behavior of the developed Ensemble model. Future study may design an Ensemble method with different machine learning model to identify the more efficient model for Mental fatigue detection.

5.4 RESEARCH CONTRIBUTIONS

Several noteworthy contributions to the field of study are made by this research. In the first place, it improves knowledge by developing a ML model that uses emotional states generated from EEG for the detection of mental fatigue in online learners. This model not only advances knowledge of the connection between emotions and mental fatigue, but it also gives educators and online learning environments a useful tool for tracking and promoting cognitive health.

This research shows that changes on the emotional state identify mental fatigue. This study also suggests a robust future investigation on mental fatigue detection using emotional metrics is required which may not only help to address mental fatigue but may improve emotional states also for the online learners. This suggests a concurrent investigation of two different cognitive states which are emotions and mental fatigue to improve both cognitive states at the same time.

This study shows that EPM are reliable features on the detection of mental fatigue. This research also identified a time period of 15 minutes, which is enough to be getting fatigued for an online learner. This research evident that Ensemble learning model are reliable compared to its base model on the detection of mental fatigue. Then, neural network model like MLP is more reliable to identify more complex pattern among the complex features like emotional states.

Additionally, this study provides practical ways for creating successful learning experiences and alleviating fatigue by presenting a comprehensive framework for mental fatigue intervention based on the research findings. A comprehensive literature review is published on existing EEG-based mental fatigue detection (Yaacob et al., 2023). This research promotes the field by publishing an extensive literature review paper (F. Hossain & Yaacob, 2022) that offers a useful synthesis of the current state of knowledge and a proposal of a mental fatigue intervention framework (F. Hossain & Yaacob, 2023a) for future studies in this area of study.

5.5 PUBLICATIONS

There are three publications from this research. The publications are following:

Hossain, F., & Yaacob, H. (2023). A proposal on a brain-computer interface model for real-time mental fatigue intervention. *Journal of Islamic, Social, Economics and Development (JISED),* 8 (55), 303 - 315.

H. Yaacob, F. Hossain, S. Shari, S. K. Khare, C. P. Ooi and U. R. Acharya, "Application of Artificial Intelligence Techniques for Brain–Computer Interface in Mental Fatigue Detection: A Systematic Review (2011–2022)," in *IEEE Access,* vol. 11, pp. 74736-74758, 2023, doi: 10.1109/ACCESS.2023.3296382.

Hossain, F., & Yaacob, H. (2022, September). Review on Signal Generation for Neurofeedback. In *2022 10th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-8). IEEE.

REFERENCES

Ai, Y., Ben Abdessalem, H., & Frasson, C. (2021). Zoo Therapy for Alzheimer's Disease with Real-Time Speech Instruction and Neurofeedback System. *Frontiers in Artificial Intelligence and Applications*, *338*, 84–94. https://doi.org/10.3233/FAIA210079

Åkerstedt, T., Knutsson, A., Westerholm, P., Theorell, T., Alfredsson, L., & Kecklund, G. (2004). Mental fatigue, work and sleep. *Journal of Psychosomatic Research*, *57*(5), 427–433. https://doi.org/10.1016/J.JPSYCHORES.2003.12.001

Akin, M. (2002). Comparison of Wavelet Transform and FFT Methods in the Analysis of EEG Signals. *Journal of Medical Systems 2002 26:3*, *26*(3), 241–247. https://doi.org/10.1023/A:1015075101937

Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. *ISRN Neuroscience*, *2014*, 1–7. https://doi.org/10.1155/2014/730218

Aliakbaryhosseinabadi, S., Farina, D., & Mrachacz-Kersting, N. (2020). Real-time neurofeedback is effective in reducing diversion of attention from a motor task in healthy individuals and patients with amyotrophic lateral sclerosis. *Journal of Neural Engineering*, *17*(3), 036017. https://doi.org/10.1088/1741-2552/AB909C

Al-Nafjan, A., Hosny, M., Al-Ohali, Y., & Al-Wabil, A. (2017). Review and Classification of Emotion Recognition Based on EEG Brain-Computer Interface System Research: A Systematic Review. *Applied Sciences 2017, Vol. 7, Page 1239*, *7*(12), 1239. https://doi.org/10.3390/APP7121239

Ansari, M. F., Edla, D. R., Dodia, S., & Kuppili, V. (2019). Brain-Computer Interface for wheelchair control operations: An approach based on Fast Fourier Transform and

On-Line Sequential Extreme Learning Machine. *Clinical Epidemiology and Global Health*, *7*(3), 274–278. https://doi.org/10.1016/J.CEGH.2018.10.007

Arns, M., Vollebregt, M. A., Palmer, D., Spooner, C., Gordon, E., Kohn, M., Clarke, S., Elliott, G. R., & Buitelaar, J. K. (2018). Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *European Neuropsychopharmacology*, *28*(8), 881–891. https://doi.org/10.1016/J.EURONEURO.2018.06.002

Arvaneh, M., Robertson, I. H., & Ward, T. E. (2019). A P300-Based Brain-Computer Interface for Improving Attention. *Frontiers in Human Neuroscience*, *12*, 524. https://doi.org/10.3389/FNHUM.2018.00524/BIBTEX

Asif, I., Javed, M., Noor, A., Tauseef, S., Abrish, ;, Abbasi, H., Muhammad, ;, Hassan, F., & Azim, E. (2023). Real-Time Neurocognitive Monitoring of Executive Functioning Tasks with and without Gamification Using Emotiv Insight Real-Time Neurocognitive Monitoring of Executive Functioning Tasks with and without Gamification Using Emotiv Insight Annals of Neurology and Neurosurgery. *Ann Neurol Neurosurg*, *2*(1), 1007. http://meddocsonline.org/

Axelsen, J. L., Kirk, U., & Staiano, W. (2020). On-the-Spot Binaural Beats and Mindfulness Reduces the Effect of Mental Fatigue. *Journal of Cognitive Enhancement*, *4*(1). https://doi.org/10.1007/s41465-019-00162-3

Aznan, N. K. N., & Yang, Y. M. (2013). Applying Kalman filter in EEG-based brain computer interface for motor imagery classification. *International Conference on ICT Convergence*, 688–690. https://doi.org/10.1109/ICTC.2013.6675451

Bailey, D. L., Maisey, M. N., Townsend, D. W., & Valk, P. E. (2005). *Positron Emission Tomography* (Vol. 2). Springer-Verlag. https://doi.org/10.1007/B136169

Baqapuri, H. I., Roes, L. D., Zvyagintsev, M., Ramadan, S., Keller, M., Roecher, E., Zweerings, J., Klasen, M., Gur, R. C., & Mathiak, K. (2021). A Novel Brain–

Computer Interface Virtual Environment for Neurofeedback During Functional MRI. *Frontiers in Neuroscience*, *14*, 1367. https://doi.org/10.3389/FNINS.2020.593854/BIBTEX

Basha, S. M., & Rajput, D. S. (2019). Survey on Evaluating the Performance of Machine Learning Algorithms: Past Contributions and Future Roadmap. *Deep Learning and Parallel Computing Environment for Bioengineering Systems*, 153–164. https://doi.org/10.1016/B978-0-12-816718-2.00016-6

Berrar, D. (2019). Cross-Validation. *Encyclopedia of Bioinformatics and Computational Biology*, *1*(April), 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Bi, L., Fan, X. A., & Liu, Y. (2013). EEG-based brain-controlled mobile robots: A survey. *IEEE Transactions on Human-Machine Systems*, *43*(2), 161–176. https://doi.org/10.1109/TSMCC.2012.2219046

Binti Mustaffa, I., Bin Zaid, Z., & Binti Mohamed Aminuddin, M. M. (2018). Significance of averaging method signal denoising. *Proceeding of 2017 International Conference on Robotics, Automation and Sciences, ICORAS 2017*, *2018-March*, 1–6. https://doi.org/10.1109/ICORAS.2017.8308040

Bol, Y., Smolders, J., Duits, A., Lange, I. M. J., Romberg-Camps, M., & Hupperts, R. (2012). Fatigue and heat sensitivity in patients with multiple sclerosis. *Acta Neurologica Scandinavica*, *126*(6), 384–389. https://doi.org/10.1111/J.1600-0404.2012.01660.X

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). Training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, 144–152. https://doi.org/10.1145/130385.130401

Brandtner, A., Antons, S., King, D. L., Potenza, M. N., Tang, Y. Y., Blycker, G. R., Brand, M., & Liebherr, M. (2022). A Preregistered, Systematic Review Considering Mindfulness-Based Interventions and Neurofeedback for Targeting Affective and

Cognitive Processes in Behavioral Addictions. *Clinical Psychology: Science and Practice*. https://doi.org/10.1037/CPS0000075

Breteler, M. H. M., Arns, M., Peters, S., Giepmans, I., & Verhoeven, L. (2010). Improvements in spelling after QEEG-based neurofeedback in dyslexia: A randomized controlled treatment study. *Applied Psychophysiology Biofeedback*, *35*(1), 5–11. https://doi.org/10.1007/S10484-009-9105-2/TABLES/2

Browne, M. W. (2000). Cross-Validation Methods. *Journal of Mathematical Psychology*, *44*(1), 108–132. https://doi.org/10.1006/JMPS.1999.1279

BRUYANT, P. P. (2002). Analytic and iterative reconstruction algorithms in SPECT. *The Journal of Nuclear Medicine (1978)*, *43*(10).

Campos da Paz, V. K., Garcia, A., Campos da Paz Neto, A., & Tomaz, C. (2018). SMR neurofeedback training facilitates working memory performance in healthy older adults: A behavioral and EEG study. *Frontiers in Behavioral Neuroscience*, *12*, 321. https://doi.org/10.3389/FNBEH.2018.00321/BIBTEX

Cella, M., & Chalder, T. (2010). Measuring fatigue in clinical and community settings. *Journal of Psychosomatic Research*, *69*(1), 17–22. https://doi.org/10.1016/J.JPSYCHORES.2009.10.007

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research*, *37*(2), 147–153. https://doi.org/10.1016/0022-3999(93)90081-P

Chen, C. L., Liao, C. Y., Chen, R. C., Tang, Y. W., & Shih, T. F. (2017). Bus drivers fatigue measurement based on monopolar EEG. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10192 LNAI*, 308–317. https://doi.org/10.1007/978-3-319-54430-4_30/COVER

Chen, K., Li, Z., Ai, Q., Liu, Q., & Wang, L. (2021). An improved CNN model based on fused time-frequency features for mental fatigue detection in BCIs. *IISA 2021 - 12th International Conference on Information, Intelligence, Systems and Applications*. https://doi.org/10.1109/IISA52424.2021.9555518

Chen, Y., & Qin, X. (2023). Student fatigue and its impact on teaching effectiveness based on online teaching. *Education and Information Technologies*, 1–24. https://doi.org/10.1007/S10639-023-12197-3/METRICS

Chilcot, J., Norton, S., Kelly, M. E., & Moss-Morris, R. (2015). The Chalder Fatigue Questionnaire is a valid and reliable measure of perceived fatigue severity in multiple sclerosis. *Http://Dx.Doi.Org/10.1177/1352458515598019*, *22*(5), 677–684. https://doi.org/10.1177/1352458515598019

Cho, H. J., Costa, E., Menezes, P. R., Chalder, T., Bhugra, D., & Wessely, S. (2007). Cross-cultural validation of the Chalder Fatigue Questionnaire in Brazilian primary care. *Journal of Psychosomatic Research*, *62*(3), 301–304. https://doi.org/10.1016/J.JPSYCHORES.2006.10.018

Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-vector networks. *Machine Learning 1995 20:3*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Craig, A., Tran, Y., Wijesuriya, N., & Nguyen, H. (2012). Regional brain wave activity changes associated with fatigue. *Psychophysiology*, *49*(4), 574–582. https://doi.org/10.1111/J.1469-8986.2011.01329.X

Cramer, J. S. (2002). The Origins of Logistic Regression. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.360300

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *NeuroImage*, *54*(4), 2808–2821. https://doi.org/10.1016/J.NEUROIMAGE.2010.10.069

Demarin, V., Morović, S., & Béné, R. (2014). Neuroplasticity. *Periodicum Biologorum*, *116*(2), 209–211.

Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, *14*(2), 241–258. https://doi.org/10.1007/S11704-019-8208-Z/METRICS

Durantin, G., Scannella, S., Gateau, T., Delorme, A., & Dehais, F. (2014). Moving Average Convergence Divergence filter preprocessing for real-time event-related peak activity onset detection : Application to fNIRS signals. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2107–2110. https://doi.org/10.1109/EMBC.2014.6944032

ELHENDY, A., BAX, J. J., & POLDERMANS, D. (2002). Dobutamine stress myocardial perfusion imaging in coronary artery disease. *The Journal of Nuclear Medicine (1978)*, *43*(12).

Emotiv. (2023). *Performance Metrics - EMOTIV*. https://www.emotiv.com/knowledge-base/performance-metrics/

Emotiv. (2024a). *Data sample object | Cortex API*. https://emotiv.gitbook.io/cortex-api/data-subscription/data-sample-object#performance-metric

Emotiv. (2024b). *Performance Metrics | EmotivPRO v3.0*. https://emotiv.gitbook.io/emotivpro-v3/data-streams/performance-metrics

Enriquez-Geppert, S., Smit, D., Pimenta, M. G., & Arns, M. (2019). Neurofeedback as a Treatment Intervention in ADHD: Current Evidence and Practice. *Current Psychiatry Reports*, *21*(6), 1–7. https://doi.org/10.1007/S11920-019-1021-4/FIGURES/1

Faruk, M. J. H., Valero, M., & Shahriar, H. (2021). An investigation on non-invasive brain-computer interfaces: Emotiv EpoC+ neuroheadset and its effectiveness. *Proceedings - 2021 IEEE 45th Annual Computers, Software, and Applications*

Conference, COMPSAC 2021, 580–589. https://doi.org/10.1109/COMPSAC51774.2021.00086/VIDEO

Feldman, D. B., & Dreher, D. E. (2012). Can Hope be Changed in 90 Minutes? Testing the Efficacy of a Single-Session Goal-Pursuit Intervention for College Students. *Journal of Happiness Studies*, *13*(4). https://doi.org/10.1007/s10902-011-9292-4

Fernández Lucas, A., Andreina Méndez, M., Hornero, R., Ortiz, T., & López-Lbor, J. J. (2010). Analysis of brain complexity and mental disorders. *Actas Espanolas de Psiquiatria*, *38*(4), 229–238. https://europepmc.org/article/med/21104468

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, *63*(2), 921–935. https://doi.org/10.1016/J.NEUROIMAGE.2012.03.049

Flach, P. (2019). Performance Evaluation in Machine Learning: The Good, the Bad, the Ugly, and the Way Forward. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 9808–9814. https://doi.org/10.1609/AAAI.V33I01.33019808

Foong, R., Ang, K. K., Zhang, Z., & Quek, C. (2019a). An iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue. *Journal of Neural Engineering*, *16*(5), 056013. https://doi.org/10.1088/1741-2552/AB255D

Friese, M., Messner, C., & Schaffner, Y. (2012). Mindfulness meditation counteracts self-control depletion. *Consciousness and Cognition*, *21*(2). https://doi.org/10.1016/j.concog.2012.01.008

Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S. (2019). EEG-Based Spatio-Temporal Convolutional Neural Network for Driver Fatigue Evaluation. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(9), 2755–2763. https://doi.org/10.1109/TNNLS.2018.2886414

Geetha, G., & Geethalakshmi, S. N. (2011). EEG De-noising using SURE Thresholding based on Wavelet Transforms. *International Journal of Computer Applications*, *24*(6), 975–8887.

Gillberg, M., Kecklund, G., & Akerstedt, T. (1994). Relations Between Performance and Subjective Ratings of Sleepiness During a Night Awake. *Sleep*, 17(3), 236–241. https://doi.org/10.1093/SLEEP/17.3.236

Gloss, D., Varma, J. K., Pringsheim, T., & Nuwer, M. R. (2016). Practice advisory: The utility of EEG theta/beta power ratio in ADHD diagnosis. *Neurology*, *87*(22), 2375–2379. https://doi.org/10.1212/WNL.0000000000003265

Gonzàlez-Alonso, J., Teller, C., Andersen, S. L., Jensen, F. B., Hyldig, T., & Nielsen, B. (1999). Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *Journal of Applied Physiology*, *86*(3), 1032–1039. https://doi.org/10.1152/JAPPL.1999.86.3.1032/ASSET/IMAGES/LARGE/JAPP05313003X.JPEG

Grillon, C., Quispe-Escudero, D., Mathur, A., & Ernst, M. (2015). Mental fatigue impairs emotion regulation. *Emotion*, *15*(3), 383–389. https://doi.org/10.1037/EMO0000058

Grossi, E., Olivieri, C., & Buscema, M. (2017). Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study. *Computer Methods and Programs in Biomedicine*, *142*, 73–79. https://doi.org/10.1016/J.CMPB.2017.02.002

Güntensperger, D., Thüring, C., Kleinjung, T., Neff, P., & Meyer, M. (2019). Investigating the efficacy of an individualized alpha/delta neurofeedback protocol in the treatment of chronic tinnitus. *Neural Plasticity*, *2019*. https://doi.org/10.1155/2019/3540898

Hariharan, M., Padmaja, G., & Rana, S. (2018). Neurofeedback as an effective intervention for academic performance in children with problems in attention and concentration.

Psychosocial Interventions for Health and Well-Being, 145–156. https://doi.org/10.1007/978-81-322-3782-2_10

Hassan, U., Ali, I. ;, Abideen, R. H. ;, Ijaz, Z. U. ;, Khan, A. Z. ;, Hassan, I. U., Ali, H., Ul Abideen, Z., Zeeshan Ijaz, A., & Khan, T. A. (2023). Towards Effective Emotion Detection: A Comprehensive Machine Learning Approach on EEG Signals. *BioMedInformatics 2023, Vol. 3, Pages 1083-1100*, *3*(4), 1083–1100. https://doi.org/10.3390/BIOMEDINFORMATICS3040065

Hendrawan, M. A., Pane, E. S., Wibawa, A. D., & Purnormo, M. H. (2018). Investigating Window Segmentation on Mental Fatigue Detection Using Single-Channel EEG. *Proceedings of 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering, ICICI-BME 2017*, 173–178. https://doi.org/10.1109/ICICI-BME.2017.8537716

Hernandez-Gonzalez, G., Bringas-Vega, M. L., Galán-Garcia, L., Bosch-Bayard, J., Lorenzo-Ceballos, Y., Melie-Garcia, L., Valdes-Urrutia, L., Cobas-Ruiz, M., Valdes-Sosa, P. A., & Cuban Human Brain Mapping Project (CHBMP). (2011). Multimodal quantitative neuroimaging databases and methods: the Cuban Human Brain Mapping Project. *Clinical EEG and Neuroscience*, *42*(3), 149–159. https://doi.org/10.1177/155005941104200303

Holman, M., & Adebesin, F. (2019). Taking the subjectivity out of UX evaluation with emotiv EPOC+. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3351108.3351139

Homan, R. W. (2015). The 10-20 Electrode System and Cerebral Location. *Http://Dx.Doi.Org/10.1080/00029238.1988.11080272*, *28*(4), 269–279. https://doi.org/10.1080/00029238.1988.11080272

Hossain, F., & Yaacob, H. (2022). *Review on Signal Generation for Neurofeedback*. 1–8. https://doi.org/10.1109/CITSM56380.2022.9935866

Hossain, F., & Yaacob, H. (2023a). A proposal on a brain-computer interface model for real-time mental fatigue intervention. *Journal of Islamic, Social, Economics and Development*, *8*, 128–1755. https://doi.org/10.55573/JISED.085530

Hossain, F., & Yaacob, H. S. (2023b). *A proposal on a brain-computer interface model for real time mental fatigue intervention / Farhad Hossain and Hamwira Sakti Yaacob*. https://www.mic3st.com/

Hossain, M. F., Yaacob, H., & Nordin, A. (2021). Development of unified neuro-affective classification tool (UNACT). *IOP Conference Series: Materials Science and Engineering*, *1077*(1), 12031.

Hossan, A., Kashem, F. Bin, Hasan, M. M., Naher, S., & Rahman, M. I. (2017). A smart system for driver's fatigue detection, remote notification and semi-automatic parking of vehicles to prevent road accidents. *1st International Conference on Medical Engineering, Health Informatics and Technology, MediTec 2016*. https://doi.org/10.1109/MEDITEC.2016.7835371

Hwang, S., Park, S., Kim, D., Lee, J., & Byun, H. (2021). Mitigating inter-subject brain signal variability for EEG-based driver fatigue state classification. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2021-June*, 990–994. https://doi.org/10.1109/ICASSP39728.2021.9414613

Jackson, C. (2015). The Chalder Fatigue Scale (CFQ 11). *Occupational Medicine*, *65*(1), 86–86. https://doi.org/10.1093/OCCMED/KQU168

Jafari, M., Shoeibi, A., Khodatars, M., Bagherzadeh, S., Shalbaf, A., García, D. L., Gorriz, J. M., & Acharya, U. R. (2023). Emotion recognition in EEG signals using deep learning methods: A review. *Computers in Biology and Medicine*, *165*, 107450. https://doi.org/10.1016/J.COMPBIOMED.2023.107450

Jang, H. J., Kim, O., Kim, S., Kim, M. S., Choi, J. A., Kim, B., Dan, H., & Jung, H. (2021). Factors Affecting Physical and Mental Fatigue among Female Hospital Nurses: The Korea Nurses' Health Study. *Healthcare*, *9*(2). https://doi.org/10.3390/HEALTHCARE9020201

Janssen, T. W. P., Geladé, K., Bink, M., van Mourik, R., Twisk, J. W. R., Maras, A., & Oosterlaan, J. (2020). Long-term effects of theta/beta neurofeedback on EEG power spectra in children with attention deficit hyperactivity disorder. *Clinical Neurophysiology*, *131*(6), 1332–1341. https://doi.org/10.1016/J.CLINPH.2020.02.020

Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, *36*(2), 2352–2359. https://doi.org/10.1016/J.ESWA.2007.12.043

Jin, H., Xiao, M., Gong, Z., & Zhao, Y. (2022). Influence of Different Protection States on the Mental Fatigue of Nurses During the COVID-19 Pandemic. *Risk Management and Healthcare Policy*, *15*, 1917–1929. https://doi.org/10.2147/RMHP.S377936

Jing, M. J., Lin, W. Q., Wang, Q., Wang, J. J., Tang, J., Jiang, E. S., Lei, Y. X., & Wang, P. X. (2016). Reliability and Construct Validity of Two Versions of Chalder Fatigue Scale among the General Population in Mainland China. *International Journal of Environmental Research and Public Health 2016, Vol. 13, Page 147*, *13*(1), 147. https://doi.org/10.3390/IJERPH13010147

Kalra, P., & Karar, V. (2022). Effect of brightness on visual fatigue during video viewing. In *Productivity with Health, Safety, and Environment* (pp. 357–363). Springer.

Kamal, M., Abo Omirah, M., Hussein, A., & Saeed, H. (2021). Assessment and characterisation of post-COVID-19 manifestations. *International Journal of Clinical Practice*, *75*(3), e13746. https://doi.org/10.1111/IJCP.13746

Kamiya, J., & Francisco, S. (2011). The First Communications About Operant Conditioning of the EEG. *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience*, *15*(1), 65–73. https://doi.org/10.1080/10874208.2011.545764

Kar, S., Bhagat, M., & Routray, A. (2010). EEG signal analysis for the assessment and quantification of driver's fatigue. *Transportation Research Part F: Traffic Psychology and Behaviour*, *13*(5), 297–306. https://doi.org/10.1016/J.TRF.2010.06.006

Katkar, V. D., & Kulkarni, S. V. (2013). A novel parallel implementation of Naive Bayesian classifier for big data. *Proceedings of the 2013 International Conference on Green Computing, Communication and Conservation of Energy, ICGCE 2013*, 847–852. https://doi.org/10.1109/ICGCE.2013.6823552

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Ko, L. W., Lai, W. K., Liang, W. G., Chuang, C. H., Lu, S. W., Lu, Y. C., Hsiung, T. Y., Wu, H. H., & Lin, C. T. (2015). Single channel wireless EEG device for real-time fatigue level detection. *Proceedings of the International Joint Conference on Neural Networks*, *2015-September*. https://doi.org/10.1109/IJCNN.2015.7280817

Kołodziej, M., Majkowski, A., & Rak, R. J. (2012). Linear discriminant analysis as EEG features reduction technique for brain-computer interfaces. *Przeglad Elektrotechniczny*, *88(33)*, 28–30.

Korats, G., Le Cam, S., Ranta, R., & Hamid, M. (2013). Applying ICA in EEG: Choice of the Window Length and of the Decorrelation Method. *Communications in Computer and Information Science*, *357 CCIS*, 269–286. https://doi.org/10.1007/978-3-642-38256-7_18/COVER

Korfmacher, A. K., Hirsch, O., Chavanon, M. L., Albrecht, B., & Christiansen, H. (2022). Self-management training vs. neurofeedback interventions for attention deficit hyperactivity disorder: Results of a randomized controlled treatment study. *Frontiers in Psychiatry*, *13*, 1804. https://doi.org/10.3389/FPSYT.2022.969351/BIBTEX

Krepel, N., Egtberts, T., Sack, A. T., Heinrich, H., Ryan, M., & Arns, M. (2020). A multicenter effectiveness trial of QEEG-informed neurofeedback in ADHD: Replication and treatment prediction. *NeuroImage: Clinical*, *28*, 102399. https://doi.org/10.1016/J.NICL.2020.102399

Kuang, H., & Qu, J. (2021). LSTM Model with Self-Attention Mechanism for EEG Based Cross-Subject Fatigue Detection. *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer, ICFTIC 2021*, 148–153. https://doi.org/10.1109/ICFTIC54370.2021.9647146

Kunasegaran, K., Ismail, A. M. H., Ramasamy, S., Gnanou, J. V., Caszo, B. A., & Chen, P. L. (2023). Understanding mental fatigue and its detection: a comparative analysis of assessments and tools. *PeerJ*, *11*, e15744. https://doi.org/10.7717/PEERJ.15744/TABLE-3

Lakshmi, M. R., Prasad, T. V, & Chandra Prakash, V. (2014). Survey on EEG Signal Processing Methods. *International Journal of Advanced Research in Computer Science and Software Engineering*, *4*(1), 2277. https://www.researchgate.net/publication/328419840

Lal, S. K. L., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, *39*(3), 313–321. https://doi.org/10.1017/S0048577201393095

Lee, D. H., Jeong, J. H., Kim, K., Yu, B. W., & Lee, S. W. (2020a). Continuous EEG Decoding of Pilots' Mental States Using Multiple Feature Block-Based

Convolutional Neural Network. *IEEE Access*, *8*, 121929–121941. https://doi.org/10.1109/ACCESS.2020.3006907

Lee, D. H., Jeong, J. H., Kim, K., Yu, B. W., & Lee, S. W. (2020b). Continuous EEG Decoding of Pilots' Mental States Using Multiple Feature Block-Based Convolutional Neural Network. *IEEE Access*, *8*, 121929–121941. https://doi.org/10.1109/ACCESS.2020.3006907

Lesser, R. P. (1986). Guideline seven: A proposal for standard montages to be used in clinical EEG. *Journal of Clinical Neurophysiology*, *3*(3 SUPPL. 1), 26–33. https://doi.org/10.1097/00004691-198610001-00008

Lewczuk, K., Wizła, M., Oleksy, T., & Wyczesany, M. (2022). Emotion Regulation, Effort and Fatigue: Complex Issues Worth Investigating. *Frontiers in Psychology*, *13*, 742557. https://doi.org/10.3389/FPSYG.2022.742557

Lipp, A., & Cohen Kadosh, K. (2020). Training the anxious brain: using fMRI-based neurofeedback to change brain activity in adolescence. *Developmental Medicine & Child Neurology*, *62*(11), 1239–1244. https://doi.org/10.1111/DMCN.14611

Lippé, S., Kovacevic, N., & McIntosh, A. R. (2009). Differential maturation of brain signal complexity in the human auditory and visual system. *Frontiers in Human Neuroscience*, *3*(NOV), 48. https://doi.org/10.3389/NEURO.09.048.2009/BIBTEX

Liu, L., Marler, M. R., Parker, B. A., Jones, V., Johnson, S., Cohen-Zion, M., Fiorentino, L., Sadler, G. R., & Ancoli-Israel, S. (2005). The relationship between fatigue and light exposure during chemotherapy. *Supportive Care in Cancer 2005 13:12*, *13*(12), 1010–1017. https://doi.org/10.1007/S00520-005-0824-5

Liu, Y. T., Lin, Y. Y., Wu, S. L., Chuang, C. H., & Lin, C. T. (2016). Brain Dynamics in Predicting Driving Fatigue Using a Recurrent Self-Evolving Fuzzy Neural Network. *IEEE Transactions on Neural Networks and Learning Systems*, *27*(2), 347–360. https://doi.org/10.1109/TNNLS.2015.2496330

Liu, Y. T., Lin, Y. Y., Wu, S. L., Hsieh, T. Y., & Lin, C. T. (2016). Assessment of Mental Fatigue: An EEG-Based Forecasting System for Driving Safety. *Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015*, 3233–3238. https://doi.org/10.1109/SMC.2015.561

Liu, Y. T., Wu, S. L., Chou, K. P., Lin, Y. Y., Lu, J., Zhang, G., Lin, W. C., & Lin, C. T. (2016). Driving fatigue prediction with pre-event electroencephalography (EEG) via a recurrent fuzzy neural network. *2016 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2016*, 2488–2494. https://doi.org/10.1109/FUZZ-IEEE.2016.7738006

Loge, J. H., Ekeberg, Ø., & Kaasa, S. (1998). Fatigue in the general norwegian population: Normative data and associations. *Journal of Psychosomatic Research*, *45*(1), 53–65. https://doi.org/10.1016/S0022-3999(97)00291-2

Loriette, C., Ziane, C., & Ben Hamed, S. (2021). Neurofeedback for cognitive enhancement and intervention and brain plasticity. *Revue Neurologique*, *177*(9), 1133–1144. https://doi.org/10.1016/J.NEUROL.2021.08.004

Lu, S., Jiang, A., Schlacht, I., Ono, A., Foing, B., Yao, X., Westland, S., & Guo, Y. (2021). The Effect on Subjective Alertness and Fatigue of Three Colour Temperatures in the Spacecraft Crew Cabin. *Lecture Notes in Networks and Systems*, *270*, 632–639. https://doi.org/10.1007/978-3-030-80012-3_74/COVER

Lubar, J. F. (1997). Neocortical Dynamics: Implications for Understanding the Role of Neurofeedback and Related Techniques for the Enhancement of Attention. *Applied Psychophysiology Biofeedback*, *22*(2), 111–126. https://doi.org/10.1023/A:1026276228832/METRICS

Maciej Serda, Becker, F. G., Cleary, M., Team, R. M., Holtermann, H., The, D., Agenda, N., Science, P., Sk, S. K., Hinnebusch, R., Hinnebusch A, R., Rabinovich, I., Olmert, Y., Uld, D. Q. G. L. Q., Ri, W. K. H. U., Lq, V., Frxqwu, W. K. H., Zklfk, E., Edvhg, L. V, …)2000. (ح, فاطمی. The theory of the STOP-emotion: On the functionality of fatigue. *Uniwersytet Śląski*, *7*(1), 45–47. https://doi.org/10.2/JQUERY.MIN.JS

Maloney, S., Axelsen, M., Stone, C., Galligan, L., Redmond, P., Brown, A., Turner, J., & Lawrence, J. (2023). Defining and exploring online engagement fatigue in a university context. *Computers and Education Open*, *4*, 100139. https://doi.org/10.1016/J.CAEO.2023.100139

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. *Journal of Applied Physiology*, *106*(3), 857–864. https://doi.org/10.1152/JAPPLPHYSIOL.91324.2008/ASSET/IMAGES/LARGE/ZDG0030984280003.JPEG

Marzbani, H., Marateb, H. R., & Mansourian, M. (2016). Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications. *Basic and Clinical Neuroscience*, *7*(2), 143. https://doi.org/10.15412/J.BCN.03070208

Matic, V., Cherian, P. J., Koolen, N., -, al, Liu, X., Lv, L., Shen, Y., Zang, B., Lin, Y., Liu, Z., Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of Neural Engineering*, *16*(3), 031001. https://doi.org/10.1088/1741-2552/AB0AB5

Mauro, T., & Cermak, S. A. (2006). *The Everything Parent's Guide to Sensory Integration Disorder: Get the Right Diagnosis, Understand Treatments, and Advocate for Your Child*. Simon and Schuster.

Mehmood, R. M., Du, R., & Lee, H. J. (2017). Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors. *IEEE Access*, *5*, 14797–14806. https://doi.org/10.1109/ACCESS.2017.2724555

Mercado, J., Espinosa-Curiel, I., Escobedo, L., & Tentori, M. (2018). Developing and evaluating a BCI video game for neurofeedback training: the case of autism.

Multimedia Tools and Applications 2018 78:10, *78*(10), 13675–13712. https://doi.org/10.1007/S11042-018-6916-2

Miley, A. Å., Kecklund, G., & Åkerstedt, T. (2016). Comparing two versions of the Karolinska Sleepiness Scale (KSS). *Sleep and Biological Rhythms*, *14*(3), 257–260. https://doi.org/10.1007/S41105-016-0048-8/FIGURES/2

Ming, Y., Wu, D., Wang, Y. K., Shi, Y., & Lin, C. T. (2021). EEG-Based Drowsiness Estimation for Driving Safety Using Deep Q-Learning. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *5*(4), 583–594. https://doi.org/10.1109/TETCI.2020.2997031

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group*, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269.

Moralista, R., Oducado, R. M., Robles, B. R., & Rosano, D. (2022). *Determinants of Zoom Fatigue among Graduate Students of Teacher Education Program*. https://papers.ssrn.com/abstract=4161936

Moretti, D. V., Babiloni, C., Binetti, G., Cassetta, E., Dal Forno, G., Ferreric, F., Ferri, R., Lanuzza, B., Miniussi, C., Nobili, F., Rodriguez, G., Salinari, S., & Rossini, P. M. (2004). Individual analysis of EEG frequency and band power in mild Alzheimer's disease. *Clinical Neurophysiology*, *115*(2), 299–308. https://doi.org/10.1016/S1388-2457(03)00345-6

Mottola, F., Blanchfield, A., Hardy, J., & Cooke, A. (2021). EEG neurofeedback improves cycling time to exhaustion. *Psychology of Sport and Exercise*, *55*, 101944. https://doi.org/10.1016/J.PSYCHSPORT.2021.101944

Mudgal, S. K., Sharma, S. K., Chaturvedi, J., & Sharma, A. (2020). Brain computer interface advancement in neurosciences: Applications and issues. *Interdisciplinary Neurosurgery*, *20*, 100694. https://doi.org/10.1016/J.INAT.2020.100694

Murphy, K., van Ginneken, B., Schilham, A. M. R., de Hoop, B. J., Gietema, H. A., & Prokop, M. (2009). A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. *Medical Image Analysis*, *13*(5), 757–770. https://doi.org/10.1016/J.MEDIA.2009.07.001

Myrden, A., & Chau, T. (2017). A Passive EEG-BCI for Single-Trial Detection of Changes in Mental State. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(4), 345–356. https://doi.org/10.1109/TNSRE.2016.2641956

Nandini, D., Yadav, J., Rani, A., & Singh, V. (2023). Design of subject independent 3D VAD emotion detection system using EEG signals and machine learning algorithms. *Biomedical Signal Processing and Control*, *85*, 104894. https://doi.org/10.1016/J.BSPC.2023.104894

Navarro Gil, M., Escolano Marco, C., Montero-Marín, J., Minguez Zafra, J., Shonin, E., & García Campayo, J. (2017). Efficacy of Neurofeedback on the Increase of Mindfulness-Related Capacities in Healthy Individuals: a Controlled Trial. *Mindfulness 2017 9:1*, *9*(1), 303–311. https://doi.org/10.1007/S12671-017-0775-1

Ngien, A., & Hogan, B. (2022). The relationship between Zoom use with the camera on and Zoom fatigue: considering self-monitoring and social interaction anxiety. *Https://Doi.Org/10.1080/1369118X.2022.2065214*. https://doi.org/10.1080/1369118X.2022.2065214

Niedermeyer, E., & da Silva, F. L. (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Noohi, S., Miraghaie liM, A., Arabi, A., & Nooripour, R. (2017). Effectiveness of neuro-feedback treatment with alpha/theta method on PTSD symptoms and their executing function. *Biomedical Research*, *28*(5), 2019–2027.

Otani, H., Kaya, M., Tamaki, A., & Watson, P. (2017). Separate and combined effects of exposure to heat stress and mental fatigue on endurance exercise capacity in the heat. *European Journal of Applied Physiology*, *117*(1), 119–129. https://doi.org/10.1007/S00421-016-3504-X/METRICS

Paranthaman, P. K., Bajaj, N., Solovey, N., & Jennings, D. (2021). Comparative Evaluation of the EEG Performance Metrics and Player Ratings on the Virtual Reality Games. *IEEE Conference on Computatonal Intelligence and Games, CIG*, *2021-August*. https://doi.org/10.1109/COG52621.2021.9619043

Pathak, M., & Jayanthy, A. K. (2017). DEVELOPMENT OF A REAL-TIME SINGLE CHANNEL BRAIN–COMPUTER INTERFACE SYSTEM FOR DETECTION OF DROWSINESS. *Https://Doi.Org/10.4015/S1016237217500193*, *29*(3). https://doi.org/10.4015/S1016237217500193

Paulo, J. R., Pires, G., & Nunes, U. J. (2021). Cross-Subject Zero Calibration Driver's Drowsiness Detection: Exploring Spatiotemporal Image Encoding of EEG Signals for Convolutional Neural Network Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 905–915. https://doi.org/10.1109/TNSRE.2021.3079505

Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., & Evans, A. C. (1997). Time-Related Changes in Neural Systems Underlying Attention and Arousal During the Performance of an Auditory Vigilance Task. *Journal of Cognitive Neuroscience*, *9*(3), 392–408. https://doi.org/10.1162/JOCN.1997.9.3.392

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents proceedings of the royal society of london, 58, 240-242. *K Pearson*.

Pearson's Correlation Coefficient. (2008). *Encyclopedia of Public Health*, 1090–1091. https://doi.org/10.1007/978-1-4020-5614-7_2569

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

Peng, Y., Wong, C. M., Wang, Z., Rosa, A. C., Wang, H. T., & Wan, F. (2021). Fatigue Detection in SSVEP-BCIs Based on Wavelet Entropy of EEG. *IEEE Access*, *9*, 114905–114913. https://doi.org/10.1109/ACCESS.2021.3100478

Perer A. Rinck. (2014). Magnetic resonance: a critical peer-reviewed introduction. In *Magnetic resonance in medicine. The basic textbook of the European magnetic resonance forum* (pp. 21–01). BoD Germany.

Polikar, R. (2012). Ensemble Learning. *Ensemble Machine Learning*, 1–34. https://doi.org/10.1007/978-1-4419-9326-7_1

Proost, M., Habay, J., De Wachter, J., De Pauw, K., Rattray, B., Meeusen, R., Roelands, B., & Van Cutsem, J. (2022). How to Tackle Mental Fatigue: A Systematic Review of Potential Countermeasures and Their Underlying Mechanisms. *Sports Medicine 2022 52:9*, *52*(9), 2129–2158. https://doi.org/10.1007/S40279-022-01678-Z

Rogala, J., Jurewicz, K., Paluch, K., Kublik, E., Cetnarski, R., & Wróbel, A. (2016). The do's and don'ts of neurofeedback training: A review of the controlled studies using healthy adults. *Frontiers in Human Neuroscience*, *10*, 301. https://doi.org/10.3389/FNHUM.2016.00301/BIBTEX

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.

Roy, R. N., Charbonnier, S., & Bonnet, S. (2014). Detection of mental fatigue using an active BCI inspired signal processing chain. *IFAC Proceedings Volumes*, *47*(3), 2963–2968. https://doi.org/10.3182/20140824-6-ZA-1003.00897

Rudroff, T., Kamholz, J., Fietsam, A. C., Deters, J. R., & Bryant, A. D. (2020). Post-COVID-19 Fatigue: Potential Contributing Factors. *Brain Sciences 2020, Vol. 10, Page 1012*, *10*(12), 1012. https://doi.org/10.3390/BRAINSCI10121012

Saad, J. F., Kohn, M. R., Clarke, S., Lagopoulos, J., & Hermens, D. F. (2015). Is the Theta/Beta EEG Marker for ADHD Inherently Flawed? *Journal of Attention Disorders*, *22*(9), 815–826. https://doi.org/10.1177/1087054715578270

Salim, J., Tandy, S., Arnindita, J. N., Wibisono, J. J., Haryanto, M. R., & Wibisono, M. G. (2022). Zoom fatigue and its risk factors in online learning during the COVID-19 pandemic. *Medical Journal of Indonesia*, *31*(1), 13–19. https://doi.org/10.13181/MJI.OA.225703

Santoyo-Mora, M., Villaseñor-Mora, C., Cardona-Torres, L. M., Martínez-Nolasco, J. J., Barranco-Gutiérrez, A. I., Padilla-Medina, J. A., & Bravo-Sánchez, M. G. (2022). COVID-19 Long-Term Effects: Is There an Impact on the Simple Reaction Time and Alternative-Forced Choice on Recovered Patients? *Brain Sciences 2022, Vol. 12, Page 1258*, *12*(9), 1258. https://doi.org/10.3390/BRAINSCI12091258

Sapta Pramana, M. F., & Arlini Puspasari, M. (2020). Evaluation on the Effects of Visual Intervention on Emotion and Fatigue while Driving Using Brainwave Indicator. *ACM International Conference Proceeding Series*, 196–201. https://doi.org/10.1145/3429551.3429554

Satterthwaite, F. E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, *2*(6), 110. https://doi.org/10.2307/3002019

Schiehser, D. M., Ayers, C. R., Liu, L., Lessig, S., Song, D. S., & Filoteo, J. V. (2013). Validation of the Modified Fatigue Impact Scale in Parkinson's disease. *Parkinsonism & Related Disorders*, *19*(3), 335–338. https://doi.org/10.1016/J.PARKRELDIS.2012.11.013

Schinkel, S., Marwan, N., & Kurths, J. (2009). Brain signal analysis based on recurrences. *Journal of Physiology-Paris*, *103*(6), 315–323. https://doi.org/10.1016/J.JPHYSPARIS.2009.05.007

Schlichta, C., Cabral, L. L., da Silva, C. K., Bigliassi, M., & Pereira, G. (2022). Exploring the Impact of Mental Fatigue and Emotional Suppression on the Performance of High-Intensity Endurance Exercise. *Https://Doi.Org/10.1177/00315125221093898*, *129*(4), 1053–1073. https://doi.org/10.1177/00315125221093898

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention*, *41*(5), 1087–1093. https://doi.org/10.1016/J.AAP.2009.06.007

Sho'ouri, N., Firoozabadi, M., & Badie, K. (2020). The effect of beta/alpha neurofeedback training on imitating brain activity patterns in visual artists. *Biomedical Signal Processing and Control*, *56*, 101661. https://doi.org/10.1016/J.BSPC.2019.101661

Shourie, N., Firoozabadi, M., & Badie, K. (2018). Fuzzy adaptive neurofeedback training: An efficient neurofeedback training procedure providing a more accurate progress rate for trainee. *Biomedical Signal Processing and Control*, *44*, 75–81. https://doi.org/10.1016/J.BSPC.2018.02.009

Sidharth, S., Samuel, A. A., Ranjana, H., Panachakel, J. T., & Parveen K, S. (2023). Emotion detection from EEG using transfer learning. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. https://doi.org/10.1109/EMBC40787.2023.10340389

Simon, J., Takács, E., Orosz, G., Berki, B., & Winkler, I. (2020). Short-term cognitive fatigue effect on auditory temporal order judgments. *Experimental Brain Research*, *238*(2), 305–319. https://doi.org/10.1007/S00221-019-05712-X/FIGURES/4

Sinclair, K. L., Ponsford, J. L., Taffe, J., Lockley, S. W., & Rajaratnam, S. M. W. (2014). Randomized controlled trial of light therapy for fatigue following traumatic brain injury. *Neurorehabilitation and Neural Repair*, *28*(4), 303–313. https://doi.org/10.1177/1545968313508472

Smolders, K. C. H. J., & de Kort, Y. A. W. (2014). Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal. *Journal of Environmental Psychology*, *39*, 77–91. https://doi.org/10.1016/J.JENVP.2013.12.010

Sörensen, S., Duberstein, P., Gill, D., & Pinquart, M. (2006). Dementia care: mental health effects, intervention strategies, and clinical implications. *The Lancet Neurology*, *5*(11), 961–973. https://doi.org/10.1016/S1474-4422(06)70599-3

Soufineyestani, M., Dowling, D., & Khan, A. (2020). Electroencephalography (EEG) Technology Applications and Available Devices. *Applied Sciences 2020, Vol. 10, Page 7453*, *10*(21), 7453. https://doi.org/10.3390/APP10217453

Soylu, Y., Ramazanoglu, F., Arslan, E., & Clemente, F. M. (2021). Effects of mental fatigue on the psychophysiological responses, kinematic profiles, and technical performance in different small-sided soccer games. *Biology of Sport*, *39*(4), 965–972. https://doi.org/10.5114/BIOLSPORT.2022.110746

Srimaharaj, W., Chaipakornwong, T., Punyahotra, V., Chaising, S., Temdee, P., Chaisricharoen, R., & Sittiprapaporn, P. (2021). Correlation of Attention and Relaxation Levels to Interactive Self-learning Classroom Indexed by eSense Meters. *Advances in Intelligent Systems and Computing*, *1288*, 827–837. https://doi.org/10.1007/978-3-030-63128-4_62/COVER

Strmiska, M., & Koudelkova, Z. (2018). Analysis of Performance Metrics Using Emotiv EPOC +. *MATEC Web of Conferences*, *210*. https://doi.org/10.1051/MATECCONF/201821004046

Subha, D. P., Joseph, P. K., Acharya U, R., & Lim, C. M. (2008). EEG Signal Analysis: A Survey. *Journal of Medical Systems 2008 34:2*, *34*(2), 195–212. https://doi.org/10.1007/S10916-008-9231-Z

Suhaimi, N. S., Mountstephens, J., & Teo, J. (2020). EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities. *Computational Intelligence and Neuroscience*, *2020*. https://doi.org/10.1155/2020/8875426

Suto, J., & Oniga, S. (2018). Music Stimuli Recognition in Electroencephalogram Signal. *Elektronika Ir Elektrotechnika*, *24*(4), 68–71. https://doi.org/10.5755/J01.EIE.24.4.21482

Tabejamaat, M., & Mohammadzade, H. (2022). Sequential nonlinear encoding: A low dimensional regression algorithm with application to EEG-based driving fatigue detection. *Scientia Iranica*, *29*(3), 1486–1505. https://doi.org/10.24200/SCI.2020.53905.3479

Takahashi, T. (2013). Complexity of spontaneous brain activity in mental disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *45*, 258–266. https://doi.org/10.1016/J.PNPBP.2012.05.001

Talukdar, U., Hazarika, S. M., & Gan, J. Q. (2020a). Adaptation of Common Spatial Patterns based on mental fatigue for motor-imagery BCI. *Biomedical Signal Processing and Control*, *58*, 101829. https://doi.org/10.1016/J.BSPC.2019.101829

Talukdar, U., Hazarika, S. M., & Gan, J. Q. (2020b). Adaptive feature extraction in EEG-based motor imagery BCI: tracking mental fatigue. *Journal of Neural Engineering*, *17*(1), 016020. https://doi.org/10.1088/1741-2552/AB53F1

Tanabe, S. I., Nishihara, N., & Haneda, M. (2007). Indoor temperature, productivity, and fatigue in office tasks. *HVAC and R Research*, *13*(4), 623–633. https://doi.org/10.1080/10789669.2007.10390975

Tanaka, M., Shigihara, Y., Ishii, A., Funakura, M., Kanai, E., & Watanabe, Y. (2012). Effect of mental fatigue on the central nervous system: an electroencephalography study. *Behavioral and Brain Functions*, *8*(1), 1–8. https://doi.org/10.1186/1744-9081-8-48/FIGURES/5

Tempelaar, D., Rienties, B., & Nguyen, Q. (2020). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLOS ONE*, *15*(6), e0233977. https://doi.org/10.1371/JOURNAL.PONE.0233977

Thakor, N. V., & Sherman, D. L. (2013). EEG Signal Processing: Theory and Applications. *Neural Engineering*, 259–303. https://doi.org/10.1007/978-1-4614-5227-0_5

Tian, S., Wang, Y., Dong, G., Pei, W., & Chen, H. (2018). Mental Fatigue Estimation Using EEG in a Vigilance Task and Resting States. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2018-July*, 1980–1983. https://doi.org/10.1109/EMBC.2018.8512666

Tuckute, G., Hansen, S. T., Kjaer, T. W., & Hansen, L. K. (2019). A framework for closed-loop neurofeedback for real-time EEG decoding. *BioRxiv*, 834713. https://doi.org/10.1101/834713

Uusitalo, M. A., & Ilmoniemi, R. J. (1997). Signal-space projection method for separating MEG or EEG into components. *Medical and Biological Engineering and Computing 1997 35:2*, *35*(2), 135–140. https://doi.org/10.1007/BF02534144

Van Cutsem, J., Van Schuerbeek, P., Pattyn, N., Raeymaekers, H., De Mey, J., Meeusen, R., & Roelands, B. (2022). A drop in cognitive performance, whodunit? Subjective mental fatigue, brain deactivation or increased parasympathetic activity? It's complicated! *Cortex*, *155*, 30–45. https://doi.org/10.1016/J.CORTEX.2022.06.006

Van Der Linden, D., & Eling, P. (2006). Mental fatigue disturbs local processing more than global processing. *Psychological Research*, *70*(5), 395–402. https://doi.org/10.1007/S00426-005-0228-7/METRICS

van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psychologica*, *113*(1), 45–65. https://doi.org/10.1016/S0001-6918(02)00150-6

Van Doren, J., Arns, M., Heinrich, H., Vollebregt, M. A., Strehl, U., & K. Loo, S. (2019). Sustained effects of neurofeedback in ADHD: a systematic review and meta-analysis. *European Child and Adolescent Psychiatry*, *28*(3), 293–305. https://doi.org/10.1007/S00787-018-1121-4/FIGURES/3

van Lutterveld, R., Houlihan, S. D., Pal, P., Sacchet, M. D., McFarlane-Blake, C., Patel, P. R., Sullivan, J. S., Ossadtchi, A., Druker, S., Bauer, C., & Brewer, J. A. (2017). Source-space EEG neurofeedback links subjective experience with brain activity during effortless awareness meditation. *NeuroImage*, *151*, 117–127. https://doi.org/10.1016/J.NEUROIMAGE.2016.02.047

Vapnik, V. N. (1997). The support vector method. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1327*, 264–271. https://doi.org/10.1007/BFB0020166/COVER

Ventouras, E. M., Monoyiou, E. A., Ktonas, P. Y., Paparrigopoulos, T., Dikeos, D. G., Uzunoglu, N. K., & Soldatos, C. R. (2005). Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: A feasibility study. *Computer Methods and Programs in Biomedicine*, *78*(3), 191–207. https://doi.org/10.1016/J.CMPB.2005.02.006

Wang, B., Xu, Z., Luo, T., & Pan, J. (2021). EEG-Based Closed-Loop Neurofeedback for Attention Monitoring and Training in Young Adults. *Journal of Healthcare Engineering*, *2021*, 5535810. https://doi.org/10.1155/2021/5535810

Wang, D., Xie, X., Tian, H., Wu, T., Liu, C., Huang, K., Gong, R., Yu, Y., Luo, T., Jiao, R., & Zhang, L. (2022). Mental fatigue and negative emotion among nurses during the COVID-19 pandemic. *Current Psychology*, *41*(11), 8123–8131. https://doi.org/10.1007/S12144-022-03468-2/FIGURES/2

WANG, Y., LI, H., YE, J., HUANG, Y., & ZHAO, J. (2017). The impact of short-term meditation on ego-depletion in college students. *Chinese Journal of Behavioral Medicine and Brain Science*, 632–635. https://doi.org/10.3760/CMA.J.ISSN.1674-6554.2017.07.012

WELCH, B. L. (1947). THE GENERALIZATION OF 'STUDENT'S' PROBLEM WHEN SEVERAL DIFFERENT POPULATION VARLANCES ARE INVOLVED. *Biometrika*, *34*(1–2), 28–35. https://doi.org/10.1093/BIOMET/34.1-2.28

Wittevrongel, B., & Van Hulle, M. M. (2017). Spatiotemporal beamforming: A transparent and unified decoding approach to synchronous visual brain-computer interfacing. *Frontiers in Neuroscience*, *11*(NOV), 630. https://doi.org/10.3389/FNINS.2017.00630/BIBTEX

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. In *Clinical Neurophysiology* (Vol. 113, Issue 6). https://doi.org/10.1016/S1388-2457(02)00057-3

Wood, G., & Kober, S. E. (2018). EEG Neurofeedback Is Under Strong Control of Psychosocial Factors. *Applied Psychophysiology Biofeedback*, *43*(4), 293–300. https://doi.org/10.1007/S10484-018-9407-3/FIGURES/3

Xie, X., Song, F., Liu, Y., Wang, S., & Yu, D. (2021). Study on the effects of display color mode and luminance contrast on visual fatigue. *IEEE Access*, *9*, 35915–35923. https://doi.org/10.1109/ACCESS.2021.3061770

Yaacob, H., Hossain, F., Shari, S., Khare, S. K., Ooi, C. P., & Acharya, U. R. (2023). Application of Artificial Intelligence Techniques for Brain-Computer Interface in Mental Fatigue Detection: A Systematic Review (2011-2022). *IEEE Access*, *11*, 74736–74758. https://doi.org/10.1109/ACCESS.2023.3296382

Yin, M., Vaughan, J. W., & Wallach, H. (2019). Understanding the effect of accuracy on trust in machine learning models. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3290605.3300509

Zeng, C., Mu, Z., & Wang, Q. (2022). Classifying Driving Fatigue by Using EEG Signals. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/1885677

Zeng, H., Yang, C., Zhang, H., Wu, Z., Zhang, J., Dai, G., Babiloni, F., Kong, W., & Chuang, L. (2019). A LightGBM-Based EEG Analysis Method for Driver Mental States Classification. *Computational Intelligence and Neuroscience*, *2019*. https://doi.org/10.1155/2019/3761203

Zhang, C., Wang, H., & Fu, R. (2014). Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, *15*(1), 168–177. https://doi.org/10.1109/TITS.2013.2275192

Zhang, X., Li, J., Liu, Y., Zhang, Z., Wang, Z., Luo, D., Zhou, X., Zhu, M., Salman, W., Hu, G., & Wang, C. (2017). Design of a Fatigue Detection System for High-Speed Trains Based on Driver Vigilance Using a Wireless Wearable EEG. *Sensors 2017, Vol. 17, Page 486*, *17*(3), 486. https://doi.org/10.3390/S17030486

Zhang, Z., Zhuo, K., Wei, W., Li, F., Yin, J., & Xu, L. (2021). Emotional Responses to the Visual Patterns of Urban Streets: Evidence from Physiological and Subjective Indicators. *International Journal of Environmental Research and Public Health 2021, Vol. 18, Page 9677*, *18*(18), 9677. https://doi.org/10.3390/IJERPH18189677

Zhao, L., & He, Y. (2013). Power Spectrum Estimation of the Welch Method Based on Imagery EEG. *Applied Mechanics and Materials*, *278–280*, 1260–1264. https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMM.278-280.1260

Zhao, Q., & Zhang, L. (2007). Temporal and spatial features of single-trial EEG for brain-computer interface. *Computational Intelligence and Neuroscience*, *2007*. https://doi.org/10.1155/2007/37695

Zhu, Y., Sun, F., Li, C., & Chow, D. H. K. (2020). Acute effects of brief mindfulness intervention coupled with carbohydrate ingestion to re-energize soccer players: A randomized crossover trial. *International Journal of Environmental Research and Public Health*, *17*(23). https://doi.org/10.3390/ijerph17239037

APPENDICES

APPENDIX A: PERMISSION TO CONDUCT A SURVEY

LEADING THE WAY
KHALĪFAH • AMĀNAH • IQRA' • RAHMATAN LIL-ĀLAMĪN | SUSTAINABILITY INSTITUTION OF THE YEAR

الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
Garden of Knowledge and Virtue

KULLIYYAH OF INFORMATION AND COMMUNICATION TECHNOLOGY

Reference No : IIUM/309/DDPG/5/8/1/ G2121261
Date : 23rd February 2023

TO WHOM IT MAY CONCERN

Dear Sir/Madam,

PERMISSION TO CONDUCT A SURVEY

May this letter reach you in the best of health.

Kindly be informed that the following is a student from the Kulliyyah of Information and Communication Technology, International Islamic University Malaysia. He seeks for a kind request to be granted permission to conduct survey for data collection from your organization in order to complete his research project.

Details of the student and the research are as follows: -

Student's Name	**: Md Farhad Hossain (G2121261)**
Programme	**: Master of Computing (Computer Science and Information Technology) (M_CST)**
Research Title	**: Automated Neurofeedback Framework for Mental Fatigue Intervention**
Main Supervisor	**: Asst. Prof. Ts. Dr. Hamwira Sakti Bin Yaacob**
Co- supervisor	**: Assoc. Prof. Dr. Faizah Idrus**

We appreciate if you could extend your kind assistance and cooperation to the above-mentioned student. He will personally discuss with you on the data requirement and the conduct of the research. Please be assured that all information provided shall be treated strictly CONFIDENTIAL and only be used for this research and academic purpose.

Should you need more information, please do not hesitate to contact our office at pa.ddpgkict@iium.edu.my / murni@iium.edu.my.

Thank you, والسلام.

PROF. DR. MURNI MAHMUD
Deputy Dean (Postgraduate & Responsible Research),
Kulliyyah of Information and Communication Technology (KICT)
International Islamic University Malaysia
Email: murni@iium.edu.my
Phone: +603-64216401(office)
c.c : Asst. Prof. Ts. Dr. Hamwira Sakti Bin Yaacob - Supervisor
: Assoc. Prof. Dr. Faizah Idrus – Co-Supervisor
: Student's file (G2121261)

KULLIYYAH OF INFORMATION AND COMMUNICATION TECHNOLOGY (KICT)
International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur
(Company No: 101067-P)
Tel: +603 6421 5601 (Ext: 5603/5605/5612) Fax: +603 6421 5179
www.iium.edu.my/kict

Figure A1 Permission To Conduct Survey For Data Collection

APPENDIX B: IIUM RESEARCH ETHICS COMMITTEE (IREC) APPROVAL

Our Ref. : IIUM/504/14/11/2/ IREC 2023-047
Date : 20 March 2023

Asst. Prof. Dr. Hamwira Yaacob (Principal Investigator)
Kulliyyah of Information and Communication Technology
IIUM Gombak Campus
53100 Gombak

Dear Asst. Prof. Dr.,

The IIUM Research Ethics Committee (IREC) has reviewed your study protocol as mentioned below:-

ID NO.	: **IREC 2023-047**
RESEARCH TITLE	: **Automated Neurofeedback Model for Mental Fatigue Intervention**
REGISTRATION DATE	: **23 Feb 2023**
CO-INVESTIGATOR	: **Assoc. Prof. Dr. Faizah Idrus**
STUDENT	: **Md Farhad Hossain (Postgraduate Student)**
STUDY SITE	: **Kulliyyah of Information and Communication Technology**
SAMPLE SIZE	: **20**
ETHICAL EXPIRY DATE	: **20 March 2024**

The IIUM Research Ethics Committee (IREC) operates in accordance to the Declaration of Helsinki, International Conference of Harmonization Good Clinical Practice Guidelines (ICH-GCP), Malaysia Good Clinical Practice Guidelines and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines

The following documents have been received and reviewed to the above study:-

1. Study Proposal/Protocol: Version 1, dated 28 Feb 2023
2. Informed Consent Form (ICF) –
   i. Information Sheet (English) – Version 1, dated 28 Feb 2023
   ii. Consent Form (English) - Version 1, dated 28 Feb 2023
3. Approval Letter from Kulliyyah of Information and Communication Technology, IIUM
4. Principal Investigator's CV

Figure B1 IIUM Research Ethics Committee (IREC) Approval Page 1

Decision by IIUM Research Ethics Committee (IREC):

(√) Approved
(   ) Disapproved

Date of Approval: 20 March 2023

The investigator(s) are required to:

a) submit the 'Continuing Review Form' 30 days before EXPIRY DATE to renew Ethical Approval.
b) notify IREC of any change in protocol and obtaining further ethical approval as appropriate.
c) report any adverse incident during the course of a study to IREC even if the incident is not directly related to the study.
d) report to the IREC within 72 hours for all internal SAEs (occurring in IIUM PI site).
e) report in a prompt manner if the information impacts the continued ethical acceptability of the trial for external SAEs (occurring in participants at other sites).
f) report any major protocol deviation occurs within 5 working days.
g) complete and submit the End of Project Report Form to the IREC Secretariat's Office.
h) All records and data subjects are CONFIDENTIAL and used only for the purposes of this study and all issues and procedures on data confidentiality must be observed.

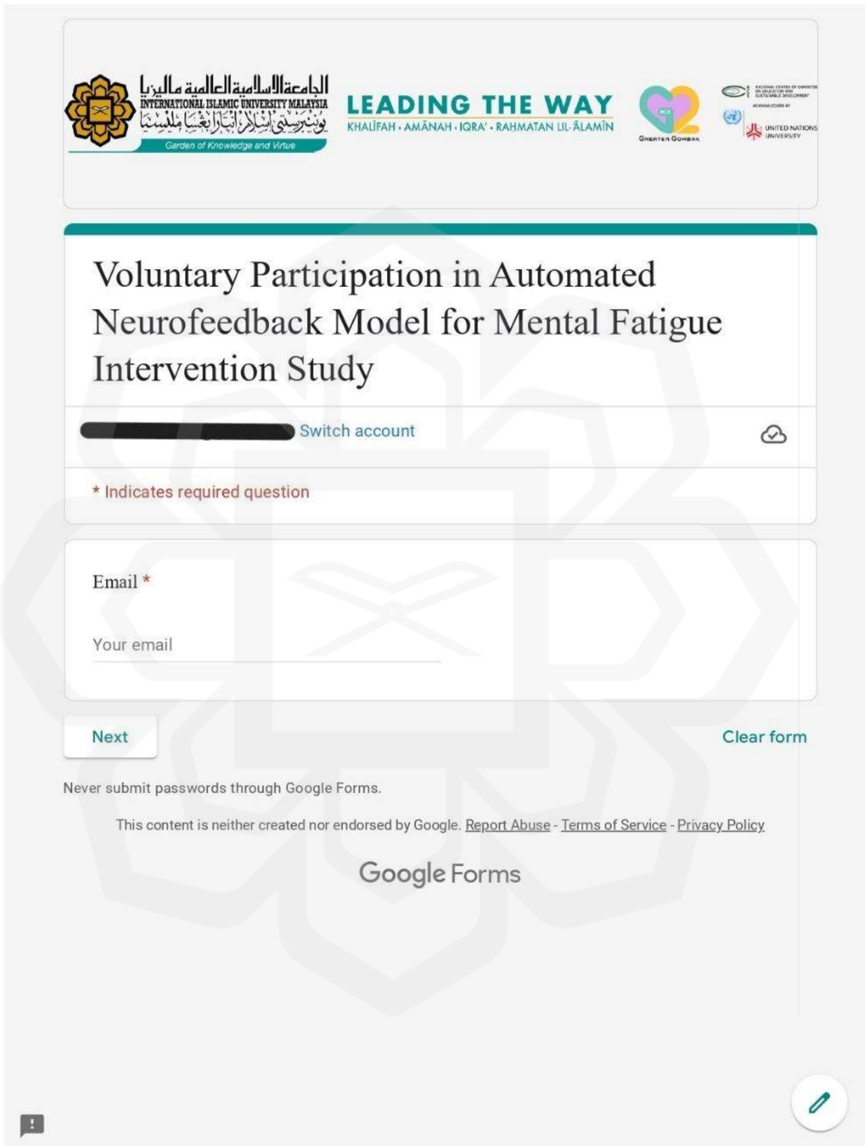
Yours sincerely,

PROF. DR. NASSER MUHAMMAD AMJAD
Chairman
IIUM Research Ethics Committee (IREC)

DISCLAIMER: The approval letter only covers the ethical aspect of your study only. Any other permission/approval to use any facilities, data or human resource should fall under applicant's responsibility.

Figure B2 IIUM Research Ethics Committee (IREC) Approval Page 2

APPENDIX C: CONSENT FORM



INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

Garden of Knowledge and Virtue

LEADING THE WAY

KHALĪFAH · AMĀNAH · IQRA' · RAHMATAN LIL-ĀLAMĪN

Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

[redacted] Switch account

* Indicates required question

Email *

Your email

Next

Clear form

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure C1 Consent Form Page 1

LEADING THE WAY

KHALĪFAH • AMĀNAH • IQRA' • RAHMATAN LIL-ĀLAMĪN

Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

[redacted] Switch account

* Indicates required question

Informed Consent

IREC ID: IREC 2023-047
Version: 1
Revision: 00
Date:

PART I: Information Sheet

I am a Md Farhad Hossain. I am a Master of Computing (Computer Science and Information Technology)-Research at the International Islamic University Malaysia's Kulliyyah of Information and Communication Technology (KICT). I am conducting my research under Dr. Hamwira Yaacob's supervision. My research project is entitled "Automated Neurofeedback Model for Mental Fatigue Intervention" as part of my research.

There may be some words/terms that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask them about me, the study. The following information is provided for you to decide whether you wish to participate in the present study.

Purpose of the Research

Figure C2 Consent Form Page 2

Mental fatigue (MF) is an experience of tiredness that develops as your brain's energy is drained. Mental fatigue is frequently caused by prolonged stress. Mental Fatigue leads to cognitive deficits: stress, anxiety, panic, loneliness, difficulty concentrating, and even depression. Mental fatigue is particularly common among students. It's more difficult to recall and retain the knowledge you've spent hours studying when your brain is fatigued. It is essential to early address mental fatigue because it leads to permanent mental disorders. When a person becomes fatigued, an intervention method can be useful.

There are many different intervention strategies available that target different kinds of problems. Interventions like physical exercise, watching videos, mental calculation tasks, changes in light brightness, adjusting temperature, and changes in monitor contrast can reduce fatigue. Neurofeedback (NF) can be an intervention technique. It uses cognitive strategies to reinforce healthy brain function. A Neurofeedback strategy can be a design of sound or video or anything that may alter brain signals to improve mental health.

The Brain-Computer Interface (BCI) can detect Mental Fatigue by acquiring electroencephalogram (EEG) signals and translating them into a concise command. The electroencephalogram (EEG) is a common, non-invasive approach for monitoring and evaluating the brain's condition. Indeed, this research works to develop an automated neurofeedback model by adapting an MF detection method and an NF modality as an MF intervention method.

Type of Research Intervention

This research will involve an EEG signal collection through Emotiv Insight 2.0 headset. There will be a video presentation. While the BCI detects your fatigued level, the monitor contrast will change accordingly to reduce your fatigue level.

Participant Selection

We are inviting students to participate in this research, who are healthy, non-smokers, and not drug-addicted.

Voluntary Participation

Your participation in this study is voluntary. You are free not to participate in or withdraw from the study at any time. Your decision will not result in any loss of benefits to which you are entitled. If you choose to participate, you may withdraw at any time by notifying me. Upon your request to withdraw, all information about you will be removed.

Procedures and Protocol

You will wear the Emotiv Insight 2.0 headset, collecting your brain rhythm with a sampling rate of 128 HZ. After that, you need to select your preferred monitor contrast. Next, 1-minute of eyes close and 1-minute eyes open task for you to prepare. Then, a YouTube tutorial video will be played for you, which long around 30 Minutes. In these 30 minutes of the learning session, if the system detects that you are fatigued then the monitor contrast will be changed to intervene to make changes to your

Figure C3 Consent Form Page 3

fatigue level. If the system does not detect then there won't be any changes in the monitor contrast. After the tutorial video, the session will end.

Duration

The research takes place only for one session around 1 hour long.

Side Effects

The EEG signal collection through Emotiv Insight 2.0 is painless. You may feel only discomfort while wearing the Headset. Participants rarely experience irritated skin on the forehead, but it is enough to observe that because it almost disappears spontaneously. However, we will follow you closely and keep track of any unwanted effects or problems during the data collection. Even, we may stop the procedure. If this is necessary, we will discuss it together with you and you will always be consulted before we move to the next step.

Risks

By participating in this research, the possibility of complications is very low. Participants rarely experience irritated skin on the forehead, but it is enough to observe that because it almost disappears spontaneously.

Benefits

There may not be any benefit for you but your participation is likely to help us find the answer to the research question. Future generations are likely to benefit.

Confidentiality

The information that we collect from this research project will be kept confidential. Information about you that will be collected during the research will be put away and no one but the researchers will be able to see it. Any information about you will have a number on it instead of your name. Only the researchers will know what your number is and we will lock that information.

Sharing the Results

Nothing that you tell us today will be shared with anybody outside the research team, and nothing will be attributed do you by name. Each participants will get a summary of the result. The result of this research will be presented at meetings or in seminar or in publications. Your identity will not be disclosed in those presentation.

Right to Refuse or Withdraw

You do not have to take part in this research if you do not wish to do so and refusing to participate will not affect you in any way. You may stop participating in the research at any time that you wish.

Who to Contact

Figure C4 Consent Form Page 4

If you have any questions, you may ask them now or later, even after the study has started. If you wish to ask questions later, you may contact any of the following:

Md Farhad Hossain, Email: emamulfarhat@gmail.com
Dr. Hamwira Yaacob, Email: hyaacob@iium.edu.my

This proposal has been reviewed and approved by IIUM Research Ethics Committee (IREC), which is a committee whose task is to make sure that research participants are protected from harm. If you wish to find out more about the IREC, email irec@iium.edu.my.

You can ask me any more questions about any part of the research study if you wish to. Do you have any questions?

PART II: Certificate of Consent

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate as a participant in this research. I also consent to provide my demographic information for this research.

Statement by the researcher/person taking consent

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands the research.

I confirm that the participant was allowed to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily. A copy of this ICF has been provided to the participant.

Do you agree to participate in this EEG signal collection process and provide demographic information? *

○ Yes

Back Next Clear form

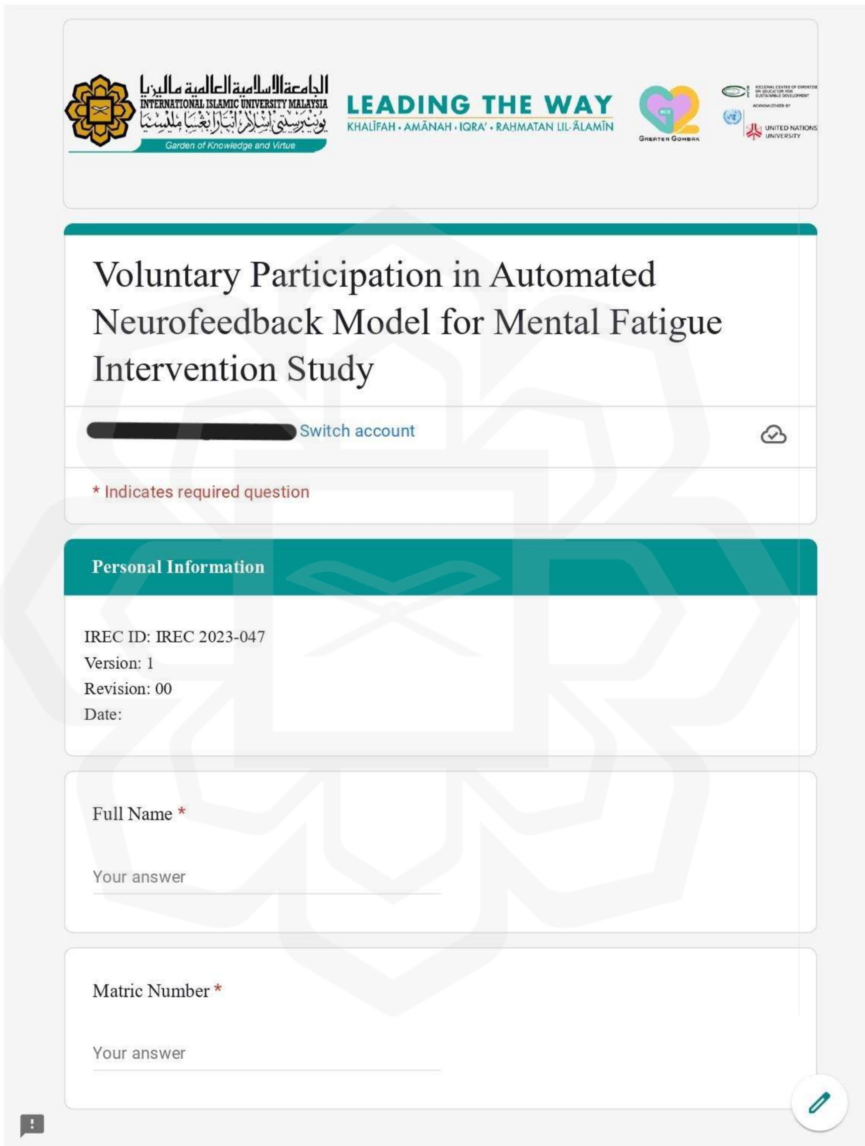
Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure C5 Consent Form Page 5

APPENDIX D: DEMOGRAPHIC INFORMATION FORM



INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

Garden of Knowledge and Virtue

LEADING THE WAY

KHALĪFAH · AMĀNAH · IQRA' · RAHMATAN LIL-ĀLAMĪN

Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

[redacted] Switch account

* Indicates required question

Personal Information

IREC ID: IREC 2023-047
Version: 1
Revision: 00
Date:

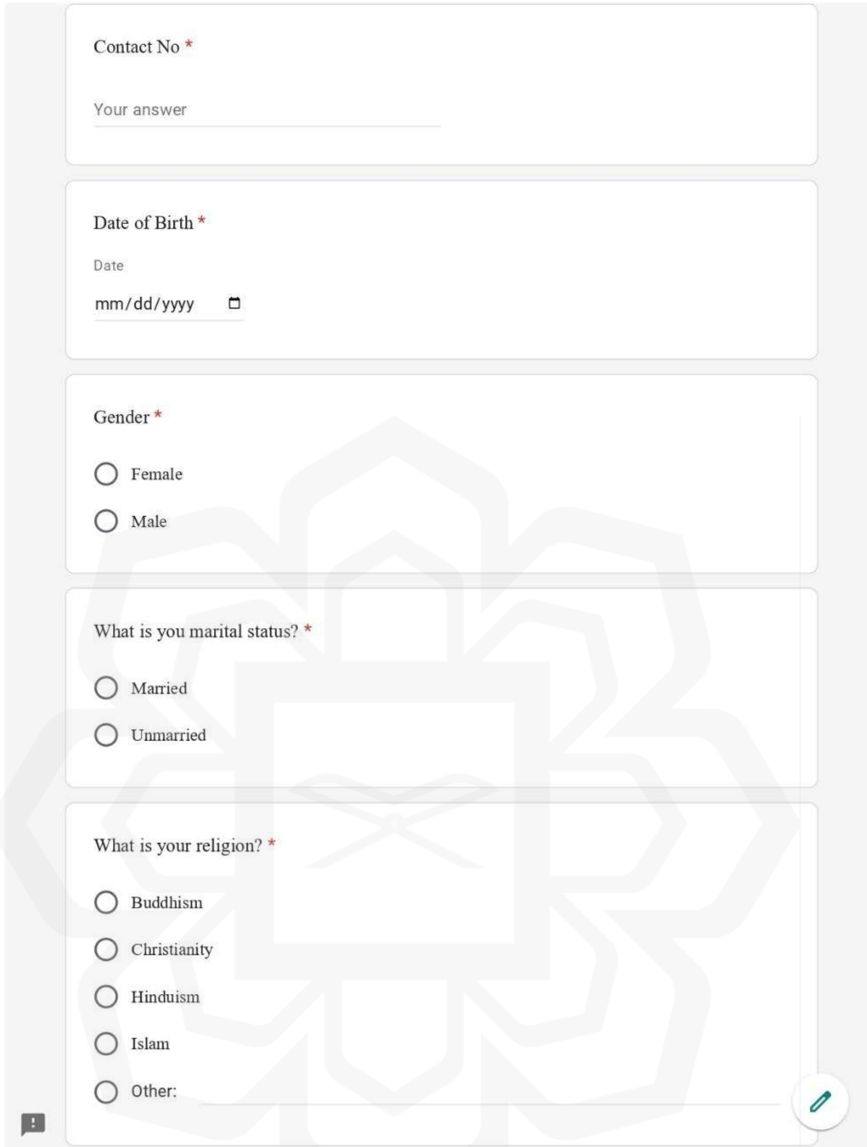
Full Name *

Your answer

Matric Number *

Your answer

Figure D1 Demographic Information Form Page 1



Contact No *

Your answer

Date of Birth *

Date

mm/dd/yyyy

Gender *

- Female
- Male

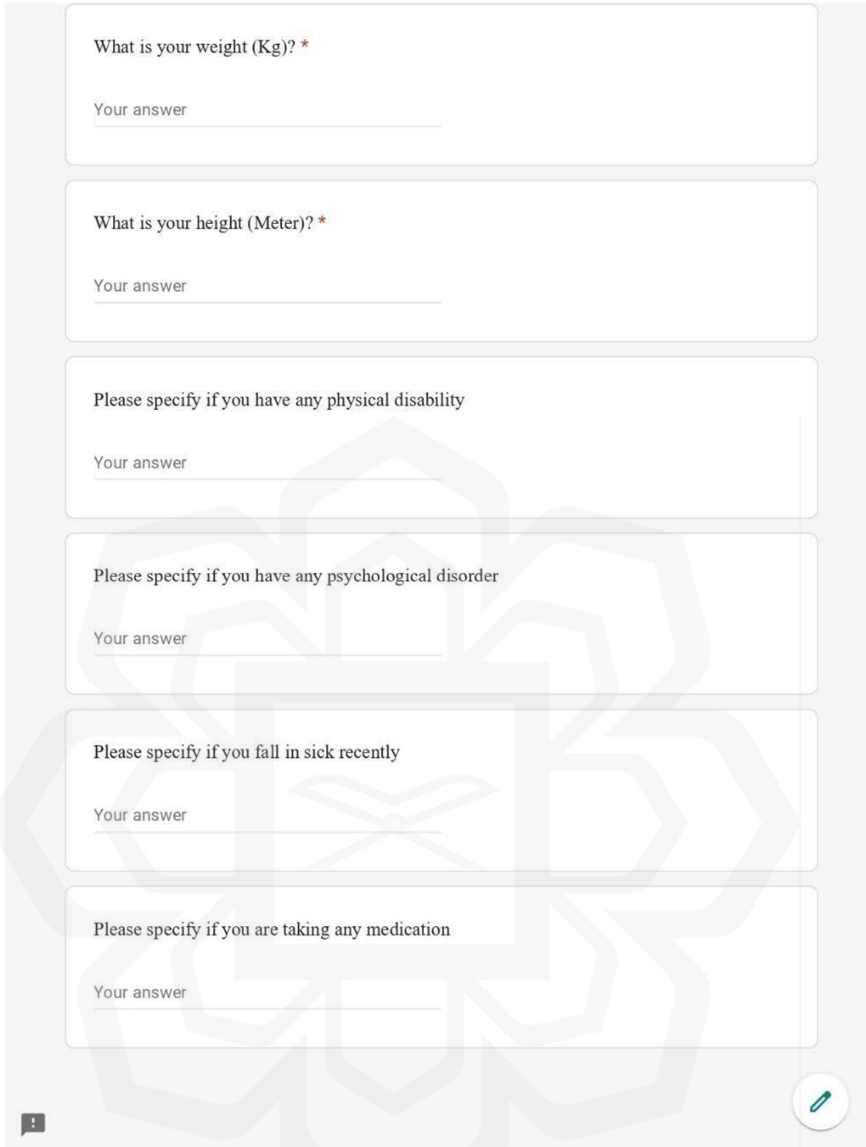
What is you marital status? *

- Married
- Unmarried

What is your religion? *

- Buddhism
- Christianity
- Hinduism
- Islam
- Other:

Figure D2 Demographic Information Form Page 2


What is your weight (Kg)? *

Your answer

What is your height (Meter)? *

Your answer

Please specify if you have any physical disability

Your answer

Please specify if you have any psychological disorder

Your answer

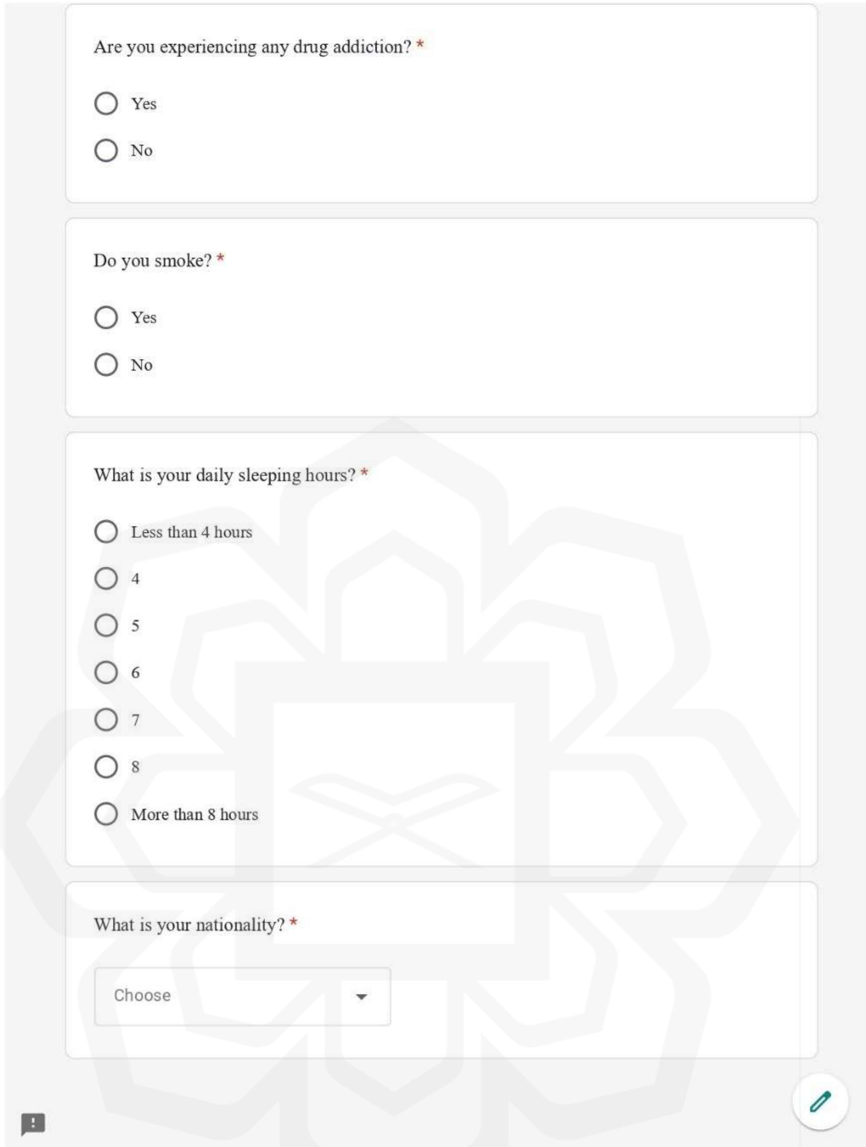
Please specify if you fall in sick recently

Your answer

Please specify if you are taking any medication

Your answer

Figure D3 Demographic Information Form Page 3



Are you experiencing any drug addiction? *

- Yes
- No

Do you smoke? *

- Yes
- No

What is your daily sleeping hours? *

- Less than 4 hours
- 4
- 5
- 6
- 7
- 8
- More than 8 hours

What is your nationality? *

Choose

Figure D4 Demographic Information Form Page 4

Which academic degree are you pursuing? *

- ◯ Bachelor
- ◯ Master
- ◯ Doctor of Philosophy (PhD)
- ◯ Other: ______

What is your current study level or year? *

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4

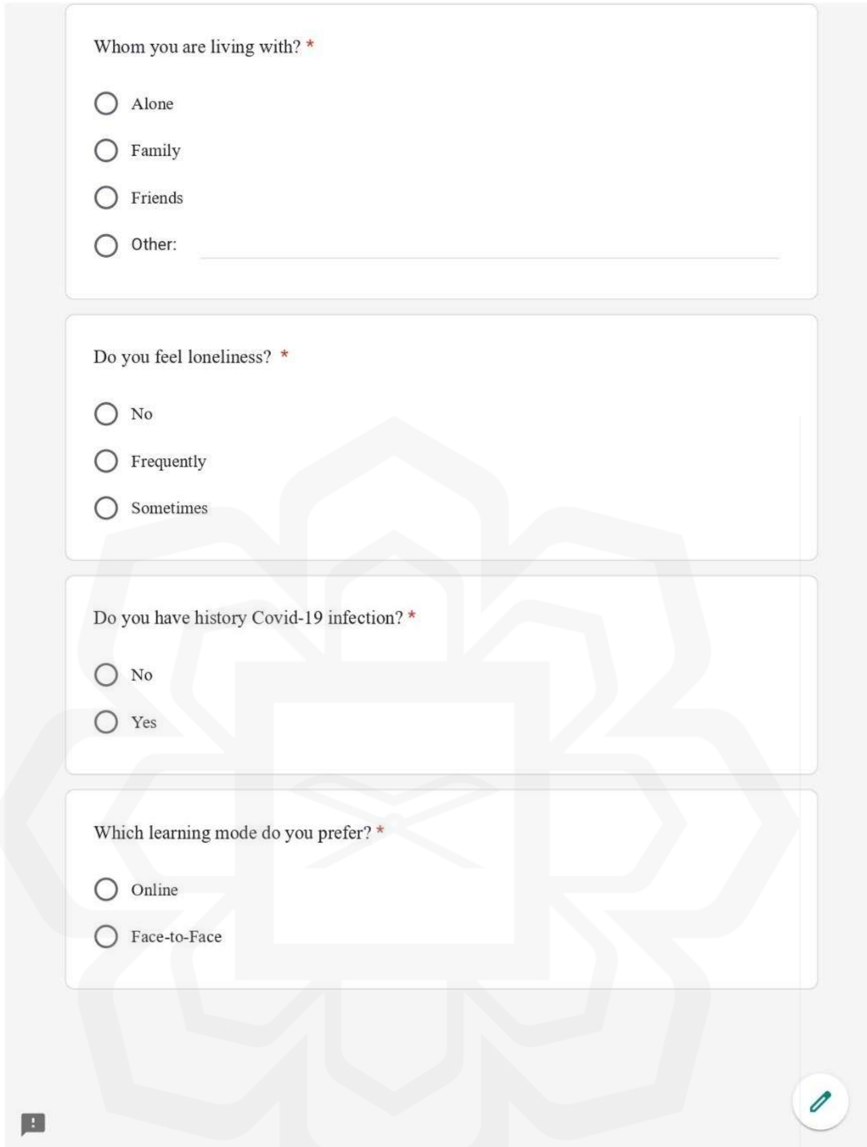
How many credit hours have you taken in this semester? *

Your answer

Are you employed? *

- ◯ No
- ◯ Part-time
- ◯ Freelance
- ◯ Full-time

Figure D5 Demographic Information Form Page 5



Whom you are living with? *

- Alone
- Family
- Friends
- Other: ______

Do you feel loneliness? *

- No
- Frequently
- Sometimes

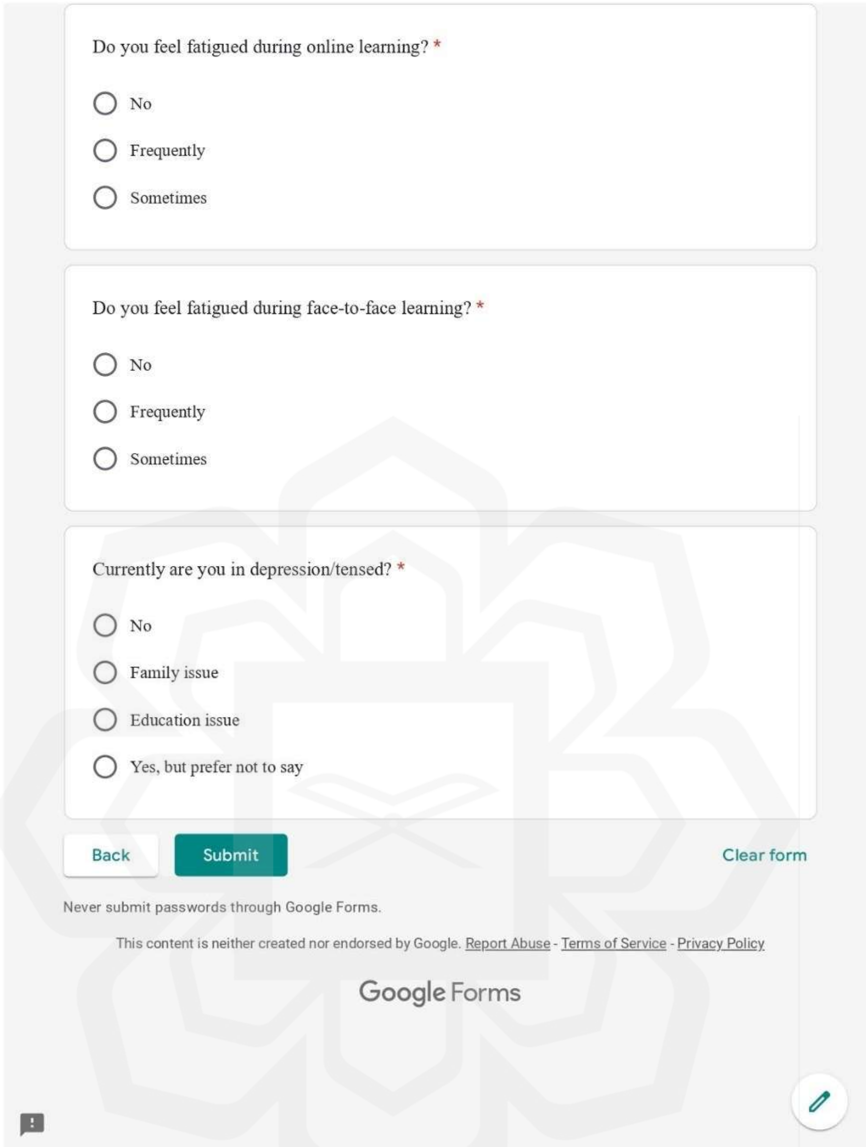
Do you have history Covid-19 infection? *

- No
- Yes

Which learning mode do you prefer? *

- Online
- Face-to-Face

Figure D6 Demographic Information Form Page 6



Do you feel fatigued during online learning? *

- No
- Frequently
- Sometimes

Do you feel fatigued during face-to-face learning? *

- No
- Frequently
- Sometimes

Currently are you in depression/tensed? *

- No
- Family issue
- Education issue
- Yes, but prefer not to say

Back Submit Clear form

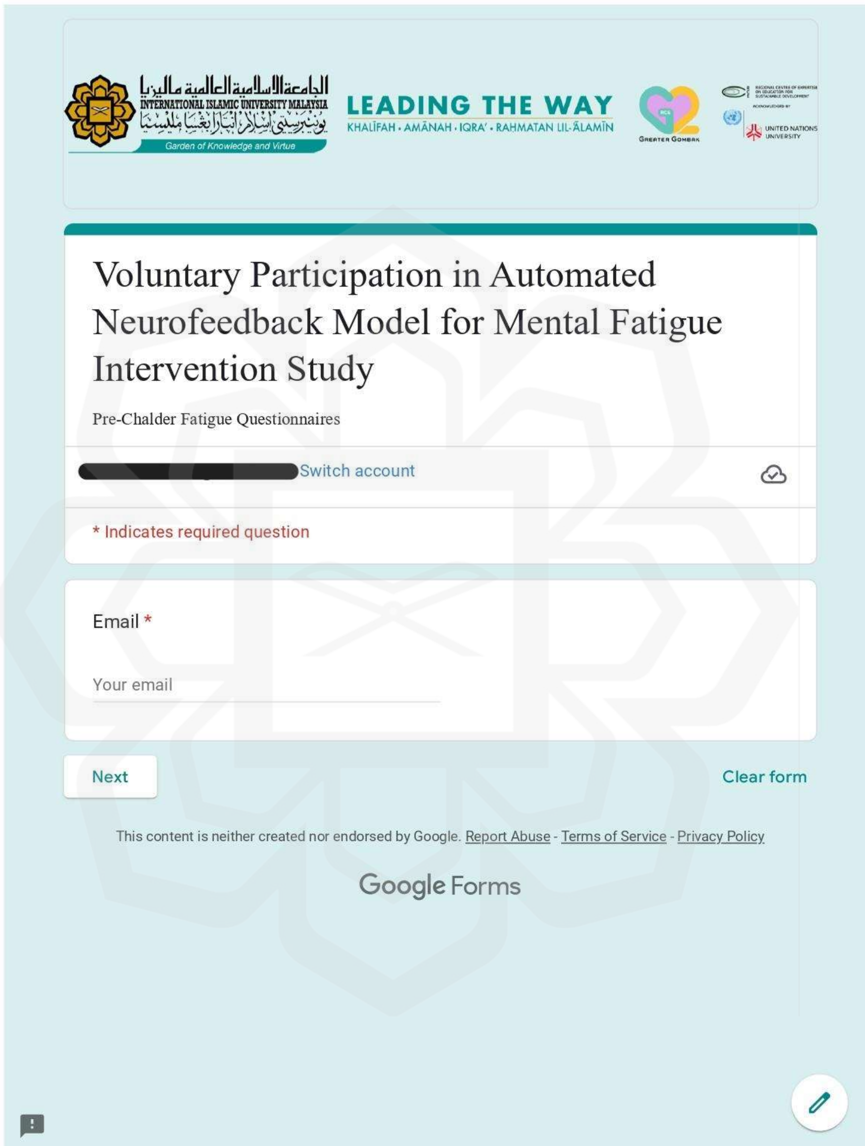
Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure D7 Demographic Information Form Page 7

APPENDIX E: PRE-CHALDER FATIGUE QUESTIONNAIRES



Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

Pre-Chalder Fatigue Questionnaires

[redacted] Switch account

* Indicates required question

Email *

Your email

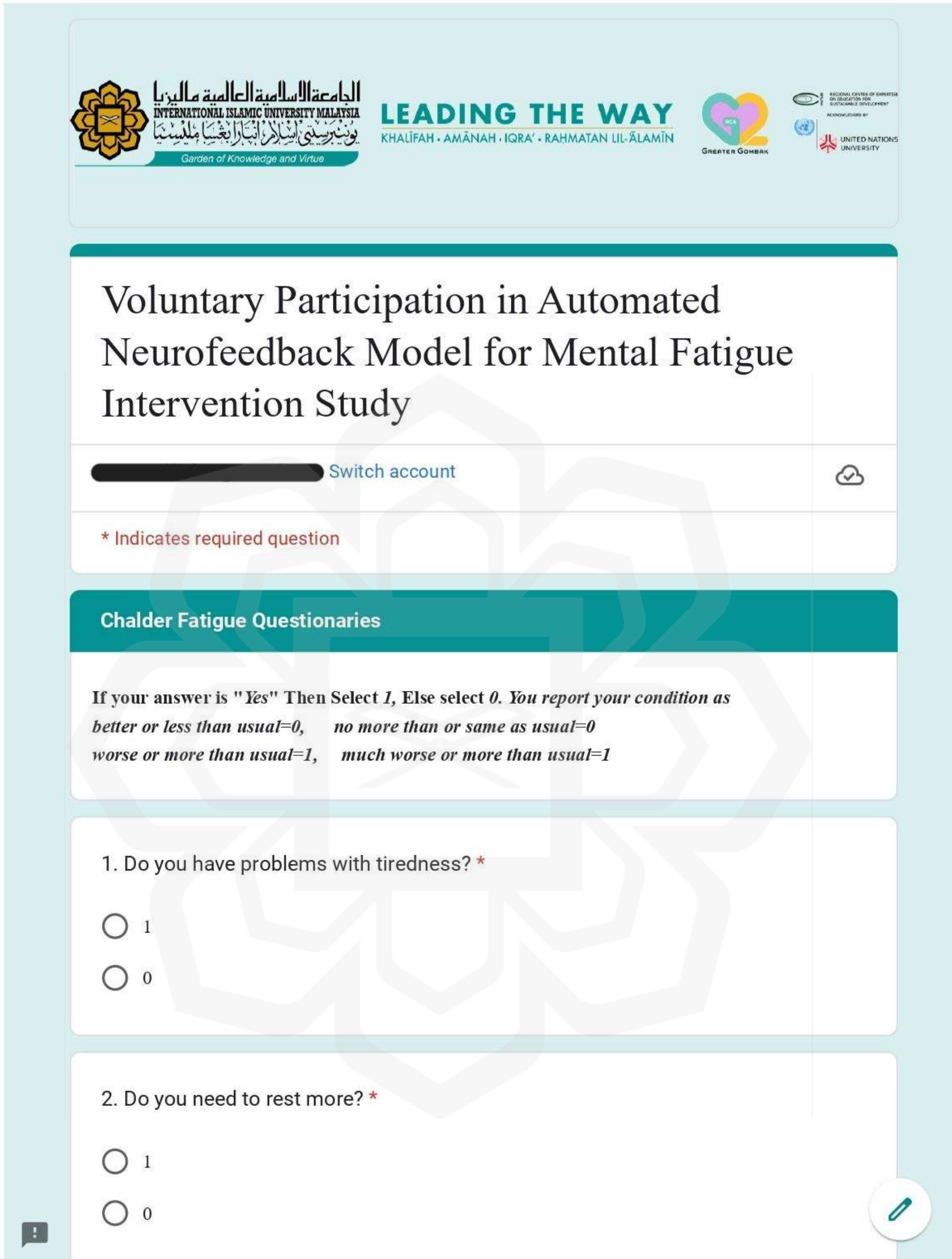
Next

Clear form

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure E1 Pre-Chalder Fatigue Questionnaires Form Page 1


LEADING THE WAY
KHALĪFAH • AMĀNAH • IQRA' • RAHMATAN LIL-ĀLAMĪN

Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

[redacted] Switch account

* Indicates required question

Chalder Fatigue Questionaries

If your answer is "*Yes*" Then Select *1*, Else select *0*. *You report your condition as better or less than usual=0, no more than or same as usual=0 worse or more than usual=1, much worse or more than usual=1*

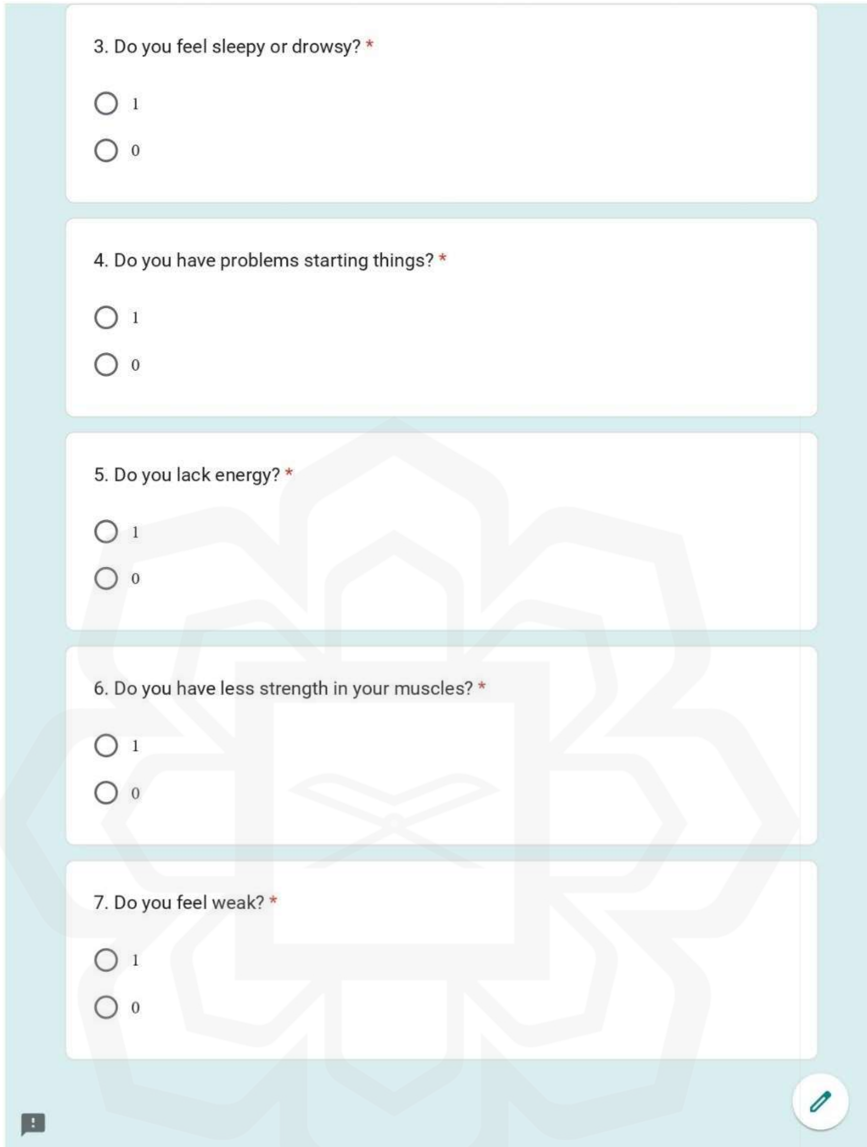
1. Do you have problems with tiredness? *

- ◯ 1
- ◯ 0

2. Do you need to rest more? *

- ◯ 1
- ◯ 0

Figure E2 Pre-Chalder Fatigue Questionnaires Form Page 2



3. Do you feel sleepy or drowsy? *

1

0

4. Do you have problems starting things? *

1

0

5. Do you lack energy? *

1

0

6. Do you have less strength in your muscles? *

1

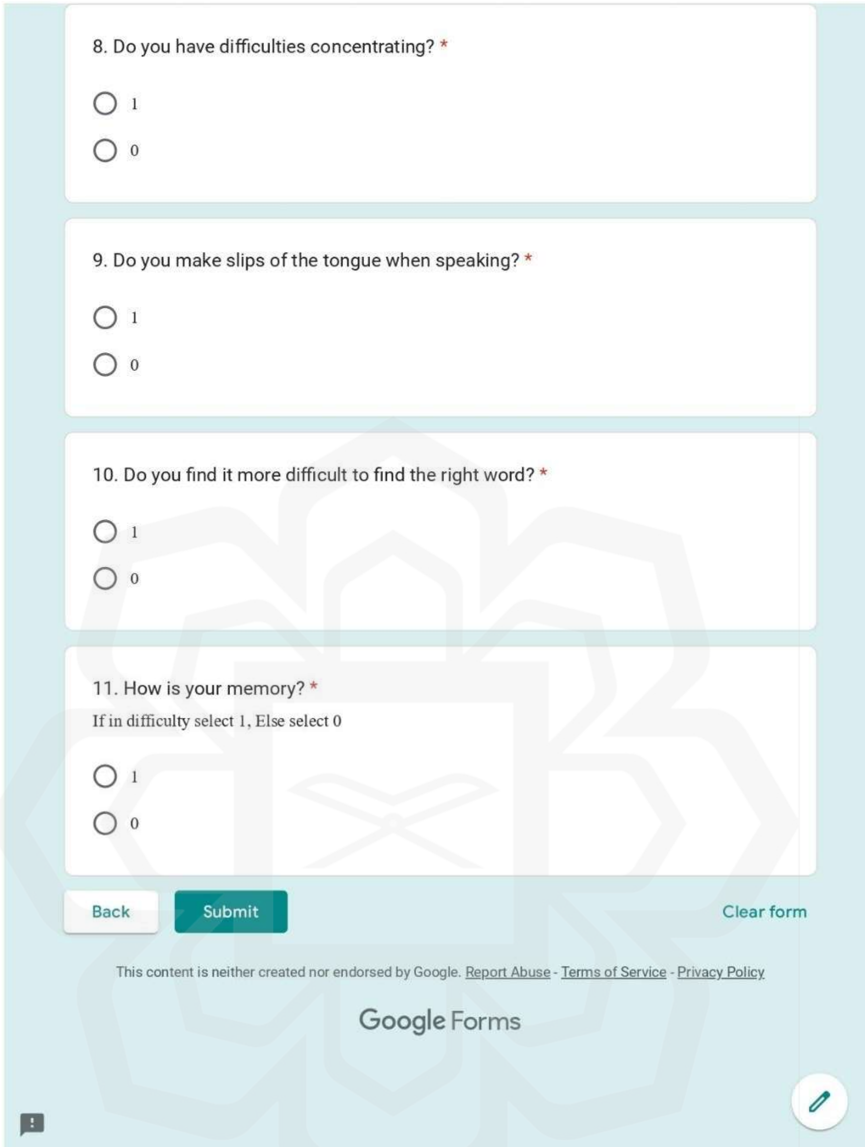
0

7. Do you feel weak? *

1

0

Figure E3 Pre-Chalder Fatigue Questionnaires Form Page 3



8. Do you have difficulties concentrating? *

- 1
- 0

9. Do you make slips of the tongue when speaking? *

- 1
- 0

10. Do you find it more difficult to find the right word? *

- 1
- 0

11. How is your memory? *

If in difficulty select 1, Else select 0

- 1
- 0

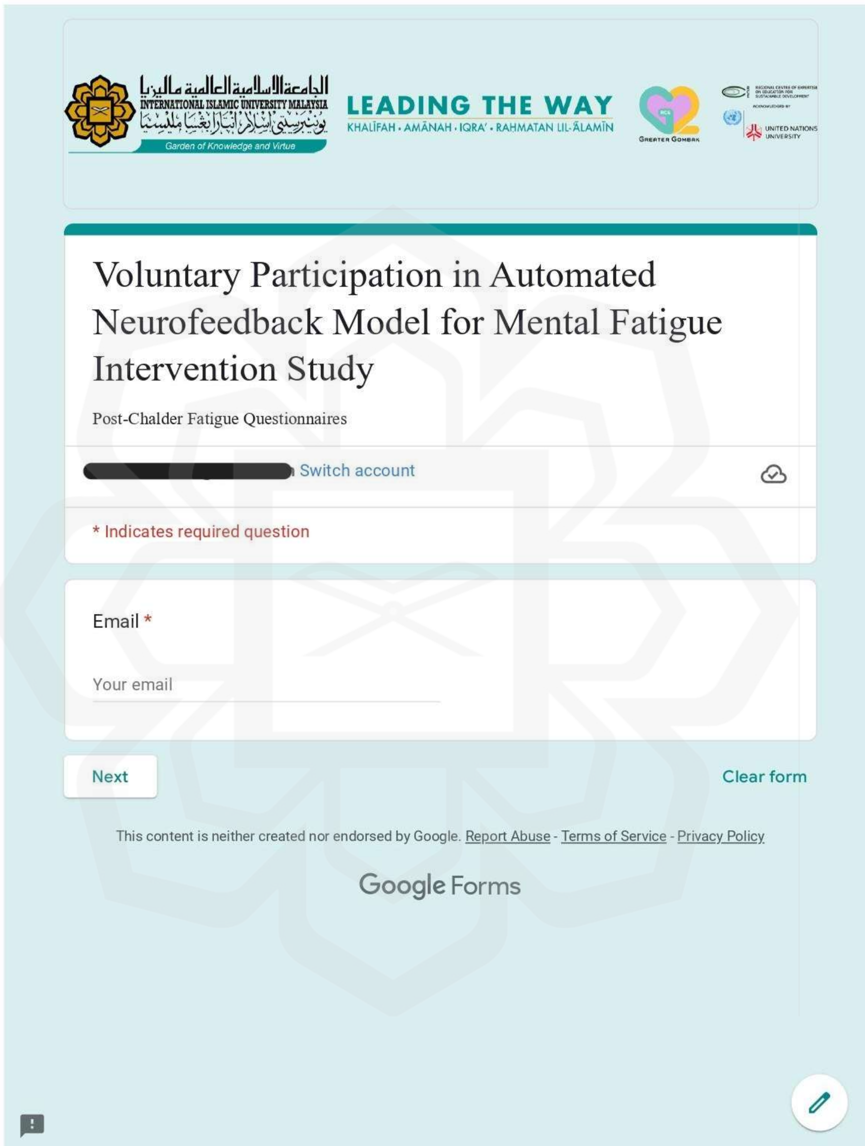
Back Submit Clear form

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure E4 Pre-Chalder Fatigue Questionnaires Form Page 4

APPENDIX F: POST-CHALDER FATIGUE QUESTIONNAIRES



LEADING THE WAY

KHALĪFAH · AMĀNAH · IQRA' · RAHMATAN LIL-ĀLAMĪN

Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

Post-Chalder Fatigue Questionnaires

[redacted] Switch account

* Indicates required question

Email *

Your email

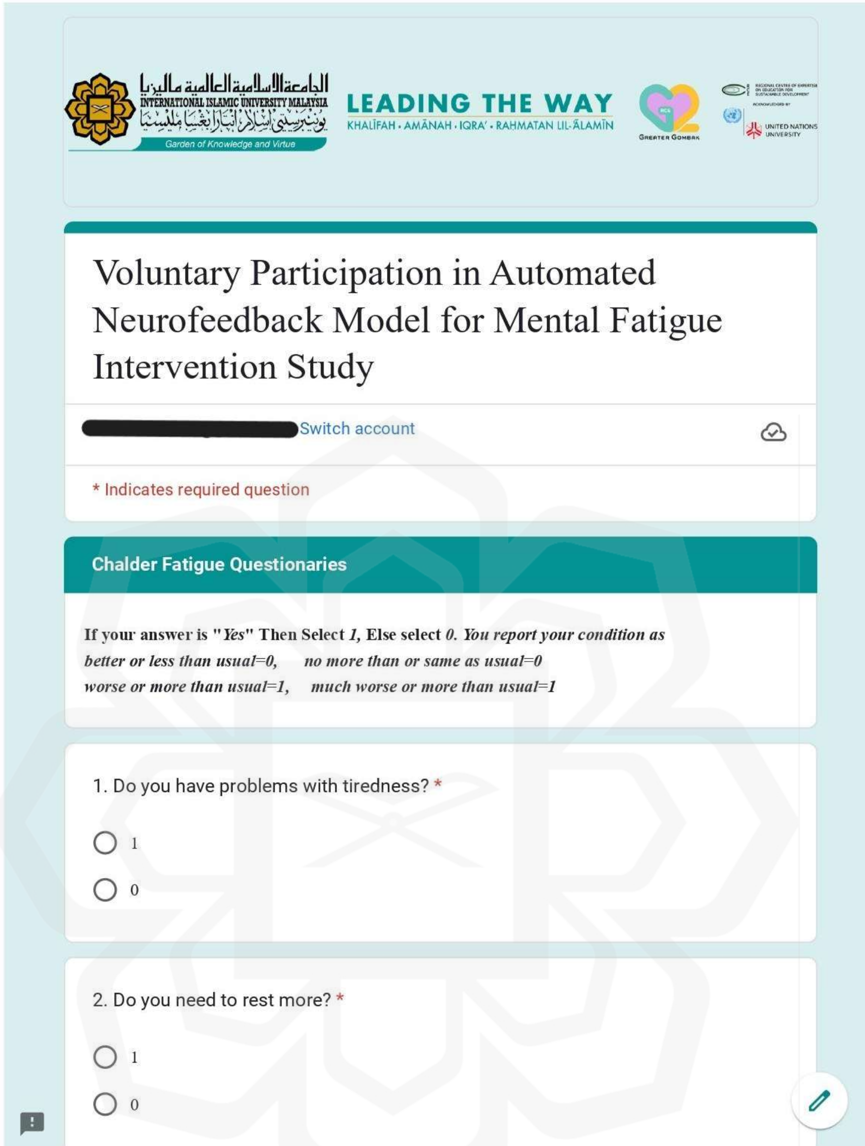
Next

Clear form

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure F1 Post-Chalder Fatigue Questionnaires Form Page 1



Voluntary Participation in Automated Neurofeedback Model for Mental Fatigue Intervention Study

[redacted] Switch account

* Indicates required question

Chalder Fatigue Questionaries

If your answer is "*Yes*" Then Select *1*, Else select *0. You report your condition as*
better or less than usual=0,   no more than or same as usual=0
worse or more than usual=1,   much worse or more than usual=1

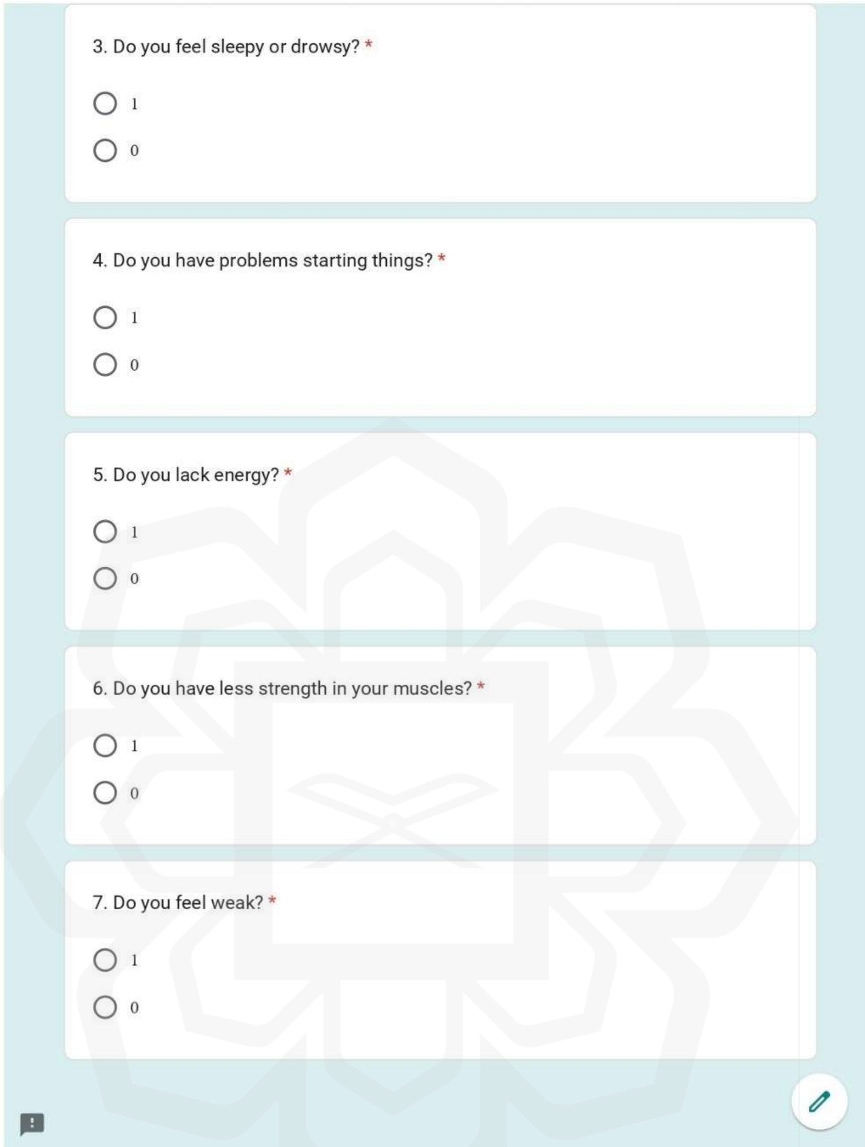
1. Do you have problems with tiredness? *

- ◯ 1
- ◯ 0

2. Do you need to rest more? *

- ◯ 1
- ◯ 0

Figure F2 Post-Chalder Fatigue Questionnaires Form Page 2



3. Do you feel sleepy or drowsy? *

- 1
- 0

4. Do you have problems starting things? *

- 1
- 0

5. Do you lack energy? *

- 1
- 0

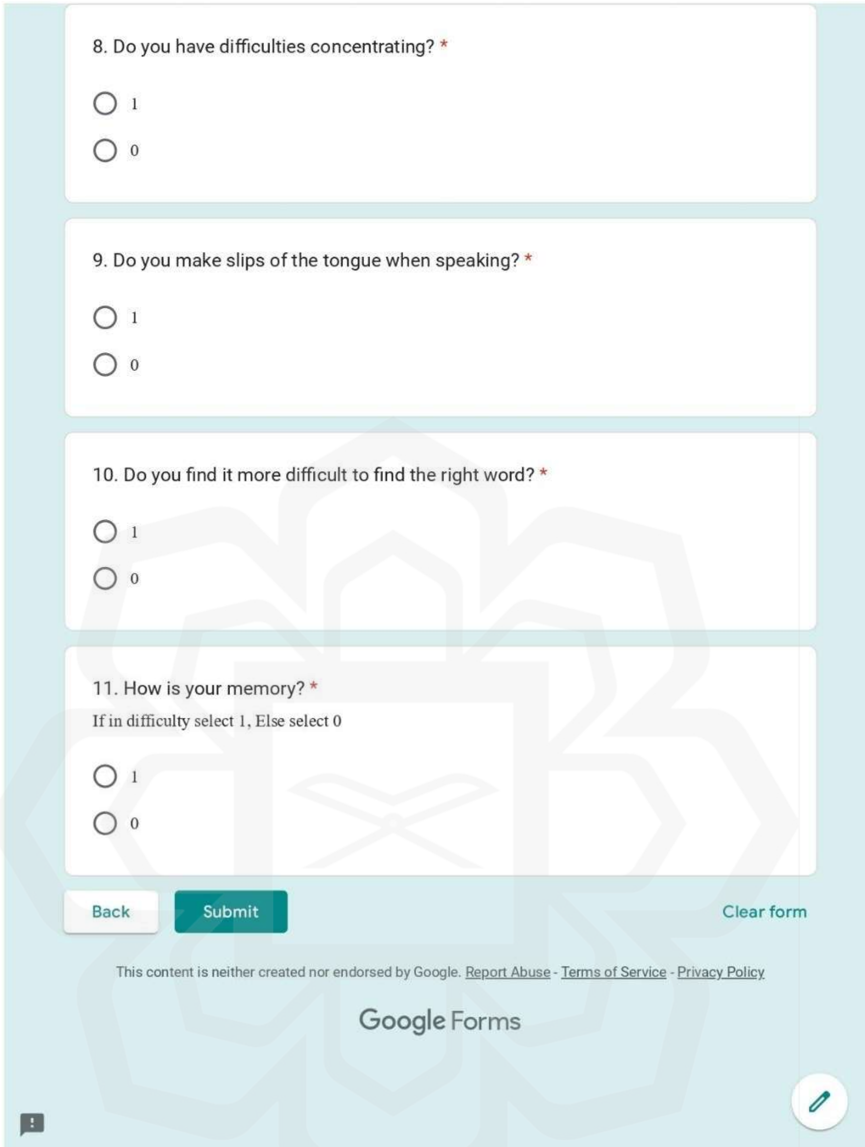
6. Do you have less strength in your muscles? *

- 1
- 0

7. Do you feel weak? *

- 1
- 0

Figure F3 Post-Chalder Fatigue Questionnaires Form Page 3



8. Do you have difficulties concentrating? *

- 1
- 0

9. Do you make slips of the tongue when speaking? *

- 1
- 0

10. Do you find it more difficult to find the right word? *

- 1
- 0

11. How is your memory? *

If in difficulty select 1, Else select 0

- 1
- 0

Back Submit Clear form

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Figure F4 Post-Chalder Fatigue Questionnaires Form Page 4

APPENDIX G: SUBJECT-DEPENDENT DETECTION MODEL PERFORMANCE

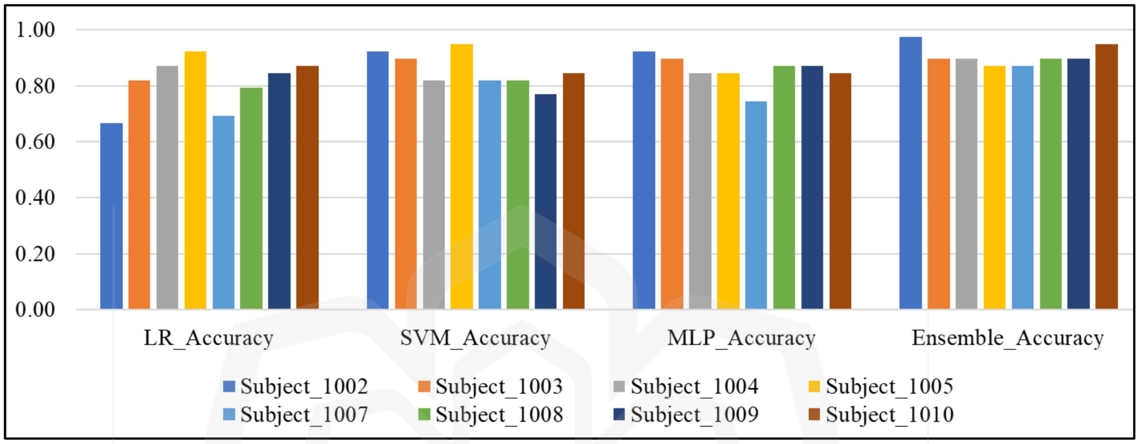


Figure G1 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using Accuracy

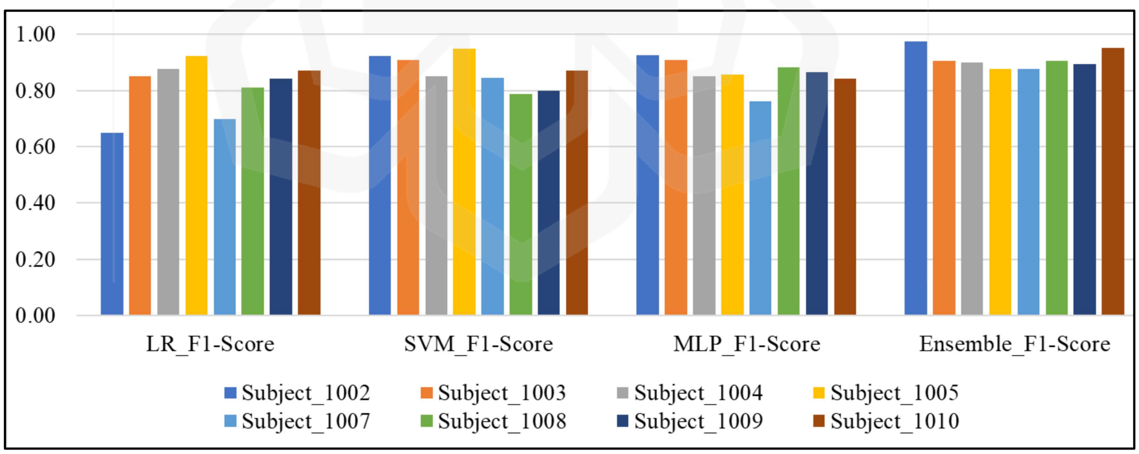


Figure G2 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using F1-Score

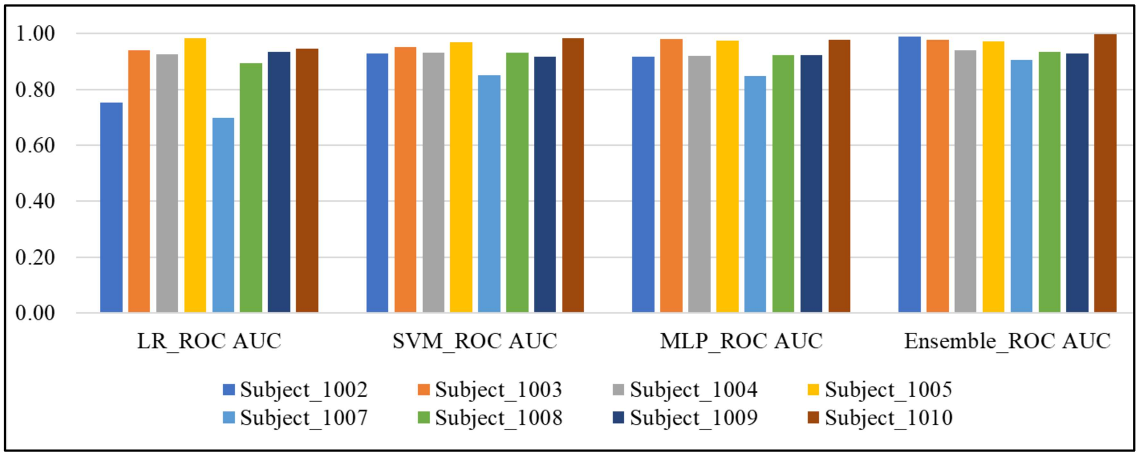


Figure G3 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using ROC AUC Score

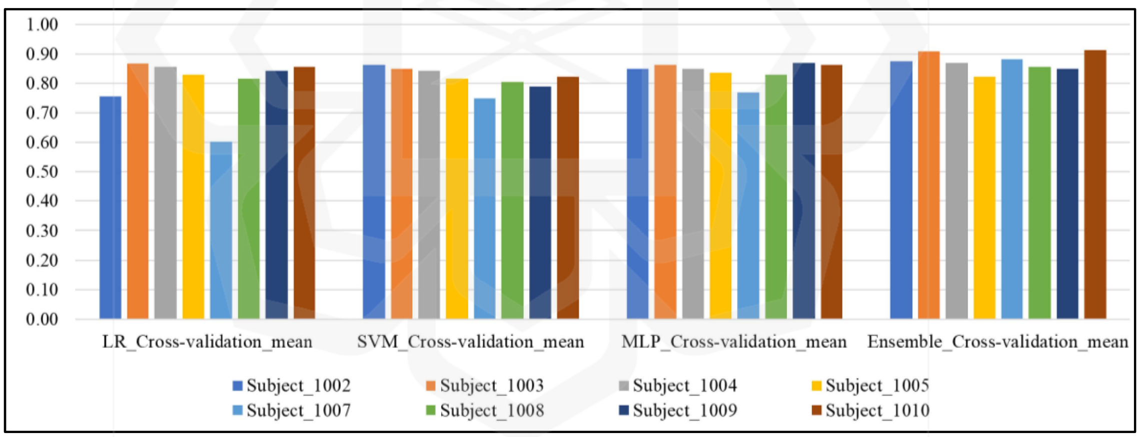


Figure G4 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using Cross-Validation Mean

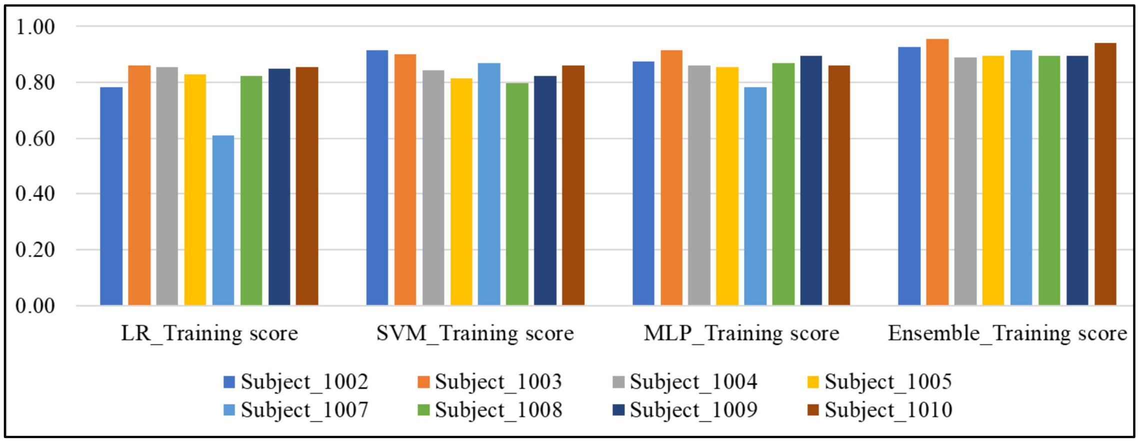


Figure G5 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using Cross-Validation Training Score

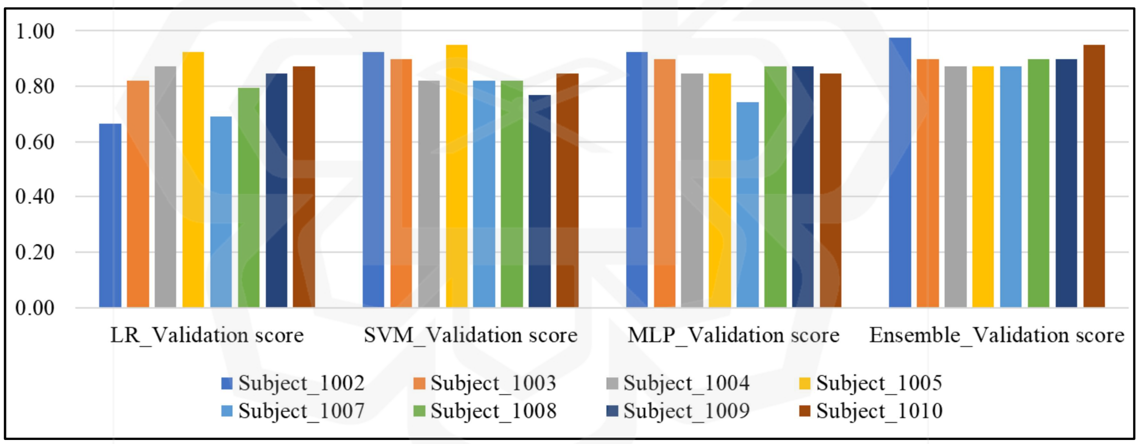


Figure G6 Subject-Dependent Mental Fatigue Detection Models Performance Comparison Using Cross-Validation Validation Score