5-2022

Multi-modal Human Fatigue Classification using Wearable Sensors for Human-Robot Teams

Likhitha Nagahanumaiah
ln2047@rit.edu

Multi-modal Human Fatigue Classification using Wearable Sensors for Human-Robot Teams

by

Likhitha Nagahanumaiah

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science
in Electrical Engineering

Supervised by

Dr. Jamison Heard
Department of Electrical and Microelectronic Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, New York
May 2022

Approved by:

Dr. Jamison Heard, Assistant Professor
Thesis Advisor, Department of Electrical and Microelectronic Engineering

---

Dr. Ferat Sahin, Department Head
Department Head, Department of Electrical and Microelectronic Engineering

---

Dr. Gill Tsouri, Professor
Committee Member, Department of Electrical and Microelectronic Engineering

---

Dr. Ferat Sahin, Department Head
Department Head, Department of Electrical and Microelectronic Engineering

Thesis Release Permission Form

Rochester Institute of Technology
Kate Gleason College of Engineering

Title:

Multi-modal Human Fatigue Classification using Wearable Sensors for Human-Robot Teams

I, Likhitha Nagahanumaiah, hereby grant permission to the Wallace Memorial Library to reproduce my thesis in whole or part.

Likhitha Nagahanumaiah

Date

Dedication

I dedicate this work to my loving family who have supported me all the time...

Acknowledgments

Firstly, I want to express my appreciation and gratitude to my advisor Dr. Jamison Heard for his guidance and emotional support throughout my Master's studies. His research expertise and high expectations pushed me to conduct high quality research and developed the fundamental skills to conduct such research.

I also want to thank the members of the my Lab members at Century Mold and MABL lab at Rochester Institute of Technology.

Lastly, I want to thank my parents, who supported me from the beginning and believed in my abilities.

Abstract

Multi-modal Human Fatigue Classification using Wearable Sensors for Human-Robot Teams

Likhitha Nagahanumaiah

Supervising Professor: Dr. Jamison Heard

Our main objective of this study is to create a fatigue detection model using real-time data by using wearable sensors. The purpose of this research is to learn more about the way humans experience fatigue in a supervisory human-machine environment. The goal of this study is to evaluate machine learning algorithms that assess fatigue detection and to use robots for adapting its interactions.

The environment itself consists of two different tasks to analyze Physical fatigue and Mental fatigue in two different task environments that are (i) Jigsaw puzzle-solving task, and (ii) Pick and Place task. Physical fatigue and mental fatigue are detected using wearable sensors: MYO armband and BioPac Bioharness.

During the experiment, the Physiological metrics used are Heart rate, respiration rate, Heart rate variability, posture, breathing wave amplitude, and EMG. All these Physiological signals are collected simultaneously in a real-time task environment. The data collected by these physiological signals are then processed and machine learning and deep learning algorithms are used for further process in building a fatigue detection model.

List of Contributions

- Modeled a Fatigue classification system using physiological metrics like heart rate, heart rate variability, EMG, respiration rate, and posture.
- The study helps in detecting both physical and mental fatigue using physiological metrics
- Approached data collection by using non- invasive wearable sensors (Zephyr Bio-harness and MYO armband).
- Compared machine learning and deep learning algorithm to classify five fatigue levels.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Physiological fatigue analysis is a more vital factor to contemplate in a Human-Robot working space. Human-Robot interaction is the study of the interaction between robots and humans. Due to the pressure involved in the work that needs continuous or continuous physical activity, the worker experiences tiredness or sleepiness which is termed fatigue. Fatigue is the more complicated factor to be determined because fatigue can be not only physical but also mental.

Fatigue affects everyone, no matter their strength, experience, or preparation. This has a direct impact on many people's physical and mental abilities, which are required to perform even basic tasks. The most common fatigue symptoms are decreased task motivation, longer response time, decreased alertness, impaired attention, poor psychometric performance, memory, and information processing issues, and poor judgment [1]. These symptoms can look simple but it could be a great loss to the human robots teams. Addressing fatigue is predominant which is ignored in most of the working environments where humans are required to interact with robots to do the work.

Fatigue is an important factor that cannot be neglected as fatigue can be both physical and mental. Physical fatigue is the subjective feeling of muscle tiredness due to intense physical activity and mental fatigue is caused due to intense cognitive activity and leads to sleepiness, lack of attention, and laziness. Physical and mental fatigue are linearly proportional as mental fatigue can affect the physical activity of a person and physical exhaustion can also result in mental fatigue. So it is important to study both mental and physical fatigue in any working environment.

Fatigue has typically been measured using subjective approaches, where a questionnaire is filled out after a task. However, these approaches are intrusive and have poor time resolution. Thus, there has been a shift in the community to focus on physiological metrics (i.e, respiration rate, electromyography (EMG), and heart rate). These metrics correlate to mental and/or physical fatigue and can be collected unintrusively using wearable sensors. Such a collection scheme allows continuous monitoring of an individual's fatigue state while minimally impacting the primary task.

Physiological metrics are often fed into machine-learning pipelines which produce a fatigue classification (e.g., high or low fatigue). However, typical approaches focus solely on physical or mental fatigue and only provide two fatigue. This thesis enhances the current research works by building a machine-learning model pipeline that classifies both mental and physical fatigue across five different levels. Data from a human-subjects experiment are used to validate the presented pipeline.

Chapter 2

Background Literature

The result of sleep loss disrupted cardiac rhythms, or an increase in workload is the reduction in the mental or physical activity of a person, which leads to gradual or sudden onset of fatigue. [2]. Fatigue is also be defined as a condition characterized by a lessened capacity of work and reduced efficiency of accomplishment. The fatigued person receives a signal from his body that the continuing activity will stop either physical activity or mental activity. Rest reduces a person's fatigue levels.

Mental fatigue is not able to maintain the ideal cognitive performance of a person. Mental fatigue gradually increases during cognitive activity and it depends on the cognitive ability of a person. It decreases physical performance by increasing drowsiness, lethargy, and direct attention fatigue (a mental mechanism). This results in decreased attention and consciousness in the workplace. Mental fatigue is dangerous during performing tasks that require constant or continuous attention and concentration. Nonetheless, most research on fatigue from a physiological perspective has focused on muscles [3]. Physical fatigue or muscle fatigue is being not able to perform optimally. It is a gradual process and depends on fitness level, age factor, sleep, and overall health. Physical fatigue is activated by an increase in serotonin levels in the central nervous system.

Fatigue has many effects on performance, physiology, cognition, and emotion [4]. Human fatigue has two major classification: mental and physical fatigue. Mental fatigue has been defined as a decrease in task outcomes that require concentration, as well as the recovery of memory-saved information [5]. Physical fatigue is defined as the phenomenon at which the person who is exercising is no longer able to exert the effort that is necessary to complete a task. [6].

Analyzing how the fatigue detection metrics are used to create a fatigue detection model which can be used to analyze fatigue level are learned in Chapters 2.1 and 2.2, which reviews common Physiological metrics and the correlation of these metrics with fatigue, and fatigue assessment algorithms respectively.

2.1 Physiological Metrics

Physiological metrics provide real-time information about a human's internal state, such as fatigue. Physiological metrics correlate with changes in human fatigue (physical or mental) and provide the foundation for creating an effective fatigue classification algorithm. There are many physiological metrics involved in analyzing a fatigue level of a person. The metrics which support fatigue analysis are presented in Table 2.1.

The next section describes the definitions of each physiological metric, and in detail how each physiological metric affects both mental and physical fatigue.

2.1.1 Correlation with Fatigue

Heart Rate is defined as the number of heartbeats in a minute or in general, it can be said as the number of heartbeats per unity time. R-R intervals(RRI) and frequency domain

Table 2.1 Common physiological metrics

Common physiological metrics
Heart rate
Heart-rate variability
Electromyography (EMG)
Respiration rate
Electrodermal activity (EDA)
Galvanic skin response (GSR)
Pulse oximetry
Photoplethysmography (PPG) pulse
Electroencephalogram(EEG)

characteristics changes with the level of mental fatigue due to physiological activities in a day [7]. Tasks have resting and task periods, the mean heart rate increases in the resting period as well as in the task period, but the difference in the mean heart rate between the resting and task period decreases. Also, the variance decreases in both cases [8].

Brown and Steven [9] state that when mental fatigue is considered in a Physical task the heart rate is relatively lower, and exercising cognitive control caused mental exhaustion and hindered later physical performance. Afternoons, the heart rate increases as the power component ratio(HRV) and low-frequency power component decreases.Evenings, when the mental fatigue is high heart rate decreases as the power component ratio(HRV) and low-frequency power component increase [7]. To conclude in general heart rate increases while doing any physical task whereas for mental activity heart rate may or may not always increase.

Heart Rate Variability(HRV) is a non-invasive method of assessing the state of autonomic nervous activity that measures the variation of time between heart rates [10]. When the power spectral of HRV is taken for the study. The results are analyzed by comparing the pre-task and post-task spectral density of HRV[11]. The detection of R waves is done by the algorithm based on wavelet transforms. Indices are based on autonomic nervous functions such as HRV to monitor changes in mental fatigue[11]. Therefore, Zhang, Chong, and Xiaolin Yu conclude that as the mental fatigue increases, the HRV decreases [11]. Mental fatigue will eventually result in varying cardiovascular functions which intern affect the heart rate and heart rate variability. Yue, Liu, et al in [7] elaborate on the Welch spectrum estimation, which can be used to obtain the spectrum diagram and high-frequency component power (HF), low-frequency component power (LF), and the ratio of high-frequency component power and low-frequency component power (LF/HF) of heart rate variability (HRV). The ratio (LF/HF) relates to the stability of the autonomic nervous system and depicts the balance between the other nervous systems which is responsible for balance in the brain(cerebellum) which in turn tells how mentally fatigue a person is. Kiryu, Motomiya, Ushiyama, and Okada et al, conduct an experiment to demonstrate physical or muscular fatigue during skiing[12]. Detailed HRV activities are analyzed using wavelet transform. The ratio (LF/HF) is addressed to analyze the Physical fatigue and the ratio decreases [12]. Therefore, HRV decreases with an increase in fatigue.

Electromyography(EMG) is a technique for assessing and recording the electrical activity of skeletal muscles. Causes for fatigue are analyzed using EMG amplitude characteristics [13]. When analyzing muscular fatigue, the high pass filter's corner frequency is 0.5 or 1 Hz. The surface detected EMG amplitude characteristics can increase even though Intracellular action potential(IAP) amplitude decreases significantly with fatigue. Due to the lengthening of the IAP profile and an increase in the negative after-potential [13]. EMG signals were collected from the biceps brachii and rectus femoris from four participants as mentioned in [14], and as fatigue set in, the spectral modification(Change in magnitude and phase of the EMG signals) shifts to a lower frequency.

The rate at which breathing occurs is referred to as the **Respiration rate**. The respiratory center sets and controls this, which is usually measured in breaths per minute. Breathing patterns change depending on the task. Under fatigued conditions, breath frequency almost remains stable, and respiration amplitude decreases [15]. The total variability of respiratory rate is unaffected by cognitive fatigue, whereas the correlated fraction decreases [16]. Therefore, the respiration rate increases with an increase in mental demand, which leads the fatigue level to increase [17].

The property of the human body that causes continuous variation in the electrical characteristics of the skin is known as **Electrodermal activity(EDA)**. The use of the EDA metric to detect muscle fatigue correlates with fatigue, providing complementary information from EMG signals. The **galvanic skin response** (GSR), which is a subset of electrodermal activity, or EDA. It is the change in sweat gland activityand it differs according to our emotions. The EDA is determined by observing variations in skin conductance (SC) at specific locations on the body where the concentration of eccrine glands is high[18]. The EDA was obtained on the nondominant hand's finger phalanxes. The index of sympathetic nervous system activity increases for a fatigued person. The mean and maximum amplitude of significant peaks of SMNA signals increases as fatigue increases. Signal patterns of GSR are analyzed using frequency analysis with a cutoff of 50 Hz [19]. It is seen that GSR has the peak power spectrum in the low-frequency band (0   0.08 Hz) during awake(not fatigued) and high-frequency band during drowsy(fatigued) [19].

Pulse oximetry measures the oxygen level (oxygen saturation) of the blood as well as changes in blood volume in the skin, resulting in a Photoplethysmograph.**Photoplethysmography (PPG)** is a simple and low-cost optical technique for detecting changes in blood volume in the microvascular bed of tissue. It is frequently used to take measurements at the skin's surface in a non-invasive manner. Fatigue is induced by a reduced supply of oxygen. The performance factor increases with an increase in blood oxygen level in the human body in normal not fatigued condition[20].PPG is also used to find the motion of the heart(calculated based on RR intervals[21]) as PPG uses lights to grasp the motion and the difference in light absorption rate in blood vessels [22] [23]. When the mental fatigue sets in heart rate found using PPG increases [21].

Electroencephalography(EEG) is an electrophysiological monitoring technique that records electrical activity on the scalp and has been shown to reflect the macroscopic activity of the brain's surface layer beneath. The electrodes are usually inserted along the scalp, making it non-invasive. The mental fatigue measured using EEG signals is determined by noticing the change in the wavelet packet feature of EEG and is extracted for every EEG data segment [11]. The results of wavelet packet entropy changes between pre-task and post-task are compared and used for analysis. As the mental fatigue onsets, the relative wavelet packet energy in the frequency band decreases, and wavelet packet entropy decreases [11].

2.2 Machine Learning Algorithms

2.2.1 Classical Machine Learning

Machine Learning is a subset of artificial intelligence and ML techniques that tries to mimic human behavior which is very complicated for a machine to understand human activity and ML algorithms try to learn from data sets or observations and try to predict the output for a new similar event. Learning with available training data set is known as supervised machine learning. Let us now see in this section the supervised machine learning algorithms that are used to classify fatigue.

Random Forest

Random Forest is a method that considers the output of multiple decisions trees to make the final prediction. The decision trees model tries to imitate the property of a tree(roots branching till node) where each decision and respective outcomes are recorded to make the best decision by checking the majority of answers from each node.

The approach by Maman, Chen, Bombard et al to explain the concept of physical fatigue detection using wearable sensors is commendable [24]. Two tasks were designed in the study, the first was the manual material handling task(MMH) and the second one is the pick-up and insertion(PI) task. The experiment lasted 6 hours (3 hours for each task), with the tasks being counterbalanced among the participants. There were 24 participants involved in this experiment, 9 females and 15 males. The mean age calculated for this experiment resulted to be 36.37 years. Five of the participants in this experiment were manufacturing workers, and the remaining were students. The condition for recruitment was the participants were needed to be in good mental and physical health, and they were screened by using a set of physical activity questionnaires. MMH task environment includes palletizing and carrying multiple weighted containers was part of the MMH task. Out of 24 participants, 9 participants were taken out during the data cleaning process, as they did not meet the requirement of the experiment. This task was for 3 hours without break and was supposed to be done with the same level of speed. Four IMUs at the ankle, hip, wrist, and torso, as well as a heart rate monitor on the chest, was used to collect data, and in-situ questionnaires were filled out every 10 minutes during the experiment. The supply pick-up and insertion task environment includes walking while carrying materials and then leaning forward to unscrew and fasten bolts at the supply box. Only 13 participants were included (instead of 15 in case 1) with reliable data after data cleaning. The sensor position is similar to that of the MMH task. The tasks were designed in such a way that one task was less tiring than the other. There are only two classification levels: fatigued and not fatigued. The data was classified as not fatigued for the first 18 minutes and fatigued for the last 10 minutes. The sensors and metrics used here were heart rate sensors, inertia measurement units (IMUS), electroencephalography (EEG), electromyography (EMG), and optical sensors. This study uses a window size of two minutes. The study extracted 42 statistical features, 10 bio-mechanical features, and 2 individual features.2 participants' data were removed due to corruption of data. The Random Forest model outperforms the other models in terms of accuracy (0.887) for this experiment.

The study looks into multilevel mental fatigue detection using the random forest[25]. The participants for this experiment were carefully selected. 12 healthy(no history of severe disease and medication) adults with an age range of 19 and 25 years were selected and also they had to have education for 11 to 15 years only. The participants who had sleep disorders were disregarded for the experiment, 8 hours sleep duration track record was required to participate in the experiment. The experiment conducted in this study was Auditory Vigilance Task (AVT), the participants were given random audio commands to follow with an interval of 500ms for each command set with a total of 50 sets. This command set gave 4 commands left, right, up, and down, the participants were required to concentrate and click these buttons within 1.5s and each session was for 3 minutes. EEG data were collected in this experiment and after each session AVT scores were calculated (percentage) and there was 5 fatigue level used for this task. The features were extracted with a window of 2 and an interval of 0.5s in-between. A total of 98160 samples were processed. Each decision tree was grown to the maximum of its depth and was weakened and these weakened trees were used to produce a strong random forest classifier. The test error rate was calculated to check the accuracy of the models, RF and recursive feature elimination scheme(RFE) together had a lower error rate and least features were obtained whereas using all RF, RFE, and heuristic initial feature ranking scheme(INIT) the error rate was maximum but obtained highest features[25],

Support Vector machine Support Vector Machine is a supervised machine learning algorithm and can be used for both classification and regression problems. SVM algorithm classifies using a hyper-plane that is used to classify multi-classes.

The study by Maman, Chen, Bombard et al mentioned in the above random forest section also estimate the result using SVM where the accuracy was 0.787 lesser than the boosting, bagging, and RF models[24].

The study by Zhang, Zheng et al are about detecting mental fatigue estimation using a support vector machine (SVM) [26].5 male right-handed participants were taken for this experiment. The relationship between mental fatigue and EEG descriptor was considered and the participants were asked to read with concentration the whole afternoon and were also asked to take notes for the readings. The EEG data were collected at 4 pm and 5 pm in the dark (closed eyes and lights off).FIR filtering was done for the signals obtained at 0.5 to 30 Hz. Two fatigue levels were considered for this experiment. SVM and Mahalanobis distance(MDBC) both classifiers were used to predict the fatigue levels SVM seems to give more accuracy around 91% was obtained using this classifier.

The real-time driver's fatigue is detected in the study by Savaş, Becerikli et al. [27]. The metrics considered here were blink rate, blink count, yawn detection, and head position detection . 10 participants 5 male and 5 female were taken for this experiment, and participants drove the training simulator. SVM model was trained for 80 percent of the data and tested on 20 percent of data with a cross-validation of 10. The SVM model was used to classify two classes: fatigue and not fatigue. The accuracy of the model was up to 97.93/

Gradient Boosted Decision tree

Boosting is a technique that combines several weak learning algorithms into a single larger algorithm. Predictions problems can be solved using the gradient boosted decision tree method. This method can be used to solve both classification and regression problems.

The physiological data were collected in the experiment conducted by Aryal, Ghahramani, Gerber et.al [28]. 12 construction workers were considered for the material handling experiment. Participants carried 15kg sandbags and walked for 10-meter for 200 trails.

The participant's helmets were fitted with infrared sensors, brainwave signal monitor, heart rate monitor, and Garmin vivofit, which collected physiological metrics like heart rate, skin temperature, and human thermoregulatory. Mental attentiveness was measured by a psychomotor vigilance test (PVT). The experiment tried to show the real-time physical fatigue detection in construction workers. The algorithms used to build the model are boosted trees and decision tree algorithms. The decision tree algorithm used 100 complex trees, 20 medium trees, and 4 simple trees. The boosted trees algorithm was employed, with medium trees serving as base learners and bagged trees serving as an ensemble of complex trees. Boosted tree algorithm performed well compared to another algorithm. 21 features were extracted from the sensors and 10-fold cross-validation was applied. Subjective rating of fatigue is collected for every 10 trials using Borg's Rating of Perceived Exertion (RPE). RPE scale range from Level 6-to 20. Level 6-11 is considered as low fatigue, level 12-14 as medium fatigue, level 15-16 as High fatigue, and level 17-20 as Very High fatigue. The accuracy is the average accuracy of 10 iterations, and the accuracy was based only on skin temperature was 9% greater accuracy compared to using only heart rate, but when both skin temperature and heart rate were taken into account the accuracy reached the maximum of 82%.

K-Nearest Neighbor(K-NN)

K-nearest-neighbor algorithm abbreviated as k-nn can also be used to classify both classification and regression problems. KNN uses a distance technique between the test data and all training data to predict the correct class for the test data. K number points closest to the test data are selected.

The [29] tries to develop an objective index for anxiety based on electroencephalogram (EEG) and photoplethysmogram (PPG) characteristics features obtained from wearable headsets and glasses. Anxiety is a psycho-physiological component related to internal tension (mental stress) and success in sport. There were two tasks given to 20 subjects which task1 was to ride at a comfortable speed, and task2 is to ride while competing. A questionnaire was asked before the two tasks to measure the anxiety level of the person. Many features were extracted from the EEG and PPG sensors, to be exact 23 EEG features and 6 PPG features. The findings of this paper revealed that the mean value and average strength of the wavelet alpha band coefficients and that of the beta band are strongly associated with the degree of anxiety. Partial auto-correlation of EEG features showed a moderate correlation with anxiety level. Both 6 PPG and 23 EEG features were used to find the classification accuracy after reducing the features to 3 levels of anxiety(low, moderate, and high) by using principal component analysis and k-nearest neighbors can achieve 62.5% accuracy across subjects. The number of nearest neighbor (K) was set to change from 1 to 10 and a three-fold cross-validation test was conducted, to reduce randomness the cross-validation was done for 100 rounds. There were only 5 data samples labeled as high anxiety, 3 were assigned for training and 2 for testing. The machine learning algorithms using PCA followed by k-nearest neighbors gave an accuracy of 62.5%.

Hidden Markov Model(HMM)

Based on probabilities The HMM is a probability model that is used to depict a stochastic process' statistical features [30]. Hu, Gong, Mu, Han, and Zhao et al. want to create a hidden Markov model to reduce driving accidents caused by drivers in a normal state. The driver's vehicle control strategy is referred to as the hidden state. The experiment takes place in a driving simulator with two straight lines and two round lines forming a road setting [30]. The algorithms used are the LBG (Linde–Buzo–Gray) VQ(Vector quantization) algorithm, Baum-Welch algorithm, and forward and backward algorithm. LBG VQ algorithm is used to quantify a large amount of data. The Baum-Welch algorithm is used to evaluate or estimate model parameters, as well as to choose and optimize HMM parameters. In driving state identification trials, signals collected in a driving simulator using forward and backward algorithms are used to test the model. The iteration directions are the only difference between forward and backward algorithms. The result analysis of [30] says that fatigue varied the distance to the center of the lane, and the steering angle. Using all the input data and running through all these algorithms and Hidden Markov Model, the driving state is identified as fatigue or not fatigue. The features of driving data were considered and after FFT the vector is given to HMM model. The log-likelihood value converged curve of driving behavior sequence with forward and backward algorithms is used to analyze the results. **Based on both probabilities and Physiological signals** Wearable sensors record the electroencephalogram (EEG), electromyogram (EMG), and respiration signals all at the same time [31]. The likelihood of fatigue is calculated using kernel distribution which estimates at various time intervals. The experiment conducted by [31] is a real-time road driving route for bus drivers.EEG signal was obtained by the brain skin surface, EMG signal from the nape of drivers' neck in differential input, and abdominal respiration signal is a belt or strap recording device which can be put on top of clothes. The experiment is analyzed using cases where only physiological signals are used and both signals, as well as probabilities of fatigue detection model, are used. This dynamic interaction between these interrelated nearby time slices is modeled using a first-order HMM model. This means that the current fatigue probability value is influenced by the previous value.

2.2.2 Neural Nets and Deep Learning

Deep Learning is a subset of machine learning that uses neural networks with many layers. Deep learning tries to mimic the neural structure of the human brain.

Multi Layer Neural nets

A multi-layer neural network is made up of more than one layer of an artificial neuron. These neural nets have one input and one output layer with multiple in-between hidden layers.

The discussion of driver fatigue is considered the main asset [32]. The authors use many classification methods and approach in this study. The approach of the study is on biological laboratory data. The nervous system is used as a reliable source of information. The relationship between these two components is represented as a biological sign by temperature, skin conductivity, and heart rate. Variation in nervous system components caused by sleep-wake activity affects Heart Rate Viability (HRV). The problem is overcome as this study presents a new HVR based on operator fatigue analysis. HRV is calculated by using ECG signals. The real-time data collected is used for fatigue analysis and classification. Fatigue is classified into two classes: low fatigue and high fatigue. They employ two supervised machine learning algorithms (Artificial Neural Network(ANN) and Support Vector Machine) that gave a high accuracy in predicting fatigue. The data were pre-processed and analyzed using a 30-sample window low pass filter. Each sample takes one minute of window size. ANN is built and trained using 80% of the dataset and 20% for testing, with a 6 node input layer, one hidden layer, and one output layer with a tangent-sigmoid transfer function. The accuracy of neural networks was found to be 88.3%.

Convolutional Neural Networks(CNN)

convolutional neural network (CNN), is a type of artificial neural network that is commonly used to analyze (process data) grid-like topology like visual images [33].

Driving is an operation that calls for great diligence. Insufficient concentration, poor vision, insufficient retrieval of knowledge, and sub-optimal anticipation are likely to cause low results in humans [34]. To improve traffic safety and driver well-being the study includes deep learning algorithms to detect arousal levels. Wrist wearable devices that have PPG sensors are used to calculate the heart rate. Physiological signals were collected by these wearable devices from 11 participants. Three classes are used for testing and training: under arousal, normal, and over-arousal. Here the author uses 7-layers convolutional neural network trained on raw physiological signals (that is, heart rate, skin temperature, and skin conductance) that outperform the neural network, and denoise auto-encoder models which are justified by comparing F-score and kappa value. The experiment uses also uses the sliding window method to extract segments of signals with a fixed 10-second phase scale. The 10, 30, 60, and 90-second windows (each containing samples of 100, 300, 600, and 900 respectively) were considered to find the optimum method. F-score and kappa value found are as follows: F-score: -0.82 vs. 0.75 and kappa: -0.64 vs. 0.53.

Recurrent neural networks(RNNs) with Long Short-Term Memory (LSTM)

Recurrent neural networks (RNN) are a type of neural network that acts as a close loop system, that is the output of the previous loop is fed as input to the current loop, same parameter, and same task steps are done in each trial for hidden and input layers so that the output should be the same.

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN). LSTM is a sequential network that remembers the data fed into the model and helps predict future outcomes.

The driver's fatigue is detected using facial information or metrics in [35]. The task environment for this study is that the participants are observed by fixing cameras and the main focus was to observe the driver's eyes the whole time as eyes give the relative features which can be used to detect fatigue driver's level. The video images were captured with 50 frames per second (fps). A real-time face detection algorithm was used to detect the driver's face in the video images and this algorithm is implemented by using SVM. RNNs with LSTM are used to detect fatigue and it consists of one input, three hidden layers, and one output layer. The LSTM model is trained using features from all of the facial parameters, such as eye parameters, yawning frequency, and other parameters obtained from vehicle steering and lane analysis. The parameters are then fed to the LSTM algorithm to train and test the model. The driver fatigue system gave an accuracy of around 97.20 percent. It is obtained with fewer epochs of around 55. This study concludes by saying that when compared to alternative feature-based techniques for classifying monthly data, RNNs with LSTM is more efficient.

Deep Long Short-Term Memory Autoencoders

Autoencoders are self-supervised learning model that is a subset of artificial neural network. These encoders are used to do efficient coding for the data which do not have labels. The autoencoders learn a type of encoding for the given data and can generally be used for dimensionality reduction. LSTM autoencoders are used in time-series sequence data, video, text, and audio types of data.

The kinematic analysis of time series data is discussed in the study [36]. IMU sensors are used to extract the data that can be used for fatigue detection in runners. A 20-minute cardiorespiratory fitness evaluation test is used to collect kinematic data from 14 runners. A sampling rate of 200 Hz is used, and IMU sensors were attached to their lower back and cervical region (lower neck). The communication protocols and data acquisition from the sensors in the wearable section of the measuring equipment were controlled by a 3-axis accelerometer (ADXL345), a 3-axis gyroscope (ITG3200), and a microprocessor (MSP430FR5969). The raw data used for the classification process were acceleration data, angular velocity data, and angular displacement data. A Second-order Butterworth filter is used for the filtering process with a cut-off frequency of 10 Hz. The algorithmic flow of this experiment includes three sub-parts i. preprocessing ii. Feature extraction iii. clustering analysis. i. In preprocessing step the data-set is split into training data and validation data and a windowing technique was used to subsequences from the time series. ii. LSTM neural network was used to extract features from the autoencoder for the feature extraction step. Spectral information was derived by applying Fast Fourier Transform (FFT) to the input sequences and then given to the LSTM model. iii. The neural network's unlabeled features were clustered using three clustering algorithms (KMeans, Agglomerative, and

DBSCAN).

2.2.3 Fatigue Detection Algorithm Discussion

This thesis intends to develop a fatigue detection algorithm that is included in a system that adapts its behavior based on fatigue levels (mental and physical fatigue). This section discusses how the reviewed algorithms do not meet the necessary criteria to achieve the goal of assessing mental and physical fatigue in real-time.

The most common physiological metrics used in experiments were EEG, ECG, EMG, and heart rate which were obtained by sensors that are more invasive in nature than the other sensors which are used to find other physiological metrics mention in 2.1. Including more classes increases the efficiency of the system.

The reviewed fatigue classification algorithms typically classify two levels of fatigue and very few algorithms classify fatigue in more than three classes (levels) like in [29] [34]. Classifying two classes is easier compared to multi-classes, but it has drawbacks of its own. It is important for machine learning or deep learning model to be able to classify more than two classes as it results in stronger and more efficient machine learning and deep learning models. It is important to classify fatigue in multi-level rather than just two-level because a person might be just about to be aroused the machine learning model might classify that the person is in a fatigued state whereas there might still be a lot of time left for the person to be fatigued. The person might be even more efficient at working for an extra hour before crossing a moderate fatigue level. Therefore, using more than two classes makes the model stronger and more efficient.

Many of the experiments mainly concentrate on drivers' fatigue which helps in assessing mental fatigue only, and other few experiments concentrate only on physical fatigue. There are no papers found that address both mental and physical fatigue from the same task. Therefore, this thesis intends to detect both physical and mental fatigue irrespective of the task. This study also considers the five-level fatigue classification method whereas other studies consider either two or three classes.

Chapter 3

Human Subject Experiment

3.1 Human Subjects

The goal of this study is to create a fatigue detection model that detects both mental and physical fatigue by capturing Physiological signals from wearable sensors. Two tasks are used to collect data corresponding to mental and physical fatigue: (i) Jigsaw puzzle-solving task, (ii) Pick and Place task. Wearable sensor data consisting of the Myo and Zephyr BioHarness is collected during each task. In-situ fatigue ratings are also collected during each task (see Table. 3.1). The data collected from the Bio Harness device are: heart rate, respiration rate, heartbeat interval, activity, posture, and breathing wave amplitude are noted down for further reference. EMG and IMU data are collected from the MYO armband device from which we get to know the direction of arm movement, muscle activity, muscle movement,8- channel EMG data, roll, pitch, and yaw.

3.2 Experiment Design

Baselines are collected at the beginning of the experiment, the participants are asked to sit still, and his/her constant physiological data are collected for 5 minutes. Participants then take a break before completing the physical or mental task, where the tasks are counter-balanced across the participants. The physical task considered is the Pick and Place task,

Table 3.1 Subjective metrics

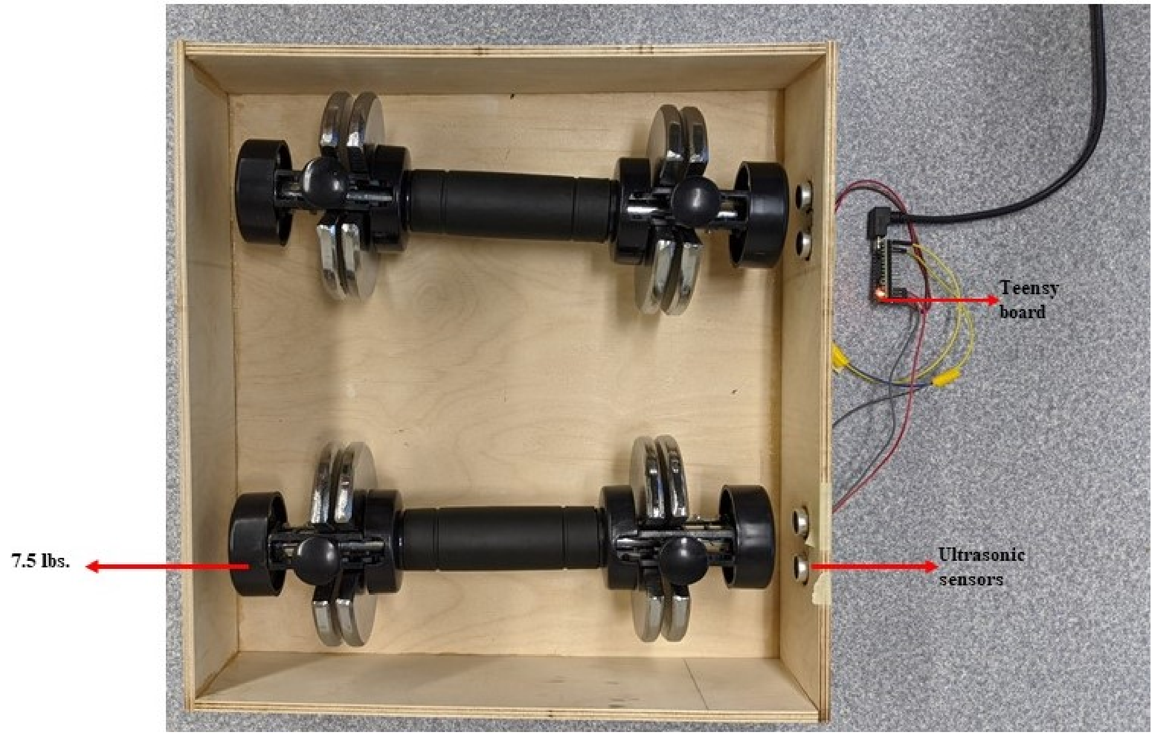
Fatigue Level		**Task Demand**		**Boredom Level**
Physical	Mental	Physical	Mental	
1- Very low	1- Very low	1- Very low	1- Very low	1- Not bored
2- Low	2- Low	2- Low	2- Low	2- Slightly bored
3- Average	3- Average	3- Average	3- Average	3- Moderately bored
4- High	4- High	4- High	4- High	4- Highly bored
5- Very High	5- Very High	5- Very High	5- Very High	5- Extremely bored

and the mental task considered is the Jigsaw puzzle-solving task. Each physical and mental task is conducted for an hour, and followed by a baseline collection and a 5-minute break. Participants complete in-situ ratings after each task round (defined in 3.2.1 and 3.2.2). Each task was designed such that each round consists of the same task demand and a round is approximately 1 minute. A consistent task demand level allows for fatigue effects to arise without being confounded by varying task demand levels.

3.2.1 Pick and Place task

Two adjustable dumbbells (set at 7.5 lbs each (see Fig. 3.1)) must be carried around a 10-meter u-shaped indoor track (see Fig. 3.2). The subject performs a curl with each arm at the beginning of the lap and walks to the end of the u-shape track and back. The weights are then placed on a standard height table. The subject then walks the track again without the weights, after which the lap is restarted with picking up the weights (see Fig. 3.2). One lap with weights and one lap without weights is considered to be one complete round, and at the end of each round, the time taken to walk with weights and without weights are recorded using timestamps from the ultrasonic sensors(teensy board) mounted on the box as shown in Fig. 3.1.

Figure 3.1 The dumbbell set-up for the physical task

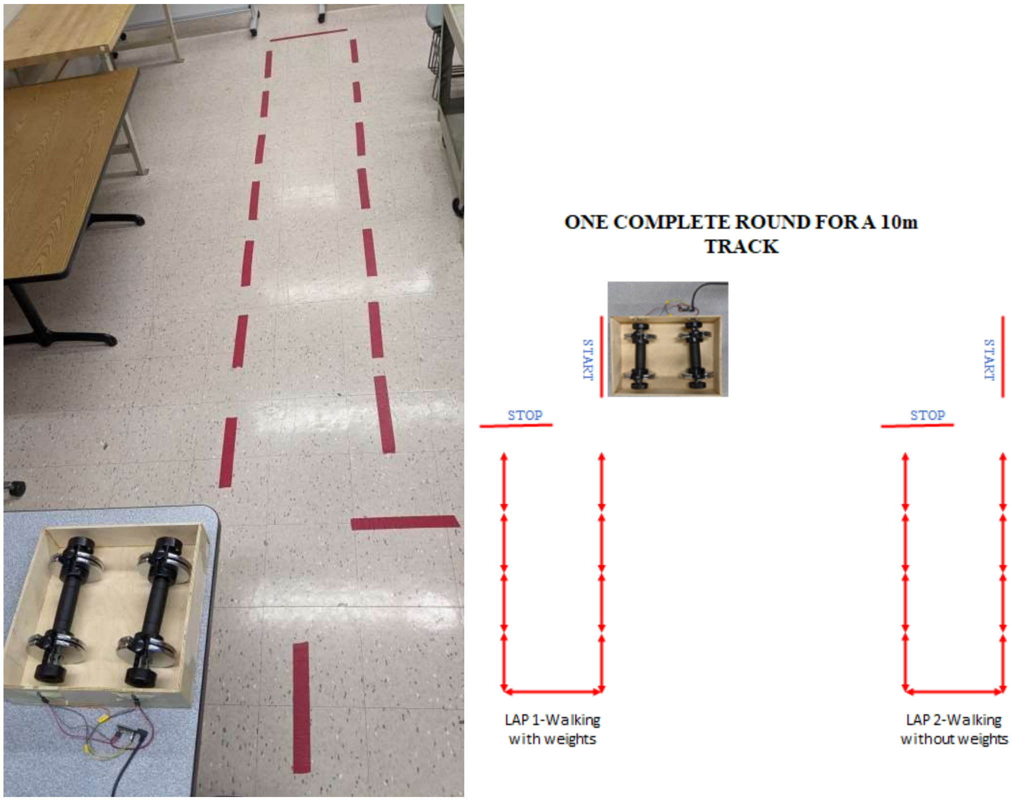


3.2.2 Puzzle-solving task

The subjects are given a jigsaw mind game puzzle (see Fig. 3.3) which can be designed with varying difficulty levels, but for this experiment, we use 16 pieces(4x4) of a jigsaw puzzle. The game is designed using the Unity game engine, where each movement of each puzzle piece is timestamped and is recorded as a correct and wrong move, game start, restart, game duration, and game exit. The subject completes the same puzzle repetitively for an hour, which elicits the desired mental fatigue levels.

Each round takes approximately 60 seconds(might vary according to subjects) to complete one puzzle, at the end of each puzzle completion, the subjects are asked a questionnaire about their fatigue level as described in Table. 3.1.

Figure 3.2 Physical task setup and description



3.2.3 Participant Information

The data is captured from 22 healthy subjects of Rochester Institute of Technology, 8 female and 4 male participants. The average age of the participants was 22.27 years. There were 11 undergrads participants, 9 Masters participants and 2 PhD participants considered for this study. The participants were not allowed in the experiment if they were not able to lift the weights, walk for an hour, pregnant or if they were sensitive to nickel and is explained in detail below in this section. The requirements for the participants to perform in this experiment are as follows: the participants need to be a non-pregnant adult (18-89 years) without a pacemaker, other metal, or any other permanent metal installations in the

Figure 3.3 Jigsaw puzzle (4x4)

chest area, or other types of implanted electronic devices. Participants with a pacemaker are unable to enroll, as there will be an incorrect response in the person's heart rate and heart-rate variability by the Bio-Harness Biopac device. Due to the physical nature of this experiment, participants are not eligible to participate in this experiment if they are pregnant, a person who cannot lift weights(15lb)/perform dumbbell curls(60 curls), and a person who cannot walk approximately 1.3 miles in an hour without any breaks. Walking and doing curls might induce slight physical strain. Participants are also expected to complete a puzzle task on the computer for an hour, which may induce eye strain. Participants may also feel stressed during the study, although this stress level is expected to be minor. These requirements will be checked after written consent has been obtained via a demographic's questionnaire and verbal questions. The experimenter will ask the participant if they have difficulty walking or seeing and explain the task environment to the participant so that the participant understands the physical and mental demands of the experiment. The experimenter will then ask the participant if they believe they can deal with the physical and mental demands and remind the participant that can withdraw from the experiment at any time. If the participant matches all the eligibility criteria only then the data will be collected else if the participant is ineligible or cannot deal with the task's physical and mental demands all existing data will be deleted and the participant will be withdrawn from the study.

Chapter 4

Methodology

4.1 Physiological metric Evaluation

This chapter tells how physiological metrics were preprocessed to get desired features and reduce the features performed by selected features containing relevant information rather than using the original features.Physiological metrics can be used in assisting the functionality of organ systems. After assessing they give the information on diseases and disabilities and in-turn helps in better medication for a person.

The physiological metrics are collected by the wearable sensors MYO armband and Zephyr Bio Harness 3 wireless chest strap. The data collected by both the MYO armbands are 8-EMG channels(EMG1-EMG8), and IMU directions(Roll, pitch, and yaw). The data collected by Zephyr Bio Harness 3 chest strap are heart rate, heart rate variability, respiration rate, posture, breathing wave amplitude, and ECG amplitude. Each physiological signal is preprocessed before inclusion in a machine-learning algorithm.

4.1.1 Data preprocessing

Data preprocessing is a technique that converts raw data into a usable format. Real-time data is frequently incomplete, inconsistent, and/or lacking in certain behaviors or trends, and it is rife with errors. Data preprocessing is a tried-and-true method for resolving such problems.

The raw EMG signals obtained from the MYO armband are sampled at the rate of 200 Hz. This signal is passed through butter worth band-stop filter with a stop-band frequency of 60 Hz. The second-order section digital filter is applied to the output to get the data in the vector form, then the absolute value of the Fast Fourier Transform of filtered data is used for further feature extraction process. Fourier transform is converting continuous-time (space) non-periodic to continuous frequency non-periodic. When only one period is considered, the decomposition is also known as the Discrete Fourier Transform (DFT). Discrete Fourier Transform (DFT), transforms a non-periodic discrete signal of finite length N,x[n], to a set of N,non-periodic transform coefficients, X[k]. Example uses are in digital signal processing using sampled values – fast implementations of DFT are called Fast Fourier Transform or FFT.

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j(\frac{2\pi}{N})kn}. \tag{4.1}$$

The segmented physiological data is preprocessed in order to fill in missing information and reduce noise. The raw EMG signals obtained from the Myo armband are sampled at the rate of 200 Hz. These signal are passed through a notch-filter at 60 Hz in order to reduce power-line noise. No filter is applied to the BioHarness data, since this data is preprocessed onboard the device. Any missing samples from the Myo or BioHarness are imputed with linear regression techniques before features are extracted.

The filtered and transformed data is then parsed for each task round (puzzle and pick and place in the mental and physical task, respectively). This parsed data is then used for necessary feature extraction. The data collected from bio-harness need not be filtered, but the data is parsed using the techniques mentioned above.

4.1.2 Feature extraction

The feature extraction or dimensionality reduction is the next step used to reduce large raw data set into meaning full data set.

The parsed data of heart rate, heart-rate variability, respiration rate, and posture from bio-harness are now considered for feature extraction. The features considered for this model are mean, standard deviation, variance, gradient, and slope. Each metric is associated with these five features and a total of 20 features from bio-harness are obtained for physical fatigue analysis and 20 features for mental fatigue analysis. The mean is the arithmetic average of a set of given numbers.

$$Arithmeticmean = \frac{(x_1 + x_2 + ... + x_n)}{n} \tag{4.2}$$

The variance is the average of the squared differences from the mean.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \tag{4.3}$$

The term “gradient” refers to a graded difference in physiological activity along an axis. A line’s slope is the ratio of how much y increases as x increases by some amount. The slope of a line tells you how steep it is, or how much y increases as x increases. The slope is constant (the same) anywhere on the line.

$$m = \frac{y_2 - y_1}{x_2 - x_1} \tag{4.4}$$

The parsed data of 8 EMG channels from the MYO armband are considered for feature extraction. The median frequency of each EMG channel is considered a feature. Median Frequency (MNF) is an average frequency calculated using the power spectrum obtained from EMG data. The median frequency f_m of a power spectrum $P(f)$ is defined as the frequency satisfying the following equation [37]:

$$\int_0^{f_m} P(f)df = \int_{f_m}^{\infty} P(f)df = \frac{1}{2}\int_0^{\infty} P(f)df \tag{4.5}$$

Therefore, a total of 36 features for each physical and mental fatigue analysis are used and are then given to feature reduction to resolve the over-fitting issue by using all the features for building a model and this process is explained in the next section.

4.1.3 Feature Reduction

Reducing the feature means reducing the number of variables without losing important information from these variables to lower the computational cost as well as to make the computer work faster. . The chosen features are known to correlate with physical and/or mental fatigue, where a total of 36 features are extracted. This number is reduced using Principal component analysis (PCA), in order to develop less complex models and help prevent overfitting.

PCA reduces dimensionality by finding projections that maximize variance (information) along an axis. This is done by computing eigen vectors and values of a dataset. Larger eigen values represent more information in the corresponding principle component; thus, the proportion of variance can be computed as a ranked ratio: $PoV = \frac{\lambda_1+\lambda_2+...+\lambda_i}{\lambda_1+\lambda_2+...+\lambda_N}$, where λ_i is the ith largest eigen value. This work found the PoV corresponding to 0.9 or explains 90% of the variance. The corresponding eigen vectors are then used to project the

42 features into a smaller dimensional space. We can also use the sklearn decomposition for PCA analysis, n-component is found by scaling and transforming the dataset values, and then the explained variance ratio is calculated and is ranked in order. The sum of the variance of features that holds 90% of the data is considered as the features which contain most information and other features are ignored. This PCA algorithm was used to reduce features 36 features down to 17 for the physical fatigue model and 19 for the mental fatigue model.

4.2 Fatigue classification Algorithms

The collected, filtered and cleaned data are used to build the physical fatigue and mental fatigue classification models. These models are built using two classical machine learning algorithms and one deep learning algorithm. The machine learning algorithms considered for this study are Random Forest(RF), Support Vector Machine (SVM), and a Long Short Term Memory(LSTM). This section reads the details of the algorithms and the specifications used for our experiment.

4.3 Classical Machine Learning Algorithms

4.3.1 Random Forest

Random forest is a classification algorithm that is used to classify problems using decision trees and the decisions are made by the majority decisions made by these trees.

The number of tree nodes and the number of decision trees are very important for using a random forest algorithm and are represented as n_estimator and max_depth. The number of trees is represented by n_estimators. Generally, the more trees there are, the better the data can be learned. However, adding a large number of trees can significantly slow down the training process, so we perform a parameter search to find the optimal tree length. The max_depth variable represents the maximum depth of each tree in the forest. The single tree depth is found by using an algorithm, first, a function for fitting trees of various depths on the training data using cross-validation(CV) of five, and tree depth is created from which we obtain CV mean, standard deviation, and accuracy scores, next we fit the trees of depth 1 to 24, finally the depth of the tree that achieves the best mean cross-validation accuracy on training data set is selected to be the optimal depth value for the algorithm.

- Mental Fatigue Model: n_estimator of 500 and a tree depth of 5 is considered to be optimal to create this model using a random forest algorithm.
- Physical Fatigue Model: n_estimator of 500 and a tree depth of 2 is considered to be optimal to create this model using a random forest algorithm.

4.3.2 Support Vector Machine

Support Vector Machines find a hyperplane that linearly separates the dataset with some slack (cost). This hyperplane is determined by support vectors that maximize the margin between the hyperplane and the edge of the examples. Similar to the random forest, a grid search was performed by changing the cost values, gamma (influence of a sample based on distance), and kernel (radial basis function or polynomial).

Selecting kernel type is as important as creating an SVM model. Kernels are used to transform the input into a readable output. There are seven popular SVM kernels that are Linear Kernel, Polynomial Kernel, Gaussian Radial Basis Function (RBF), Sigmoid Kernel, Gaussian Kernel, Bessel function kernel, ANOVA kernel. The data-set for this study is not linearly separable, therefore we consider using radial basis function (RBF) kernel and polynomial (poly) kernel. The parameters for these kernels are C value, gamma value, degree, and probability. RBF is the default and most widely used kernel as it resembles the Gaussian distribution.

The best fit parameter values are found by running an algorithm with a set of C values and gamma values and PCA of 18, best parameters are C= 10, gamma =0.07 with true probability for kernel= RBF. The best parameters for kernel = poly are C=10, gamma = 0.05 with true probability.

4.4 Deep Learning

4.4.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN). LSTM is a sequential network that remembers the data fed into the model and helps predict future outcomes. The applications in which LSTMs are used are for weather forecasts or in predicting or auto typing words to complete sentences in emails and many more. LSTMs perform somewhat better when it comes to memorizing specific patterns. LSTM, like any other NN, can have multiple hidden layers, and as it passes through each layer, the relevant information is retained while the irrelevant information is discarded in every cell. Therefore, we use LSTM in this study to predict the fatigue level of an individual. LSTMs have both long-term and short-term memory so to deal with this gates are used to help the memorizing process. The four gates are forgotten, learn, remember and use gates. The forget gate forgets the unnecessary information and new information are learned from short term memory and is given as input, this information is fed to long term memory so that the data keeps adding to the current information, and finally use gate is used to forecast the outcome of the current event using all the information from the memory and the output is again saved to the memory.

The x_train and x_test data set are scaled and transformed, and the y_train and y_test are hard encoded for 5 classes. The physical and mental fatigue model is built sequentially with one LSTM layer, two hidden layers, and one output layer classifying 5 classes. The LSTM layer consists of fifty neurons, a tanh activation function with a dropout value of 0.2, and a sigmoid recurrent activation function with a recurrent dropout value of 0.2. There are two hidden layers both with sixty-four neurons and a sigmoid activation function. The output layer of the model has 5 neurons where each neuron represents a fatigue level (1-5) and is given to an activation function of softmax which classifies the class according to the maximum probability obtained from each neuron. The LSTM model is trained with a loss function of categorical cross-entropy and the ADAM optimizer with an initial learning rate of 0.001. This model is then fit with x_train and y_train values with an epoch of 70, validation split of 0.1, and a batch size of 32.

4.5 Research questions and hypothesis

This section talks about what to expect from this study. The research questions and the hypothesis are mentioned below.

Research Question 1: How do physiological signals differ between mental and physical fatigue for each mental and physical task?

Hypothesis 1: Significant difference will exist between the physiological data and the fatigue levels. The Physiological signals such as heart rate, heart rate variability, respiration rate, posture, and EMG middle frequency are considered for both mental tasks and physical tasks. The heart rate is expected to increase for both tasks, whereas the heart rate variability decreases for both tasks. The respiration rate is expected to increase for the physical task and decrease for the mental task. These findings are verified with the actual results in the next chapter.

Hypothesis 2: All physiological data will correlate with fatigue levels.

Research Question 2: How effectively classical machine Learning models should classify five fatigue levels?

Hypothesis 1: The accuracy of the physical task model and mental task model with random forest classifier and SVM classifier using the Leave one out (LOO) method is expected to be above 80%.

Hypothesis 2: The Confusion for level 4 and 5 is expected to be less as there are very little data and for level 1,2, and 3 the confusion is expected to be higher as there is huge data for each of these fatigue levels.

Research Question 3: How effectively classical Deep Learning models should classify five fatigue levels?

Hypothesis 1: The accuracy of the physical task model and mental task model with the LSTM classifier using the Leave one out (LOO) method is expected to be above 75%.

Hypothesis 2: The Confusion for level 4 and 5 is expected to be less as there are very little data and for level 1,2, and 3 the confusion is expected to be higher as there is huge data for each of these fatigue levels.

4.6 Summary

The algorithms used in this study: are Random Forest, Support Vector Machine, and LSTM. These algorithms are discussed individually in this chapter with their respective parameter specification which helps in building an optimal model for this study. The next chapter talks about the validation methods and the validation accuracy of each fatigue model and discusses the experimental validation outputs.

Chapter 5

Results and Discussions

5.1 Experimental Validation

The experiment validation gives better knowledge of the correlation between features, and differences in the dependent variable for two independent groups. This study uses the Mann-Whitney U test to check the correlation between features between two independent groups or classes. There are five groups from 1-5 and the correlation between one dependent variable from one group is compared with the same dependent variable from the other group. The correlation and validation between the same dependent variables themselves are calculated using the pandas Pearson correlation method.

Pearson Correlation: Pearson's correlation coefficient is calculated by dividing the covariance of the two variables by the product of the standard deviations of each data sample. This can also be said as a measure of linear correlation between two sets of data, it is the normalization of the co-variance between the two variables to yield an interpretable score, and it is represented in the equation below:

$$\gamma = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{5.1}$$

Mann-Whitney U test: The Mann–Whitney U test is a non-parametric test of the null hypothesis that the likelihood of X being higher than Y for randomly picked values X and Y from two populations is equal to the chance of Y being greater than X. It compares whether the dependent variable's distribution is the same for the two groups, implying that they are from the same population.

5.1.1 Heart-Rate

The mean heart rate and the standard deviation are calculated for each level of mental and physical fatigue. The mean heart rate and the standard deviation of the heart rate of the physical fatigue task are more compared to that of the mental fatigue task. The maximum heart rate value is found to be at level 3 (88.12) for mental fatigue and level 2 (110.16) for physical fatigue. A detailed overview of mean and standard deviation by Fatigue Type and Subjective Level is provided in Table. 5.1.

Table 5.1 Heart-Rate Descriptive Statistics by Fatigue Type and Subjective Level

Fatigue Type	Fatigue Level	Descriptive Statistics (Mean(StDev).)	Count
Mental	1	77.57(4.64).	112
	2	85.90 (3.47).	190
	3	88.18 (4.23).	251
	4	79.80 (4.95).	129
	5	81.11 (4.05).	44
Physical	1	102.15 (4.80).	256
	2	110.16 (5.22).	222
	3	105.81 (5.24).	111
	4	105.55 (3.53).	48
	5	105.18 (4.19).	32

The Mann-Whitney results for heart rate with pairwise comparisons for mental fatigue level and physical fatigue level are provided in Table 5.2 and 5.3 respectively. The table gives an overview of the p-values of the Mann-Whitney U-test analysis and if the p-value is below the agreed risk of 5 percent (0.05), the null hypothesis can be rejected else if the p-value is greater than 0.05 we the null hypothesis is not rejected. For the (4 vs 5) pair and

(2 vs 3) pair the null hypothesis is not rejected as the p-value is more than 0.05 for Mental Fatigue. For the (4 vs 5),(1 vs 3),(1 vs 4),(2 vs 4), and (2 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Physical Fatigue. Therefore, we cannot conclude that a significant difference exists between these groups.

Table 5.2 Heart-Rate Mann-Whitney Pairwise Comparisons for Mental Fatigue. **Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	**				
3	**	0.167			
4	**	*	**		
5	**	0.136	*	0.673	

Table 5.3 Heart-Rate Mann-Whitney Pairwise Comparisons for Physical Fatigue. **Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	*				
3	0.090	**			
4	0.079	0.816	*		
5	*	0.517	*	0.105	

Heart rate was negatively correlated with trial duration (r=-0.2308) and positively correlated with mental fatigue (r=0.0377). Heart rate was negatively correlated both with trial duration (r=-0.179) and with physical fatigue (r=-0.00777).

5.1.2 Heart-Rate Variability

The mean and standard deviation of heart rate variability is calculated for each level of mental and physical fatigue. Mean heart rate variability and standard deviation of heart rate variability of mental fatigue type is more compared to that of physical fatigue type, as heart rate variability is inversely proportional to that of heart rate, therefore, as the heart-rate increases heart rate variability decreases. The maximum heart rate variability value is found to be at level 1 (60.847) for mental fatigue and level 1 (36.208) for physical fatigue. A detailed overview of mean and standard deviation by Fatigue Type and Subjective Level is provided in Table 5.4.

Table 5.4 Heart-Rate Variability Descriptive Statistics by Fatigue Type and Subjective Level

Fatigue Type	Fatigue Level	Descriptive Statistics (Mean(StDev).)	Count
Mental	1	60.84(2.12).	112
	2	42.52 (1.40).	190
	3	46.10 (1.18).	251
	4	57.04 (1.35).	129
	5	53.88 (1.12).	44
Physical	1	36.20 (1.23).	256
	2	28.65 (1.18).	222
	3	33.30 (1.51).	111
	4	24.34 (0.83).	48
	5	27.65 (1.28).	32

The Mann-Whitney results for heart rate variability with pairwise comparisons for mental fatigue level and physical fatigue level are provided in Tables 5.5 and 5.6, respectively. For the (1 vs 4) pair and (4 vs 5) pair the null hypothesis is not rejected as the p-value is more than 0.05 for Mental Fatigue. For the (4 vs 5),(1 vs 5),(1 vs 4),(2 vs 3), and (2 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Physical Fatigue. Therefore, we cannot conclude that a significant difference exists between these groups.

Table 5.5 Heart-Rate Variability Mann-Whitney Pairwise Comparisons for Mental Fatigue.**Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	**				
3	**	*			
4	0.136	**	**		
5	*	**	**	0.336	

Heart-rate variability was positively correlated with both trial duration (r=0.1922) and with mental fatigue (r=0.0283). Heart-Rate Variability was positively correlated with trial

Table 5.6 Heart-Rate Variability Mann-Whitney Pairwise Comparisons for Physical Fatigue.**Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	*				
3	**	0.108			
4	0.964	**	**		
5	0.248	0.830	*	0.097	

duration (r=0.1574) and negatively correlated with physical fatigue (r=-0.0901).

5.1.3 Respiration-Rate

The mean and standard deviation of Respiration-Rate is calculated for each level of mental and physical fatigue. The mean Respiration-Rate and standard deviation of Respiration-Rate of physical fatigue type is more compared to that of mental fatigue type. Heart rate and respiration rate are directly proportional, respiration rate increases as heart rate increases. The maximum Respiration-Rate value is found to be at level 1 (26.875) for physical fatigue and level 5 (16.762) for mental fatigue. A detailed overview of mean and standard deviation by Fatigue Type and Subjective Level is provided in Table 5.7.

Table 5.7 Respiration-Rate Descriptive Statistics by Fatigue Type and Subjective Level

Fatigue Type	**Fatigue Level**	**Descriptive Statistics (Mean(StDev).)**	**Count**
Mental	1	15.19(1.28).	112
	2	16.23 (0.83).	190
	3	16.40 (0.92).	251
	4	15.60 (0.93).	129
	5	16.76 (0.90).	44
Physical	1	26.87 (2.51).	256
	2	21.75 (1.98).	222
	3	24.04 (2.60).	111
	4	25.87 (3.042).	48
	5	25.53 (2.77).	32

The Mann-Whitney results for respiration rate with pairwise comparisons for mental fatigue level and physical fatigue level are provided in Tables 5.8 and 5.9, respectively. For the (1 vs 4),(1 vs 5),(2 vs 3),(2 vs 5),(3 vs 5), and (4 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Mental Fatigue. For the (4 vs 5),(1 vs 5), and (3 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Physical Fatigue. Therefore, we cannot conclude that a significant difference exists between these groups.

Table 5.8 Respiration-Rate Mann-Whitney Pairwise Comparisons for Mental Fatigue.**Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	*				
3	*	0.768			
4	0.849	*	*		
5	0.255	0.797	0.836	0.359	

Table 5.9 Respiration-Rate Mann-Whitney Pairwise Comparisons for Physical Fatigue. **Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	**				
3	**	**			
4	*	**	*		
5	0.071	**	0.165	0.410	

Respiration rate is negatively correlated with trial duration (r=-0.1654) and positively correlated with mental fatigue (r=0.05789). Respiration rate was positively correlated with trial duration (r=0.20563) and negatively correlated with physical fatigue (r=-0.0670).

5.1.4 Posture Magnitude

The mean and standard deviation of Posture is calculated for each level of mental and physical fatigue. Mean Posture and standard deviation of Posture Magnitude of physical fatigue type is more compared to that of mental fatigue type. There are more postures in the physical type compared to that of the mental type. The maximum Posture value is found to be at level 2 (-3.416763) for physical fatigue and level 2 (-4.081288) for mental fatigue. A detailed overview of mean and standard deviation by Fatigue Type and Subjective Level is provided in Table 5.10.

Table 5.10 Posture Descriptive Statistics by Fatigue Type and Subjective Level

Fatigue Type	Fatigue Level	Descriptive Statistics (Mean(StDev).)	Count
Mental	1	-10.00(1.34).	112
	2	-4.08 (0.77).	190
	3	-6.81 (1.00).	251
	4	-19.99 (2.60).	129
	5	-27.80 (0.71).	44
Physical	1	-10.57 (21.69).	256
	2	-3.41 (18.02).	222
	3	-9.14 (14.63).	111
	4	-9.60 (15.59).	48
	5	-12.30 (12.13).	32

The Mann-Whitney results for Posture Magnitude with pairwise comparisons for mental fatigue level and physical fatigue level are provided in Tables 5.11 and 5.12, respectively. For the (1 vs 3),(2 vs 3), and (4 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Mental Fatigue. For the (4 vs 5),(1 vs 5),(2 vs 4),(2 vs 5),(3 vs 4), and (3 vs 5) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Physical Fatigue. Therefore, we cannot conclude that a significant difference exists between these groups.

Table 5.11 Posture Mann-Whitney Pairwise Comparisons for Mental Fatigue. **Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	*				
3	0.270	0.133			
4	**	**	**		
5	**	**	**	0.131	

Table 5.12 Posture Mann-Whitney Pairwise Comparisons for Physical Fatigue. **Note:** * indicates the difference between the groups, * p<0.05, ** p<0.001

Level	1	2	3	4	5
1					
2	**				
3	*	**			
4	*	0.133	0.253		
5	0.794	0.069	0.646	0.410	

Posture was negatively correlated both with trial duration (r=-0.115329) and with mental fatigue (r=-0.255575). Posture was negatively correlated both with trial duration (r=-0.155628) and with physical fatigue (r=-0.050882).

5.1.5 EMG Average Middle Frequency across All Channels

The mean and standard deviation of EMG Average Middle Frequency is calculated for each level of mental and physical fatigue. Mean EMG Average Middle Frequency and standard deviation of EMG Average Middle Frequency of mental fatigue type is more compared to that of physical fatigue type. The maximum EMG Average Middle-Frequency value is found to be at level 2 (50.615146) for physical fatigue and level 1 (50.126116) for mental fatigue. A detailed overview of mean and standard deviation by Fatigue Type and Subjective Level is provided in Table 5.13.

Table 5.13 EMG Middle-Frequency Descriptive Statistics by Fatigue Type and Subjective Level

Fatigue Type	**Fatigue Level**	**Descriptive Statistics (Mean(StDev).)**	**Count**
Mental	1	50.12(2.75).	112
	2	49.44 (3.05).	190
	3	50.03 (3.05).	251
	4	50.00 (2.77).	129
	5	49.18 (3.63).	44
Physical	1	50.14 (1.88).	256
	2	50.61 (2.22).	222
	3	49.60 (2.55).	111
	4	48.71 (2.38).	48
	5	48.32 (2.21).	32

The Mann-Whitney results for EMG Average Middle Frequency with pairwise comparisons for mental fatigue level and physical fatigue level are provided in Tables 5.14 and 5.15, respectively. For the (1 vs 3),(1 vs 4),(2 vs 3),(2 vs 5), and (3 vs 4) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Mental Fatigue. For the (4 vs 5),(1 vs 2), and (2 vs 3) pairs the null hypothesis is not rejected as the p-value is more than 0.05 for Physical Fatigue. Therefore, we cannot conclude that a significant difference exists between these groups.

Table 5.14 EMG Middle-Frequency Mann-Whitney Pairwise Comparisons for Mental Fatigue. **Note:** * indicates the difference between the groups, * $p<0.05$, ** $p<0.001$

Level	1	2	3	4	5
1					
2	**				
3	0.073	0.084			
4	0.358	*	0.424		
5	**	0.300	**	**	

Table 5.15 EMG Middle-Frequency Mann-Whitney Pairwise Comparisons for Physical Fatigue. **Note:** * indicates significant difference, * $p<0.05$, ** $p<0.001$

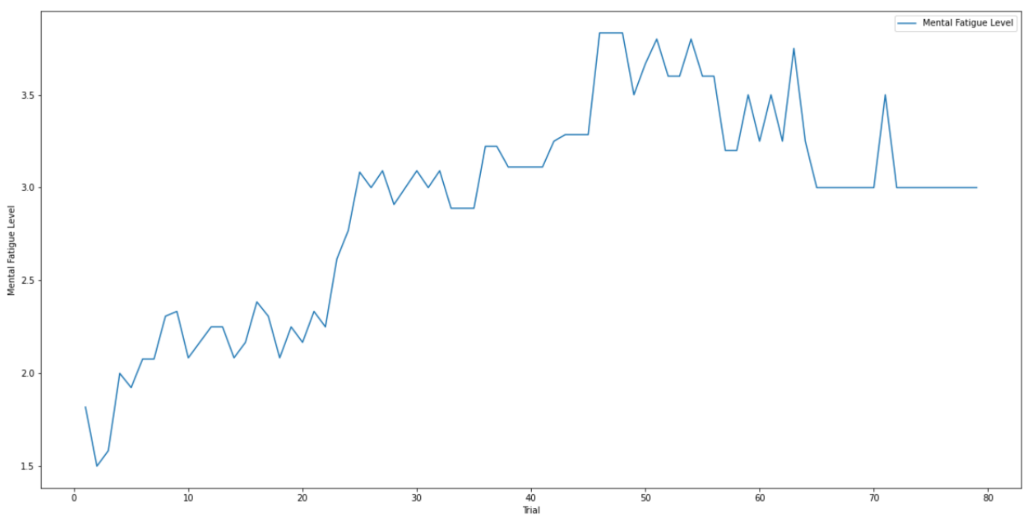
Level	1	2	3	4	5
1					
2	0.066				
3	**	0.090			
4	**	**	*		
5	**	**	*	0.286	

Average EMG Middle-Frequency was positively correlated with trial duration (r=0.5722) and negatively correlated with mental fatigue (r=-0.205077). Average EMG Middle-Frequency was positively correlated with trial duration (r=0.153649) and negatively correlated with physical fatigue (r=-0.208064).

5.1.6 Subjective Ratings

The Fig 5.1.& 5.2. presents the plot of Mental Fatigue level and Physical fatigue level by the number of trials, respectively. The mental plot shows that mental fatigue level increases with an increase in the number of trials but decreases around 65_th trial as there are very little data for level 5. The physical plot shows that physical fatigue level increases with an increase in the number of trials but decreases drastically around 55_th trial as there are very little data for level 4 and 5.

Figure 5.1 Plot of average Mental Fatigue level by trial number



The Fig 5.3.& 5.4. presents the plot of trial duration of Mental Fatigue and Physical fatigue by the number of trials, respectively. The mental plot shows that trial duration decreases linearly with an increase in the number of trials. The physical plot shows that trial duration decreases with an increase in the number of trials but the decrease in trial duration is so linear than the mental plot.

Figure 5.2 Plot of average Physical Fatigue level by trial number

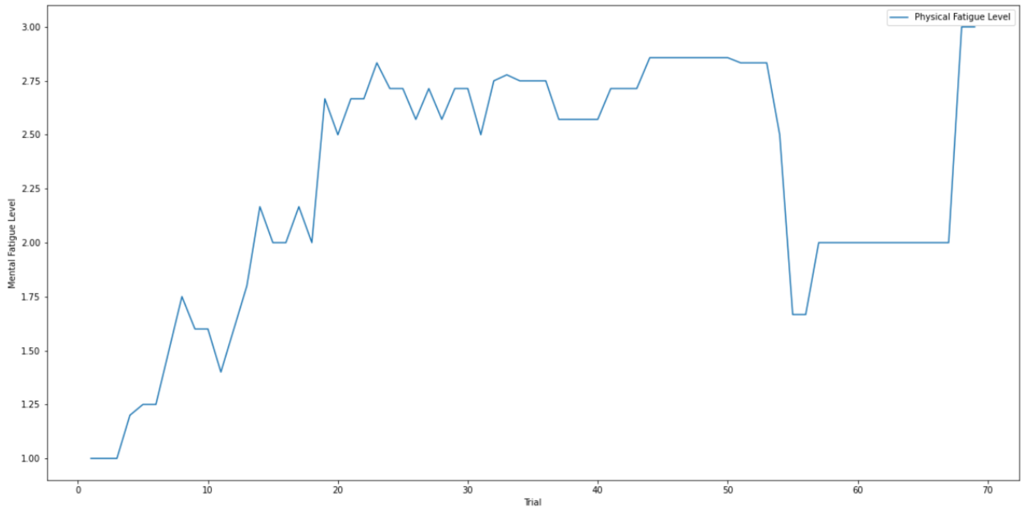


Figure 5.3 Plot of trial duration by trial number for average Mental Fatigue

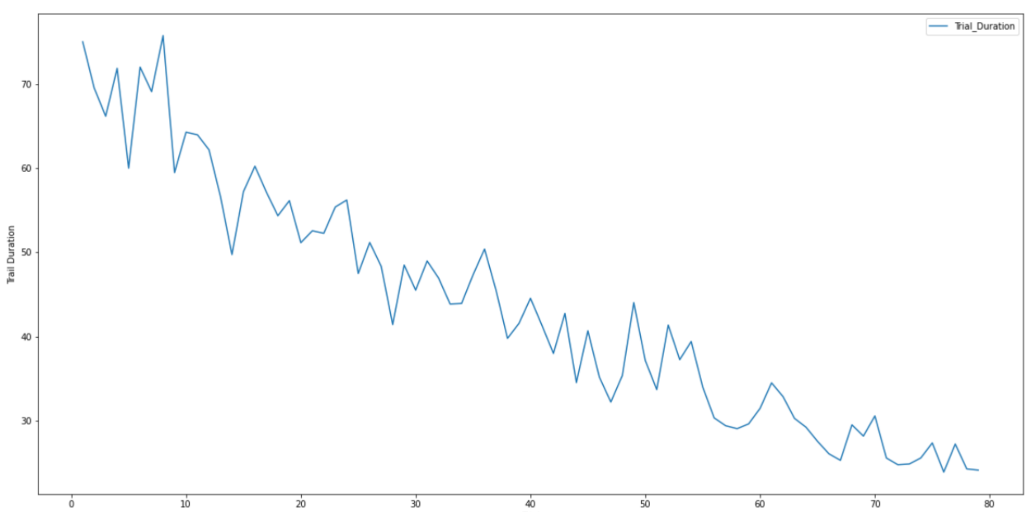
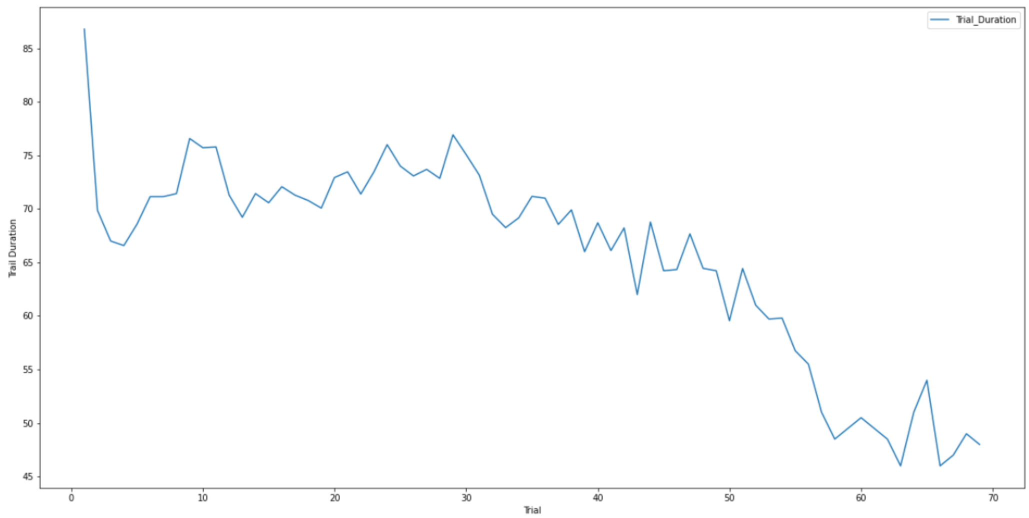


Figure 5.4 Plot of trial duration by trial number for average Physical Fatigue



The correlation of average mental fatigue with trial duration is negatively correlated (r=-0.759923). The correlation of average physical fatigue with trial duration is negatively correlated (r=-0.16206).

5.2 Classical Machine Learning

The random forest and support vector machine models are evaluated using the Leave One Subject Out cross-validation(LOSOCV) method.

LOOCV: The Leave-One-Out Cross-Validation, or LOOCV, the procedure is used to estimate the performance of machine learning algorithms when they are used to make predictions on data not used to train the model. To see how well the predictions are made we use leave one out cross validation methodology.Generally, we can evaluate the performance of a model also by using mean squared error (MSE). Leave one subject out(LOSO/LOOCV) CV trains a model on all but one participant and validates on the remaining participant.

This approach is repeated such that each participant is tested on once. The LOSO CV results provide insight into a model's population generalizability (how well it performs on an previously unseen human).

Table 5.16 Random Forest and Support Vector Machine Classification Accuracy by Participant and Mental Fatigue

Participant	**Mental Fatigue**	
	RF	SVM
P1	0.33	0.29
P2	0.34	0.22
P3	0.08	0.08
P4	0.13	0.23
P6	0.38	0.39
P7	0.16	0.16
P8	0.42	0.5
P9	0.4	0.3
P10	0.33	0.5
P11	0.20	0.27
P12	0.18	0.13
P13	0.23	0.36
P14	0.04	0.04
P17	0.11	0.11
P18	0.19	0.22
P19	0.22	0.19
P20	0.26	0.13
P21	0.05	0.09
P22	0.08	0.08
Avg.	0.23	0.24

Table 5.17 Random Forest and Support Vector Machine Classification Accuracy by Participant and Physical Fatigue

Participant	**Physical Fatigue**	
	RF	SVM
P2	0.58	0.78
P3	0.35	0.38
P6	0.20	0.35
P7	0.04	0.0
P8	0.54	0.54
P10	0.0	0.0
P12	0.09	0.05
P13	0.14	0.17
P15	0.0	0.0
P17	0.05	0.0
P19	0.02	0.23
P20	0.25	0.28
Avg.	0.25	0.27

The confusion matrix is an accurate count of actual value to predicted value and is represented in a matrix form. We have 5 classes in our study when the actual value of class 1 is equal to the predicted value of class 1 it is known as true positive(TP). When the actual value of class 2 is equal to the predicted value of class 2 it is known as true negative(TN) and goes on till class 3,4,5 that is TN value for a class will be the sum of values of corresponding rows except for the TP value. When class 1 is predicted as class 2 it is known as a false negative(FN), for a multi-class, the FN value of a class will be the total of the values of all columns and rows except those of the class for which the values are being calculated, and when class 2 is predicted as class 1 it is known false positive(FP), for a multi-class the FP value of a class will be the sum of values of the corresponding column except for the TP value.

The RF-confusion matrix of physical fatigue as provided in figure 5.7. has class 1 and class 2 with maximum true positive values compared to that of the remaining 3 classes. The RF-confusion matrix of mental fatigue as provided in figure 5.5. has class 3 and class 2 with maximum true positive values. The SVM-confusion matrix of physical fatigue as provided in figure 5.8. has class 1 and class 2 with maximum true positive(30,24) values compared to that of the remaining 3 classes. The SVM-confusion matrix of mental fatigue as provided in figure 5.6 has class 3 and class 2 with maximum true positive values.

The f1-score, precision and recall of random forest, and SVM for physical and mental classification model are represented in Table 5.19 and 5.18 respectively. A recall is defined as the ratio of the total number of correctly classified true positive classes divide by the total number of positive classes(TP+FN). The RF- recall shows that class 3(0.65) has been classified correctly compared to the other 4 classes for mental fatigue classification. The RF- recall shows that class 1 has been classified correctly compared to the other 4 classes for physical fatigue classification. The SVM- recall shows that class 2 has been classified correctly compared to the other 4 classes for mental fatigue classification. The SVM- recall shows that class 1 has been classified correctly compared to the other 4 classes for physical fatigue classification. Precision is the ratio of the total number of correctly classified positive classes(TP) divided by the total number of predicted positive classes(TP+FP). The RF-precision shows that classes 3 and 4 are more precise than the other 3 classes for mental fatigue. The RF- precision shows that class 4 is more precise than the other 4 classes for physical fatigue. The SVM- precision shows that class 3 is more precise than the other 4 classes for mental fatigue. The SVM- precision shows that class 1 is more precise than the other 4 classes for physical fatigue. F-1 score is used to compare one class with other classes and is defined as the ratio of the product of recall and precision values divided by the total sum of recall and precision values. The RF-f1 score shows that class 2 has more scores compared to the other 4 classes for mental fatigue. The RF-f1 score shows that class 1 has more scores compared to the other 4 classes for physical fatigue. The SVM-f1 score shows that class 3 has more scores compared to the other 4 classes for mental fatigue. The SVM-f1 score shows that class 1 has more scores compared to the other 4 classes for physical fatigue.

Table 5.18 F-1 score,precision and recall for Mental Fatigue.

Metric	**Random Forest**	**SVM**
F1 score	[0.07,0.33,0.28,0.10,0.00]	[0.05,0.31,0.32,0.02,0.00]
Precision	[0.08, 0.29,0.24,0.14,0.00]	[0.09, 0.28,0.24,0.18,0.00]
Recall	[0.06, 0.40,0.33,0.08,0.00]	[0.03, 0.36,0.49,0.01,0.00]

Figure 5.5 Random forest Confusion matrix- Mental fatigue

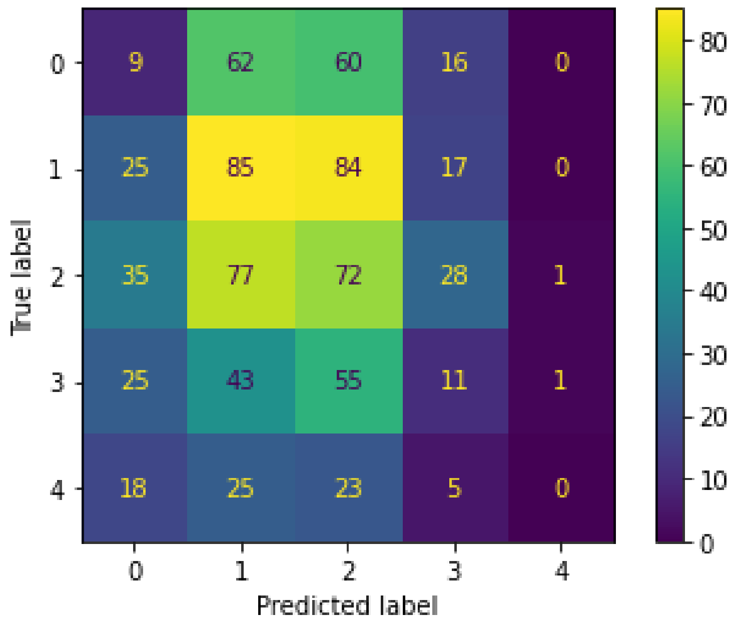


Figure 5.6 SVM Confusion matrix- Mental fatigue

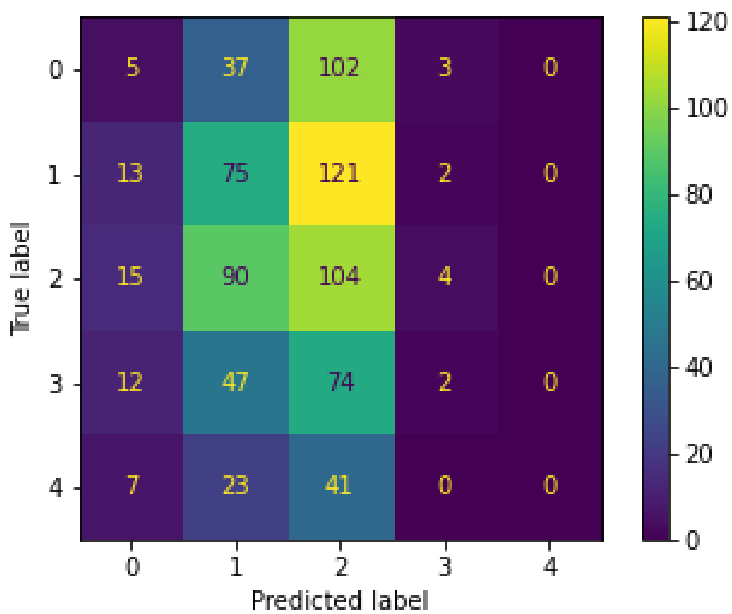


Figure 5.7 Random forest Confusion matrix- Physical fatigue

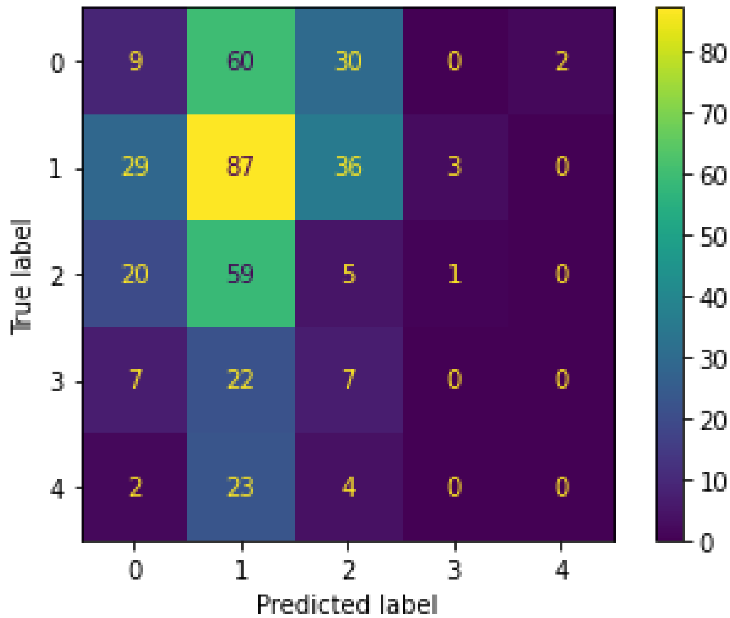


Figure 5.8 SVM Confusion matrix- Physical fatigue

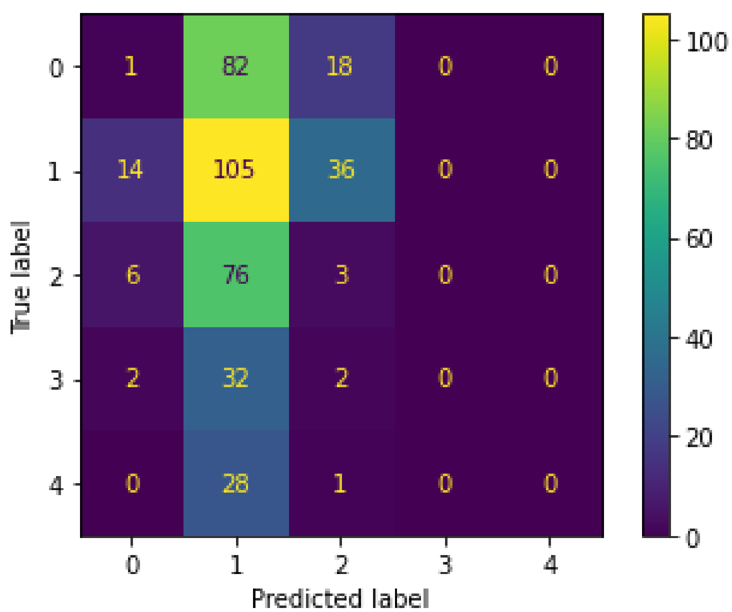


Table 5.19 F-1 score,precision and recall for Physical Fatigue.

Metric	Random Forest	SVM
F1 score	[0.11,0.43, 0.06, 0.00,0.00]	[0.02,0.44,0.04,0.00,0.00]
Precision	[0.13,0.35,0.06,0.00,0.00]	[0.04,0.33,0.05,0.00,0.00]
Recall	[0.09,0.56,0.06,0.00,0.00]	[0.1,0.68,0.04, 0.00,0.00]

The average accuracy of the Mental fatigue model using the Leave one subject out (LOSO) method for the random forest is 23% and for the support vector machine is 24% as is shown in table 5.16 based on which participant is left out. The average accuracy of the Physical fatigue model using the Leave one subject out (LOSO) method for the random forest is 25% and for the support vector machine is 27% as is shown in table 5.17 based on which participant is left out. Due to sensor failure, many participants were dropped as the heart rate was 0 and which in turn affected the model accuracy.

5.2.1 Train Physical and Test Mental

The Random Forest and SVM models are trained for physical task data considering physiological signals as x-train and physical fatigue level as y-train. This trained model is tested for mental task data considering physiological signals as x-test and the prediction accuracy is calculated using physical fatigue level as y-test.

The Random Forest model gave an accuracy of 22.56 % and the SVM model gave an accuracy of 25.69 % . The confusion matrix of random forest and SVM classification model are shown in figure 5.9 and 5.10 respectively. The RF-confusion matrix of physical fatigue has class 1 and class 2 with maximum true positive values compared to that of the remaining 3 classes. The SVM-confusion matrix has class 1 and class 2 with maximum true positive values compared to that of the remaining 3 classes.

Figure 5.9 Train Physical and Test Mental- Random forest Confusion matrix

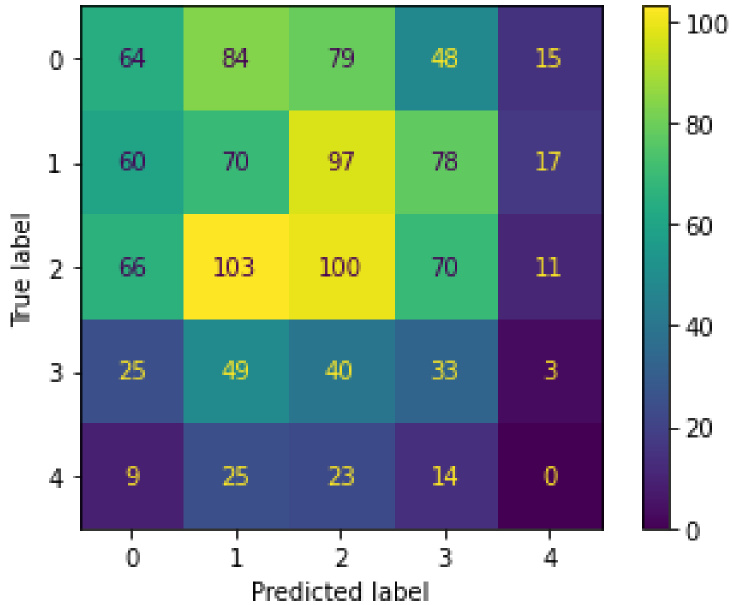
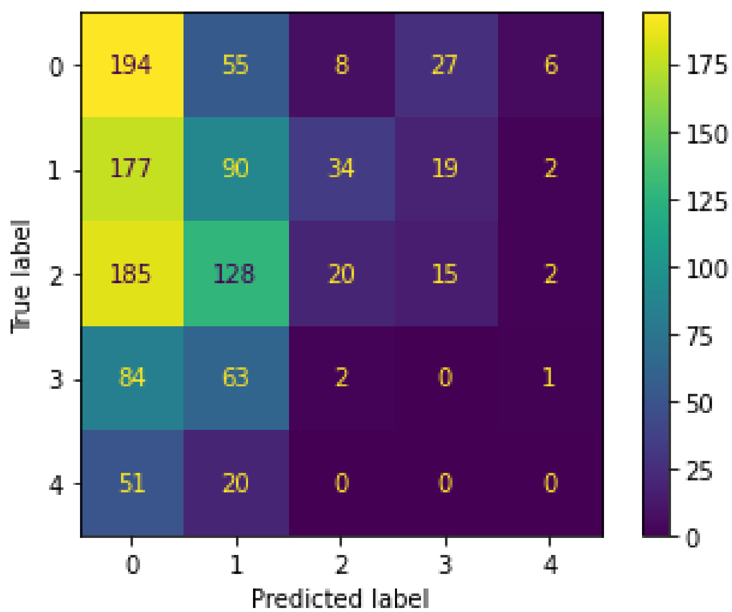


Figure 5.10 Train Physical and Test Mental- SVM Confusion matrix



5.2.2 Train Mental and Test Physical

The Random Forest and SVM models are trained for mental task data considering physiological signals as x-train and mental fatigue level as y-train. This trained model is tested for physical task data considering physiological signals as x-test and the prediction accuracy is calculated using mental fatigue level as y-test.

The Random Forest model gave an accuracy of 18.76 % and the SVM model gave an accuracy of 41.67 %. The confusion matrix of random forest and SVM classification model are shown in figure 5.11 and 5.12 respectively. The RF-confusion matrix of physical fatigue has class 3 and class 2 with maximum true positive values compared to that of the remaining 3 classes. The SVM-confusion matrix has class 3 and class 2 with maximum true positive values compared to that of the remaining 3 classes.

Figure 5.11 Train Mental and Test Physical- Random forest Confusion matrix

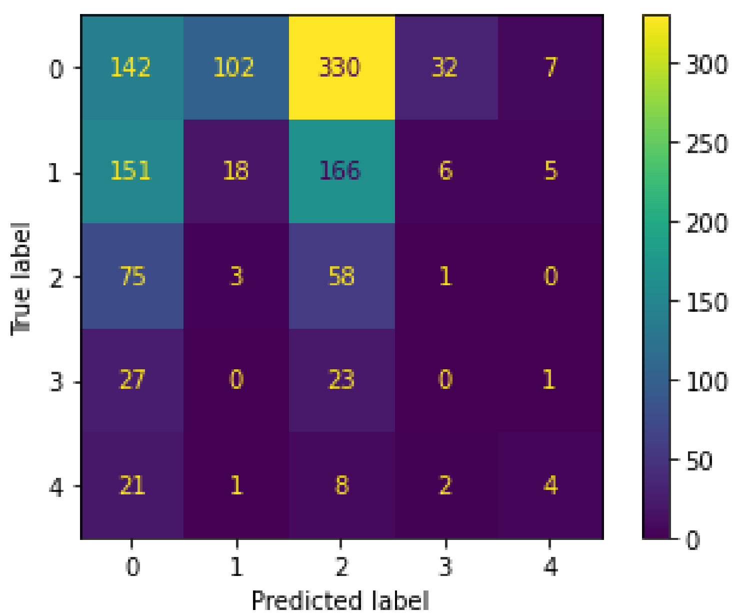
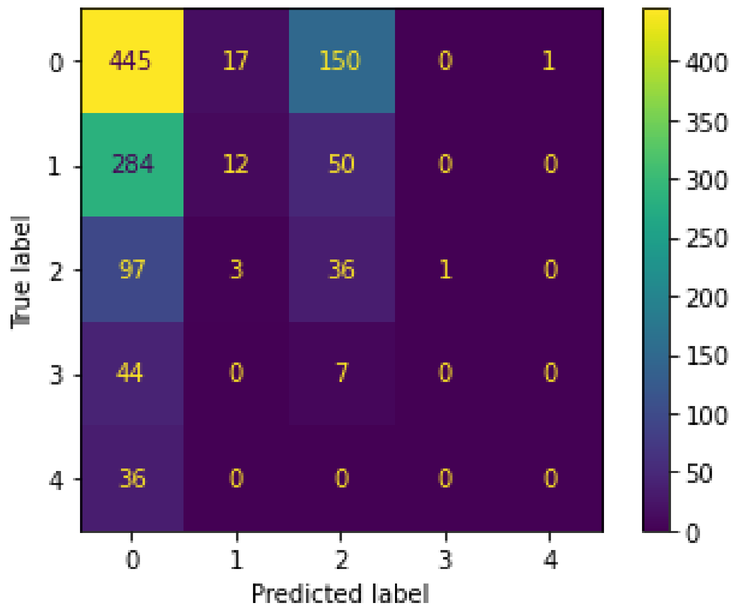


Figure 5.12 Train Mental and Test Physical- SVM Confusion matrix



5.3 Deep Learning

The cross-validation of the LSTM model is done using the Leave one out cross-validation method. The procedure to calculate the validation accuracy, confusion matrix, and accuracy score is the same as explained in RF and SVM except the model built and used here is an LSTM model. The LOOCV leaves one participant for testing and training for the remaining participants and this procedure is calculated n-times where n is the number of participants in a study. The f1-score, precision, and recall of LSTM for physical and mental classification model are represented in Table 5.20and 5.21 respectively.

The LSTM- confusion matrix of mental fatigue and physical fatigue is shown in figure 5.13. and figure 5.14. respectively. The LSTM-confusion matrix of mental fatigue has class 3 and class 2 with maximum true positive values. The LSTM-confusion matrix of physical fatigue has class 1 and class 2 with maximum true positive values compared to that of the remaining 3 classes. The LSTM- recall shows that class 2 has been classified correctly compared to the other 4 classes for mental fatigue classification. The LSTM-recall shows that class 1 has been classified correctly compared to the other 4 classes for physical fatigue classification. The LSTM- precision shows that class 1 is more precise than the other 4 classes for mental fatigue. The LSTM- precision shows that class 1 is more precise than the other 4 classes for physical fatigue. The LSTM-f1 score shows that class 3 has more scores compared to the other 4 classes for mental fatigue. The LSTM-f1 score shows that class 1 has more scores compared to the other 4 classes for physical fatigue.

Table 5.20 F-1 score,precision and recall for Mental Fatigue.

Metric	LSTM
F1 score	[0.09,0.34,0.31,0.20,0.04]
Precision	[0.13,0.30,0.28,0.20,0.08]
Recall	[0.06,0.39,0.34,0.20,0.04]

Table 5.21 F-1 score,precision and recall for Physical Fatigue.

Metric	LSTM
F1 score	[0.15,0.31,0.10,0.00,0.00]
Precision	[0.15,0.30,0.09,0.00,0.00]
Recall	[0.14,0.33,0.11,0.00,0.00]

The average accuracy of the Mental fatigue model using the Leave one subject out (LOSO) method for LSTM is 23% as is shown in table 5.22 based on which participant is left out. The average accuracy of the Physical fatigue model using the Leave one subject out (LOSO) method for LSTM is 19% as is shown in table 5.23 based on which participant is left out. Due to sensor failure, many participants were dropped as the heart rate was 0 and which in turn affected the model accuracy.

Figure 5.13 LSTM Confusion matrix- Mental fatigue

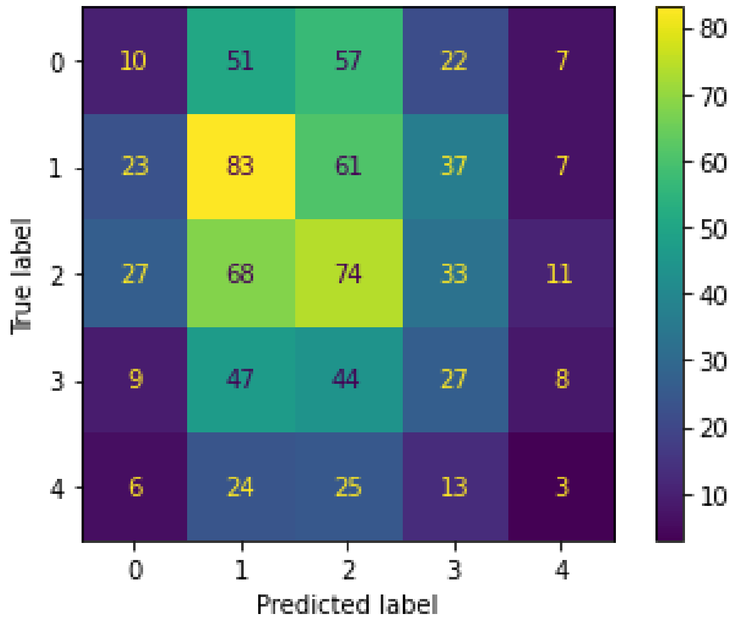


Figure 5.14 LSTM Confusion matrix- Physical fatigue

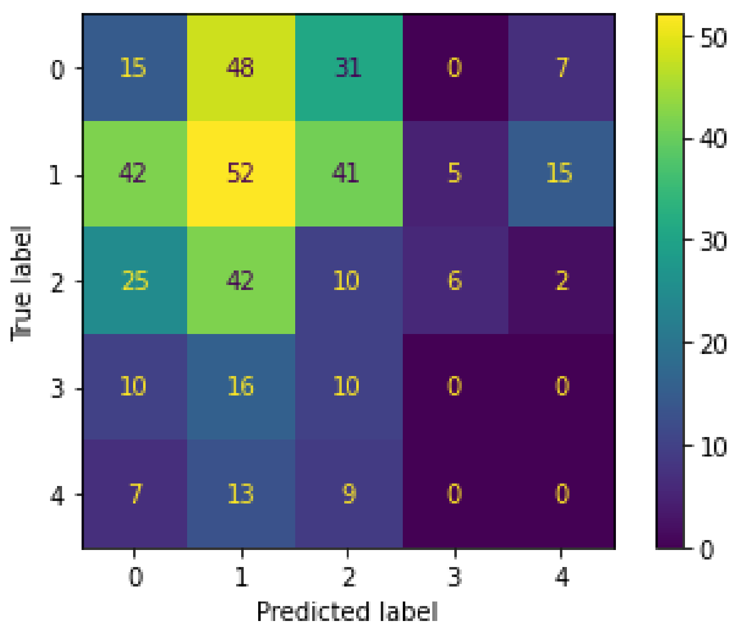


Table 5.22 LSTM Classification Accuracy by Participant and Mental Fatigue

Participant	Mental Fatigue
P1	0.35
P2	0.37
P3	0.18
P4	0.29
P6	0.32
P7	0.19
P8	0.38
P9	0.5
P10	0.40
P11	0.19
P12	0.21
P13	0.23
P14	0.04
P17	0.11
P18	0.20
P19	0.20
P20	0.26
P21	0.09
P22	0.15
Avg.	0.23

Table 5.23 LSTM Classification Accuracy by Participant and Physical Fatigue

Participant	Physical Fatigue
P2	0.33
P3	0.24
P6	0.15
P7	0.06
P8	0.32
P10	0.0
P12	0.14
P13	0.09
P15	0.20
P17	0.02
P19	0.02
P20	0.40
Avg.	0.19

5.4 Binary fatigue Classification

The collected physiological information is windowed for 1-minute with a 10-second stride. This segmented data is then preprocessed to reduce noise and support feature extraction.

The extracted features are then used in a machine-learning algorithm that classifies mental or physical fatigue as low or high. A total of 36 features were extracted after preprocessing the data. After PCA the features were reduced to 17 features for the physical model and 18 features for the mental model.

5.4.1 Leave One subject out cross validation

The machine learning models considered for this study are Random Forest(RF), and Support Vector Machine (SVM). Each model was trained to classify either physical or mental fatigue as low (0) or high (1). The low labels were from the first 5-minute resting baseline collection and the high labels were from the last 5-minutes from the mental or physical fatigue conditions. The intuition is that a person will not be fatigued when they first come into the experiment but will be fatigued at the end of each fatigue condition. LOSO CV trains a model on all but one participant and validates on the remaining participant. This approach is repeated such that each participant is tested once.

Random Forest Model: The best parameter used for the random forest classification of mental fatigue are: 500 trees and a tree depth of 10 for the mental model and the confusion matrix is shown in Figure. 5.17 and the f1-score, precision,support and recall is provided in Table. 5.24. The accuracy of the Mental model is 86.2%. The best parameter used for the random forest classification of mental fatigue is: 500 trees and a tree depth of 12 for the mental model and the confusion matrix is shown in Figure. 5.20 and the f1-score, precision,support and recall is provided in Table. 5.24. The accuracy of the Physical model is 84.4%.

Figure 5.17 Mental fatigue Confusion matrix

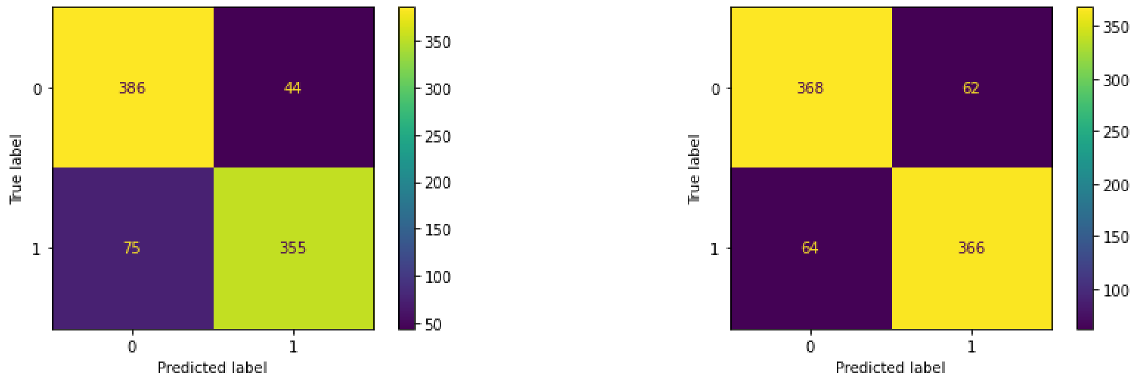
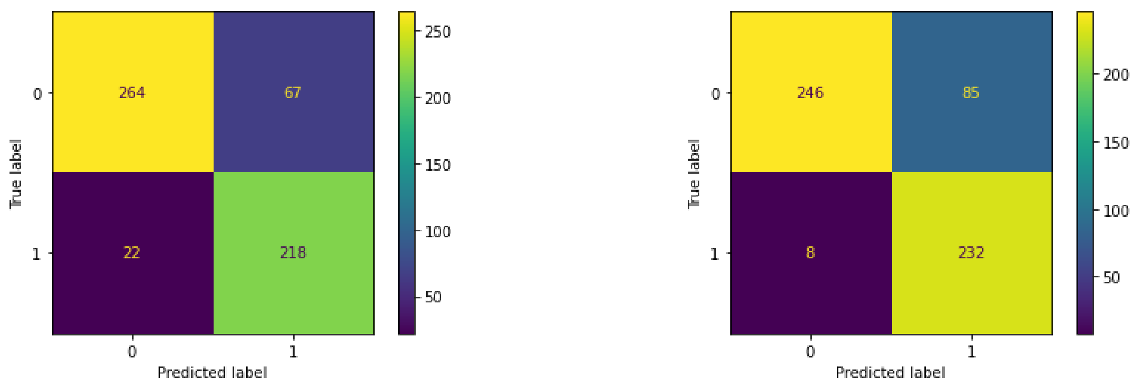


Figure 5.20 Physical fatigue Confusion matrix



Support Vector Machine Model: The best parameter used for the support vector machine classification of mental fatigue are: the polynomial kernel and a gamma value of 0.07 for the mental model and the confusion matrix is shown in Figure. 5.17 and the f1-score, precision,support and recall is provided in Table. 5.24. The accuracy of the Mental model is 85.3%. The best parameter used for the support vector machine classification of physical fatigue is: a polynomial kernel and a gamma value of 0.07 for the mental model and the confusion matrix is shown in Figure. 5.20 and the f1-score, precision,support and recall is provided in Table. 5.24. The accuracy of the Physical model is 83.7%.

Table 5.24 LOSO CV Results by Model Type and Fatigue Type.

Model	Fatigue	Precision	Recall	F-1 score	Support	Accuracy(%)
Random Forest	Mental	0.86	0.86	0.86	860	86.2
	Physical	0.86	0.84	0.85	571	84.4
Support Vector Machine	Mental	0.85	0.85	0.85	860	85.3
	Physical	0.87	0.84	0.84	571	83.7

5.4.2 Cross fatigue validation

The Cross fatigue validation paradigm examines how a model performs classifying any fatigue type by training on data corresponding to a baseline and one fatigue condition and testing on the remaining condition. For example, a model may be trained on the baseline condition and the last 5-minutes of the mental fatigue condition. The model is then evaluated on the last 5-minutes of the physical fatigue condition. There are two cases: Train physical-test mental and train mental-test physical.

Random Forest Model: The best parameter used for the random forest classification is: 200 trees and the number of features at every split is done using the sqrt function for train physical - test mental model and the confusion matrix is shown in Figure. 5.23 and the f1-score, precision,support and recall is provided in Table. 5.25. The accuracy of the Mental model is 50.7%. The best parameter used for the random forest classification is: 200 trees, tree depth of 220, and the number of features at every split is done using the auto function for train mental - test physical model, and the confusion matrix is shown in Figure. 5.26 and the f1-score, precision,support and recall is provided in Table. 5.25. The accuracy of the Mental model is 54.2%.

Support Vector Machine Model: The best parameter used for the random forest classification are: radial basis function(RBF) kernel, C-component of 30, gamma value of 0.005,

Figure 5.23 Train physical test mental Confusion matrix

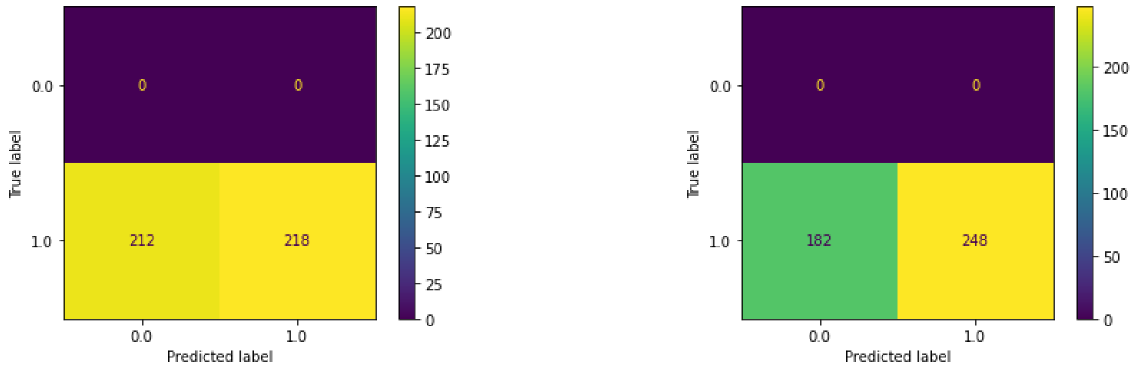


Figure 5.26 Train mental test physical Confusion matrix

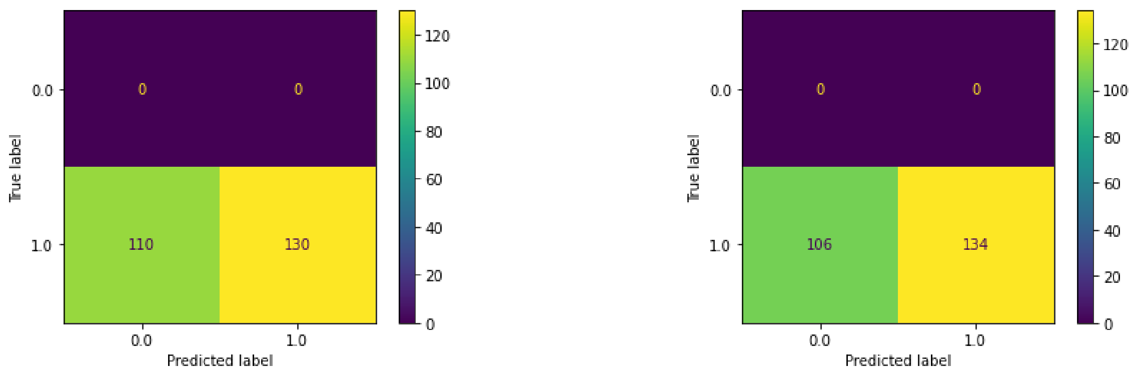


Table 5.25 Cross-Fatigue Classification Results by Model Type and Fatigue Type.

Model	Fatigue	Precision	Recall	F-1 score	Support	Accuracy(%)
Random Forest	Physical	1.00	0.51	0.67	430	50.7
	Mental	1.00	0.54	0.70	240	54.2
Support Vector Machine	Physical	1.00	0.51	0.67	430	50.7
	Mental	1.00	0.54	0.70	240	54.2

and true probability for train physical - test mental model, and the confusion matrix is shown in Figure. 5.23 and the f1-score, precision,support and recall is provided in Table. 5.25. The accuracy of the Mental model is 50.7%. The best parameter used for the random forest classification is: radial basis function(RBF) kernel, C-component of 30, gamma value of 0.005, and with true probability for train mental - test physical model and the confusion matrix is shown in Figure. 5.26 and the f1-score, precision,support and recall is provided in Table. 5.25. The accuracy of the Mental model is 54.2%.

Chapter 6

Conclusion and Future Work

6.1 Conclusion and Discussion

The real-time physiological metrics such as heart rate, heart rate variability, respiration rate, posture magnitude, and average EMG median frequency statistical results are analyzed in this session. Each of these metrics is used to obtain the statics by fatigue type and subjective fatigue levels (that is mental and physical fatigue analysis) by calculating the mean and standard deviation of the metrics. The analysis shows that the as mean the heart rate increases the heart rate variability decreases, for example, the fatigue level 1 of the physical fatigue type has a mean heart rate of 102, level2- 110, and level 3- 105, the respective heart rate variability values are level 1- 36, level 2- 28 and level 3- 33. The heart rate with the highest mean value has the least heart rate variability mean value. Therefore, HR and HRV are inversely proportional to the onset of fatigue. The increased respiration rate decreases the standard deviation of heart rate as fatigue level increases. The EMG median frequency decreases as the fatigue level increases. The Mann-Whitney U-test is used to calculate the difference between the groups concerning each metric, it found that group 4 and group 5 always fail to reject the null hypothesis except for the average EMG median frequency metric. The subjective rating of average fatigue levels increases with trials and, trial duration decreases with an increase in trial numbers. Mental and physical fatigue levels are negatively correlated with trail duration with (r=-0.759923)and (r= -0.16206) respectively.

Machine Learning: The machine learning results are validated using the Leave one out cross-validation (LOOCV) method for each mental and physical model. The train test split is done using the LOO split. The accuracies are mentioned in Tables 5.16. and 5.17. The accuracy using the Random Forest of Mental classification model shows that the model gives the best accuracy of 0.42 when participant 8 is considered as a test participant, and the Support Vector Machine gives the best accuracy of 0.5 when participant 10 is considered as a test participant. The accuracy using the Random Forest of Physical classification model shows that the model gives the best accuracy of 0.58 when participant 2 is considered as a test participant, and the Support Vector Machine gives the best accuracy of 0.78 when participant 2 is considered as a test participant. This section discusses one more method where the RF and SVM models are trained and tested in two ways, case 1: the models are trained for the mental task to detect mental fatigue level and tested on physical task to detect the mental fatigue level, case 2: the models are trained for the physical task to detect physical fatigue level and tested on mental task to detect the physical fatigue level. The accuracy of both cases is less than 50%.

Deep Learning: The deep learning results of the LSTM model are also validated using the Leave one out cross-validation (LOOCV) method for each mental and physical model. The accuracies are mentioned in Table 5.22 and 5.23. The accuracy using LSTM of the Mental classification model shows that the model gives the best accuracy of 0.40 when participant 10 is considered as a test participant. The accuracy using LSTM of the Physical classification model shows that the model gives the best accuracy of 0.32 when participant 2 is considered as a test participant.

Binary classification: Two machine learning models were developed to classify mental and physical fatigue as high or low using data collected during a human-subjects experiment. Overall, both models were able to classify a previously unseen person's fatigue level (validated using Leave-One-Subject-Out Cross-Validation) using cardiac, respiration, and electromyography information. This result indicates that both fatigue components can be incorporated into a human-robot system to help monitor a human's performance level. The binary classification models gave better accuracy that is more than 80% which proves that the data collected is valid and can be used to improve accuracy for the classification of 5 fatigue levels.

The ability to improve task performance in a high-intensity work environment by adapting to the fatigue level has gained considerable research interest. This study focuses both on identifying mental fatigue and physical fatigue without being task-specific. This thesis developed a model which detects both physical and mental fatigue levels. The experiments were designed in which both mental and physical abilities were required to complete the tasks. The machine learning and deep learning algorithms were validated using the experimental data obtained from the tasks. A fatigue prediction model used machine learning and a deep learning algorithm to predict the fatigue level. The developed model can be deployed into a system that can be applied across different task domains. The system may be able to allocate a physically demanding task to a human if the human has been classified as mentally fatigued (or vice-versa). Similarly, the system may invoke autonomy for the current task to help mitigate the human's fatigued state.

6.1.1 Limitations

This thesis developed a multi-modal fatigue classification model which is used to detect both physical and mental fatigue in any working environment. The major limitation of this study is that all participants used for this experiment were college students. The participants were asked to do a repetitive task which lead to boredom which might have been confused with fatigue level during the subjective rating. The sensor failure and assessing a limited set of sample sizes affected the model accuracy. The fatigue classification algorithm failed to detect the 4th and 5th fatigue level efficiently. Therefore more data supporting this problem was required. All of the fatigue algorithms are limited to the subjective rating of fatigue level for experiments. None of the algorithms achieve more than 80% accuracy in this study. But to prove the valid dataset a new model was built where the dataset was divided into two classes and the results achieved were more than 80%. Thus, said algorithms are not completely suitable for assessing fatigue in this task environment, more accuracy might be achieved using more data and a higher epoch and removing more noise from the data.

6.2 Future Work

The proposed algorithm can be improved to create a better fatigue classification model which can be further used human robots team. This fatigue detecting model can be used in a working environment where continuous mental and physical attention is required to complete the task. For example, in an industrial environment where a worker needs to work with a robot to complete a task continuously for an hour or two where a minute mistake might incur a great loss to the company, in such situations we need a model which inputs the robot the real-time physiological data of the worker and predict the fatigue level and change the speed of the work or take over the work according to the fatigue level of the worker. This classification model will decrease many human-caused disasters and control the future disaster which might affect on a small or large scale the worker and the working environment.

Bibliography

[1] K. S. Pasupathy and L. M. Barker, "Impact of fatigue on performance in registered nurses: Data mining and implications for practice," *Journal for Healthcare Quality*, vol. 34, no. 5, pp. 22–30, 2012.

[2] H. Van Dongen, G. Maislin, J. M. Mullington, and D. F. Dinges, "The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation," *Sleep*, vol. 26, no. 2, pp. 117–126, 2003.

[3] J.-J. Wan, Z. Qin, P.-Y. Wang, Y. Sun, and X. Liu, "Muscle fatigue: general understanding and treatment," *Experimental & molecular medicine*, vol. 49, no. 10, pp. e384–e384, Oct 2017, 28983090[pmid]. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28983090

[4] V. J. Gawron, J. French, and D. Funke, "An overview of fatigue," *Stress, workload, and fatigue*, 2001.

[5] J. A. Stern, D. Boyer, and D. Schroeder, "Blink rate: a possible measure of fatigue," *Human factors*, vol. 36, no. 2, pp. 285–297, 1994.

[6] L. S. Caldwell and J. M. Lyddan, "Serial isometric fatigue functions with variable intertrial intervals," *Journal of Motor Behavior*, vol. 3, no. 1, pp. 17–30, 1971, pMID: 23941345. [Online]. Available: https://doi.org/10.1080/00222895.1971.10734888

[7] Y. Yue, D. Liu, S. Fu, and X. Zhou, "Heart rate and heart rate variability as classification features for mental fatigue using short-term ppg signals via smartphones instead of ecg recordings," in *2021 13th International Conference on Communication Software and Networks (ICCSN)*, 2021, pp. 370–376.

[8] T. Gohara, H. Mizuta, I. Takeuchi, O. Tsuda, K. Yana, T. Yanai, Y. Yamamoto, and N. Kishi, "Heart rate variability change induced by the mental stress: the effect of accumulated fatigue," in *Proceedings of the 1996 Fifteenth Southern Biomedical Engineering Conference*, 1996, pp. 367–369.

[9] D. M. Brown and S. R. Bray, “Heart rate biofeedback attenuates effects of mental fatigue on exercise performance,” *Psychology of Sport and Exercise*, vol. 41, pp. 70–79, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1469029218303960

[10] S. Akselrod, D. Gordon, F. Ubel, D. Shannon, A. Berger, and R. Cohen, “Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control,” *Science*, vol. 213, no. 4504, pp. 220–222, 1981. [Online]. Available: https://science.sciencemag.org/content/213/4504/220

[11] C. Zhang and X. Yu, “Estimating mental fatigue based on electroencephalogram and heart rate variability,” *Polish Journal of Medical Physics And Engineering*, vol. 16, no. 2, p. 67, 2010.

[12] T. Kiryu, N. Motomiya, Y. Ushiyama, and M. Okada, “Evaluation of fatigue using heart rate variability and myoelectric signals during skiing,” in *Proceedings of the First Joint BMES/EMBS Conference. 1999 IEEE Engineering in Medicine and Biology 21st Annual Conference and the 1999 Annual Fall Meeting of the Biomedical Engineering Society (Cat. N*, vol. 1, 1999, pp. 585 vol.1–.

[13] N. Dimitrova and G. Dimitrov, “Interpretation of emg changes with fatigue: facts, pitfalls, and fallacies,” *Journal of Electromyography and Kinesiology*, vol. 13, no. 1, pp. 13–36, 2003.

[14] T. Sadoyama and H. Miyano, “Frequency analysis of surface emg to evaluation of muscle fatigue,” *European Journal of Applied Physiology and Occupational Physiology*, vol. 47, no. 3, pp. 239–246, 1981.

[15] Z. Dong, M. Zhang, J. Sun, T. Cao, R. Liu, Q. Wang, and Danliu, “A fatigue driving detection method based on frequency modulated continuous wave radar,” in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, 2021, pp. 670–675.

[16] M. Grassmann, E. Vlemincx, A. Von Leupoldt, J. M. Mittelstädt, and O. Van den Bergh, “Respiratory changes in response to cognitive load: a systematic review.” *Neural plasticity*, 2016.

[17] D. Wang, P. Shen, T. Wang, and Z. Xiao, “Fatigue detection of vehicular driver through skin conductance, pulse oximetry and respiration: A random forest classifier,” in *2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN)*, 2017, pp. 1162–1166.

[18] A. Greco, A. Guidi, F. Felici, A. Leo, E. Ricciardi, M. Bianchi, A. Bicchi, L. Citi, G. Valenza, and E. P. Scilingo, "Muscle fatigue assessment through electrodermal activity analysis during isometric contraction," in *2017 39th annual international conference of the Ieee engineering in medicine and biology society (Embc)*. IEEE, 2017, pp. 398–401.

[19] L. Boon-Leng, L. Dae-Seok, and L. Boon-Giin, "Mobile-based wearable-type of driver fatigue detection by gsr and emg," in *TENCON 2015-2015 IEEE Region 10 Conference*. IEEE, 2015, pp. 1–4.

[20] D. Wang, P. Shen, T. Wang, and Z. Xiao, "Fatigue detection of vehicular driver through skin conductance, pulse oximetry and respiration: A random forest classifier," in *2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN)*. IEEE, 2017, pp. 1162–1166.

[21] Y. Yue, D. Liu, S. Fu, and X. Zhou, "Heart rate and heart rate variability as classification features for mental fatigue using short-term ppg signals via smartphones instead of ecg recordings," in *2021 13th International Conference on Communication Software and Networks (ICCSN)*. IEEE, 2021, pp. 370–376.

[22] V. Jeyhani, S. Mahdiani, M. Peltokangas, and A. Vehkaoja, "Comparison of hrv parameters derived from photoplethysmography and electrocardiography signals," in *2015 37th annual international conference of the ieee engineering in medicine and biology society (EMBC)*. IEEE, 2015, pp. 5952–5955.

[23] J. Lee, J. Kim, and M. Shin, "Correlation analysis between electrocardiography (ecg) and photoplethysmogram (ppg) data for driver's drowsiness detection using noise replacement method," *Procedia computer science*, vol. 116, pp. 421–426, 2017.

[24] Z. S. Maman, Y.-J. Chen, A. Baghdadi, S. Lombardo, L. A. Cavuoto, and F. M. Megahed, "A data analytic framework for physical fatigue management using wearable sensors," *Expert Systems with Applications*, vol. 155, p. 113405, 2020.

[25] K.-Q. Shen, C.-J. Ong, X.-P. Li, Z. Hui, and E. P. V. Wilder-Smith, "A feature selection method for multilevel mental fatigue eeg classification," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 7, pp. 1231–1237, 2007.

[26] C. Zhang, C. Zheng, X. Yu, and Y. Ou, "Mental fatigue estimation based on multichannel linear descriptors and support vector machine with optimizing parameters," in *Third International Conference on Natural Computation (ICNC 2007)*, vol. 2, 2007, pp. 309–313.

[27] B. K. Savaş and Y. Becerikli, “Real time driver fatigue detection based on svm algorithm,” in *2018 6th International Conference on Control Engineering Information Technology (CEIT)*, 2018, pp. 1–4.

[28] A. Aryal, A. Ghahramani, and B. Becerik-Gerber, “Monitoring fatigue in construction workers using physiological measurements,” *Automation in Construction*, vol. 82, pp. 154–165, 2017.

[29] Y. Zheng, T. C. Wong, B. H. Leung, and C. C. Poon, “Unobtrusive and multimodal wearable sensing to quantify anxiety,” *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3689–3696, 2016.

[30] D. li Hu, G. cheng Gong, Z. chun Mu, C. wu Han, and X. hua Zhao, “Modeling research on driver fatigue,” in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, vol. 4, 2010, pp. V4–158–V4–162.

[31] R. Fu, H. Wang, and W. Zhao, “Dynamic driver fatigue detection using hidden markov model in real driving condition,” *Expert Systems with Applications*, vol. 63, pp. 397–411, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417416303293

[32] H. Al-Libawy, A. Al-Ataby, W. Al-Nuaimy, and M. A. Al-Taee, “Hrv-based operator fatigue analysis and classification using wearable sensors,” in *2016 13th International Multi-Conference on Systems, Signals & Devices (SSD)*. IEEE, 2016, pp. 268–273.

[33] M. Valueva, N. Nagornov, P. Lyakhov, G. Valuev, and N. Chervyakov, “Application of the residue number system to reduce hardware costs of the convolutional neural network implementation,” *Mathematics and Computers in Simulation*, vol. 177, pp. 232–243, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378475420301580

[34] A. Saeed, S. Trajanovski, M. Van Keulen, and J. Van Erp, “Deep physiological arousal detection in a driving simulator using wearable sensors,” in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2017, pp. 486–493.

[35] V. Yarlagadda, S. G. Koolagudi, M. Kumar M V, and S. Donepudi, “Driver drowsiness detection using facial parameters and rnns with lstm,” in *2020 IEEE 17th India Council International Conference (INDICON)*, 2020, pp. 1–7.

[36] K. Balaskas and K. Siozios, “Fatigue detection using deep long short-term memory autoencoders,” in *2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 2021, pp. 1–4.

[37] A. Georgakis, L. K. Stergioulas, and G. Giakas, “Fatigue analysis of the surface emg signal in isometric constant force contractions using the averaged instantaneous frequency,” *IEEE transactions on biomedical engineering*, vol. 50, no. 2, pp. 262–265, 2003.

.1 Appendix

The 32 features are included for correlation of this experiment. The features are:

heart_rate_grad, heart_rate_mean, heart_rate_slope, heart_rate_std, heart_rate_var, posture_rate_grad, posture_rate_mean, posture_rate_slope, posture_rate_std, posture_rate_var, respiration_rate_grad, respiration_rate_mean, respiration_rate_slope, respiration_rate_std, respiration_rate_var, EMG_1_medianfreq_r, EMG_2_medianfreq_r, EMG_3_medianfreq_r, EMG_4_-medianfreq_r, EMG_5_medianfreq_r, EMG_6_medianfreq_r, EMG_7_medianfreq_r, EMG_8-_medianfreq_r, EMG_1_medianfreq_l, EMG_2_medianfreq_l, EMG_3_medianfreq_l, EMG_4-_medianfreq_l, EMG_5_medianfreq_l, EMG_6_medianfreq_l, EMG_7_medianfreq_l, EMG_8-_medianfreq_l, Trial_Duration. The Pearson correlation of physical model is shown in figure below (fig 4): The Pearson correlation of mental model is shown in figure below (fig 5): The heat map of correlation between features are represented in the fig 4 & 5, the color encoding is mentioned in the image, the color white represents that the features are highly correlated represented by value 1 and the black represents the features are highly not correlated represented value -1. The highly correlated features are removed from the data-set to decrease the complexity of the algorithm, thus increasing the risk of errors. The correlation value generally range from (-1,1). We see that the diagonal value of correlation is one in both the models as it is self correlation.The correlation is moderate in mental fatigue model compared to that of physical fatigue model.

	Trial_Duration	Physical Fatigue Level
Trial_Duration	1	0.167722385
heart_rate_grad	0.024034	0.005407575
heart_rate_mean	-0.14294	0.211429817
heart_rate_slope	0.002938	0.006158822
heart_rate_std	0.199625	-0.15200995
heart_rate_var	0.149622	-0.136617702
posture_grad	0.069644	-0.006414259
posture_mean	-0.23391	-0.09797874
posture_slope	0.082718	-0.057823936
posture_std	-0.50377	-0.244508235
posture_var	-0.46005	-0.282523684
respiration_rate_grad	-0.00082	-0.022652063
respiration_rate_mean	0.044905	0.205537376
respiration_rate_slope	0.00748	-0.020777379
respiration_rate_std	0.345842	0.205803033
respiration_rate_var	0.334056	0.207186515
EMG_1_medianfreq_r	-0.17662	-0.157670249
EMG_2_medianfreq_r	-0.08811	-0.171573736
EMG_3_medianfreq_r	0.099651	-0.028367707
EMG_4_medianfreq_r	-0.07046	0.000245065
EMG_5_medianfreq_r	-0.17802	-0.094957798
EMG_6_medianfreq_r	-0.16632	-0.032275406
EMG_7_medianfreq_r	-0.10947	-0.152885078
EMG_8_medianfreq_r	-0.0898	-0.133358025
EMG_1_medianfreq_l	-0.3288	-0.093113053
EMG_2_medianfreq_l	-0.22786	-0.161217439
EMG_3_medianfreq_l	-0.10303	-0.100481988
EMG_4_medianfreq_l	-0.15624	-0.081715037
EMG_5_medianfreq_l	-0.15912	-0.03343019
EMG_6_medianfreq_l	-0.13553	-0.112137149
EMG_7_medianfreq_l	-0.18841	-0.053465367
EMG_8_medianfreq_l	-0.23192	-0.004894926
Physical Fatigue Level	0.167722	1

heart_rate	gradient	mean	slope	standard deviation	variance
Trial_Duration	0.024034	-0.14294	0.002938	0.199625	0.149622
heart_rate_grad	1	0.026	0.955948	-0.11573	-0.12321
heart_rate_mean	0.026	1	-0.00374	-0.48821	-0.41491
heart_rate_slope	0.955948	-0.00374	1	-0.1046	-0.1227
heart_rate_std	-0.11573	-0.48821	-0.1046	1	0.949972
heart_rate_var	-0.12321	-0.41491	-0.1227	0.949972	1
posture_grad	-0.06506	-0.01887	-0.07424	-0.0096	-0.01569
posture_mean	-0.00873	0.158424	-0.01534	-0.03647	0.00061
posture_slope	-0.05995	0.013818	-0.07127	-0.00252	0.000757
posture_std	-0.03093	0.153163	-0.01439	-0.18619	-0.10762
posture_var	-0.03689	0.169081	-0.01795	-0.19727	-0.12233
respiration_rate_grad	-0.06493	0.004034	-0.03822	-0.00415	-0.01003
respiration_rate_mean	0.016042	0.3873	0.00976	-0.31924	-0.2366
respiration_rate_slope	-0.04145	0.005561	-0.01773	-0.00374	-0.01276
respiration_rate_std	0.002157	0.231938	-0.0061	-0.12483	-0.09417
respiration_rate_var	0.006252	0.169036	0.001639	-0.11466	-0.09531
EMG_1_medianfreq_r	0.023895	0.020913	0.006222	0.02257	0.014692
EMG_2_medianfreq_r	-0.03177	0.091021	-0.02943	-0.0746	-0.05235
EMG_3_medianfreq_r	-0.06995	0.051083	-0.07135	0.011908	0.024445
EMG_4_medianfreq_r	-0.05798	0.037804	-0.06843	-0.01807	-0.01153
EMG_5_medianfreq_r	-0.07971	0.066273	-0.0803	0.003836	0.012945
EMG_6_medianfreq_r	-0.07004	-0.03678	-0.06388	0.018107	0.016613
EMG_7_medianfreq_r	-0.00437	0.054244	-0.01998	-0.00122	-0.01687
EMG_8_medianfreq_r	-0.02522	0.040395	-0.04169	-0.02572	-0.04082
EMG_1_medianfreq_l	0.02779	0.082183	0.035459	-0.10875	-0.09783
EMG_2_medianfreq_l	0.001816	-0.10138	0.00868	-0.02374	-0.02983
EMG_3_medianfreq_l	-0.06093	-0.12433	-0.04159	-0.0322	-0.03368
EMG_4_medianfreq_l	-0.04615	-0.00099	-0.03191	-0.08676	-0.0953
EMG_5_medianfreq_l	-0.02151	0.020715	-0.01317	-0.10076	-0.09526
EMG_6_medianfreq_l	-0.04968	-0.07165	-0.03092	-0.12647	-0.09535
EMG_7_medianfreq_l	-0.02301	0.110287	-0.03055	-0.14432	-0.12305
EMG_8_medianfreq_l	-0.01638	0.028378	0.005743	-0.07685	-0.06502
Physical Fatigue Level	0.005408	0.21143	0.006159	-0.15201	-0.13662

posture	gradient	mean	slope	stand deviation	variance
Trial_Duration	0.069644	-0.23391	0.082718	-0.50377	-0.46005
heart_rate_grad	-0.06506	-0.00873	-0.05995	-0.03093	-0.03689
heart_rate_mean	-0.01887	0.158424	0.013818	0.153163	0.169081
heart_rate_slope	-0.07424	-0.01534	-0.07127	-0.01439	-0.01795
heart_rate_std	-0.0096	-0.03647	-0.00252	-0.18619	-0.19727
heart_rate_var	-0.01569	0.00061	0.000757	-0.10762	-0.12233
posture_grad	1	0.006711	0.958675	0.081671	0.088799
posture_mean	0.006711	1	0.072154	0.377661	0.279674
posture_slope	0.958675	0.072154	1	0.131131	0.130804
posture_std	0.081671	0.377661	0.131131	1	0.973828
posture_var	0.088799	0.279674	0.130804	0.973828	1
respiration_rate_grad	0.053937	-0.01063	0.038866	0.000307	0.002836
respiration_rate_mean	-0.07741	-0.31094	-0.05158	-0.01065	0.003344
respiration_rate_slope	0.058462	-0.01702	0.046676	-0.003	0.000612
respiration_rate_std	0.026142	-0.21678	0.039865	-0.09907	-0.06117
respiration_rate_var	0.057407	-0.2537	0.06184	-0.14366	-0.09364
EMG_1_medianfreq_r	-0.09172	0.242311	-0.0722	0.102449	0.080167
EMG_2_medianfreq_r	-0.1182	0.068121	-0.11478	0.133624	0.151591
EMG_3_medianfreq_r	-0.02884	-0.04149	-0.02244	0.036268	0.0403
EMG_4_medianfreq_r	-0.1856	0.08493	-0.17578	0.079907	0.085516
EMG_5_medianfreq_r	-0.16243	0.25376	-0.14613	0.192931	0.15836
EMG_6_medianfreq_r	-0.05274	0.188519	-0.06731	0.090708	0.064342
EMG_7_medianfreq_r	-0.02884	0.190305	-0.01557	0.094723	0.093736
EMG_8_medianfreq_r	-0.0399	0.234695	-0.027	-0.00089	-0.00077
EMG_1_medianfreq_l	-0.17941	0.172206	-0.14468	0.247818	0.229491
EMG_2_medianfreq_l	-0.12534	-0.01258	-0.14315	0.130433	0.149334
EMG_3_medianfreq_l	-0.04026	-0.02704	-0.03392	0.007535	0.016586
EMG_4_medianfreq_l	-0.0358	0.102655	-0.0244	0.005049	-0.01943
EMG_5_medianfreq_l	-0.05503	0.140789	-0.04696	0.039023	0.012503
EMG_6_medianfreq_l	-0.03053	-0.01572	-0.05152	0.179562	0.182122
EMG_7_medianfreq_l	-0.11155	0.148355	-0.08822	0.170352	0.143026
EMG_8_medianfreq_l	-0.20333	0.219942	-0.17456	0.177912	0.147921
Physical Fatigue Level	-0.00641	-0.09798	-0.05782	-0.24451	-0.28252

respiration_rate	gradient	mean	slope	standard deviation	variance
Trial_Duration	-0.00082	0.044905	0.00748	0.345842	0.334056
heart_rate_grad	-0.06493	0.016042	-0.04145	0.002157	0.006252
heart_rate_mean	0.004034	0.3873	0.005561	0.231938	0.169036
heart_rate_slope	-0.03822	0.00976	-0.01773	-0.0061	0.001639
heart_rate_std	-0.00415	-0.31924	-0.00374	-0.12483	-0.11466
heart_rate_var	-0.01003	-0.2366	-0.01276	-0.09417	-0.09531
posture_grad	0.053937	-0.07741	0.058462	0.026142	0.057407
posture_mean	-0.01063	-0.31094	-0.01702	-0.21678	-0.2537
posture_slope	0.038866	-0.05158	0.046676	0.039865	0.06184
posture_std	0.000307	-0.01065	-0.003	-0.09907	-0.14366
posture_var	0.002836	0.003344	0.000612	-0.06117	-0.09364
respiration_rate_grad	1	0.005006	0.975209	0.270896	0.326074
respiration_rate_mean	0.005006	1	0.015907	0.515097	0.43569
respiration_rate_slope	0.975209	0.015907	1	0.319924	0.390639
respiration_rate_std	0.270896	0.515097	0.319924	1	0.93546
respiration_rate_var	0.326074	0.43569	0.390639	0.93546	1
EMG_1_medianfreq_r	0.042742	-0.15529	0.032359	-0.08878	-0.09991
EMG_2_medianfreq_r	-0.02925	0.063039	-0.0429	0.062226	0.01892
EMG_3_medianfreq_r	0.007962	0.128645	-0.00177	0.106838	0.065389
EMG_4_medianfreq_r	-0.0406	-0.0032	-0.04106	-0.03958	-0.04784
EMG_5_medianfreq_r	0.031633	-0.08894	0.024852	-0.09008	-0.12317
EMG_6_medianfreq_r	0.019799	-0.17779	0.003113	-0.13665	-0.13552
EMG_7_medianfreq_r	-0.03762	-0.1487	-0.05303	-0.10152	-0.10994
EMG_8_medianfreq_r	-0.02516	-0.21366	-0.0289	-0.16851	-0.15716
EMG_1_medianfreq_l	-0.0305	0.030902	-0.04244	-0.10543	-0.10587
EMG_2_medianfreq_l	-0.04032	-0.01839	-0.04308	-0.17224	-0.13255
EMG_3_medianfreq_l	-0.0461	0.101987	-0.05777	-0.1061	-0.09321
EMG_4_medianfreq_l	-0.05517	0.069194	-0.06341	-0.107	-0.11812
EMG_5_medianfreq_l	0.01873	0.071915	0.011803	-0.08264	-0.07825
EMG_6_medianfreq_l	0.023006	0.130592	0.015135	-0.01896	-0.0027
EMG_7_medianfreq_l	-0.01219	0.11971	-0.0303	-0.01514	-0.02305
EMG_8_medianfreq_l	0.004981	0.013894	-0.01666	-0.10844	-0.11071
Physical Fatigue Level	-0.02265	0.205537	-0.02078	0.205803	0.207187