TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technology
Department of Software Science

Olesja Senkiv 162967IAPM

FATIGUE RECOGNITION MODELLING

Master's thesis

Supervisors: Sven Nõmm, PhD
Aaro Toomela, PhD

Tallinn 2018

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

Olesja Senkiv 162967IAPM

VÄSIMUSE TUVASTAMISE MODELEERIMINE

Magistritöö

Juhendajad: Sven Nõmm, PhD
Aaro Toomela, PhD

Tallinn 2018

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Olesja Senkiv

07.05.2018

Abstract

Fatigue is a natural condition of the human organism after it has lost its energy on an activity. During last century, the state of fatigue has been studied separately and in conjunction with numerous chronic illnesses. However, the issue of fatigue is topical and poorly understood because of its subjectivity. The focus of the present thesis is modelling human fatigue, using a digitalised fine motor skill test as an easy-to-use solution. The main goal is to determine a set of significant parameters that would be used for modelling of fatigue recognition and to make it possible to distinguish between fatigued and non-fatigued individuals.

This thesis is written in English and is 38 pages long, including 8 chapters, 15 figures and 12 tables.

Annotatsioon

Väsimus on inimorganismi loomulik seisund pärast oma jõu kaotamist millegi tegemise tõttu. Viimase sajandi jooksul on uuritud väsimuse seisundit omaette ja mitmete krooniliste haiguste korral. Sellest hoolimata on väsimuse teema aktuaalne ja seda mõistetakse halvasti selle subjektiivsuse tõttu. Käesoleva väitekirja keskmes on inimeste väsimuse modeleerimine, kasutades motoorsete võimete digiteeritud testi kergesti kasutatava lahendusena. Peamine eesmärk on määrata kindlaks oluliste parameetrite kogum, mida kasutada väsimuse äratundmise modeleerimiseks ning mis võimaldaks eristada väsinud ja väsimata üksikisikuid.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 38 leheküljel, 8 peatükki, 15 joonist, 12 tabelit.

List of abbreviations and terms

JSON	JavaSciprt Object Notation – data exchange format
CSV	Comma-Separated Values
PyCharm IDE	Integrated Development Environment
Amazon S3	Amazon Simple Storage Service – data storage in the cloud
DTW	Dynamic Time Warping
MM	Motion Mass
VAS	Visual Analog Scale
E	Euclidean distance
V_m	Velocities mass
A_m	Accelerations mass
J_m	Jerks mass
t	Time
dist_div_am	Ratio of Euclidean distance to acceleration mass
am_div_dist	Ratio of acceleration mass to Euclidean distance
s3, s6, s9, s12	Spiral separation ways respectively – 3 o'clock, 6 o'clock, 9 o'clock, 12 o'clock in the daytime
DT	Decision tree classifier
LR	Logistic regression classifier
SVM	Support Vector Machine classifier
kNN	K Nearest Neighbours classifier

Table of contents

List of figures

List of tables

1 Introduction

The focus of this thesis is to research the influence of fatigue on human fine motor functions. Since there is no information in the literature that physical fatigue and mental fatigue are involved in the same motor functions, they are considered separately in this work.

The fast pace of modern life is keeping people in high tension which naturally causes the lack of energy. At the same time, for people working as pilots or truck drivers, for instance, the state of fatigue is in principle unacceptable and it is, therefore, important to create preventive measures. Fatigue has been recently identified as one of the major factors to cause incidents, efficiency loss and other similar problems which disrupt many areas of human machine interaction [1]. Modern approach to the working environment is to eliminate low productivity periods of an employee. An ability to recognize people's fatigue may improve the quality of the working environment and make people's lives safer.

During the recent time fatigue detection and modelling gain a lot of attention. Nevertheless, relatively few results are devoted to the applicability of fine motor test to tackle the problem. Initially developed to diagnose and model cognitive impairments, like those cause by neurodegenerative disease, fine motor tests provide unique inside in the state of human motor functions both, on the levels of planning and execution of limb motions. The working hypothesis of the present research is that fatigue should affect human abilities to plan and execute motions [2], in the similar way like neurodegenerative diseases [3] or in opposite way to the learning processes [4].

The main goal of this thesis is to research the applicability of fine motor test for distinguishing fatigued and non-fatigued individuals. In case positive result is archived during the research, the aim is to find out a set of parameters that can be used to build a predictive model of differentiation of the fatigued state.

Given the novelty of the topic and the usefulness of the present research in the study of the problem of fatigue, the investigation which forms the basis of this work is supposed to become the first step in modelling an easy-to-use solution capable of evaluating the level of fatigue in employees who work using their fine motor skills such as dentists, surgeons, welders, etc.

The following structure is used for the thesis. In Chapter 2, the influence of fatigue on motor-cognitive functions is presented and spiral drawing test is explained to be an appropriate method for screening of motor-cognitive interference. Formally, the problem is presented in Chapter 3 and the next chapter provides literature overview. Tools, experimental settings and research methodology are discussed in Chapter 5. Chapter 6 presents and interprets the achieved results. It is followed by a short discussion of the results and concluding remarks are drawn in the last chapter.

2 Background

Perception of fatigue is subjective. There is no exact definition of fatigue because it is overlapped between symptoms of various illnesses and can be also noticed in a healthy person. Therefore, fatigability is individual and the period of recovering from fatigue in healthy individuals is different as well. Fatigue is best defined as the difficulty in initiating or sustaining voluntary activities [5]. *It occurs due to the impairment of one or several physiological processes, which enable the contractile proteins to generate force* [6].

Before fatigue modelling, it is necessary to understand the essence of the phenomenon. Two kinds of fatigue should be considered – mental (cognitive) fatigue and physical (muscle) fatigue. As an example, the case of mental fatigue can be manifested in the difficulty to concentrate on performing a task, e.g., reading a document may take twice as much time and even cause drowsiness. Physical or muscle fatigue, best explained as inability to continue running after a hard training, may include muscle weakness. Both of the fatigue types have their own symptoms, but both also have an influence on the motor unit. According to the type of exhaustion, the latter may be caused by different components of the central nervous system. In the scope of the current study the entire mechanism of fatigue is not going to be explained. It suffices to know that motor unit is the functional unit of movement. When one performs some kind of action, fibres of one's muscles are innervated by a motor neuron which transmits impulses from the brain that are triggered by those regions. Violation of the conductivity of those impulses from the brain to the muscle fibre can result in fatigue. Based on a research article about the effect of physical exhaustion on cognitive functioning, the impact of physical fatigue on cognitive functions can be assumed [7, 8]. Moreover, mental fatigue can also dramatically influence those functions by decreasing mental energy [9].

With the previous explanation, understanding of the influence of fatigue on motor-cognitive functions has been archived. The following part of this section will focus on what can be used for screening of motor-cognitive interference and be appropriate for differentiating between fatigued and non-fatigued individuals. Applicability of the spiral drawing test to fatigue modelling is in the focus of the present research. Spiral drawing test was chosen due to its popularity among practitioners and due to the fact that it was among the first tests to have been digitised. The latest studies provide different opportunities to use commonly used features to describe and interpret achieved results.

Drawing of the Archimedean spiral, the so-called spirography, is commonly used in neurological diagnostic tests to quantify motor activity. Spiral analysis is a clinically validated method that gives objective evaluation for such disorders as Parkinson disease or tremor disorder. Therefore, spiral analysis can be used to study the details of normal motor control. In its classical (non-computerised) version, the spiral drawing test is performed by means of pen and paper. Tested individual is demonstrated the etalon of spiral drawn on the paper. To conduct the test, a testee is asked to follow the contour of the spiral with the pen. Different ways of conducting the test are described in literature. Contour following may be done clockwise or counter clockwise. Instead of following the contour, a testee may be asked to draw the spiral in the white space limited by the contour, so that the pen would not touch the etalon drawing. These variations depend on the tested case and on the practitioner. Assessment of a testee is done by the practitioner on the basis of visual observation of the testing process and its results. In the digitised (computerised) form of the test, paper is replaced by the screen of the tablet computer, and pencil is replaced by the stylus pen. Obviously, the digitised version of the test allows to observe and record more parameters more precisely, including those invisible for the naked eye, like pressure on the pen, accelerations and velocities at each point, etc. Some recent contributions suggest up to 90 parameters to have been recorded and analysed [10].

3 Problem statement

Fatigue has been acknowledged as a fundamental factor to consider in the efficient management of human resource and also as a major cause of many incidents. Fatigue is a reason for people's unpredicted behaviour and low productivity. Since the perception of fatigue is subjective, it is mostly measured using different subjective methods as questionnaires or sleeping diaries, but objective methods also exist. An overview of available assessment tools for the fatigue measurement are discussed in the study of fatigue in multiple sclerosis [11]. However, existing measuring methods are too complex and cannot be used for the evaluation of fatigue in an everyday working environment. Ideally, every employee during the working day could determine his or her less productive periods and even dangers that are caused by increased fatigue. Ability to recognize fatigue may improve working environment and make it safer.

Recent advances in information technology have enabled advances in medicine. Faster and comprehensive data collection allowed clinicians to conduct cause-and-effect analysis and also contribute to early diagnosing of several disorders. For example, nowadays, digitised such handwriting assessments as spiral drawing test are used to timely detect such neurological disorders as Parkinson disease [5]. Handwriting definitely suggests cognitive impairment and may be relevant to fatigue evaluation.

In the scope of "Research-Based Software Development Project: Startup" (ITX8549) course, fatigue recognition tool infrastructure has been developed for Android mobile users that made possible the collection of training data and its storage. Moreover, test group data have been analysed and prediction models have been composed. However, the result of prediction was unexpectedly low. Generally, the initial version of the tool has its weaknesses and needs to be analysed and improved, so that fatigue can be detected.
Finally, it was decided to continue the research of the current problem under controlled environment, where the following main conditions meant for the purity of the experiment were set: tests were conducted under supervision and with the use of the same device. Formally, the problem is defined as follows: to construct a method to distinguish between

the spiral drawings done by individuals experiencing fatigue from those done by non-fatigued individuals. This leads to a number of subproblems to be solved:

- data acquisition: motion capture during spiral drawing test
- feature extraction and selection
- classifier selection training

4 Related work

The state of fatigue has been studied separately and also in the context of numerous chronic illnesses during the past century [11, 5, 8]. Still, the topic of fatigue remains acute and poorly understood because of its subjectivity. Nevertheless, there have been attempts to detect physical and mental fatigue using objective methods. For example, Volkswagen worked out a monitoring system for drivers which follows their blink rate, nodding and eye pupil position. If any symptoms of fatigue have occurred, the system will alert a driver with alarms [12]. In case of physical fatigue, electromyography (EMG) is widely used and established as a relevant technique of assessing muscle function through placement of electrodes on the skin [13]. Since existing practices are expensive, complicated and some of them can only be applied in clinical research, there is a need to create a separate easy-to-use application, especially for those areas where human fatigue may cause great damage.

Fatigue has an indisputable impact on human cognitive and/or motor function [7]. Volkswagen (VW) researchers acknowledged human error to be a casual factor in many road accidents [12]. One of the components of human error is general physical and mental fatigue which affects cognitive processes. In other words, when drivers are tired, their cognitive function is so much impaired that it does not enable them to react to a situation and this results in an accident.

What if fatigue affects the same functions of motion planning and execution as neurological disease? There is a huge amount of literature on the topic of fatigue, describing numerous experiments and explaining the nature of muscle fatigue. Moreover, mental fatigue has been investigated as a subjective component using various questionnaires and scales. Different types of fatigue and its causes received an extensive treatment [14]. Nevertheless, the possibility of fine motor tests to be used for fatigue recognition and modelling received much less attention [15]. In the [16] review of physical and cognitive consequences of fatigue based on several articles referred to in this thesis, authors propose that fatigue appears to influence cognitive functions which interfere with executive function of motor performance and should be further investigated.

The common approach to evaluating motor function is to apply a neurological examination such as fine motor tests. Spiral drawing test or spirography is widely used by clinicians for the study of upper limb motor dysfunctions in patients with essential tremor, Parkinson’s disease, and related disorders [17]. Spirography has its primary advantage of being appropriate for the implementation on a wide variety of devices [18], including relatively small ones. In the study [19], feasibility of spirography resulted in 86% of classification for assessment of motor function in Parkinson’s disease.

Various studies [17, 20, 21] focused on digitalisation of the handwriting process to quantify the kinematic parameters that can be applied in further analysis. In the research [22], a set of kinematic parameters were offered to measure quantity and smoothness of human limb motions. Subsequently this method was presented as an alternative approach to distinguishing movements in patients with Parkinson’s disease [3]. Reliability of using kinematic features in the case of fatigue has been addressed in several studies. The research [23] of the effect of mental fatigue on speed–accuracy trade-off shows increase of movement duration due to mental fatigue in case of whatever task difficulty.

5 Methodology

5.1 Tools

In the course of this work, different techniques were used to achieve the main goals. Data acquisition process was constructed using an application that was preinstalled on iPad Pro 9.7-inch tablet, equipped with native stylus (Apple Pen). This application has been developed by the members of the working group and allows to collect the data describing position of the pen tip, orientation of the pen and its pressure on the screen 200 times per second. Based on this data, different parameters describing movements of the pen tip are computed.

The collected data was stored in Amazon S3 and made available to pull it in JSON format for further analysis. Analysis of the extracted features is made by using Python 3 programming language in PyCharm IDE development tool. The choice of the language is based on its wide range of libraries for data manipulation, statistical and visual analysis. It is necessary to mention the most prominent libraries that are used during the research work. Plotting functionality is ensured by Matplotlib and statistical calculations by SciPy and Numpy. To solve classification problem, scikit-learn machine learning library is used.

5.2 Design of experiment

In the scope of this work, both mental and physical states of fatigue are researched in healthy subjects. Altered handwriting suggests cognitive impairment and may be relevant to fatigue evaluation. There are several fine motor tests to do it. The initial realisation asks a subject to overdraw a spiral clockwise using a background pattern, as it is illustrated in Figure 1.

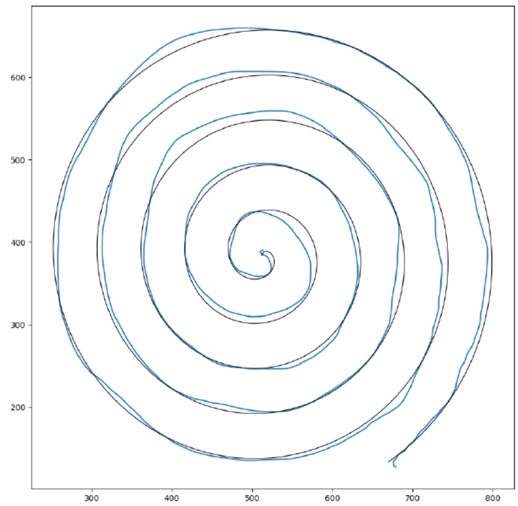

Figure 1. Etalon spiral and spiral drawn by a testee

Three kinds of experiments are going to be conducted – the first is supposed to measure mental fatigue (Experiment A), the second and the third ones – physical fatigue (Experiments B1 and B2). The basic idea of an experiment is to let a subject to draw a handwriting sample.

A. A group of subjects is measured with a tablet in the office during normal working day for three times, or every 2 hours. Working in the office is a pretty much exhausting process and supposed to give a result in mental fatigue research. Since fatigue is a subjective perception of tiredness, a Visual Analogue Scale (VAS) [24] depicted in Figure 2, has been considered as a simple subjective fatigue assessment method to estimate cognitive fatigue changes during the working day. The VAS consists of a straight line with the endpoints defining “Energetic, no fatigue” and “Worst possible fatigue”. Participants were asked before each spiral drawing test to point their current felt of fatigue on the line between two endpoints.

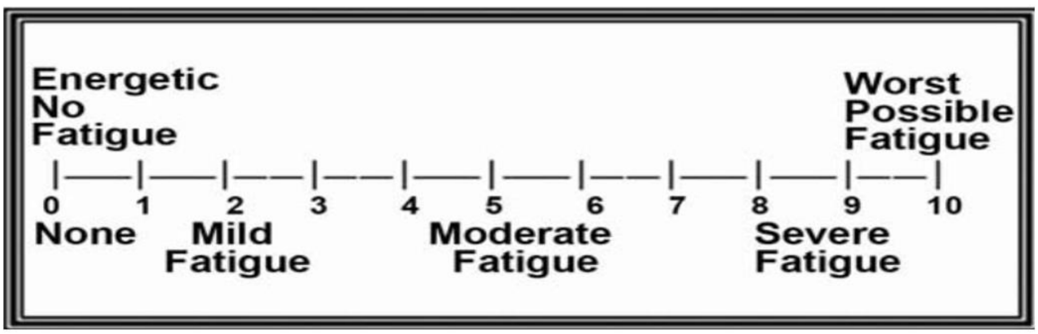

Figure 2. Visual Analog Test [25]

Before lunchtime, at 12 o'clock in the daytime, a subject will be tested for the first time. Lack of energy can be restored by having a rest or even during food consumption. So, a little walk outside of office and lunchtime in a nice company makes the working day in the office possible. This hypothesis will be checked after 2 hours and it is expected to see that the subject will become less tired or at least remain on the same level of tiredness. The last attempt at measuring fatigue will be performed before the end of working day.

B. First, the researcher is more focused on hand muscle fatigue and asks a subject to exercise with a dumbbell until the subject feels real inability in his or her muscles to continue the exercise. Two samples are derived from this test – before and after the exercise. The second test measures physical exhaustion during a difficult rock-climbing training. Three samples will be collected before the physical activity and three samples after it.

Additional criteria for experiments as listed below are found to be useful for the purity of the experiment:

- The age of subjects in the test group is in the range from 25 to 50.
- A healthy individual without diagnosed cognitive disorders is supposed to be tested.
- A drawing must consist of a continuous line or polyline.
- Task execution time must be less than one minute for each trial.
- A subject performs the test with his or her working hand.
- Subjects do not experience a severe headache at the moment of experiment.

5.3 Feature extraction

Drawing samples can be fed into a classifier itself for building a fatigue recognition model only in the case of deep learning, but this result is difficult to interpret. For the cases of medical problems, it is important to give clinicians understandable result. This calls for the use of decision tree and logistic regression classifiers which can handle as an input only numerical variable in case of scikit-learn Python library. There is a common practice to extract some specific attributes or so-called features from those samples using mathematical principles which can explain the difference between fatigued and non-fatigued individuals.

In the present research, there is no knowledge what features are meaningful. However, two kinds of features are supposed to be calculated – motion mass and similarity–based ones. Firstly, principles of the calculation of those features will be explained, while initial attributes are presented. The raw data of each handwriting sample derived from a tablet consists of five values illustrated in Figure 3: pen tip positions x(t) and y(t), pen tip pressure p(t) and pen orientation angle altitude ($0^{o} – 90^{o}$) and azimuth ($0^{o} – 359^{o}$).

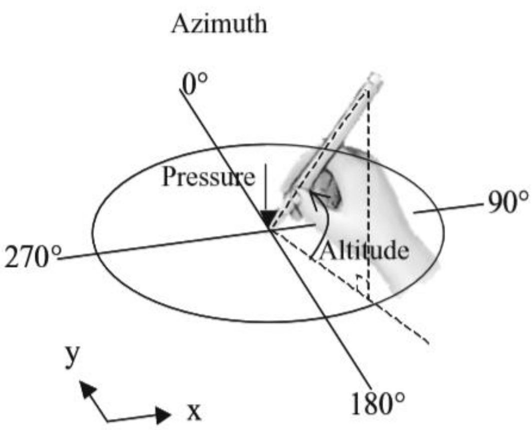

Figure 3. The raw data of each handwriting sample derived from a tablet [26].

For every subject, data is stored as a data matrix demonstrated in Figure 4 and, therefore, used for feature calculation. To avoid invalid data in this research, data pre-processing process has been performed and visually incomplete spiral samples have been excluded.

	a	l	p	t	x	y
0	-0.042552	0.718501	0.333333	5.402090e+08	670.7344	136.3477
1	-0.042552	0.718501	0.333333	5.402090e+08	670.7344	136.3477
...	...	...	...	...	...	...
8252	-0.049632	1.008373	1.332682	5.402090e+08	511.0938	383.2383
8253	-0.049632	1.008373	0.000000	5.402090e+08	511.6875	382.3711

Figure 4. Example of the initial data

5.3.1 Motion Mass features

In the present research, *Motion Mass* (MM) parameters, describing amount and smoothness of motion, were adopted to extract features for further analysis to distinguish between tired and non-tired subjects. Similar approach was previously applied in the analysis of human limb motions and this approach was acknowledged to be sufficient for differentiating between movements of patients with Parkinson's disease and healthy controls [3]. Basically, it is assumed that fatigue is linked to motion functions of execution and planning which means that MM approach could be also appropriate for the analysis of spiral drawings and give us the set of significant features. Because there is no information about relevant parameters for the stated problem, different strategies for spiral splitting are going to be tried and motion mass parameters will be calculated separately for retrieved parts. Actual splitting strategy will be decided upon during the statistical analysis of MMs.

Spiral drawing and each retrieved segment are considered as one single stroke S that is presented with collections of many points that are ordered in time. The way passed by stylus from the starting point of stroke to the final point of the spiral or its segment is described as follows:

$$S = \{S_1, \dots, S_{n-1}\} \tag{1}$$

where the path between two points S_i or line segment and n is the number of registered points.

The notion of MM parameters is denoted for the line segment as follows:

$$M_s = \{T_m, V_m, A_m, J_m, t\} \tag{2}$$

where T_m stands for Trajectory Mass, V_m stands for Velocity Mass, A_m stands for Acceleration Mass, J_m is Jerk Mass and t is time.

Trajectory Mass parameter is excluded because the sum of Euclidean distances of each line segment is calculated separately and it equals to the sum of trajectory masses of each line segment.

MM parameters of each registered stylus movement are calculated as shown below:

$$t_i = t_{end} - t_{start} \tag{3}$$

$$T_{m_i} = \sum_{i=1}^{n} E_i \tag{4}$$

$$V_{m_i} = \sum_{i=1}^{n} \frac{E_i}{t_i} \tag{5}$$

$$A_{m_i} = \sum_{i=1}^{n} \frac{V_{m_i}}{t_i} \tag{6}$$

$$J_{m_i} = \sum_{i=1}^{n} \frac{A_{m_i}}{t_i} \tag{7}$$

where E is Euclidean distance that is defined as follows:

$$E_i = \sqrt{(x_i - x_{i-1})^2 + \left(y_j - y_{j-1}\right)^2} \tag{8}$$

Finally, the set of sums of MM parameters is calculated for each stroke.

5.3.2 Similarity -based features

In addition to the abovementioned feature extraction approach, a commonly used *Dynamic Time Warping distance* (DTW) technique is applied for assessing the similarity between two time series. In other words, there are two hand-drawn spirals that were produced with different speed and can have different length, but still have a similar pattern. The basic principle of DTW algorithm is recursive search for minimal distances between two trajectories. According to standard DTW algorithm time-complexity $O(N^2)$, open-source Python library 'FastDTW' with linear time and memory complexity for calculation of the following features was chosen: dtw_time, dtw_position, dtw_speed, dtw_acceleration, dtw_distance_div_acceleration, dtw_acceleration_divided_by_distance.

5.4 Statistical evaluation

After extracting features statistical evaluation is going to be conducted to figure out the set of significant features. Such statistical concept as hypothesis testing is widely used and applied for every extracted feature to check *null hypothesis*. For the current problem, that selected feature is not indicative and does not show any association between test group samples. In other words, the conducted experiment does not provide a solution for the current problem. Therefore, tired and non-tired conditions of subjects are not differentiated.

To be more precise, two-sided p-value approach is applied and this value is calculated using the open-source python library 'SciPy'. To observe the means of two independent samples from different populations or test groups, a built-in function *ttest_ind(a, b,* *axis=0, equal_var=False, nan_policy='propagate')* is used. For this *Two-Sample t-test for Equal Means*, 5% threshold for *p-value* is applied. If *p-value* is smaller than a selected threshold, then the *null hypothesis* of equal means for an exact feature will be rejected. In other words, the feature is indicative of fatigue.

A statistical hypothesis test pointed out those features which mean values of non-tired and tired subjects are differentiated. Therefore, to better evaluate feature relevance to one of the class labels and select only significant ones for the classification process, the Fisher scoring method is going to be applied, which allows to measure the discriminatory power of the feature. The larger this number, the more evidence that this feature is sensitive to the classification algorithm.

For calculating the Fisher score *(F)* of a feature, the ration of the interclass separation to intraclass separation is defined as follows:

$$F = \frac{\sum_{j=1}^{k} p_j(\mu_j - \mu)^2}{\sum_{j=1}^{k} p_j {\sigma_j}^2} \tag{9}$$

where μ_j, σ_j and p_j are the mean, standard deviation and fraction of data points belonging to class j for the feature being evaluated and μ is its global mean [27].

5.5 Classification

This thesis is focusing on dealing with classification problem that is a part of a modern trend in informational technology known as data mining. In short, data mining is the practice of discovering patterns and trends in a large set of data.

Talking about classification, the basic idea is to accurately predict or identify to which category or class a new observation belongs, based on the knowledge about group members. For the current problem, it is needed to assign a subject's state of fatigue into "Tired" and "Non-Tired" classes. To solve classification problem, techniques described in the book "Data Classification: Algorithms and Applications" by Charu C. Aggarwal are going to be used. Basically, all the process consists of two phases – the training phase and the model validation phase. In the training phase a randomly chosen subset of data is used to create the model, and after that the rest of the data is used to validate created model.

In classifiers building it is expected to reach high accuracy of prediction and, if possible, interpretable result. Basically, two linear methods are going to be tested – logistic regression (LR) and Support Vector Machine (SVC linear), and two nonlinear methods – Decision Tree (DT), and K Nearest Neighbours (*k*NN). First, it needs to be observed if the problem is linearly separable and can be easily and faster solved by linear methods. LR is selected because it is widely used in medical research and works well in case of a problem with many features. Linear SVC is used in handwriting recognition application. Nonlinear methods are usually applied when linear classification is not giving expected accuracy. DT method can find interactions between features and is good for easy interpretation of the result. *k*NN is reasonable to try because there is much noise in the dataset, and similarity-based learning may improve predictive accuracy.

To assess predictive performance of models, the following techniques are going to be used:

- cross-validation
- confusion matrix

Cross-validation approach makes it possible to evaluate whether features that separate the training set well have been identified successfully and can be further applied for "real-world" data. Otherwise, this knowledge can be used for tuning of the model or proving that this learning method is not productive. Furthermore, to evaluate quality of the model, confusion matrix is going to be composed.

6 Analysis

In the previous chapter, methods have been described which should be appropriate to achieve the goals of this master thesis. This chapter will discuss the results of implementation of each method in detail and give the initial estimation for methods chosen and the experiment form for fatigue recognition problem. The chapter is organised in the way of implementation workflow. First, the set of features were acquired from spiral drawing test and formed into separate datasets to analyse the problem from different angles. To create models for fatigue recognition, the sets of significant and highly correlated features were selected using statistical hypothesis test and Fisher's score, respectively. At the end of this chapter, the results of machine learning will be presented.

6.1 Feature extraction

The basic idea of this phase is to extract as many features as possible from spiral samples. A wide variety of features is necessary because there is no knowledge about relevant features for the current problem. Overall 260 features of two groups were extracted from the spiral sample. In the following part of this section, data separation process of experiments will be explained, and each group of features is going to be reviewed in detail. During data acquisition process, collected data were divided into five test groups depending on an experiment form. Moreover, features that belong to each group were calculated and divided into separate datasets. Each row of the dataset had information whether its features belong to tired or non-tired subjects. The group forming principle is provided in Table 1:

Table 1. The group forming principle

	Number of tired (T) and non-tired (NT) subjects		Description
	T	**NT**	
Group I	49	39	Group contains spiral samples of all three experiments, A, B1, B2. In case of B2, attempt 1 result belongs to NT and 2,3 to T
Group II	18	12	Group contains spiral samples of B2 experiment. Attempt 1 result corresponds to NT and 2,3 to T
Group III	10	12	Group contains spiral samples of B2 experiment. Attempt 1 result corresponds to NT and 3 to T
Group IV	17	17	Group contains spiral samples of B1 experiment. Some results were invalid and removed from the research
Group V	10	20	Group contains spiral samples of B2 experiment. Attempt 1,2 results correspond to NT and 3 to T

This approach was found to be useful in the case of two types of experiments. One is based on physical exhaustion and another one on mental exhaustion. There is no knowledge whether the same motor units suffer from a different kind of exhaustion.

Initially each spiral sample data was stored in JSON format into a separate file and was extracted from this file using built-in data parsing method from Pandas library. Each sample or spiral drawing consisted of one single stroke that was presented with a collection of many points. Generally, each point had the following attributes – x-axis and y-axis position, pressure, azimuth, altitude and time. Therefore, to check the correctness of drawing and exclude faulty samples, drawings were preliminarily visualised using x and y coordinates.

After extracting attributes of the point, mean values of pressure, azimuth and altitude are calculated. Next, Motion Mass group features are computed. The spiral does not have any corners or intersection causing the testee to slow drawing down or stop. Nevertheless, it become apparent that behaviour of the drawing process differs very much between different sectors of the spiral. This gave an idea to split the spiral into sectors and generate full set of the features for each sector. Four different splitting depicted in Figure 5:

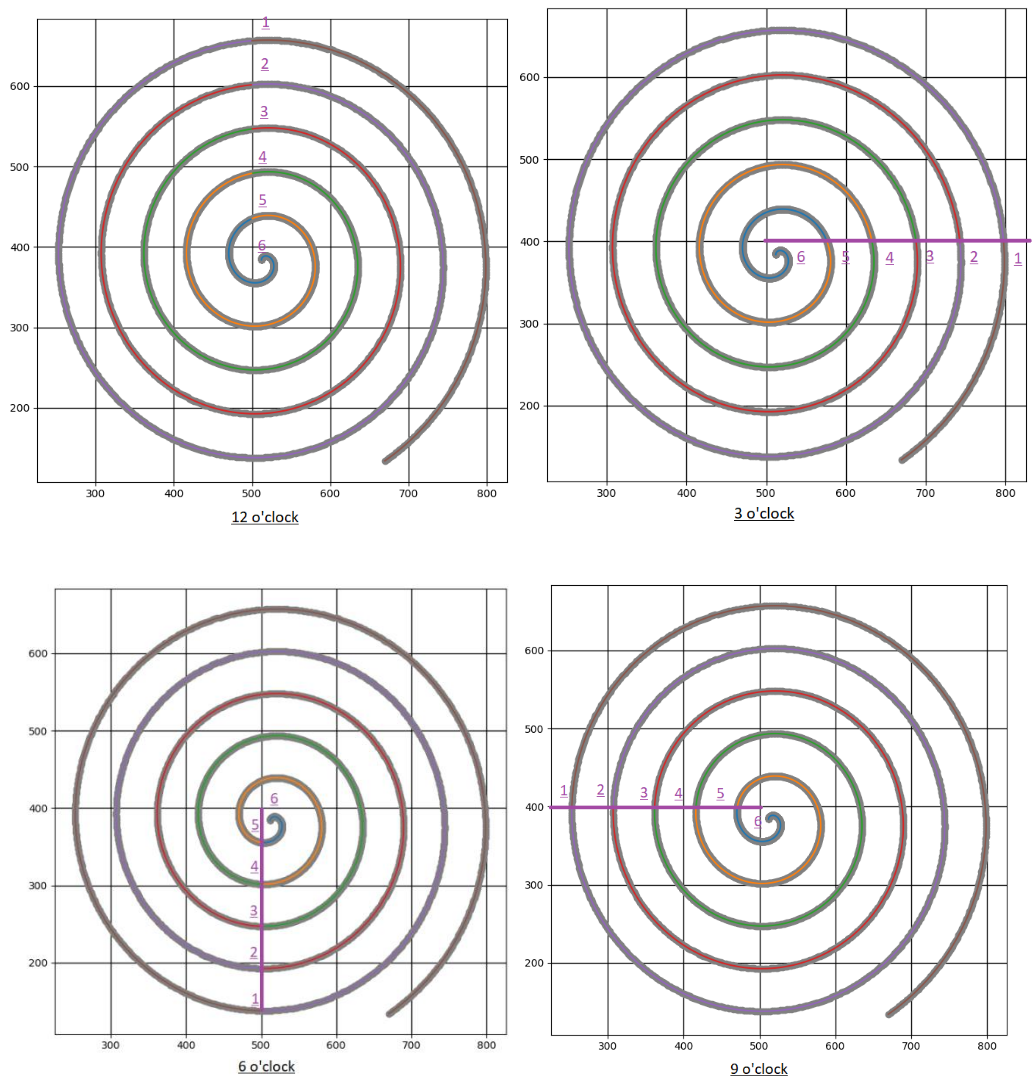

Figure 5. Possible ways of splitting a spiral

Each splitting is defined by the orientation of splitting line, three o'clock, six o'clock nine o'clock and 12 o'clock and denoted correspondingly as *s3*, *s6*, *s9* and *s12*. Each splitting produces six segments referred as *segment*. All the MM parameters were computed for each segment of each splitting leading relatively large feature set for initial selection. The following notation system is adopted. The feature name consists of three parts: splitting name, segment number and finally parameter name. For example, s3_segment4_jerks refer to the accumulated jerks observed over fourth segment of the spiral in splitting defined by three o'clock splitting line.

Currently there are 25 strokes – the full spiral stroke plus 1–6 spiral segments multiplied by 4 for each separation way as it is shown in Figure 5. For each of those strokes the following MM features were calculated:

- Euclidean distance (*E)*
- Velocity mass (V_m)
- Acceleration mass (A_m)
- Jerk mass (J_m)
- Time deltas (*t*)
- Ratio of Euclidean distance to acceleration mass (*dist_div_am*)
- Ratio of acceleration mass to Euclidean distance (*am_div_dist*)

In the current issue, there is no knowledge about relevant features. In this situation, any intuitively used techniques are accepted if they are applied equally to each sample. For example, there were attempts to use several strategies to find out significant features. Firstly, there were attempts to change spiral intersection points with a line. And secondly, features' values were divided by: a) Euclidean distance, b) logarithm of a trajectory mass, c) squared trajectory mass values. To evaluate the significance of features in the case of current issue, statistical analysis was applied to them, but its result will be discussed in the following section. Summing up the above, the so-called "clock" separation ways and dividing features' values by distance were included into the final analysis.

The last group of features were added into the final assessed set are similarity-based features: dtw_time, dtw_position, dtw_speed, dtw_acceleration, dtw_dist_div_acc, dtw_am_div_dist.

After extracting all the features for all five test groups, they are stored in separate *CSV* file. Moreover, each row carries knowledge whether a set of features belongs to tired or non-tired subjects. In the following sections gathered features would be used for the further analysis and machine learning model training.

6.2 Statistical analysis

Comparison of the averages of the experimental features for all datasets was separately performed using Student' t-test. A p < 0,05 was considered statistically significant.

Instead of explaining the results for all datasets, only results of Group 2 and 4 as the most notable representatives of mental and physical fatigue are explicated in the current thesis. The result of statistical analysis in the case of office workers describing mental fatigue (Group 2) was impressive and almost half of features were highlighted, which probably can be used as a relevant predictor for creating a model in machine learning. Table 12 in Appendix 2 provides an example of the most significant features filtered by the $p < 0.009$. Moreover, average scores of two features located at the top of the table were significantly different between two groups ($p < 0.003$).

According to the results presented in Table 12, the most efficient way of spiral separation is on 3 o'clock, and its first and fourth segments pointed out differences between the group of tired and non-tired subjects. The result of statistical analysis showed one more specific result. Execution time of drawing spiral or its segment is supposed to differentiate tired and non-tired spirals. But more remarkably, not only time features can be used in fatigue modelling but also jerk, acceleration and Euclidian distance.

Statistical analysis of physically tired individuals after climbing training (Group 4) detected only one relevant feature – distance listed in Table 2. The distance of execution of the second segment of spiral, according to 12 o'clock separation way, were different for tired and non-tired subjects.

Table 2. Significant feature according to the result of Group 4

Features	**T-stat**	**P-value**
s12_part2_distances	-2.074354	0.046179

Unexpectedly, t-test analysis did not find the set of dtw features in case of both groups potentially significant, whereas the popularity of the *Dynamic Time Warping* approach in assessing similarity between two datasets has been increasing over the years.

6.3 Feature selection

The statistical hypothesis test pointed out those features which mean values of non-tired and tired subsets are distinguished. Therefore, to better evaluate feature relevance to one of the class labels and select only significant ones for the classification process, Fisher scoring method is going to be applied, which measures the discriminatory power of the feature, where *F* denotes Fisher's score. The larger this number, the more evidence that the feature is sensitive to the classification algorithm. The results of highly correlated features for mental fatigue dataset are demonstrated in Table 3.

Table 3. An example of highly correlated features with Fisher score larger than 0.4 for Group 2 (mental fatigue).

Features	**Fisher's score**
s3_segment3_time_deltas	0.800138
s3_segment3_time_deltas_ratio_to_length	0.797788
s6_segment2_time_deltas	0.779537
s6_segment2_time_deltas_ratio_to_length	0.777582
…	…
time_deltas	0.488000
s3_segment3_am_div_dist	0.487915
s3_segment4_jerks	0.483130
s6_segment2_am_div_dist	0.482549
s3_segment4_jerks_ratio_to_length	0.478012
s12_segment2_am_div_dist	0.476241
s9_segment2_am_div_dist	0.453036
s3_segment4_accelerations	0.438980

According to Fisher's score results, the execution time of spiral and its segments might be a good predictor. However, other features independent of time, such as jerk and acceleration masses, are also applicable for fatigue recognition modelling.

After the results of both test – *t-test* and Fisher scoring were calculated, it is necessary to decide upon the right number of features to be applied in the learning algorithm directly because a learning model tends to overfit in case of many parameters. In other words, a model can predict with high accuracy only on training data but will obviously fail with unseen instances.

Since Group 2 indicated 126 significant features but only 1/3 were signed as highly correlated ($F > 0.5$), only the most significant features ($p < 0.006$) were selected for training of models. At the same time, one of the goals of this thesis is to build the time-independent model. Accordingly, a demand to Fisher's score was degraded and accepted $F > 0.4$.

The results of Fisher's score for Group 4 are presented in Table 4 and demonstrate a low number:

Table 4. An example of Fisher's score for Group 4 (mental fatigue).

Features	**Fisher's score**
s12_segment1_time_deltas_ratio_to_length	0.028320
s12_segment1_time_deltas	0.027588
s9_segment1_time_deltas	0.022652
s9_segment1_time_deltas_ratio_to_length	0.021249
…	…
s9_segment1_dist_div_am	0.141105
s12_segment2_distances	0.138539
s6_segment1_ dist_div_am	0.136719
s3_segment1_ dist_div_am	0.132978

6.4 Model selection

Four machine learning methods were investigated: *k* nearest neighbours (*k*-NN), support vector machines (SVM), logistic regression (LR) and decision tree (DT). To access the predictive performance of chosen methods, cross-validation technique (k-fold=3) and confusion matrix were used. Overall, two types of models were built for each dataset: a) temporal, b) non-temporal.

6.4.1 Mental fatigue modelling

Based on *t-test* and Fisher's score results, purely temporal features were selected as predictors and, therefore, used for training the first set of predictive models. Temporal features are drawing time for the third segment in three o'clock splitting, ratio of time to the Euclidian distance of the drawn line and the same parameters for the second segment of six o'clock splitting. Separation of the individuals affected by fatigue and control group described by the first three features is depicted in Figure 6, where red dots represent individuals affected by mental fatigue and blue dots belong to the control group:

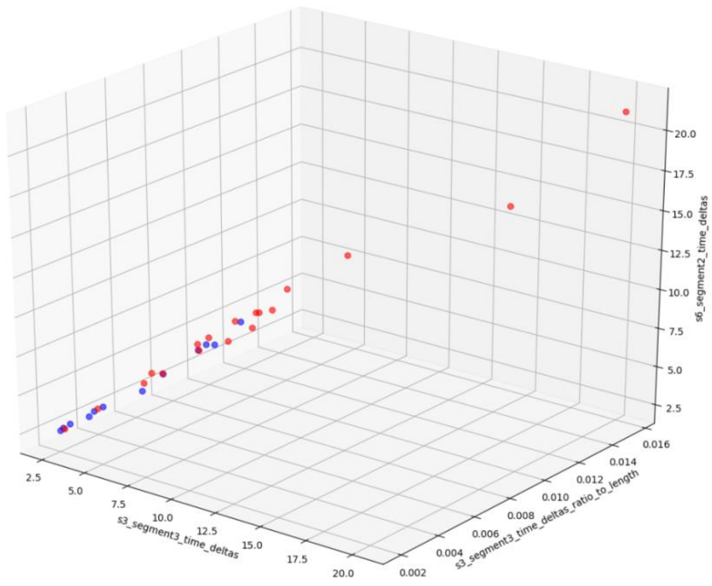

Figure 6. Separation of two groups described by temporal features only (mental fatigue).

For the second set of models, only non-temporal features were selected. These features are ratio of the acceleration mass to the Euclidean distance *am_div_dist* of the third segment in three o'clock splitting, jerk mass of the fourth segment in the same splitting, ratio of the acceleration mass to the Euclidean distance *am_div_dist* in the same splitting and ratio of jerk mass to the Euclidean distance. Separation of the individuals affected by mental fatigue and control group described by the first three features is depicted in Figure 7.

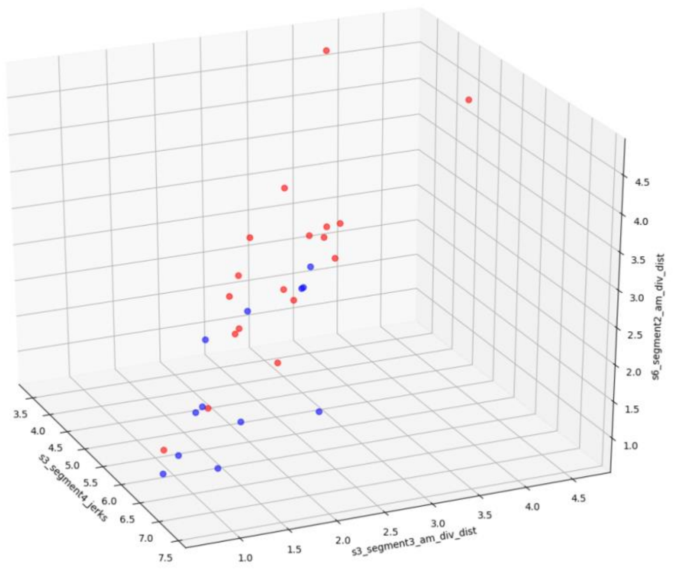

Figure 7. Separation of two groups described by kinematic features only (mental fatigue). Y and Z axis values scaled by 1e+12 and X – 1e+16.

For both sets it was attempted to train all four classifiers based on two, three and four predictors. The results of cross-validation for the trained classifiers are presented in Table 5. For all chosen methods, temporal models showed lower accuracy results comparing to non-temporal models. In spite of the lower Fisher's score, classifiers trained for the kinematic features in many cases have higher accuracy compared to those trained on the basis of temporal features only. There were attempts to tune models, which resulted in increased levels of accuracy, and final scores are presented in the last column of Table 5:

Table 5. Models training results for the case of mental fatigue.

Models	**s3_segment3_time**	**s3_segment3_time _ratio_to_length**	**s6_segment2_time**	**s6_ segment2_time _ratio_to_length**	**s3_ segment 3 _am_div_dist**	**s3_ segment 4_jerks**	**s6_ segment 2 _am_div_dist**	**s3_ segment4_jerks _ratio_to_length**	**Accuracy**	**Tuned model accuracy**
DT	✓	✓							0.754	0.786
DT	✓	✓	✓						0.754	0.786
DT	✓	✓	✓	✓					0.762	0.794
DT					✓	✓			0.762	0.841
DT					✓	✓	✓		0.762	0.841
DT					✓	✓	✓	✓	0.762	0.841
SVC	✓	✓							0.659	0.794
SVC	✓	✓	✓						0.659	0.794
SVC	✓	✓	✓	✓					0.659	0.794
SVC					✓	✓			0.754	0.897
SVC					✓	✓	✓		0.706	0.841
SVC					✓	✓	✓	✓	0.706	0.841
LR	✓	✓							0.659	0.794
LR	✓	✓	✓						0.659	0.794
LR	✓	✓	✓	✓					0.659	0.794
LR					✓	✓			0.754	0.897
LR					✓	✓	✓		0.754	0.897
LR					✓	✓	✓	✓	0.754	0.897
*k*NN	✓	✓							0.659	0.690
*k*NN	✓	✓	✓						0.603	0.690
*k*NN	✓	✓	✓	✓					0.603	0.643
*k*NN					✓	✓			0.762	0.897
*k*NN					✓	✓	✓		0.762	0.897
*k*NN					✓	✓	✓	✓	0.762	0.944

Confusion matrices for each best classifier in its class are presented in Table 6 – Table 9, where F denotes an individual affected by fatigue and C – a non-tired individual:

Table 6. Confusion matrix for the classifier for the most accurate kNN model (mental fatigue).

	Predicted (F)	**Predicted (C)**
Actual (F)	3	1
Actual (C)	2	4

Table 7. Confusion matrix for the classifier for the most accurate SVM model (mental fatigue).

	Predicted (F)	**Predicted (C)**
Actual (F)	3	1
Actual (C)	1	5

Table 8. Confusion matrix for the classifier for the most accurate LR model (mental fatigue).

	Predicted (F)	**Predicted (C)**
Actual (F)	3	1
Actual (C)	1	5

Table 9. Confusion matrix for the classifier for the most accurate DT model (mental fatigue).

	Predicted (F)	**Predicted (C)**
Actual (F)	3	1
Actual (C)	2	4

The following analysis of kinematic feature – jerk, and VAS score as a subjective assessment tool of fatigue is proposed to give some understanding for incorrectly predicted instances. In Figure 8 and Figure 9 are depicted, respectively, spiral samples of non-fatigued and fatigued individuals, where orange dots symbolise areas with high values of jerks, more than 2.5e+14, and green dots – areas with low values, less than 5e+9. In other words, formed heat map characterises smoothness of movement.

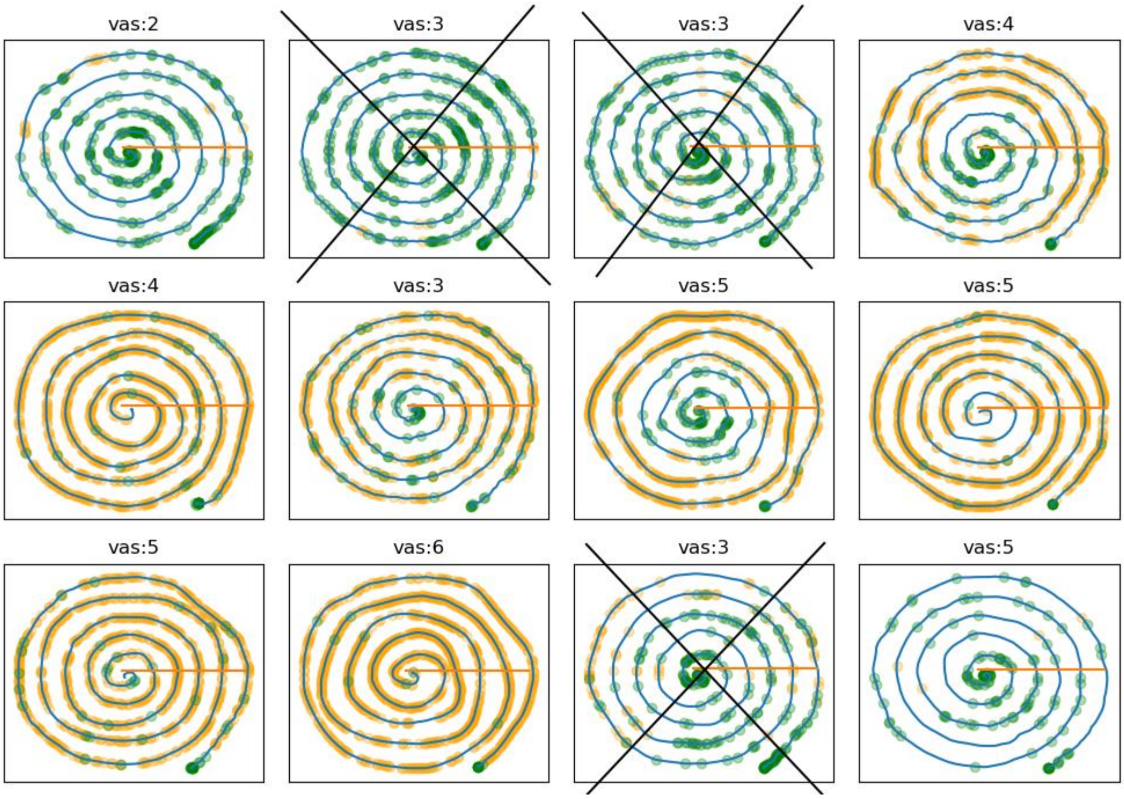

Figure 8. Spiral samples of non-fatigued individuals.

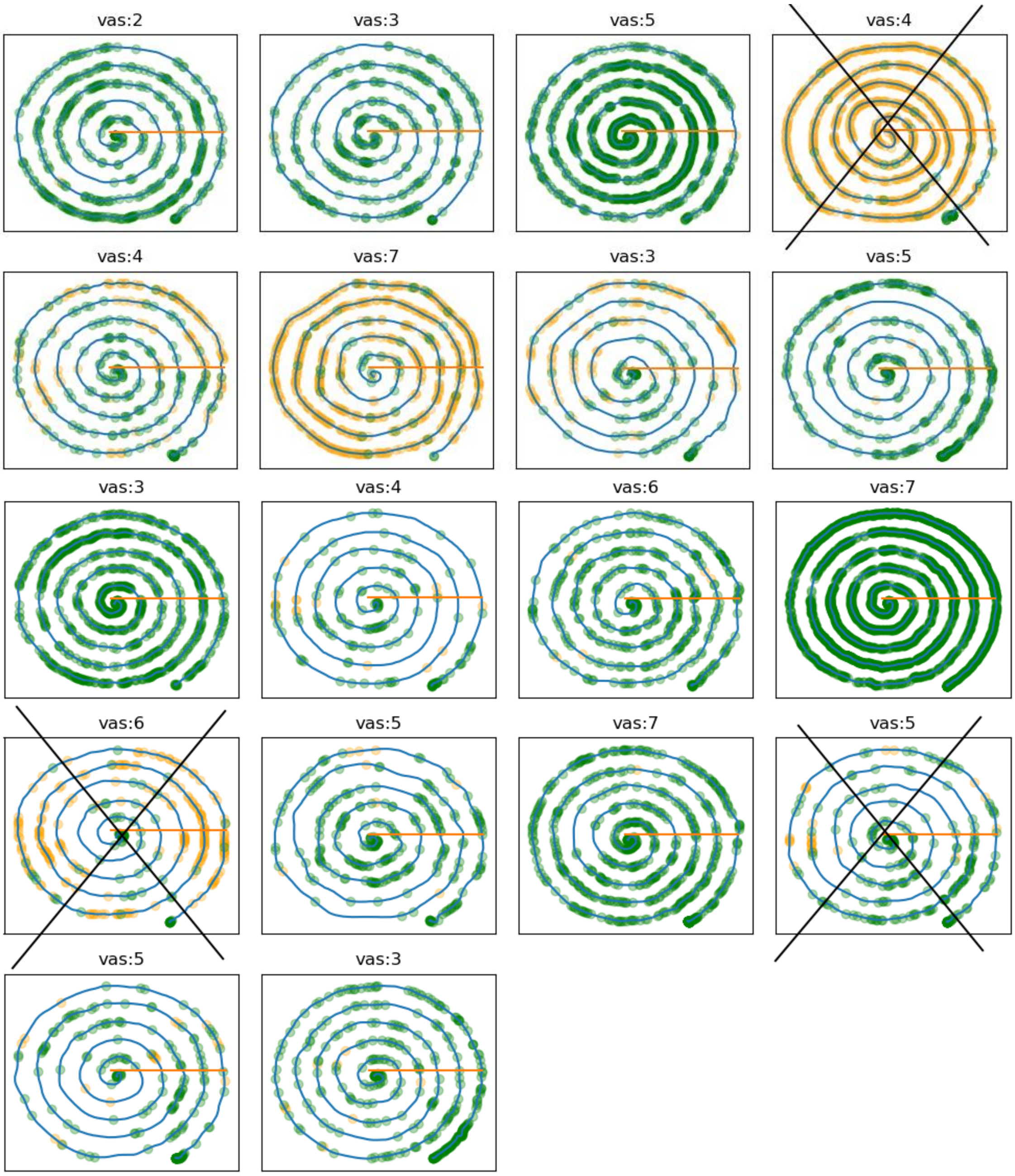

Figure 9. Spiral samples of fatigued individuals.

According to the results of training and validation, incorrectly predicted samples are crossed out. Subjective perception of fatigue showed relatively low difference between fatigued (average VAS score - 4) and non-fatigued (average VAS score - 4,7) individuals. Overall, it is supposed to have outliers in both sets, but more samples could give a defining statement.

To sum up training results for the case of mental fatigue, decision tree and *k*-NN models demonstrated higher levels of the accuracy. The common practice to understand classifier behaviour is using a decision boundary technique. More specifically, decision boundary illustrates how a classifier will decide whether observed point belongs to one or other class. For the case of mental fatigue, data spreading and decision boundaries for decision tree and *k*NN classifiers are demonstrated in Figure 10. Data spreading and decision boundaries for decision tree (right) and k-NN (left) classifiers., where black dots belong to control individual and white dots to fatigued individual. If any instance occurs in the shaded area, this means classifier will define it as fatigued. For example, some instance has jerk value for the fourth segment in three o'clock splitting 7e+16 and ratio of acceleration mass to Euclidean distance 2e+12. *k*NN and DT classifiers will label the instance as fatigued. Decision boundary of DT presents additional interest because it has its shaded area in non-fatigued area. According to personal communication with another student who has similar case, this needs to be further analysed.

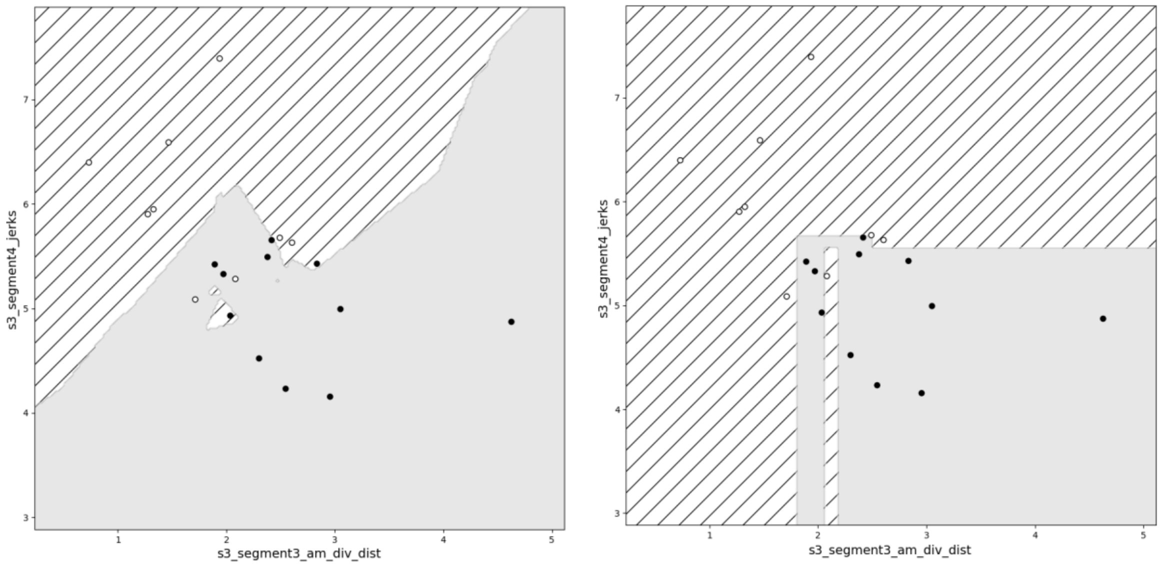

Figure 10. Data spreading and decision boundaries for decision tree (right) and k-NN (left) classifiers. X-axis values scaled by 1e+12 and Y-axis values scaled by 1e+16

As an example of why decision tree classifier is useful for the medical problems is that it can be visually interpreted. The graph of the decision tree classifier for the two best kinematic features is depicted in Figure 11, where full logic of decision in each step is clear.

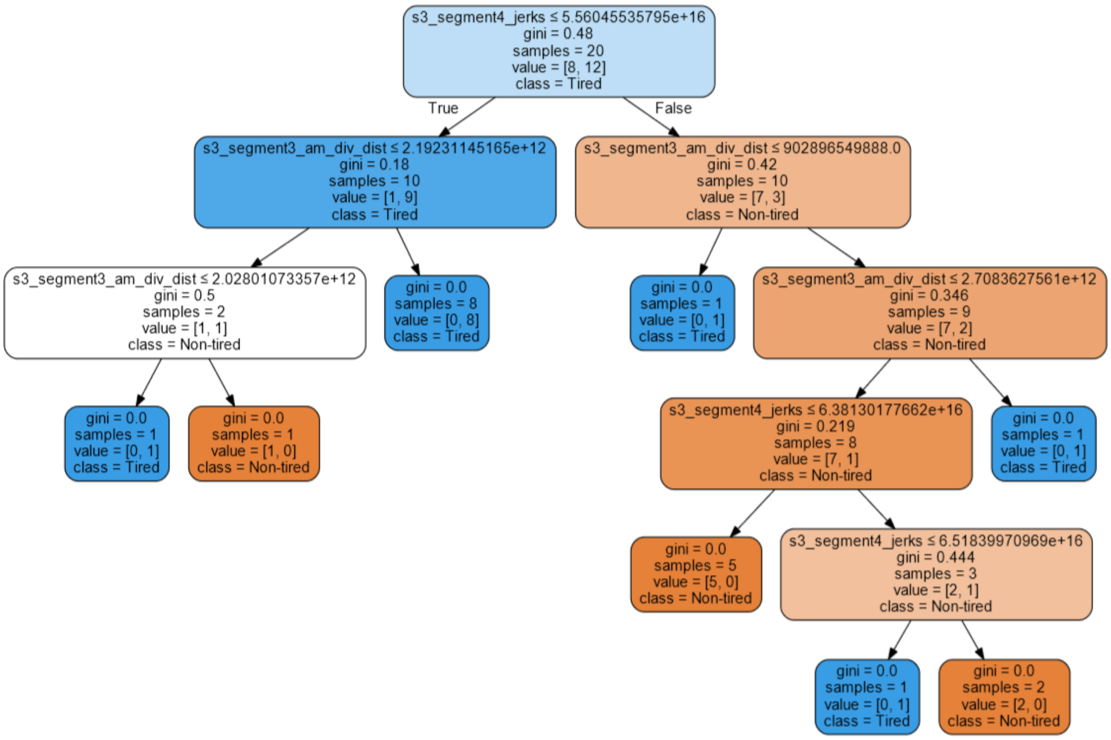

Figure 11. Example of the graph of decision tree with the best two kinematic features.

6.4.2 Physical fatigue modelling

Using the same principle of selecting predictive models for mental fatigue, the set of models for the case of physical (muscle) fatigue were constructed. The only difference is that *t-test* highlighted one feature as significant and, therefore, predictors were selected purely on the basis of the Fisher's score. Temporal features are time and ratio of time to the length of the drawn line for the first segment in 12 o'clock in the daytime splitting, and the same parameters for the first segment of nine o'clock splitting. Separation of the individuals affected by physical fatigue and control group described by the first three features is depicted in Figure 12, where red dots represent individuals affected by fatigue and blue dots belong to the control group:

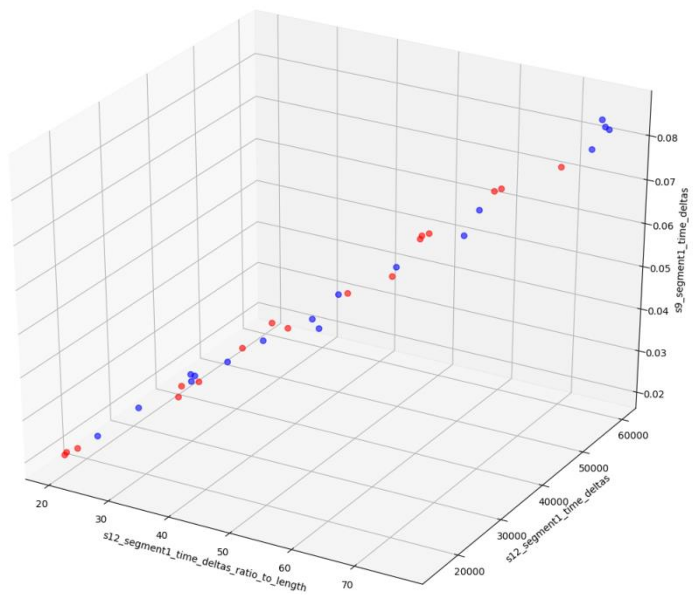

Figure 12. Separation of two groups described by temporal features only (physical fatigue).

For the second set of models, only non-temporal features were selected. These features are ratio of the Euclidean distance to the acceleration mass *dist_div_am* of the first segment in nine o'clock splitting, the same parameter for the first segment in six and three o'clock splitting and the length of the second segment in 12 o'clock in the daytime splitting. Similarly, separation of the individuals affected by physical fatigue and control group described by the first three non-temporal features is depicted in Figure 13:

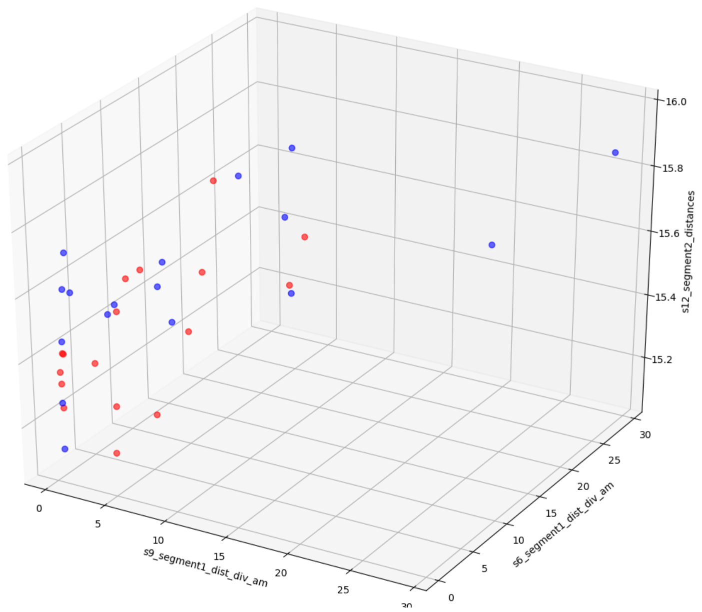

Figure 13. Separation of two groups described by kinematic features only (physical fatigue).

Results of cross-validation for the trained classifiers are presented in Table 10 and it demonstrates for almost all classifiers relatively low levels of accuracy close to random labelling. Therefore, only for the best classifier — logistic regression, confusion matrix is presented in Table 11:

Table 10. Models training results for the case of physical fatigue (Group 4).

Models	**s12_segment1_time_ratio_to_length**	**s12_segment1_time**	**s9_segment1_time**	**s9_segment1_time_ratio_to_length**	**s9_segment1_dist_div_am**	**s12_segment2_distances**	**s6_segment1_dist_div_am**	**s3_segment1_dist_div_am**	**Accuracy**
DT	✔	✔							0.439
DT	✔	✔	✔						0.506
DT	✔	✔	✔	✔					0.539
DT					✔	✔			0.622
DT					✔	✔	✔		0.567
DT					✔	✔	✔	✔	0.622
LR	✔	✔							0.500
LR	✔	✔	✔						0.561
LR	✔	✔	✔	✔					0.561
LR					✔	✔			0.656
LR					✔	✔	✔		0.656
LR					✔	✔	✔	✔	0.656
KNN	✔	✔							0.533
KNN	✔	✔	✔						0.494
KNN	✔	✔	✔	✔					0.533
KNN					✔	✔			0.556
KNN					✔	✔	✔		0.528
KNN					✔	✔	✔	✔	0.528
SVC	✔	✔							0.589
SVC	✔	✔	✔						0.589

SVC	✔	✔	✔	✔					0.589
SVC					✔	✔			0.556
SVC					✔	✔	✔		0.589
SVC					✔	✔	✔	✔	0.589

Table 11. Confusion matrix for the most accurate LR model (physical fatigue).

	Predicted (F)	**Predicted (C)**
Actual (F)	5	1
Actual (C)	2	4

Actual accuracies are not presenting any interest to conduct further analysis for selecting models for the case of physical fatigue.

7 Discussion

Since the issue in question has not been investigated before using fine motor test, the main problem was to find an appropriate and available approach to research it. Unquestionably, acquisition of the relative data took more time than expected, because there was no knowledge as to which experiments would give differentiating data for fatigued and control individuals. By means of trial suitable experiments were conducted but gathering of huge datasets was limited in time and resources. Difficulty in the collection of data arose from the errors in the application, which in turn led to the loss of some of the collected samples. Therefore, the number of samples is too modest for making indisputable conclusions. However, it still proves the need to research the current problem.

Analysis of the results showed that, among all the experiments, only the experiment with employees in the office was an appropriate method for investigating fatigue, more precisely – mental fatigue. Moreover, the result of machine learning among all trained classifiers gave the highest accuracy of 0.762. Based on the results of validation, the data collected in the confusion matrices confirmed the ability of models to predict with similar accuracy. Running classifiers with different parameters, the so-called tuning the models, gave improved accuracies. Furthermore, in case of *k*NN classifier, the result of 0.9 was achieved, which means that nine in ten spirals will be correctly predicted. However, validation of the models gave lower precisions that can be explained by the presence of outliers visible in the spreading diagrams and in the jerks heatmap of spirals. Overall, the result can be interpreted as positive because the working hypothesis of the present research is confirmed, which is that mental fatigue affects temporal and kinematic parameters describing spiral drawing test.

Another point to note is the importance of kinematic parameters. In some cases, lesser number of kinematic parameters provided the possibility to train more accurate classifiers. This clearly demonstrates that cognitive fatigue affects fine motor functions. Finally, among the kinematic parameters acceleration masses and jerk masses have higher

Fisher's score and were placed close to the roots of the decision trees. This is a clear indication of the fact that cognitive fatigue affects smoothness of fine motor motions.

For the case of physical fatigue, achieved by experiment with climbers, the result can be characterised as random because the result of training for all the classifiers did not exceed 0.65. Statistical evaluation of the averages did not show significant features, and in the case of Fisher's analysis scores were also low, which did not allow to find a suitable predictor for physical fatigue. Analysing the result, it can be assumed that the following factors could have a negative effect: small amount of data, an experiment was conducted in the evening after a working day and general fatigue by that time might have been already at its peak. It is worth noting that the second experiment of physical fatigue with dumbbells also failed at the stage of statistical analysis without selecting any of significant features. Since the approach to studying the problem of fatigue is new, one should not immediately chase the collection of a huge amount of data, but it is worth determining the method to solving the problem and setting demonstrative experiments. Undoubtedly, it is worthwhile to try other fine motor tests capable of evaluating the functions of motion planning and execution. Perhaps results of this research will encourage scientists to further explore the possibility of estimating fatigue with the help of fine motor tests.

8 Summary

The focus of the present master thesis is on a study of the ability of digitalised fine motor skills test to determine fatigue at the levels of planning and execution of limb motions. A commonly used spiral drawing test was applied in the current work for the screening of motor-cognitive interference due to a research influence of fatigue on those functions. Physical and mental fatigue were considered separately. The main goals were to research applicability of fine motor test for distinguishing individuals affected by fatigue and a control group, and to find a set of parameters differentiating those groups so that a fatigue recognition model could be built.

For the possibility of studying this problem, it was necessary to set up experiments in a controlled environment, which were able to catch the higher level of fatigue compared to the control group. Thus, mental fatigue was assessed within a normal working day in an office. For evaluation of physical fatigue, the first group of testees were asked to exercise with a dumbbell and the second group was tested before and after a hard rock-climbing training. To conduct statistical analysis so that it would be possible to evaluate an association between the test group samples fatigued and non-fatigued, a set of specific attributes, or so-called features, were extracted from spiral samples using mathematical principles. To calculate those features, two approaches were applied: Motion Mass and DTW. In total, five datasets were formed, but according to statistical analysis, further research was applied only for two datasets.

Generally, the results reported in the present work demonstrated that the spiral drawing test may be used to detect mental fatigue. First, it was demonstrated that the proposed experimental setting is sensitive enough to capture the difference between fine motor motions of two groups. Importance of the temporal features was achieved, and it was established that kinematic features describing smoothness of motions allow constructing classifiers of a higher accuracy. This is a clear demonstration of the fact that cognitive fatigue affects the smoothness of finger motions.

For the case of physical fatigue testing, there was not found any proof of the working hypothesis. However, it could mean that the experimental setting is inappropriate and a further testee should be rested to keep purity of the experiment.

The approach used in this work was not previously applied to fatigue modelling, but the results of the current research confirm the need for further analysis, which could lead to the creation of an easy-to-use solution. This solution provides a special interest to those areas of human machine interaction where fine motor skills are involved. Furthermore, it may improve the quality of a working environment, decrease the number of incidents, and exclude the subjective factor of fatigue.

References

[1] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia and F. Babiloni, “Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness,” 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0149763412001704?via%3Dihub. [Accessed 28 04 2018].

[2] J. F. Mathers and M. A. Grealy, “Motor control strategies and the effects of fatigue on golf putting performance,” 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3888943/. [Accessed 28 04 2018].

[3] S. Nõmm, A. Toomela, M. Vaske, D. Uvarov and P. Taba, “An Alternative Approach to Distinguish Movements of Parkinson Disease Patients,” 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S240589631632136X. [Accessed 28 04 2018].

[4] S. Nomm, A. Toomela and J. Borushko, “Alternative approach to model changes of human motor functions,” 2013. [Online]. Available: https://ieeexplore.ieee.org/document/6779841/. [Accessed 28 04 2018].

[5] A. B. P. Chaudhuri, “Fatigue in Neurologic Disorders,” 2014. [Online]. Available: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(04)15794-2/fulltext. [Accessed 28 04 2018].

[6] H. M. Abd-Elfattah, F. H. Abdelazeim and S. Elshennawy, “Physical and cognitive consequences of fatigue: A review,” 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4522584/. [Accessed 28 04 2018].

[7] F. YA1, F. A, V. H. A and R. M., “Effect of physical exhaustion on cognitive functioning,” 1997. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/9132722. [Accessed 28 04 2018].

[8] A. Abraham and V. E. Drory, “Fatigue in motor neuron disease,” 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960896612006220. [Accessed 28 04 2018].

[9] L. Caixeta, G. M. N. d. S. Júnior, V. d. M. Caixeta, C. H. R. Reimer and P. V. B. e. Azevedo, “Occupational health, cognitive disorders and occupational neuropsychology,” 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5619330/. [Accessed 28 04 2018].

[10] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal and M. Faundez-Zanuy, “Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease,” 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365716000063. [Accessed 28 04 2018].

[11] S. S. Ayache and M. A. Chalah, “Fatigue in multiple sclerosis - Insights into evaluation and management.,” 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0987705316303847. [Accessed 28 04 2018].

[12] N. Owen, H. King and M. Lamb, “Literature Review of Race Driver Fatigue Measurement in Endurance Motorsport,” 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877705815015179. [Accessed 28 04 2018].

[13] R. RR, D. C. S, S. A and M. C., “Methods for the assessment of peripheral muscle fatigue and its energy and metabolic determinants in COPD,” 2009. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20011849. [Accessed 28 04 2018].

[14] S. S. H. Nazari, A. Moradi and K. Rahmani, “A systematic review of the effect of various interventions on reducing fatigue and sleepiness while driving,” 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1008127517301530. [Accessed 28 04 2018].

[15] O. Rasouli, E. A. Fors, P. C. Borchgrevink, F. Öhberg and A.-K. Stensdotter, “Gross and fine motor function in fibromyalgia and chronic fatigue syndrome.,” 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5304994/. [Accessed 28 04 2018].

[16] H. M. Abd-Elfattah, F. H. Abdelazeim and S. Elshennawy, “Physical and cognitive consequences of fatigue: A review,” 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4522584/. [Accessed 28 04 2018].

[17] J. A. Sisti, B. Christophe, A. R. Seville, A. L. A. Garton, V. P. Gupta, A. J. Bandin, Q. Yu and S. L. Pullman, “Computerized spiral analysis using the iPad,” 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027016302692. [Accessed 28 04 2018].

[18] C. Marquardt and N.Mai, “A computational procedure for movement analysis in handwriting,” 1994. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165027094900531. [Accessed 28 04 2018].

[19] A. Sadikov, V. Groznik, M. Možina, J. Žabkar, D. Nyholm, M. Memedi, I. Bratko and D. Georgiev, “Feasibility of spirography features for objective assessment of motor function in Parkinson's disease,” 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365717301616. [Accessed 28 04 2018].

[20] S. Aghanavesi, D. Nyholm, M. Senek, F. Bergquist and M. Memedi, “A smartphone-based system to quantify dexterity in Parkinson's disease patients,” 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352914817300230. [Accessed 28 04 2018].

[21] F. Miralles, S. Tarongí and A. Espino, “Quantification of the drawing of an Archimedes spiral through the analysis of its digitized picture,” 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027005002955. [Accessed 28 04 2018].

[22] A. Toomela and S. Nõmm, “An alternative approach to measure quantity and smoothness of the human limb motion,” *Estonian Journal of Engeneering,* vol. 19, no. 4, 2013.

[23] V. Rozand, F. Lebon, C. Papaxanthis and R. Lepers, “Effect of mental fatigue on speed–accuracy trade-off,” 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306452215003152. [Accessed 28 04 2018].

[24] T. Mollayeva, T. Kendzerska, S. Mollayeva, C. M. Shapiro, A. Colantonio and J. D. Cassidy, “A systematic review of fatigue in patients with traumatic brain injury: The course, predictors and consequences,” 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0149763414002772. [Accessed 28 04 2018].

[25] “Novel method for measurement of fatigue in multiple sclerosis: Real-Time Digital Fatigue Score,” [Online]. Available: https://www.rehab.research.va.gov/jour/10/475/kim.html. [Accessed 1 5 2018].

[26] D. Muramatsu and T. Matsumoto, “Effectiveness of Pen Pressure, Azimuth, and Altitude Features for Online Signature Verification,” 2007. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-3-540-74549-5_53. [Accessed 28 04 2018].

[27] C. C. Aggarwal, “Datat classification,” in *Data Mining: The Textbook*, 2015, p. 290.

Appendix 1 – Visualisation of kinematic features

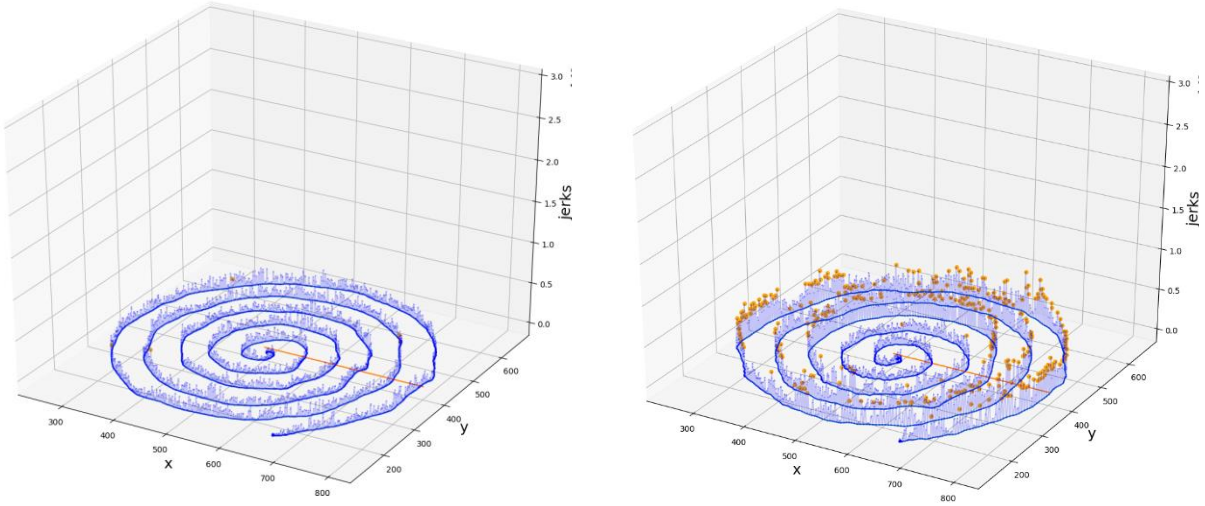

Figure 14. 3D plotting example of spiral of fatigued (left) and non-fatigued (right) individuals and their jerks at each point for the case of mental fatigue. Orange dots symbolise jerks values more than 25e+13.

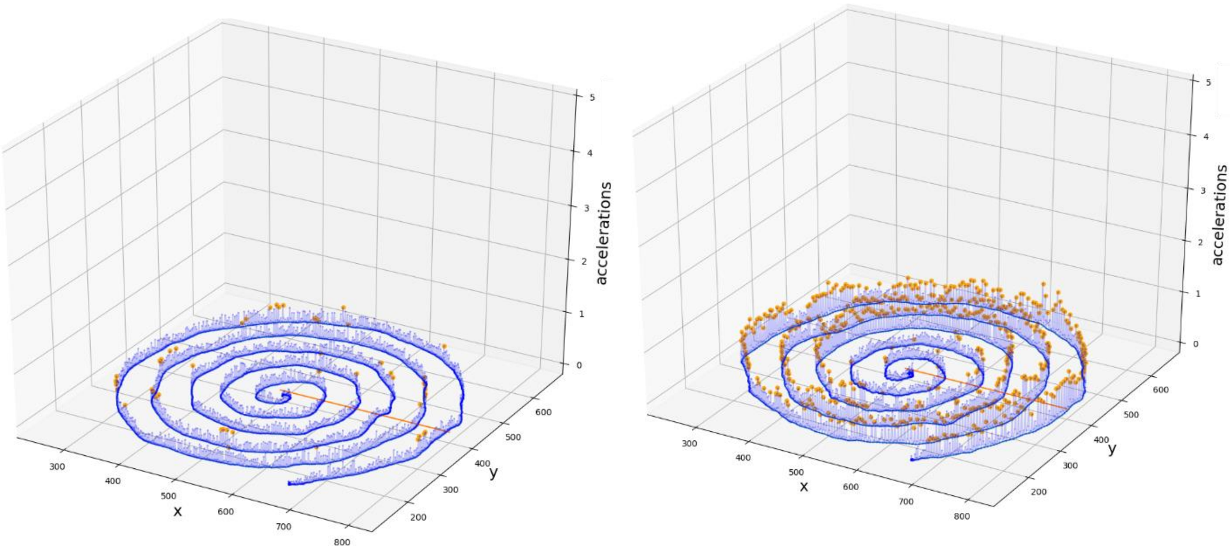

Figure 15. 3D plotting example of spiral of fatigued (left) and non-fatigued (right) individuals and their accelerations at each point for the case of mental fatigue. Orange dots symbolise accelerations values more than 3e+9.

Appendix 2 – Statistical evaluation

Table 12. An example of the most significant (p-value < 0.009) features sorted by lowest p-value according to the result of Group 2 (mental fatigue). s3, s6, s9, s12 prefixes identify the separation method of spiral presented in Figure 5.

Features	**T-stat**	**P-value**
s3_segment4_jerks	3.409646	0.002174
s3_segment1_time_deltas	-3.304345	0.002632
s3_segment4_jerks_ratio_to_length	3.264040	0.003019
s3_segment1_time_deltas_ratio_to_length	-3.225199	0.003226
s3_segment4_accelerations	3.250555	0.003232
s6_segment2_time_deltas_ratio_to_length	-3.203842	0.003464
s9_segment2_time_deltas_ratio_to_length	-3.189824	0.003518
s6_segment2_time_deltas	-3.173548	0.003750
s9_segment2_time_deltas	-3.158003	0.003818
s6_segment2_am_div_dist	3.153335	0.003896
s3_segment3_am_div_dist	-3.133149	0.004079
s12_segment2_time_deltas_ratio_to_length	-3.106208	0.004357
s12_segment2_time_deltas	-3.073505	0.004737
s9_segment2_am_div_dist	-3.048328	0.005055
s3_segment4_accelerations_ratio_to_length	3.050275	0.005079
s3_segment3_time_deltas_ratio_to_length	-3.056446	0.005098
s12_segment3_time_deltas_ratio_to_length	-3.039240	0.005259
s12_segment2_am_div_dist	-3.020729	0.005362
s3_segment3_time_deltas	-3.033574	0.005409
s3_segment1_am_div_dist	-3.028745	0.005453
s12_segment3_time_deltas	-3.014757	0.005605
s3_segment4_velocities	2.972400	0.006151
s9_segment3_time_deltas_ratio_to_length	-2.952964	0.006474
s12_segment4_jerks	2.961577	0.006691
s3_segment2_time_deltas_ratio_to_length	-2.913906	0.007004
s9_segment3_time_deltas	-2.918968	0.007047
s12_segment3_am_div_dist	-2.902107	0.007306
s3_segment2_time_deltas	-2.874833	0.007720
s12_segment1_am_div_dist	-2.880816	0.007749
s3_segment2_am_div_dist	-2.855989	0.008103
s9_segment1_am_div_dist	-2.840008	0.008360
s6_segment1_am_div_dist	-2.831713	0.008582
s6_segment1_time_deltas_ratio_to_length	-2.825350	0.008722
s12_segment1_time_deltas_ratio_to_length	-2.827754	0.008772
s3_segment1_jerks	2.851876	0.008897