ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS

Hala ABDULKADER

PREDICTING THE MAXIMAL OXYGEN UPTAKE USING DEEP LEARNING

DEPARTMENT OF COMPUTER ENGINEERING

ADANA, 2019

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PREDICTING MAXIMAL OXYGEN UPTAKE USING DEEP LEARNING

Hala Abdulkader

MSc THESIS

DEPARTMENT OF COMPUTER ENGINEERING

We certify that the thesis titled above was reviewed and approved for the award of the degree of the Master of Science by the board of jury on 01/10/2019

………………………	…………………………………	…………………………
Prof. Dr. M. Fatih AKAY	Asst. Prof. Dr. Serkan KARTAL	Asst. Prof. Dr. Onur ÜLGEN
SUPERVISOR	MEMBER	MEMBER

This MSc Thesis is written at the Department of Institute of Natural And Applied Sciences of Çukurova University.

Prof. Dr. Mustafa GÖK
Director
Institute of Natural and Applied Sciences

This study was supported by Çukurova University Scientific Research Projects Unit. Project Number: FYL-2018-10517

ABSTRACT

MASTER THESIS

PREDICTING THE MAXIMAL OXYGEN UPTAKE USING DEEP LEARNING

HALA ABDULKADER

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCE
DEPARTMENT OF COMPUTER ENGINEERING

Supervisor : Prof. Dr. Mehmet Fatih AKAY
Year: 2019, Pages: 63
Jury: : Prof. Dr. Mehmet Fatih AKAY
: Asst. Prof. Dr. Serkan KARATAL
: Asst. Prof. Dr. Onur ÜLGEN

Maximal oxygen uptake, or VO_2max, is an external parameter that is affected by things like how many red blood cells the body has, how adapted the muscles are to distance running, and how much blood the heart can pump. It is measured as milliliters of oxygen used in one minute per kilogram of body weight. In a laboratory, it is calculated by measuring the volume (V) of oxygen (O_2) that the body consumes while running on a treadmill which is the most accurate way. However, because of the serious drawbacks of direct measurement, a lot of studies have been conducted using machine learning methods to predict VO_2max. The purpose of this study is to build new VO_2max prediction models using deep learning (DL). The dataset has been split into training and test data using 70-30%, 80-20% split ratio, and 10-fold cross-validation. For comparison purposes, VO_2max prediction models based on Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Single Decision Tree (SDT) have also been developed. The performance of the prediction models has been evaluated using Standard Error of Estimate (SEE) and Multiple Correlation Coefficient (R). As a conclusion, DL can be used safely in VO_2max prediction domain.

Key Words: Deep Learning, Multi-Layer Perceptron, Support Vector Machine, Single Decision tree, Maximal Oxygen Uptake, Machine Learning.

ÖZ

YÜKSEK LİSANS TEZI

DERİN ÖĞRENME İLE MAKSİMUM OKSİGEN TÜKETİMİ TAHMİNİ

Hala ABDULKADER

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Prof. Dr. Mehmet Fatih AKAY
Yıl: 2019, Sayfa: 63
Jüri : Prof. Dr. Mehmet Fatih AKAY
: Dr. Öğr. Üyesi Serkan KARTAL
: Dr. Öğr. Üyesi Onur ÜLGEN

Kalp-solunum uygunluğu; solunum, kardiyovasküler ve iskelet-kas sistemlerine bağlı olduğu için sağlık ve fiziksel uygunluğun belirlenmesinde kullanılan önemli bir bileşendir. Kalp-solunum uygunluğunu tespit etmek için kullanılan standart yöntem maksimal egzersiz testi sırasında maksimum oksijen tüketiminin (VO_2max) direk olarak ölçülmesidir. VO_2max testi pahalı laboratuvar malzemesi, eğitimli teknisyenler, yüksek motivasyonlu bireyler, hatta bazı durumlarda bir doktorun desteğini gerektiren ve zaman alan bir testtir. Maksimal VO_2max testinin dezavantajlarından dolayı, literatürde bu teste alternatif olarak birçok VO_2max tahmin modeli önerilmiştir. Bu çalışmada, Derin Öğrenme Deep Learning (DP) yöntemi ile bireylerin VO_2max değerini tahmin etmek üzere yeni modeller geliştirilecektir. Karşılaştırma amaçlı Destek Vektör Makinesi Support Vector Machine (SVM), Çok Katmanlı Algılayıcı Multilayer Perceptron (MLP) ve Tek Karar Ağacı Single Decision Tree (SDT) tabanlı tahmin modelleri de geliştirilecektir. Tahmin modellerinin performansı, 10 katlı çapraz doğrulama kullanılarak, standart tahmin hatası (Standard Error of Estimate – SEE) ve çoklu korelasyon katsayısı (Multiple Correlation Coefficient – R) hesaplanarak değerlen¬dirilecektir. DP tabanlı VO_2max tahmin modellerinin diğer makine öğrenme tabanlı tahmin yöntemlerine göre daha performanslı çalışacağı beklenmektedir.

Anahtar Kelimeler: Makine Öğrenmesi, Destek Vektör Makinesi, Çok Katmanlı Algılayıcı, Tek Karar Ağacı Maksimum Oksijen Tüketimi, Tahmin.

GENİŞLETİLMİŞ ÖZET

VO_2max, vücudun oksijeni kullanılacağı noktalara taşıyabilme ve kullanabilme kapasitesidir. Maximal oksijen tüketimi, maximal oksijen alımı, oksijen alımı tepe noktası veya aerobik kapasite olarak da anılır. Bu son terim aslında ipucu vermektedir; VO_2max kişinin oksijen kullanarak yaptığı aktivitelerdeki kapasitesini yani fiziksel fitness'ını gösterir. Peki VO_2max terimini oluşturan harf ve rakamların anlamları nedir? (V) Volume yani hacim, (O_2) Oksijen, max ise maksimumdur; burdan da tanıma gidebilirdik aslında; vücudun tüketebileceği maksimum oksijen hacmi. Bu değer net ifade edilmek istendiğinde litre/dakika (dakikada tüketilen litre cinsinden hacim) ya da göreceli olarak ifade edilmek istendiğinde de mililitre/kilogram/dakika (dakikada kilogram başına tüketilen mililitre cinsinden hacim) şeklinde kullanılır.

Kalp-solunum uygunluğu; solunum, kardiyovasküler ve iskelet-kas sistemlerine bağlı olduğu için sağlık ve fiziksel uygunluğun belirlenmesinde kullanılan önemli bir bileşendir. Kalp-solunum uygunluğunu tespit etmek için kullanılan standart yöntem maksimal egzersiz testi sırasında maksimum oksijen tüketiminin (VO2max) direk olarak ölçülmesidir. VO2max testi pahalı laboratuvar malzemesi, eğitimli teknisyenler, yüksek motivasyonlu bireyler, hatta bazı durumlarda bir doktorun desteğini gerektiren ve zaman alan bir testtir. Maksimal VO2max testinin dezavantajlarından dolayı, literatürde bu teste alternatif olarak birçok VO2max tahmin modeli önerilmiştir. Bu çalışmada, Derin Öğrenme Deep Learning (DP) yöntemi ile bireylerin VO2max değerini tahmin etmek üzere yeni modeller geliştirilecektir. Karşılaştırma amaçlı Destek Vektör Makinesi Support Vector Machine (SVM), Çok Katmanlı Algılayıcı Multilayer Perceptron (MLP) ve Tek Karar Ağacı Single Decision Tree (SDT) tabanlı tahmin modelleri de geliştirilecektir. Tahmin modellerinin performansı, 10 katlı çapraz doğrulama kullanılarak, standart tahmin hatası (Standard Error of Estimate – SEE) ve çoklu korelasyon katsayısı (Multiple Correlation Coefficient – R) hesaplanarak

değerlendirilecektir. DP tabanlı VO2max tahmin modellerinin diğer makine öğrenme tabanlı tahmin yöntemlerine göre daha performanslı çalışacağı beklenmektedir.

VO_2max tespitinde en etkin olan ve en doğru sonucu veren yöntem yüze takılan bir maske ile bisiklet üzerinde veya koşu bandında yapılan maksimum efor testidir. Gittikçe artan zorlukta dirence karşı pedal çeviren sporcunun soluduğu ve dışarı verdiği havadaki oksijen ve karbondioksit miktarları bu maske aracılığıyla ölçülür. Testin çok uzamaması önemlidir (10-15 dk civarı), çünkü sporcunun en üst seviyeye ulaşmadan bacaklarının yorulmaması iyi olur. Egzersiz sırasında efor arttıkça tüketilen oksijen (ya da üretilen karbondioksit) miktarı artar. Ancak VO_2max seviyesine ulaşıldığında harcanan güç artsa da tüketilen oksijen miktarı değişmemeye başlar. Bu testin doktor kontrolünde yapılması çok önemlidir, çünkü sporcu testin sonuna doğru tamamen tükenebilir. VO_2max'ın ölçülmesinde maksimal egzersiz testi yüksek düzeyde doğruluğa sahip olmasına karşın, bu testin bazı sınırlamaları da bulunmaktadır. Maksimal VO_2max testi pahalı laboratuar malzemesi gerektiren zaman alıcı bir testtir. Testi gerçekleştirecek ve verileri yorumlayacak deneyimli eleman¬lara ihtiyaç duyulur. Bunlara ek olarak, denek sayısının fazla olduğu durumlarda tüm denekler için maksimal VO_2max testinin gerçekleştirilmesi pratik olarak mümkün değildir.

Bu projenin amacı; DP yöntemi ile bireylerin VO_2max değerini tahmin etmek üzere yeni modeller geliştirmektir. DP tabanlı VO_2max tahmin modellerinin diğer makine öğrenme tabanlı tahmin yöntemlerine göre daha performanslı çalışacağı beklenmektedir. Bu çalışmalardan elde edilen sonuçlara göre DL tabanlı tahmin modelleri ile SVM, MLP ve SDT tabanlı tahmin modellerine kıyasla daha yüksek performans sağlanmıştır. tez kapsamında iki farklı veri seti kullanılacaktır. Birinci veri seti içerisinde maksimal, submaksimal ve egzersize bağlı olmayan değişkenler, ikinci veri setinde ise, maksimal ve egzersize bağlı olmayan değişkenler bulunmaktadır.

Endüstri ve akademik çevrelerdeki veri bilimciler görüntü sınıflandırma, video analizi, konuşma tanıma ve doğal dil öğrenme süreci dahil olmak üzere çeşitli uygulamalarda çığır açan gelişmeler elde etmek üzere makineyle öğrenmede GPU'ları (Grafik İşlemci Ünitesi) kullanmaktadır. Özellikle, büyük miktarlarda etiketlenmiş eğitim verilerinden özellik saptama yapabilen sistemler oluşturmak için ileri teknoloji, çok seviyeli "derin" sinir ağların kullanılması olan DP, önemli derecede yatırım ve araştırmanın yapıldığı bir alandır. Makineyle öğrenme yıllardır kullanılan bir yöntem olmasına rağmen, iki yeni yeni trend makineyle öğrenmenin yaygın bir şekilde kullanılmasına yol açmıştır: çok büyük miktarlarda eğitim verisi ile GPU hesaplama ile elde edilen güçlü ve verimli paralel hesaplama. GPU'lar, çok daha büyük eğitim setleri kullanarak bu derin nöral ağları çok daha kısa sürelerde ve çok daha az veri merkezi altyapısı kullanarak eğitmek için kullanılmaktadır. GPU'lar aynı zamanda, çok daha fazla veri hacmi ve daha az güç ve altyapı destekleyerek, bulut içinde sınıflandırma ve tahmin yapmak için bu eğitilmiş makineyle öğrenme modellerini çalıştırmak için kullanılmaktadır.

SVM, yüksek doğruluğundan dolayı birçok uygulama alanında yaygın olarak kullanılan bir son teknoloji ürünü regresyon yöntemidir. Yüksek doğruluk, gen ekspresyonu gibi yüksek boyutlu verilerle çalışabilmesi, çeşitli kaynakların verilerini modelleme esnekliği gibi özellikleri nedeniyle Biyoinformatik (ve diğer disiplinlerde) yaygın olarak kullanılır. SVM tabanlı modellerin performansını belirleyen en önemli parametreler cost (C), epsilon (ε), gamma (γ) değerleridir.

MLP basit, iyi sunum kabiliyeti, erişilebilirliği ve yapısal esnekliği, çok sayıda programlama algoritması ile en çok kullanılan sinir ağı mimari programıdır. MLP ileri beslemeli sinir ağlarındandır ve evrensel yaklayışçılar geri yayılım ağları algoritması ile programlanırlar. Giriş datasını istenen bir cevaba dönüştürebilirler bu nedenle genellikle model sınıflandırılmasında kullanılır. Bir yada iki gizli katman ile fiilen giriş-çıkış haritası oluştururlar. Genellikle, bir MLP 3 katmandan oluşur: Giriş katmanı, Çıkış katmanı ve Gizli katman. Bu ağda her nöron bir sonraki katmandaki tüm nöronlara bağlıdır, bu nedenle MLP tamamen bağlı ağlardır. Gizli katmanda kullanılan nöron sayısı, öğrenme katsayısı ve momentum MLP tabanlı modellerin performansını etkileyen önemli parametreler arasındadır. SDT, karar alternatiflerinin düğümler halinde ve kendilerine ait olasılık değerleri belirtilerek ifade edildiği bir şekildir. Düğümlere ilişkin hesaplanan beklenen değerler şekle işlenir. SDT, sade ve kolay bir yapısı olmasına rağmen iyi sonuçlar verebilen bir yöntemdir ve çeşitli alanlardaki SDT ile yapılan deneyler, yeterince büyük numunelerde bile iyi performans gösterdiğini göstermektedir. Bu yöntem ile geliştirilen tahmin modellerinin performansını etkileyen parametreler minimum satır sayısı ve minimum ağ sayısıdır. Geliştirilecek VO_2max tahmin modellerinin performansı 10-katlı çağraz doğrulama kullanılarak ve RMSE performans metriği hesaplanarak değerlendirilecektir. DP tabanlı tahmin modellerine ait RMSE değerlerinin 3-7 ml/kg/min-1 arasında olması. DP tabanlı tahmin modellerinin diğer makine öğrenmesi tabanlı tahmin modellerine göre daha üstün performans sergilemesi. Temel olarak elde edilen sonuçlara bakıldığında başta DP olmak üzere makine öğrenme metodlarının kabul edilebilir ve doğru sonuçlar vermesi nedeniyle VO_2max tahmininde kullanılabileceği kanıtlanmıştır.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Prof. Dr. Mehmet Fatih AKAY, for his motivation, patience, and enthusiasm. His guidance helped me during all the thesis research and writing work.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF ABBREVIATIONS

DL	:	Deep Learning
MLP	:	Multi-Layer Perceptron
SVM	:	Support Vector Machine
SDT	:	Single Decision Tree
VO_2max	:	Maximal Oxygen Uptake
RMSE	:	Root Mean Square Error
R	:	Multiple correlation Coefficient
W	:	Weight
G	:	Gender
H	:	Height
BMI	:	Body Mass Index

1 INTRODUCTION

1.1 Overview of Maximal Oxygen Uptake

Maximal Oxygen Uptake, or VO_2max, is defined as the highest rate of oxygen consumption attainable during maximal or exhaustive exercise. It is measured as millilitres of oxygen used in one minute per kilogram of body weight. It is one factor that may help to determine an individual's capacity to perform sustained exercise. VO_2max score is generally considered by exercise physiologists as one of the best indicators of the person's cardiovascular fitness and aerobic endurance. This measure originates from pioneering research by A. A. Hill and colleagues that not only introduced the concept but also attempted to explain its physiological mechanisms. Hill commented, "In running the oxygen intake may attain its maximum and remain constant merely because it cannot go any higher owing to the limitations of the circulatory and respiratory system." Hill also attempted to explain the onset of fatigue: "At the higher speeds the requirement of the body for oxygen cannot be satisfied lactic acid accumulates, a continuously increasing oxygen debt being incurred, fatigue and exhaustion setting in." These conclusions were accepted by exercise physiologists for the next 70 years, and until recently the concept of VO_2max was not critically evaluated by research.

Measuring VO_2max accurately requires an all-out effort (usually on a treadmill or bicycle) performed under a strict protocol in a sports performance lab. These protocols involve a specific increase in the speed and intensity of the exercise, collection, and measurement of the volume and oxygen concentration of inhaled and exhaled air. This determines how much oxygen the individual's using. There is a specific point at which oxygen consumption plateaus even if the exercise intensity increases. This plateau marks the VO_2max.

Figure 1.1. VO_2max treadmill test

VO_2max can also be obtained using tests such as a bleep test, step test or Blake test. Bleep test involves running shuttles between two cones keeping in time with an ever increasing pace of bleeps. Once you are unable to keep up, then the level you reached will correspond with an estimate of VO_2max. This estimate will be deduced from measuring a large number of athletes accurately in a laboratory and then observing how high they achieve on a bleep test. It is important that whatever test will be applied it should be relevant to your sport or activity. For example, if ones are a middle distance runner, then doing a VO_2max test on an exercise bike will not be as relevant as doing it on a treadmill, or if a rower then performing it on a treadmill will not be relevant, while in step test it requires three minutes and a step tool with up and down movement, Blake test requires more effort, it needs a 10 minutes warms up, and running as far as possible in 15 minutes, then recording the total distance achieved in 15 minutes to the nearest 10 meters.

1.2 Motivation and Purpose of the Thesis

Although directly measuring the maximal oxygen uptake during graded, maximal exertion exercise on a treadmill or cycle ergometer in the laboratory (Maximal Exercise Test) gives the most accurate results. On the other hand, this method has some difficulties like the list below:

- Maximal tests are expensive to administer and can be risky because the subject reaches the point of exhaustion in terms of heart rate, and some older or higher risk individuals should not perform the test without a medical supervision
- The requirement of the expensive gas analysis and ventilation equipment
- Trained staff is needed in order to perform VO_2max measurements and interpret the test results
- Also, when the number of the subject is large, it is not feasible to apply VO_2max tests to all subjects

The need to assess aerobic capacity in the general public has led to the development of various exercise and non-exercise prediction models. Such models are effective for use in large epidemiological cohorts in which an exercise test to predict or measure VO_2max would be impractical.

VO_2max can be also predicted using exercise test that require maximal exertion on treadmill. Although it may be efficacious to use an exercise test requiring maximal effort in young, fit, and willing Pratik-time, are practical in a variety of settings. Sub-maximal exercise testing provides the administrator an opportunity to observe responses to exercise and to teach participants about selection of an appropriate intensity of exercise. For that reason, there was a quite need for more studies that can develop new prediction models which are needed to decide which predictor variables have the most effect on predicting an accurate VO_2max result.

The purpose of this thesis is to build new VO_2max prediction models using deep learning. The predictor variables which have been used to develop the prediction models include physiological variables, non-exercises, sub-maximal, and maximal models. For comparison purposes, VO_2max prediction models based on MLP, SVM, and SDT also have been developed. By carrying out 10-fold cross-validation and using 70-30%, 80-20% split ratios the error rate has been calculated. Using *RMSE* and *R*, the prediction errors have been measured.

1.3 Literature Review

Table 1.1 summarizes some important studies in the literature that used maximal models to estimate VO_2max.

Table 1.1. Summary of Maximal models for prediction of VO_2max

Study	Year	Method	Predictor Variables	R	SEE (mL kg^{-1} min^{-1})
Akay et al[10]	2015	SVM	Age, Gender, Height, Weight, Velocity	0.95	0.53
Wang et al[11]	2017	MLR	Age, Gender, BF%, Velocity	0.92	1.38
Akay et al[17]	2017	SVM	Age, Gender, Height, Weight, HR_{max}, Time, Speed, Grade	0.77	4.87
Akay et al[17]	2017	DTF	Age, Gender, Height, Weight, HR_{max}, Time, Speed, Grade	0.70	5.62
Akay et al[17]	2017	RBF	Age, Gender, Height, Weight, HR_{max}, Time, Speed, Grade	0.51	7.24
Akay et al[18]	2017	RBF	Gender, Age, Height, Weight, HR_{max}, Time, Speed, Grade	0.51	7.24
Akay et al[18]	2017	DTF	Gender, Age, Height, Weight, HR_{max}, Time, Speed, Grade	0.70	5.62
Akay et al[18]	2017	SVM	Gender, Age, Height, Weight, HR_{max}, Time, Speed, Grade	0.77	4.87
Akay et al[18]	2017	GRNN	Gender, Age, Height, Weight, HR_{max}, Time, Speed, Grade	0.81	4.51

BMI; body mass index, **PACER**; progressive aerobic cardiovascular endurance run, **BF%**; body fat percentage, **RER**; respiratory exchange ratio from treadmill test, **RPE**; self-reported rating of perceived exertion from treadmill test, **HRmax**; maximal heart rate, **HRLT**; heart rate at lactate threshold, **SR**; shuttle run, **NLP**; number of laps performed, **AHR**; after heart rate, **PR**; protocol.

Table 1.2 summarizes some important studies in the literature that used sub-maximal models to estimate VO_2max.

Table 1.2. Summary of sub-maximal models for prediction of VO_2max

Study	Year	Method	Predictor Variables	R	SEE (mL kg^{-1} min^{-1})
Tönis et al[23]	2012	MLR	Gender, Age, Height, Length, HR, AD	0.90	2.05
Cao et al[24]	2013	MLR	Gender, Age, 3MWD, BF%	0.84	4.57
Akay et al[22]	2015	GRNN	HRLT, HR, Weight, Height, Gender, Age	0.92	2.98
Akay et al[22]	2015	GEP	HRLT, HR, Weight, Height, Gender, Age	0.87	3.72
Akay et al[22]	2015	GMDH	HRLT, HR, Weight, Height, Gender, Age	0.88	4.02
Akay al[22]	2015	SDT	HRLT, HR, Weight, Height, Gender, Age	0.86	4.39
Abut et al[22]	2015	CCN	Gender, Age, Height, Weight, MIN3, HR3	0.89	2.69
Abut et al[22]	2015	GEP	Gender, Age, Height, Weight, MIN3, HR3	0.83	3.32
Abut et al[22]	2015	SDT	Gender, Age, Height, Weight, MIN3, HR3	0.75	3.89

HR; heart rate, **BMI**; body mass index, **MIN3**; elapsed exercise time at the end of 1.5 miles, **HR3**; exercise heart rate at the end of 1.5 miles, **AD**; accelerometer date, **3MED**; 3-minutes' walk distance, **BF%;** body fat percentage.

Table 1.3 summarizes some important studies in the literature that used non-exercise models to estimate VO_2max.

Table 1.3. Summary of non-exercise models for VO_2max prediction

Study	Year	Method	Predictor Variables	R	SEE (mL kg^{-1} min^{-1})
Akay et al[26]	2015	SVM	Gender, BMI, RPE-8	0.93	19.91
Abut et al[37]	2015	SVM	Gender, Age, Height, Weight, ES, Stage, HR, PFA, PAR	0.93	3.14
Abut et al[37]	2015	DTF	Gender, Age, Height, Weight, ES, Stage, HR, PFA, PAR	0.90	3.80
Abut et al[37]	2015	SVM	Gender, Age, Height, Weight, ES, Stage, HR, PFA, PAR	0.93	3.14
Abut et al[37]	2015	SDT	Gender, Age, Weight, Height, ES, Stage, HR, PFA, PAR	0.86	4.60
Akay et al[26]	2016	LMM	BF%, PFA	0.92	-
Cao et al[27]	2017	MFANN	Age, Gender, BMI, RPE--7, RPE-8, PAR	0.92	10.61
Schembre et al[26]	2017	MLR	Gender, RPE-8	0.93	10.80
Akay et al[26]	2016	LMM	BF%, PFA	0.92	-

BMI; body mass index, **SC**; step count, **VPA**; vigorous physical activity, **VA**; vigorous activity, **L-PA**; leisure-time physical activity, **W-PA**; work-related physical activity, **PFA**; perceived functional ability, **BSA**; body source area, **PAR**; physical activity rating, **BF%**; body fat percentage, **LBM**; lean body mass, **AC**; activity code, **LMM**: Lasso Multi-Marker.

1.4 Overview of the Datasets

In order to carry out this study, the first dataset includes data of 439 subjects that were required from the Amherst and Worcester, Massachusetts, communities, and successfully completed a maximal walking treadmill test.

Table 1.4. Statistical information about the 1st dataset

Variables	**Mean**	**Stand. Dev.**
Gender	0.52	0.50
Age (year)	46.60	16.66
Height (m)	1.68	11.41
Weight (Kg)	72.9	14.80
BMI (Kg/m_2)	25.46	4.10
BF	23.39	8.44
AC	4.46	2.04
LBM	55.68	12.30
MAXHR (bpm)	178.96	15.59
MAXRER	1.17	0.06
MAXRPE	18.29	1.42
MAXTIMEs	16.16	3.39
VO_2max (mL Kg^{-1} min^{-1})	38.61	10.37

BF; body fat, **AC**; activity code, **LBM**; lean body mass, **MXHR**; maximal heart rate, **MXRPE**; maximal respiratory exchange ratio, **MXRPE**; a maximal self-reported rating of perceived exertion, **MXTIME**; time to exhaustion(seconds).

The second dataset includes one hundred participants were recruited from the Y-Be-Fit Wellness Program at Brigham Young University and employees from the LDS Hospital in Salt Lake City, Utah. Table 1.4 shows statistical information about the first dataset.

Table 1.5. Shows statistical information about the 2nd dataset

Variables	**Mean**	**Stand. Dev.**
Gender	0.5	0.50
Age (year)	36.23	13.15
Height (m)	1.71	0.08
Weight (Kg)	75.36	15.18
BMI (Kg/m^2)	25.59	4.10
SMendingspeed	5.24	1.00
SMHR	147.73	12.29
SMstage	2.11	0.48
MXendingspeed	5.31	1.04
MXHR	185.16	13.18
MXRPE	19.01	0.78
MXGRD	9.38	2.71
MXRER	1.16	0.05
VO_2max ($mL\ Kg^{-1}\ min^{-1}$)	41.38	9.09

2 OVERVIEW OF METHOD

2.1 Deep Neural Network Learning

Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans; learn by example. Deep learning is the key technology behind driverless cars, enabling them to recognize a stop sign or to distinguish a pedestrian from a lamppost. It is the key to voice control in consumer devices like phones, tablets, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It is achieving results that were not possible before.

In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labelled data and neural network architectures that contain many layers.

Recent advances in deep learning have improved to the point where deep learning outperforms humans in some tasks like classifying objects in images. While deep learning was first theorized in the 1980s, there are two main reasons it has only recently become useful; deep learning requires large amounts of labelled data. For example, driverless car development requires millions of images and thousands of hours of video. Deep learning requires substantial computing power; High-performance GPUs have a parallel architecture that is efficient for deep learning. When combined with clusters or cloud computing, this enables development teams to reduce training time for a deep learning network from weeks to hours or less.

Most deep learning methods use neural network architecture, which is why deep learning models are often referred to as deep neural networks. The term “deep” usually refers to the number of hidden layers in the neural network. Traditional neural networks only contain 2-3 hidden layers, while deep networks can have as many as 150. Deep learning models are trained by using large sets of labelled data and neural network architectures that learn features directly from the data without the need for manual feature extraction.

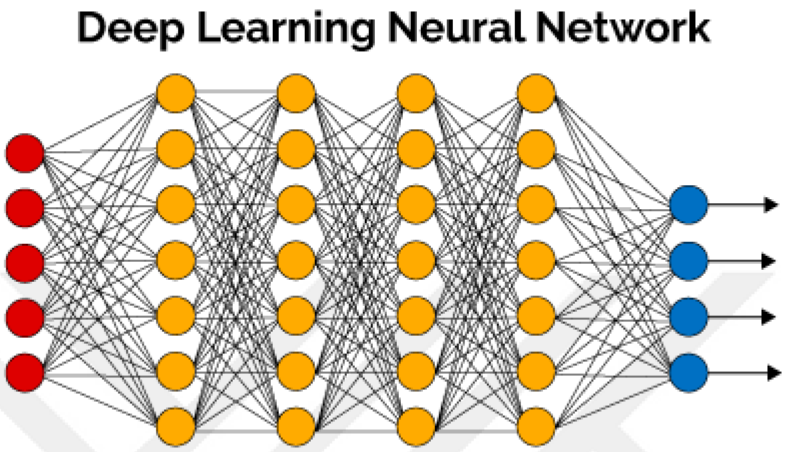


Figure 2.1. Deep learning structure

Figure 2.1: Neural networks, which are organized in layers consisting of a set of interconnected nodes. Networks can have tens or hundreds of hidden layers.

Training: the procedure used to carry out the learning process is called training (or learning) strategy. The training strategy is applied to the neural network in order to obtain the best possible loss. The type of training is determined by the way in which the adjustment of the parameters in the neural network takes place.

Quasi-Network method: these methods, instead of calculating the Hessian directly and then evaluating its inverse, build up an approximation to the inverse Hessian at each iteration of the algorithm. This approximation is computed using only information on the first derivatives of the loss function. The Hessian matrix is composed of the second partial derivatives of the loss function. The main idea behind the quasi-Newton method is to approximate the inverse Hessian by another matrix G, using only the first partial derivatives of the loss function. Then, the quasi-Newton formula can be expressed as:

$$w_{i+1} = w_i - (G_i \cdot g_i) \cdot \eta_i, \quad i=0,1,...$$

The training rate η can either be set to a fixed value or found by line minimization. The inverse Hessian Approximation G has different flavors. Two of the most used are the Davidson-Fletcher-Powell formula (DFP) and the Broyden-Fletcher-Goldfarb-Shanno formula (BFGS). The activity diagram of the Quasi-Newton training process is illustrated below. Improvement of the parameters is performed by first obtaining a Quasi-Newton training direction and then finding a satisfactory training rate. The quasi-newton method is used here for training. It is based on Newton's method but does not require the calculation of second derivatives. Instead, the Quasi-Newton method computes an approximation of the inverse Hessian at each iteration of the algorithm, by only using gradient information. The activity diagram of the training process with the Quasi-Newton method can be represented as given in Figure 2.2.

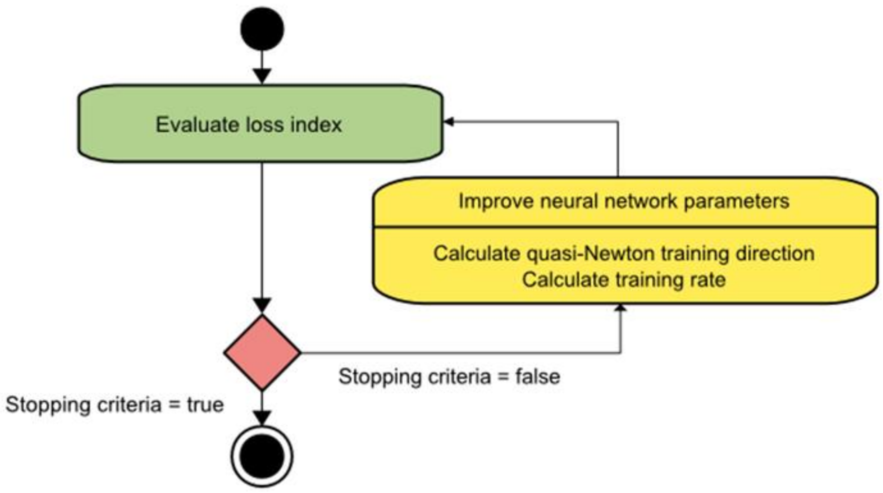


Figure 2.2. Quasi-Newton training diagram

This is the default method to use in most cases: It is faster than gradient descent and conjugate gradient, and the exact Hessian does not need to be computed and inverted.

Testing errors: this task measures all the losses of the model. It takes into account every used instance and evaluates the model for each use.

2.2 Support Vector Machine

SVM is a supervised machine learning algorithm invented by Vladimir N. Vapnik and Alexey Chervonenkis in 1963 and introduced by Boser, Guyon, and Vapnik in 1992. It is one of the most effective classifiers, there is a strong mathematical intuition behind the support vector machine which makes this method able to handle certain cases where there is non-linearity by using non-linear basis functions or the kernel functions. It has a clever way to prevent overfitting, and it can work with a relatively large number of features without requiring too much computation. Its popularity has grown because of the accuracy in classification and prediction.

2.3 Multi-Layer Perceptron

It is a perceptron with multiple layers, it has one input layer, one output layer, and many hidden layers. The perceptron consists of weights, activation, and computation function, it takes the weighted sum of inputs and outputs. The inputs and connection weights are typically real values. The input values are presented to the perception, and if the predicted output is the same as the desired output, then the performance is considered satisfactory and no changes to the weight are made. However, if the output does not match the desired output, then the weights need to be changed to reduce the error. It is as forward passing something from one layer to the next (propagate inputs by adding all the weighted inputs and the computing outputs using sigmoid threshold). In MLP the signal can move from the input layer to the output layer in one direction only that explains Feedforward. By changing the number of the hidden layers the complexity of MLP changes. It can be presented as given in figure 2.4.

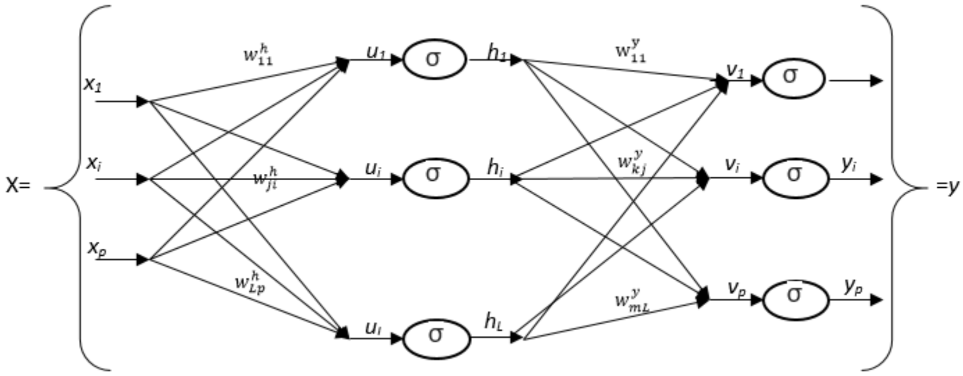


Figure 2.3. MLP network

In this network, we can see three layers (from left to right); input layer, hidden layer, an output layer, and each layer has three perceptrons.

Input layer: it introduces the input values into the network. In the input layer, there is no activation function or other processes.

Hidden layer: arriving at a perceptron in the hidden layer, it multiplies the value of the perception by its weight (w_{ji}), and it adds the results together to produce the combined value u_j.

$$U_j=\sum_{i=0}^{P} w_{ji}x_i$$

(u_j) the weighted sum is fed into a transfer function σ, which gives an output value h_j.

Output Layer: like the hidden layer, each hidden layer perceptron value is multiplied by a weight (w_{kj}), and the results are added together to produce the value v_j.

2.4 Single Decision Tree

A tree-structured classifier in a form of tree graph which has nodes and branches, a root node, children, and leafs which do not have children. It has two types of nodes; decision nodes, and leaf nodes. The decision nodes specify a test based to decide which branch direction to go. The test is usually done on the value of a feature or attribute of the instance. The leaf nodes indicate the classification of an example or the value of an example. A decision tree can be used in regression and classification. However, it is more popularly used in classification. The SDT work starts with the root of the tree and based on the value of the test it goes to the corresponding branch and continues until the leaf node, then gives the predicted value. The space of the decision tree is too big for a systematic search. So searching for a good tree needs to decide wither to stop or continue, stop means return the value for the target feature or a distribution over target feature values, continue and choose a test to split on means for each value of the test, build a subtree for those examples with these value for the test. The framework of a basic decision tree algorithm, it is called a top-down induction OD decision trees ID3.

- At the current node choose the best decision attribute
- Assign A as the decision attribute
- For each value of A create a new descendant
- Sort training examples to leaf node according to the attribute value of the branch
- If training example is perfectly classified stop, else iterate over new leaf nodes

3 DEVELOPMENT OF PREDICTION MODELS

3.1 Deep Learning Prediction Models

Using back-propagation, the prediction models have been created, in order to have an effective and accurate result. This DL based prediction model accuracy is enhanced by varying the layers' number and the unit number for each layer. The relationship between the neurons number of the hidden layer and the prediction models performance is quietly closed. Sometimes a loss of information could happen because of the small number of neurons, on the other hand using a large number of neurons will increase the model complexity. Still, the number of hidden layers' neurons cannot be specified in advance. In general, we can get the optimal neurons number by trying and error processing depending on the problem complexity. Table 3.1 shows the parameter's value for the DL models.

Table 3.1. List of the parameter's value of DL prediction models

Parameter	Value
Number of neurons in the hidden layer	[6 – 50]
Number of the hidden layers	[3 – 30]
Momentum	[0 – 1]
Maximum iteration	1000

The number of hidden layers can’t be specified in advanced, it depends on the dataset size and the problem complexity. In this work DL showed different performance due to the different dataset; as the first dataset [439 subjects] is larger than the second one [100 subjects].

In the first dataset larger number of hidden layer have been used up to 30 hidden layers especially with maximal models.

On the other hand, with the second dataset up to 9 hidden layers have been used with maximal model.

Increasing the hidden layers and neurons numbers in small dataset could increase the network complexity and causes undesired output and underfitting problem. Also using quasi network function in small dataset with large number of hidden layers causes a stuck point during the algorithm running time. For that reason, DL with large dataset gives ability to use large number of hidden layers, otherwise it would not give optimized results.

3.2 Support Vector Machine Prediction Models

The capable of extracting the optimal solution using SVM is improved with the small training set size. Many techniques such as grid search and swarm optimization are employed to find the optimal values, yet the grid search technique had been used in this study. It is very effective in optimizing the value of *C*, ε, and gamma on a medium dataset. The value of the parameters is changed within a determined interval and the three parameters which show the lowest rate is retained. The idea behind this is simple and it relies on a trial and error process. The values of the parameters are varied within a predefined range in the grid search, and the value *C, ε,* and γ yielding the maximum prediction performance are selected. The limit values used for the grid search method were selected according to the recommendations that says trying exponentially growing of sequence of *C* and γ *i*s an effective way to determine the optimal values. Similarly; the *ε,* values were chosen so that percentage of support vectors in the respective SVM-based models is about 50% of the number of total samples Table 3.2 shows the parameter value for SVM models.

Table 3.2. List of the parameter's value for SVM models

Parameter	**Value**
Cost (*C*)	[1^{-1} – 5000]
Epsilon (ε)	[0.01 – 150]
Gamma (γ)	[1^{-3} – 50]
Kernel function	RBF

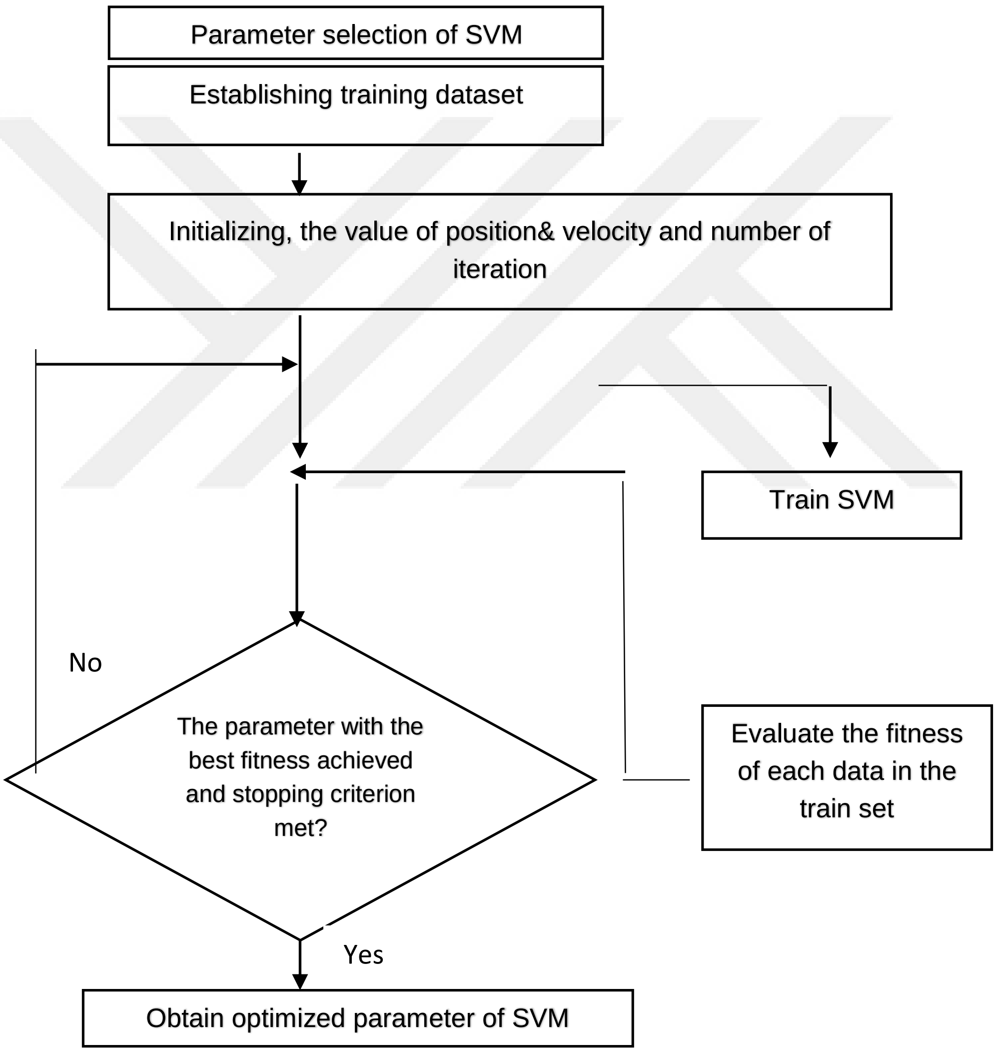


Figure 3.1. SVM selection process flow chart

3.3 Multi-Layer Perceptron Prediction Models

Using back propagation, the prediction models have been created. In order to have effective and accurate results. This MLP based prediction model accuracy is enhanced by varying the unit number for each layer. The relationship between the neurons number of the hidden layer and the prediction models performance is quietly closed. The learning strategy minimizes a global error function for the set of input data. Single hidden layer had been used and $(1+e^{-x})^{-1}$ for the transfer function, a learning rate and momentum rate. The input layer was randomly selected; the output layer uses 1- number of classes encoding. For the output neurons $y_{i,}$ = $[c_{i,} c_i$ $, c_{i,} c_i]$ where $c_i=1$, if the respective input spectrum belongs to the corresponding class and $c_i=0$, otherwise the data is encoded in a similar fashion, rather than randomly initializing the neuron weights, they are initialized so that the slopes of the neurons are randomly placed within the domain of the input spectra

Table 3.3 shows the parameter value for the MLP models.

Table 3.3. List of the parameter value for the MLP models

Parameter	**Value**
Number of neurons in the hidden layer	[2 – 20]
Learning rate	[0 - 1]
Momentum	[0 - 1]
Maximum iteration	10000

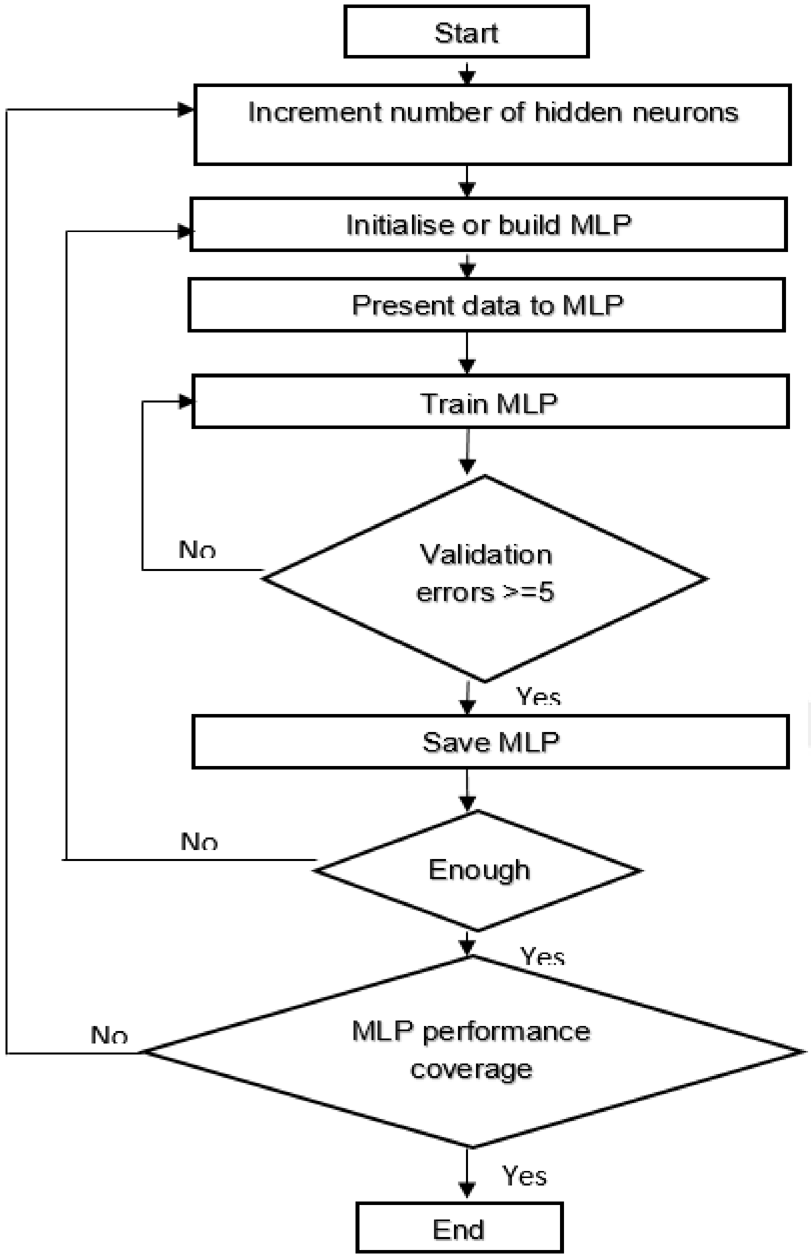


Figure 2.4. MLP training process flow chart

3.4 Single Decision Tree Prediction Models

It is primarily composed of training data, test data, a heuristic evaluation function, and a stopping criterion function. It uses recursion from the divide and conquer data structure to induce data from the root node to downward. Ultimately, decision tree represents the gold standard for partition based models. However, size, complexity, and data sparseness lead to misclassification of elements when tree is large. Here are three main parameters that are responsible for the SDT performance; the minimum size node to split, the minimum row in a node, and the maximum tree levels. The minimum size node to split indicates that the group of node should not be split if there were fewer rows than the specified value, the minimum rows in anode determines the minimum number of the rows that could fall in the node after splitting, the maximum tree levels determines that the maximum number of the three levels that are created when the building begins. In order to get an effective SDT-based prediction model, the above-mentioned parameters should have the optimal value. Table 3.4 shows the parameter values for SDT models.

Table 3.4. List of the parameter value for SDT models

Parameter	Value
Minimum size node to split	[10 – 30]
Minimum rows in a node	[5 – 25]
Maximum tree levels	[5 – 10]

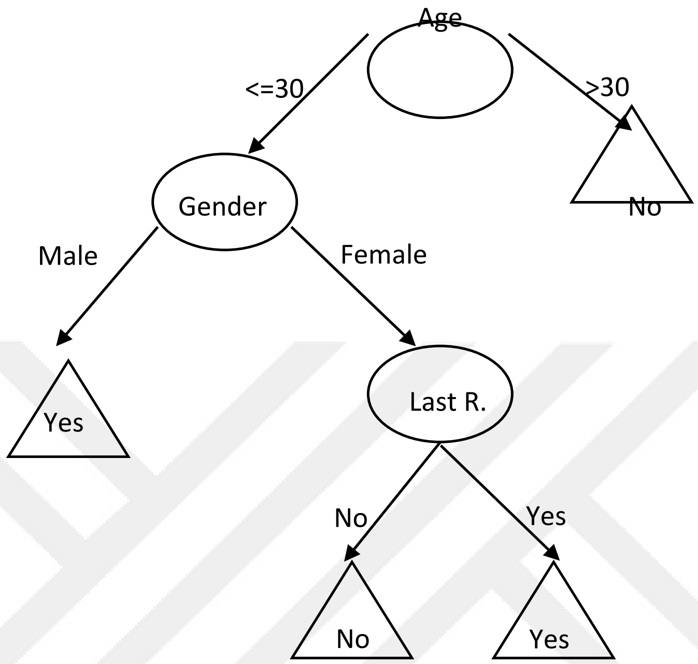


Figure 3.2. Example of SDT algorithm workflow

4 RESULT AND DISCUSSIONS

4.1 Approach of the Thesis

Using the combination of these predictor variables: physiological variables, Maximal, Sub-maximal, and Non-exercise; the models have been created. There are two categories, in the first one Height and Weight have been used, while in the other BMI predictor variable has been used instead of Height and Weight. In order to increase the performance and the accuracy of the prediction, 80 – 20%, 70 – 30% split ratio and 10-fold cross validation have been used to compute *RMSE*. The equation of *RMSE* is given below.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y - Y')}$$

While Y is the measured value, Y' is the predictor value.

Table 4.1 to Table 4.5 shows the prediction models.

Table 4.1. Overview of prediction models with Sub-maximal predictor variables

Models	**Predictor Variables**				
M1	Age & Gender	Height & Weight	SM-ES		
M2	Age & Gender	Height & Weight	SM-HR		
M3	Age & Gender	Height & Weight	SM-S		
M4	Age & Gender	Height & Weight	SM-ES	SM-HR	
M5	Age & Gender	Height & Weight	SM-ES	SM-S	
M6	Age & Gender	Height & Weight	SM-HR	SM-S	
M7	Age & Gender	Height & Weight	SM-ES	SM-HR	SM-S
M8	Age & Gender	BMI	SM-ES		
M9	Age & Gender	BMI	SM-HR		
M10	Age & Gender	BMI	SM-S		
M11	Age & Gender	BMI	SM-ES	SM-HR	
M12	Age & Gender	BMI	SM-ES	SM-S	
M13	Age & Gender	BMI	SM-HR	SM-S	
M14	Age & Gender	BMI	SM-ES	SM-HR	SM-S

Table 4.2. Overview of prediction models with maximal predictor variables

M	Predictor Variables					
M1	Age & Gender	Height & Weight	MX-ES			
M2	Age & Gender	Height & Weight	MX-HR			
M3	Age & Gender	Height & Weight	MX-RPE			
M4	Age & Gender	Height & Weight	MX-GRD			
M5	Age & Gender	Height & Weight	MX-ES	MX-HR		
M6	Age & Gender	Height & Weight	MX-ES	MX-RPE		
M7	Age & Gender	Height & Weight	MX-ES			
M8	Age & Gender	Height & Weight	MX-HR	MX-RPE		
M9	Age & Gender	Height & Weight	MX-HR			
M10	Age & Gender	Height & Weight	MX-RPE			
M11	Age & Gender	Height & Weight	MX-ES	MX-HR	MX-RPE	
M12	Age & Gender	Height & Weight	MX-ES	MX-HR	MX-GRD	
M13	Age & Gender	Height & Weight	MX-ES	MX-RPE	MX-GRD	
M14	Age & Gender	Height & Weight	MX-HR	MX-RPE	MX-GRD	
M15	Age & Gender	Height & Weight	MX-ES	MX-HR	MX-RPE	MX-GRD
M16	Age & Gender	BMI	MX-ES			
M17	Age & Gender	BMI	MX-HR			
M18	Age & Gender	BMI	MX-RPE			
M19	Age & Gender	BMI	MX-GRD			
M20	Age & Gender	BMI	MX-ES	MX-HR		
M21	Age & Gender	BMI	MX-ES	MX-RPE		
M22	Age & Gender	BMI	MX-ES			
M23	Age & Gender	BMI	MX-HR	MX-RPE		
M24	Age & Gender	BMI	MX-HR			
M25	Age & Gender	BMI	MX-RPE			
M26	Age & Gender	BMI	MX-ES	MX-HR	MX-RPE	
M27	Age & Gender	BMI	MX-ES	MX-HR	MX-GRD	
M28	Age & Gender	BMI	MX-ES	MX-RPE	MX-GRD	
M29	Age & Gender	BMI	MX-HR	MX-RPE	MX-GRD	
M30	Age & Gender	BMI	MX-ES	MX-HR	MX-RPE	MX-GRD

Table 4.3. Overview of prediction models with Non-exercise predictor variables

Models	**Predictor Variables**				
M1	Age & Gender	Height & Weight	PFA1		
M2	Age & Gender	Height & Weight	PFA2		
M3	Age & Gender	Height & Weight	PAR		
M4	Age & Gender	Height & Weight	PFA1	PFA2	
M5	Age & Gender	Height & Weight	PFA1	PAR	
M6	Age & Gender	Height & Weight	PFA2	PAR	
M7	Age & Gender	Height & Weight	PFA1	PFA2	PAR
M8	Age & Gender	BMI	PFA1		
M9	Age & Gender	BMI	PFA2		
M10	Age & Gender	BMI	PAR		
M11	Age & Gender	BMI	PFA1	PFA2	
M12	Age & Gender	BMI	PFA1	PAR	
M13	Age & Gender	BMI	PFA2	PAR	
M14	Age & Gender	BMI	PFA1	PFA2	PAR

Table 4.4. Overview of prediction models with maximal predictor variables

M	**Predictor Variables**					
M1	Age & Gender	Height & Weight	MX-HR			
M2	Age & Gender	Height & Weight	MX-RER			
M3	Age & Gender	Height & Weight	MX-RPE			
M4	Age & Gender	Height & Weight	MXTIME			
M5	Age & Gender	Height & Weight	MX-HR	MX-RER		
M6	Age & Gender	Height & Weight	MX-HR	MX-RPE		
M7	Age & Gender	Height & Weight	MX-HR	MXTIME		
M8	Age & Gender	Height & Weight	MX-RER	MX-RPE		
M9	Age & Gender	Height & Weight	MX-RER	MXTIME		
M10	Age & Gender	Height & Weight	MX-RPE	MXTIME		
M11	Age & Gender	Height & Weight	MX-HR	MX-RER	MX-RPE	
M12	Age & Gender	Height & Weight	MX-HR	MX-RER	MXTIME	
M13	Age & Gender	Height & Weight	MX-HR	MX-RPE	MXTIME	
M14	Age & Gender	Height & Weight	MX-RER	MX-RPE	MXTIME	
M15	Age & Gender	Height & Weight	MX-HR	MX-RER	MX-RPE	MXTIME
M16	Age & Gender	BMI	MX-HR			
M17	Age & Gender	BMI	MX-RER			
M18	Age & Gender	BMI	MX-RPE			
M19	Age & Gender	BMI	MXTIME			
M20	Age & Gender	BMI	MX-HR	MX-RER		
M21	Age & Gender	BMI	MX-HR	MX-RPE		
M22	Age & Gender	BMI	MX-HR	MXTIME		
M23	Age & Gender	BMI	MX-RER	MX-RPE		
M24	Age & Gender	BMI	MX-RER	MXTIME		
M25	Age & Gender	BMI	MX-RPE	MXTIME		
M26	Age & Gender	BMI	MX-HR	MX-RER	MX-RPE	
M27	Age & Gender	BMI	MX-HR	MX-RER	MXTIME	
M28	Age & Gender	BMI	MX-HR	MX-RPE	MXTIME	
M29	Age & Gender	BMI	MX-RER	MX-RPE	MXTIME	
M30	Age & Gender	BMI	MX-HR	MX-RER	MX-RPE	MXTIME

Table 4.5. Overview of prediction models with Non-exercise predictor variables

Models	Predictor Variables			
M1	Age & Gender	Height & Weight	AC	
M2	Age & Gender	Height & Weight	BF	
M3	Age & Gender	Height & Weight	AC	BF
M4	Age & Gender	BMI	AC	
M5	Age & Gender	BMI	BF	
M6	Age & Gender	BMI	AC	BF

4.2 Results

Table 4.6. Test results for Sub-Maximal Prediction models.

Models	DL	SVM	MLP	SDT
	10-fold	**10-fold**	**10-fold**	**10-fold**
M 1	3.42	4.19	4.26	5.66
M 2	4.03	5.7	4.85	6.51
M 3	4.18	4.79	4.89	6.81
M 4	3.4	4.14	4.31	5.49
M 5	3.51	4.37	4.43	5.66
M 6	3.45	4.89	4.87	6.89
M 7	3.36	4.42	4.51	5.49
M 8	3.46	4.24	4.13	5.42
M 9	3.5	4.92	4.91	5.88
M 10	4.2	4.82	4.88	6.04
M 11	3.39	4.16	4.23	5.55
M 12	3.3	4.26	4.3	5.39
M 13	3.4	4.78	4.84	6.1
M 14	3.47	4.31	4.39	5.44

Models	DL	SVM	MLP	SDT
	80%	**80%**	**80%**	**80%**
M 1	3.57	3.83	3.61	3.99
M 2	4.08	4.6	4.52	7.23
M 3	3.34	3.9	3.89	7.23
M 4	3.36	3.57	3.62	3.99
M 5	3.19	3.58	3.81	3.99
M 6	3.74	3.85	4.64	7.23
M 7	3.31	3.47	3.92	3.99
M 8	3.09	5.28	3.38	4.34
M 9	3.84	4.24	4.76	5.43
M 10	2.93	3.83	3.84	6.03

M 11	3.37	3.65	3.6	4.34
M 12	3.45	3.53	3.63	4.34
M 13	3.37	3.57	4.05	5.43
M 14	2.54	3.39	3.69	4.34

Models	**DL**	**SVM**	**MLP**	**SDT**
	70%	**70%**	**70%**	**70%**
M 1	2.88	5.22	5	5.24
M 2	4.92	5.21	5.08	7.31
M 3	4.05	5.62	5.54	7.11
M 4	4.35	4.99	5.17	5.8
M 5	4.71	5.24	5.13	5.24
M 6	4.84	5.63	5.45	7.11
M 7	4.56	5.18	5.31	5.8
M 8	4.1	4.81	4.97	5.04
M 9	5.06	5.4	5.12	7.32
M 10	5.37	5.44	5.57	7.39
M 11	3.67	4.75	5.02	5.58
M 12	4.37	5.04	4.84	5.04
M 13	5.14	5.4	6.26	7.39
M 14	4.66	5.07	4.86	5.58

From these Tables we can see that the best performance has been given from split ratio 80%-20% with 2.54*R* including: physiological variables with BMI, SM-ES, SM-HR, SM-S, while in split ratio 70%-30% it is 2.8*R* for: physiological variables with (Height, Weight), and SM-ES, in 10-fold it is 3.3*R* for: physiological variables with BMI, SM-ES, SM-S.

Table 4.7. Test results for Maximal prediction models

Models	DL	SVM	MLP	SDT
	10-fold	10-fold	10-fold	10-fold
M 1	3.65	4.27	4.47	5.05
M 2	4.34	5.07	4.94	6.85
M 3	4.11	5.12	5.23	6.41
M 4	4.12	4.92	5.05	6.15
M 5	3.43	4.22	4.34	5.25
M 6	3.43	4.12	4.27	5
M 7	2.62	3.1	3.09	4.86
M 8	4.03	5.19	5.06	6.76
M 9	3.86	4.95	4.92	6.64
M 10	4.16	4.99	5.04	6.34
M 11	3.67	4.25	4.25	5.22
M 12	2.65	3.24	3.25	5.09
M 13	2.73	3.12	3.25	4.9
M 14	4.01	5.04	5.17	6.79
M 15	2.59	3.15	3.32	5.09
M 16	3.4	3.98	4.13	4.75
M 17	4.5	4.98	5.04	6.38
M 18	4.2	5.05	4.95	5.75
M 19	4.1	4.85	4.99	5.84
M 20	3.18	4.19	4.18	4.99
M 21	3.57	4.18	4.15	4.81
M 22	2.8	3.1	3.22	4.88
M 23	4	5.19	5	6.36
M 24	4.6	4.91	4.91	6.32
M 25	4.1	5.02	4.9	5.91
M 26	3.6	4.13	4.19	5.05
M 27	2.5	3.15	3.17	5.1
M 28	2.49	3.28	3.29	4.92
M 29	4.3	4.97	4.99	6.41
M 30	2.3	3.28	3.24	5.1

In 10-fold the best performance has been given from 2.3R for: physiological variables with (Height, Weight), MX-HR, MX-RPE, MX-GRD.

Models	**DL**	**SVM**	**MLP**	**SDT**
	80%	**80%**	**80%**	**80%**
M 1	3.36	3.42	3.66	4.11
M 2	3.6	6.5	4.49	7.03
M 3	4.01	4.57	4.65	7.25
M 4	4.27	4.32	4.49	7.31
M 5	3.15	3.66	3.66	4.97
M 6	3.09	3.56	3.33	3.87
M 7	2.07	3.15	2.56	3.73
M 8	3.92	4.78	4.88	7.03
M 9	3.52	4.21	4.62	7.03
M 10	3.79	4.15	4.28	7.31
M 11	3.24	3.4	3.54	5.13
M 12	2.15	3.14	2.58	4.03
M 13	2.51	2.53	2.64	3.73
M 14	3.69	4.21	4.9	7.03
M 15	2.4	2.52	3.21	4.03
M 16	3.02	3.45	3.34	4.27
M 17	4.14	4.37	4.44	6.03
M 18	3.75	5.36	4.32	6.03
M 19	4.02	4.15	4.38	6.03
M 20	3.39	3.75	3.46	4.97
M 21	3.27	3.5	3.36	4.25
M 22	2.87	3.2	3.19	3.82
M 23	3.27	4.34	4.19	6.03
M 24	3.59	4.21	5.01	6.03
M 25	3.61	4.23	4.48	6.03
M 26	2.87	3.37	3.38	4.97
M 27	2.23	3.14	2.5	4.03
M 28	2.46	2.62	2.76	3.82
M 29	3.84	4.28	4.78	6.03
M 30	2.31	2.62	2.54	4.03

The best performance has been given from split ratio 80%-20% with 2.07R including: physiological variables with (Height, Weight).

In split ratio 70%-30% the best performance has been given from 2.08 *R* also for: physiological variables with (Height, Weight).

Models	DL	SVM	MLP	SDT
	70%	**70%**	**70%**	**70%**
M 1	4.37	4.82	4.56	5.65
M 2	4.45	5.39	5.66	7.31
M 3	5.01	6.28	5.33	7.31
M 4	4.59	5.16	5.14	5.63
M 5	3.14	4.82	5.04	5.63
M 6	4.17	5.16	4.88	5.67
M 7	2.08	2.91	3.1	5.65
M 8	5.16	6.32	5.77	7.31
M 9	5.03	5.05	5.24	7.31
M 10	5.18	5.92	6.06	7.31
M 11	4.1	4.79	4.81	5.63
M 12	2.74	3.08	2.99	5.35
M 13	3.06	3.41	3.42	5.67
M 14	4.95	5.89	6.31	7.31
M 15	2.67	3.5	3.63	5.36
M 16	4.31	4.53	4.44	5.7
M 17	4.93	5.27	5.13	7.32
M 18	5.04	5.85	5.45	7.32
M 19	4.61	5.11	4.75	7.32
M 20	4.27	4.6	4.56	5.73
M 21	4.29	5.25	4.71	5.7
M 22	2.21	2.84	2.62	5.7
M 23	4.66	5.41	5.52	7.32
M 24	4.12	5.05	5.64	7.32
M 25	4.19	5.43	5.06	7.32
M 26	4.3	5.14	4.74	5.73
M 27	2.44	2.74	2.49	5.36
M 28	2.3	3.34	3.15	5.7
M 29	3.96	5.73	5.23	7.32
M 30	3.1	3.38	3.24	5.36

Table 4.8. Test results for non-exercise prediction models.

Models	**DL**	**SVM**	**MLP**	**SDT**
	10-fold	**10-fold**	**10-fold**	**10-fold**
M 1	2.95	4.01	3.99	5.65
M 2	3.94	4.27	4.56	6.33
M 3	3.29	4.1	4.1	6.99
M 4	3.38	4	4.08	5.44
M 5	3.03	3.71	3.71	5.73
M 6	3.3	3.93	4.2	6.43
M 7	3.17	3.79	3.94	5.63
M 8	2.99	3.99	4.16	5.42
M 9	3.8	4.51	4.67	5.55
M 10	3.1	4.21	4.04	6.17
M 11	3.2	4	3.96	5.33
M 12	3	3.73	3.74	5.68
M 13	3.3	4.04	4.03	5.72
M 14	2.99	3.62	3.63	5.69

Models	**DL**	**SVM**	**MLP**	**SDT**
	80%	**80%**	**80%**	**80%**
M 1	3.2	3.4	4.02	6.51
M 2	3.99	4	4.25	5.36
M 3	2.7	3.38	3.25	7.28
M 4	3.15	3.42	3.85	6.51
M 5	2.9	3.71	3.4	6.51
M 6	3.03	3.27	3.58	5.43
M 7	3.06	3.4	3.86	6.51
M 8	2.41	3.93	3.42	6.51
M 9	3.79	3.9	3.93	5.58
M 10	2.82	3.6	3.34	5.22
M 11	3.24	3.38	3.65	6.51
M 12	2.5	3.21	3.84	6.51
M 13	3.34	3.53	3.38	6.56
M 14	2.61	3.38	3.18	6.51

Models	DL	SVM	MLP	SDT
	70%	**70%**	**70%**	**70%**
M 1	3.67	4.06	3.82	4.54
M 2	4.02	4.53	4.5	6
M 3	3.67	4.2	3.81	6.85
M 4	3.48	3.82	3.79	4.37
M 5	3.24	3.6	3.87	4.86
M 6	3.71	4.09	3.88	6
M 7	3.52	3.63	3.83	4.63
M 8	3.47	3.83	3.69	4.8
M 9	4.22	4.33	4.43	5.94
M 10	3.89	4.24	3.96	7.32
M 11	3.25	3.75	3.61	4.63
M 12	3.19	3.66	3.57	4.86
M 13	3.34	4.03	3.89	5.91
M 14	2.95	3.62	3.66	4.63

From these Tables we can see that the best performance has been given from split ratio 80%-20% with 2.41R including: physiological variables with BMI, PFA-1 while in split ratio 70%-30% it is 2.95R including; physiological variables with (Height, Weight), PFA-1, in 10-fold it is 2.95R including physiological variables with BMI, PFA-1, PFA-2, PAR.

Table 4.9. Test results for Maximal prediction models

Models	DL	SVM	MLP	SDT
	10-fold	10-fold	10-fold	10-fold
M 1	5.2	5.75	5.97	6.67
M 2	5.05	5.2	5.57	6.34
M 3	5.1	5.65	5.88	6.73
M 4	4.25	4.49	4.76	5.68
M 5	4.95	5.21	5.47	6.07
M 6	4.9	5.2	5.72	6.03
M 7	4.1	4.27	4.62	5.54
M 8	5.2	5.63	5.87	6.51
M 9	4.35	4.44	4.62	5.64
M 10	3.99	4.53	4.58	5.63
M 11	5.12	5.2	5.66	6.31
M 12	4.09	4.29	4.5	5.71
M 13	4.38	4.39	4.58	5.73
M 14	4.17	4.2	4.48	5.68
M 15	4.03	4.24	4.59	5.77
M 16	5.3	5.82	5.95	6.51
M 17	5.1	5.37	5.56	6.34
M 18	5	5.78	5.91	6.54
M 19	4	4.5	4.6	5.65
M 20	5.2	5.37	5.52	6.23
M 21	5.03	5.4	5.6	6.22
M 22	4.3	4.36	4.47	5.65
M 23	5.1	5.75	5.92	6.38
M 24	4.2	4.45	4.5	5.64
M 25	3.99	4.54	4.59	5.57
M 26	5.05	5.39	5.65	6.42
M 27	4	4.28	4.61	5.64
M 28	4.4	4.44	4.66	5.74
M 29	4.13	4.2	4.47	5.68
M 30	4.01	4.22	4.6	5.67

In 10-fold the best performance has been given from 3.99R including: physiological variables with (Height, Weight), MX-RPE, MX-TIME.

Models	DL	SVM	MLP	SDT
	80%	**80%**	**80%**	**80%**
M 1	5.72	6.49	6.21	7.1
M 2	5.66	5.84	6.02	6.97
M 3	6.15	6.5	6.34	7.8
M 4	4.63	5.1	4.85	5.76
M 5	5.66	5.98	6.18	6.96
M 6	5.85	6.56	6.51	7.8
M 7	3.99	5.18	4.85	5.87
M 8	5.24	6.21	5.81	7.27
M 9	4.32	4.97	4.88	5.83
M 10	4.62	4.77	4.75	6.04
M 11	5.13	5.95	6.24	7.27
M 12	4.24	4.79	5.09	5.92
M 13	4.58	4.87	4.81	6.1
M 14	4.26	4.66	4.99	6.09
M 15	4.39	4.75	4.73	6.17
M 16	6.09	6.42	6.59	7.6
M 17	5.83	6	6.35	7.46
M 18	6.03	6.71	6.69	7.64
M 19	4.4	4.92	5.08	6.08
M 20	5.7	5.97	6.42	7.45
M 21	5.93	6.59	6.52	7.64
M 22	4.61	5.02	4.8	6.08
M 23	5.47	6.29	6.09	7.5
M 24	4.07	4.9	4.75	6.08
M 25	4.33	4.73	4.8	6.23
M 26	5.62	6.19	6.17	7.59
M 27	4.4	4.85	4.68	6.08
M 28	4.38	4.71	4.73	6.23
M 29	4.37	4.77	4.75	6.23
M 30	4.47	4.6	4.83	6.23

In 80%-20% the best performance has been given from 3.99R with: physiological variables with (Height, Weight), MX-HE, MX-TIME.

Models	DL	SVM	MLP	SDT
	70%	**70%**	**70%**	**70%**
M 1	5.62	6.03	5.89	6.75
M 2	5.67	5.78	6.04	6.62
M 3	5.86	6.01	6.65	6.68
M 4	4.34	4.62	4.34	5.45
M 5	5.45	5.64	5.82	6.62
M 6	5.51	5.92	5.84	6.75
M 7	3.93	4.34	4.44	5.39
M 8	5.25	5.72	6.05	6.62
M 9	4.29	4.59	4.29	5.43
M 10	4.33	4.46	4.45	5.45
M 11	5.71	5.74	5.92	6.62
M 12	4.39	4.66	4.39	5.58
M 13	4.25	4.47	4.3	5.39
M 14	4.01	4.7	4.54	5.43
M 15	4.31	4.66	4.4	5.58
M 16	5.82	6.17	6.08	6.65
M 17	5.66	5.88	6.16	6.46
M 18	6.02	6.08	6.33	6.51
M 19	4.26	4.4	4.35	6.07
M 20	5.76	5.82	6.22	6.86
M 21	5.92	6.08	6.18	6.65
M 22	4.2	4.6	4.21	6.02
M 23	5.71	5.94	6.15	6.54
M 24	4.15	4.6	4.23	5.82
M 25	4.15	4.51	4.25	6.05
M 26	5.5	5.83	6.26	6.85
M 27	4.3	4.42	4.3	6.02
M 28	4.19	4.37	4.32	6.02
M 29	4.07	4.46	4.04	5.82
M 30	4.12	4.49	4.15	6.02

The best performance has been given from split ratio 70%-30% with 3.93R including: physiological variables with (Height, Weight), MX-HR, MX-TIME.

Table 4.10. Test results for non-exercise prediction models

Models	DL	SVM	MLP	SDT
	10-fold	10-fold	10-fold	10-fold
M 1	4.95	4.96	5.23	5.95
M 2	5.01	5.21	5.58	5.88
M 3	4.7	4.71	4.99	5.46
M 4	4.86	5.02	5.37	6
M 5	5.02	5.22	5.52	5.66
M 6	4.67	4.69	5.04	5.24

Models	DL	SVM	MLP	SDT
	80%	80%	80%	80%
M 1	4.97	5.6	5.52	6.54
M 2	5.07	6.02	6.26	6.52
M 3	4.92	5.32	5.38	6.34
M 4	4.77	5.67	5.93	6.57
M 5	5.33	6.05	6.41	6.15
M 6	4.39	5.32	5.46	6.4

Models	DL	SVM	MLP	SDT
	70%	70%	70%	70%
M 1	4.82	5.48	5.45	6.28
M 2	5.24	5.53	5.79	6.01
M 3	4.83	5.21	5.4	5.56
M 4	4.92	5.5	5.89	6.43
M 5	5.27	5.43	5.68	5.89
M 6	4.97	5.13	5.28	5.57

From these Tables we can see that the best performance has been given from split ratio 80%-20% with 4.39R including: physiological variables with BMI, AC, BF, while in 10-fold it is 4.67R including: physiological variables with BMI, AC, BF, in 70%-30% it is 4.82R with: physiological variables with (Height, Weight), AC.

Figure 4.1 through figure 4.17 show the prediction models' Arithmetical mean

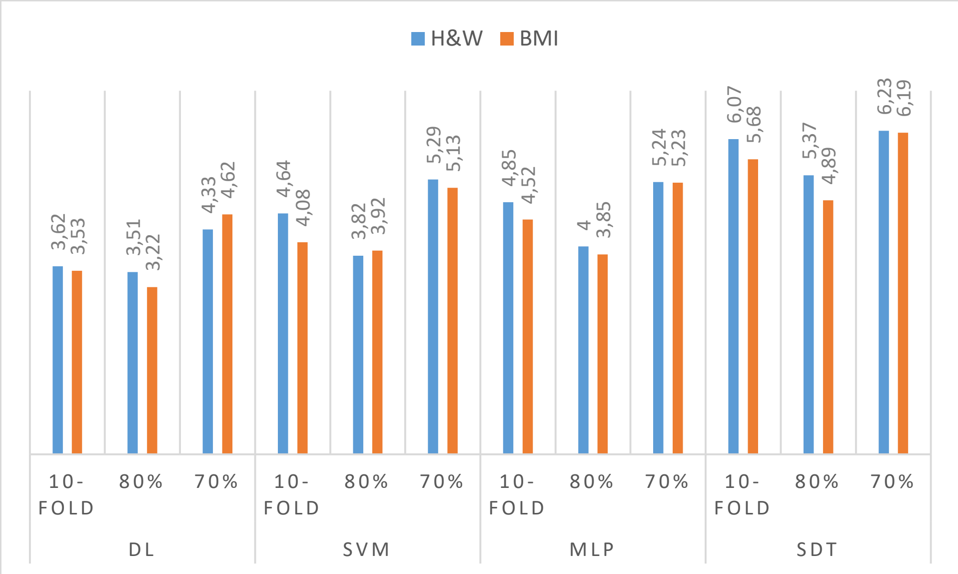


Figure 4.1. Arithmetical mean in RMSE's of Sub_Maximal predictor variables

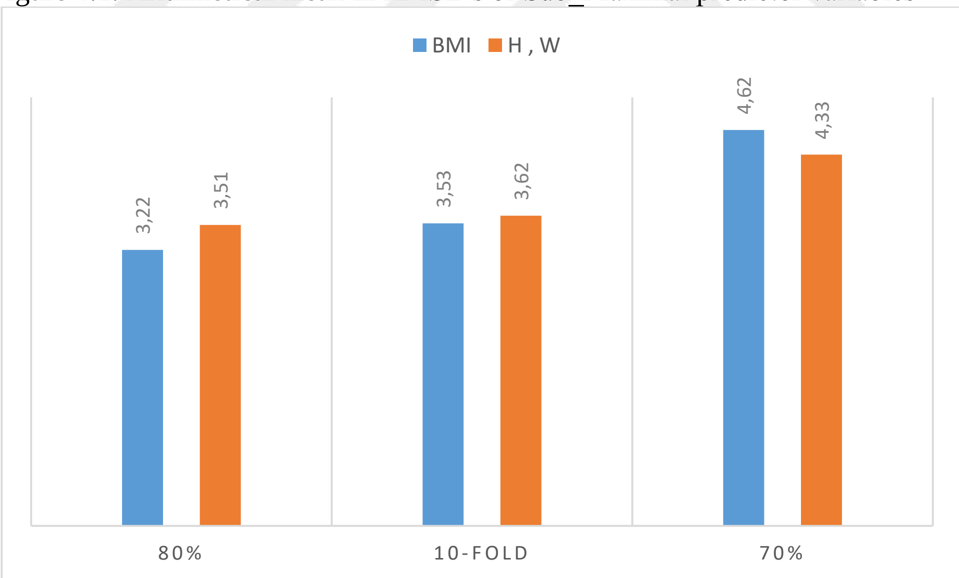
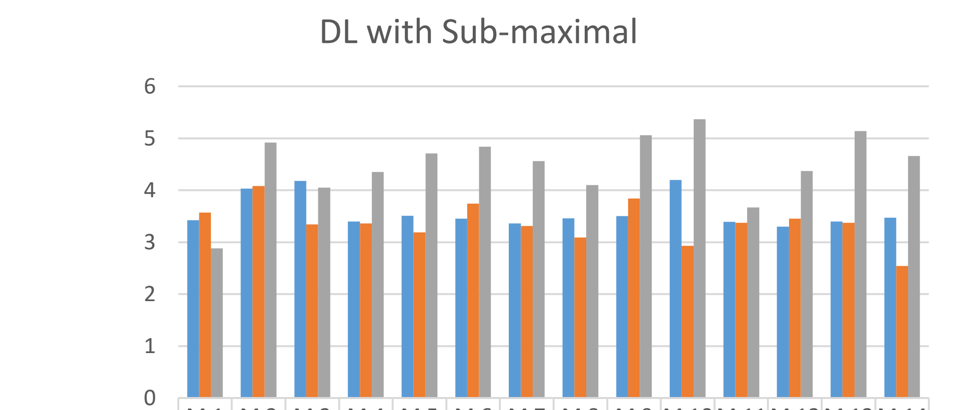


Figure 4.2. Arithmetical mean in RMSE's of DL methods with Sub-Maximal predictor variables



	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14
DL 10-fold	3,42	4,03	4,18	3,4	3,51	3,45	3,36	3,46	3,5	4,2	3,39	3,3	3,4	3,47
DL 80%	3,57	4,08	3,34	3,36	3,19	3,74	3,31	3,09	3,84	2,93	3,37	3,45	3,37	2,54
DL 70%	2,88	4,92	4,05	4,35	4,71	4,84	4,56	4,1	5,06	5,37	3,67	4,37	5,14	4,66

Figure 4.3. Arithmetical mean in RMSE's with all DL models

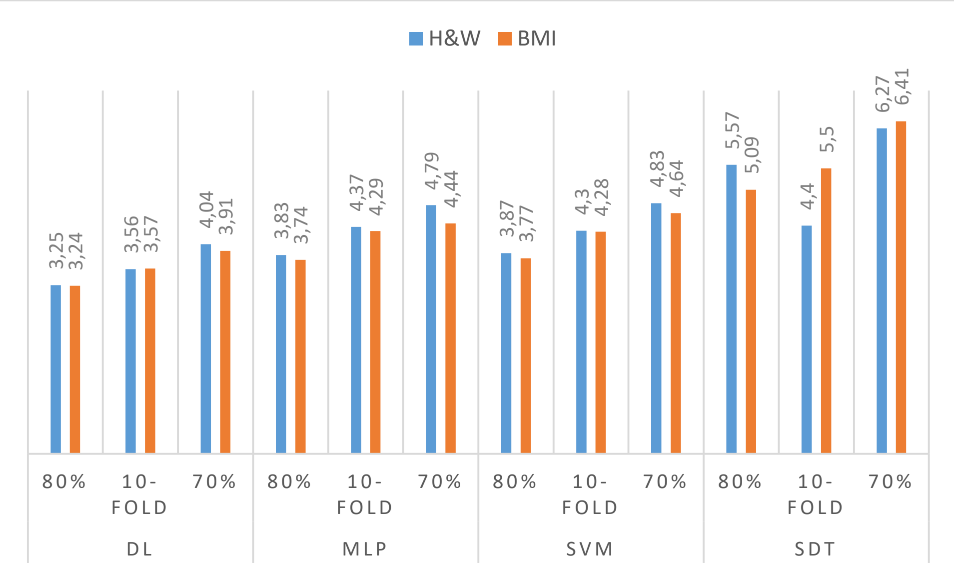


Figure 4.4. Arithmetical mean in RMSE's of Maximal predictor variables

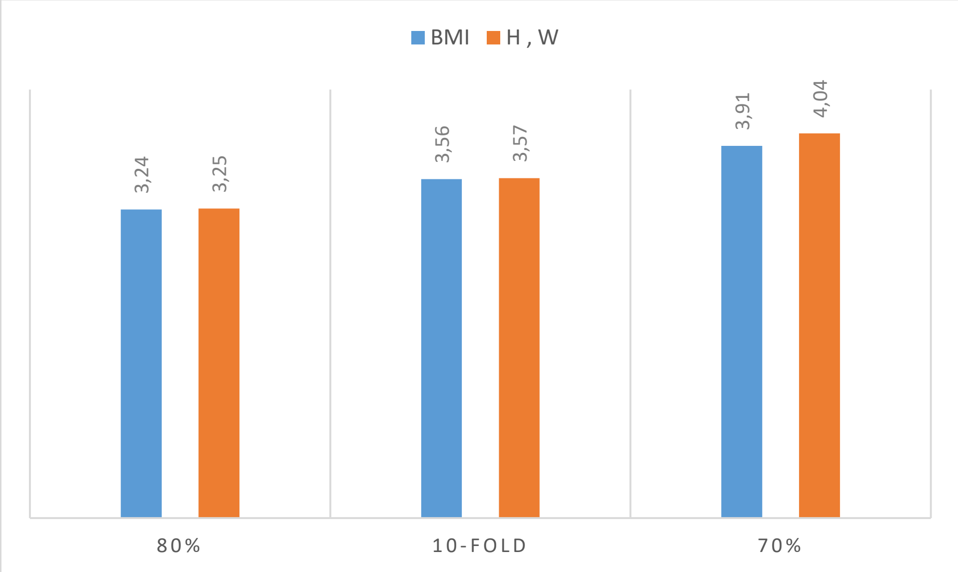
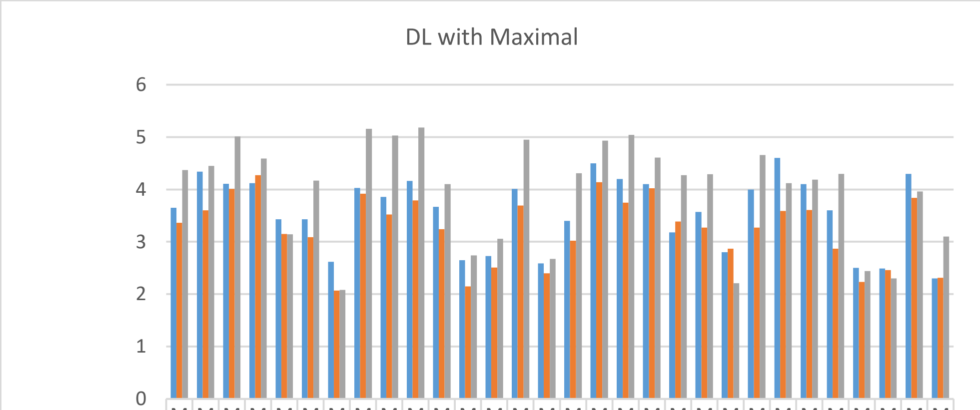


Figure 4.5. Arithmetical mean in RMSE's of DL methods with Maximal predictor variables



	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 12	M 22	M 23	M 24	M 25	M 26	M 27	M 28	M 29	M 30
DL 10-fold	4	4	4	4	3	3	3	4	4	4	4	3	3	4	3	3	5	4	4	3	4	3	4	5	4	4	3	2	4	2
DL 80%	3	4	4	4	3	3	2	4	4	4	3	2	3	4	2	3	4	4	4	3	3	3	3	4	4	3	2	2	4	2
DL 70%	4	4	5	5	3	4	2	5	5	5	4	3	3	5	3	4	5	5	5	4	4	2	5	4	4	4	2	2	4	3

Figure 4.6. Arithmetical mean in RMSE's of DL with all models

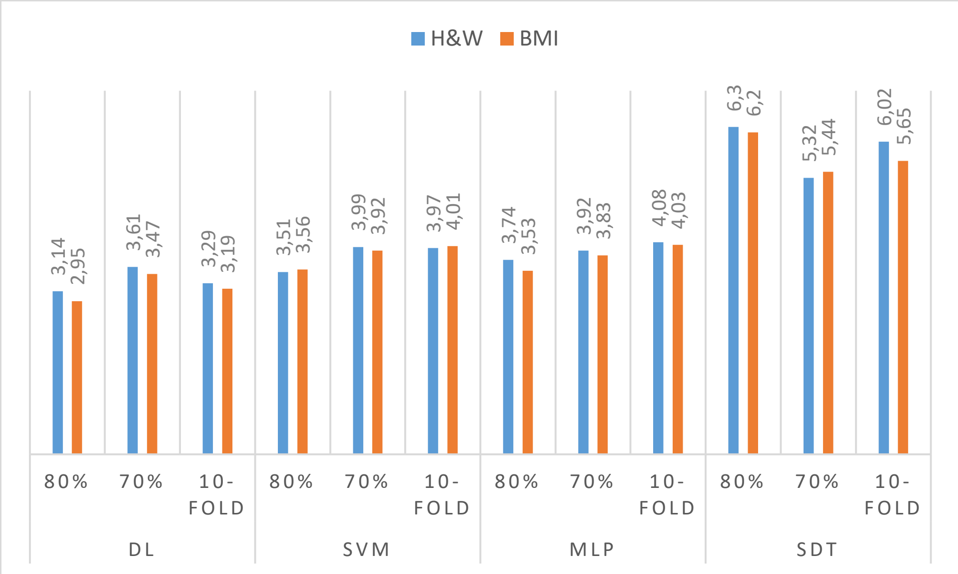


Figure 4.7. Arithmetical mean in RMSE's of Non-Exercise Predictor variables

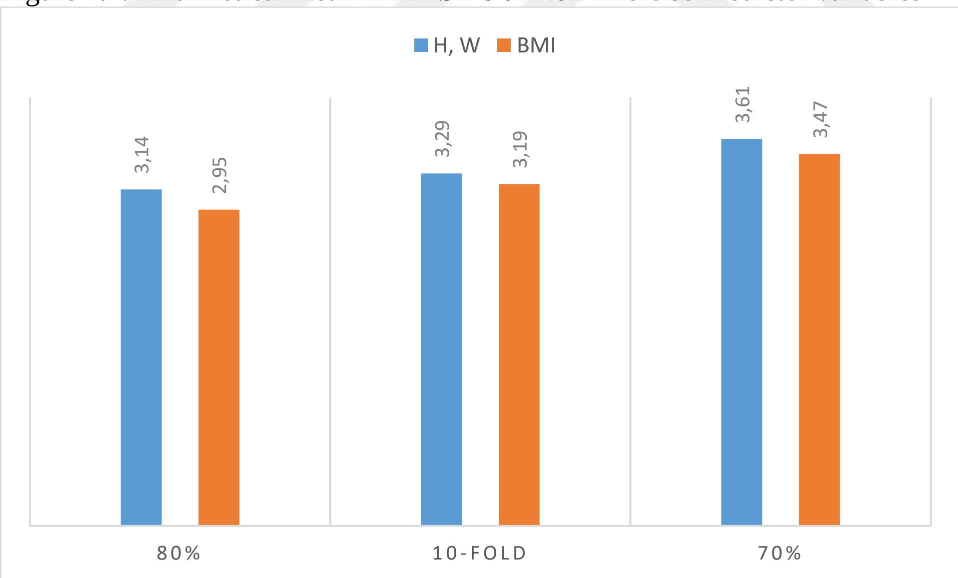
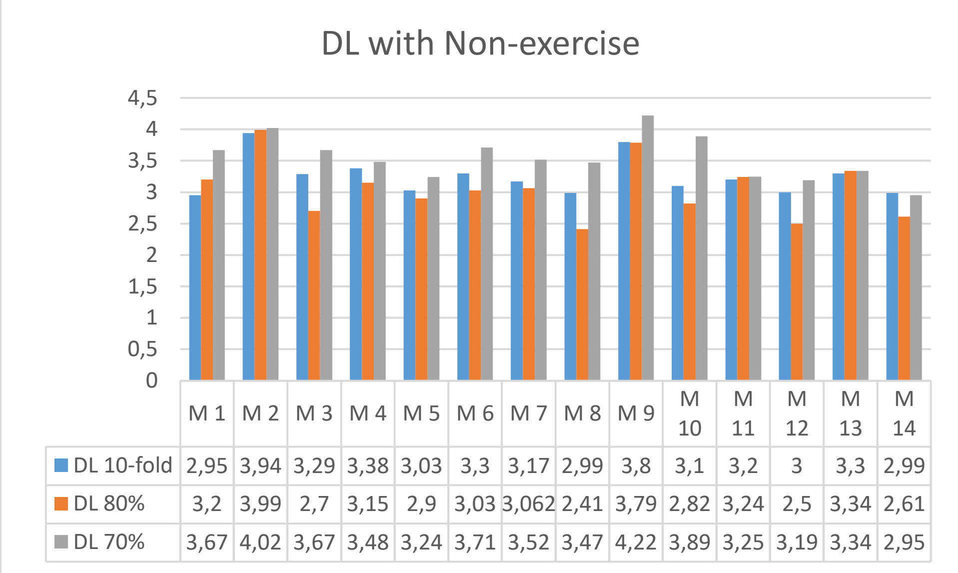


Figure 4.8. Arithmetical mean in RMSE's of DL method with Non-exercise predictor variables



	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14
DL 10-fold	2,95	3,94	3,29	3,38	3,03	3,3	3,17	2,99	3,8	3,1	3,2	3	3,3	2,99
DL 80%	3,2	3,99	2,7	3,15	2,9	3,03	3,062	2,41	3,79	2,82	3,24	2,5	3,34	2,61
DL 70%	3,67	4,02	3,67	3,48	3,24	3,71	3,52	3,47	4,22	3,89	3,25	3,19	3,34	2,95

Figure 4.9. Arithmetical mean in RMSE's of DL with all models

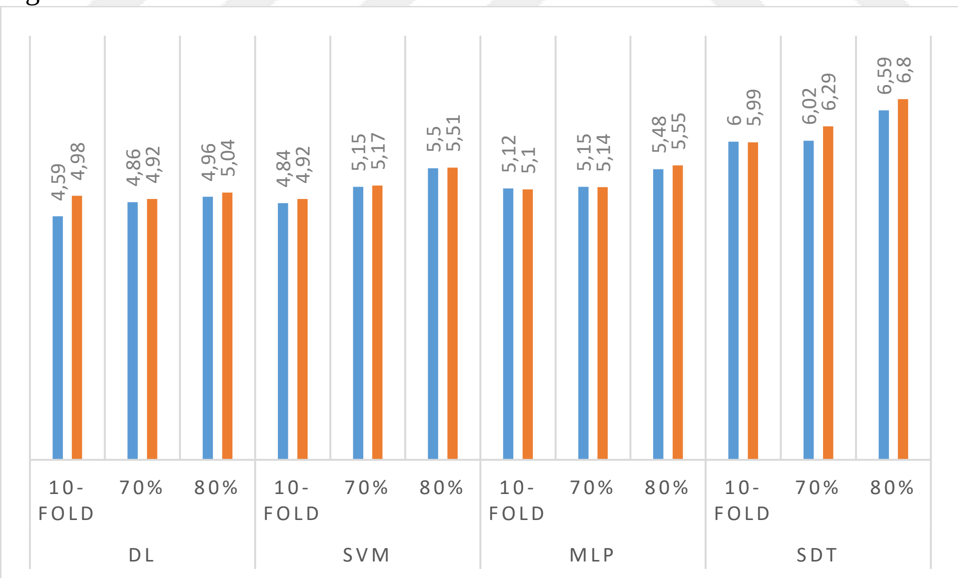


Figure 4.10. Arithmetical mean in RMSE's of Maximal predictor variables

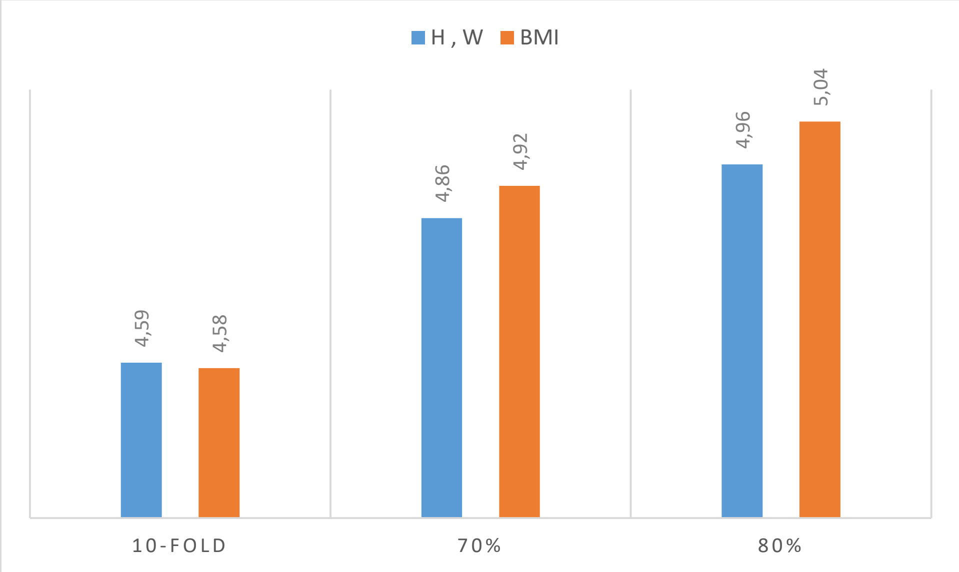
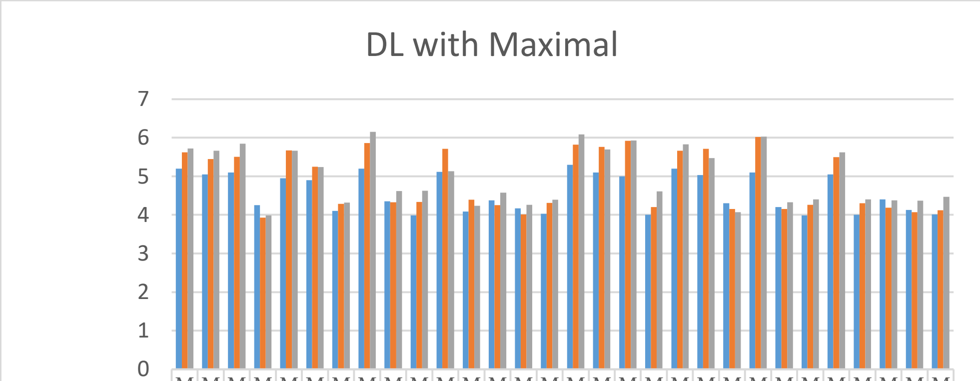


Figure 4.11. Arithmetical mean in RMSE's of DL method with Maximal predictor variables



	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 12	M 22	M 23	M 24	M 25	M 26	M 27	M 28	M 29	M 30
DL 10-fold	5	5	5	4	5	5	4	5	4	4	5	4	4	4	4	5	5	5	4	5	5	4	5	4	4	5	4	4	4	4
DL 70%	6	5	6	4	6	5	4	6	4	4	6	4	4	4	4	6	6	6	4	6	6	4	6	4	4	6	4	4	4	4
DL 80%	6	6	6	4	6	5	4	6	5	5	5	4	5	4	4	6	6	6	5	6	5	4	6	4	4	6	4	4	4	4

Figure 4.12. Arithmetical mean in RMSE's of DL with all models

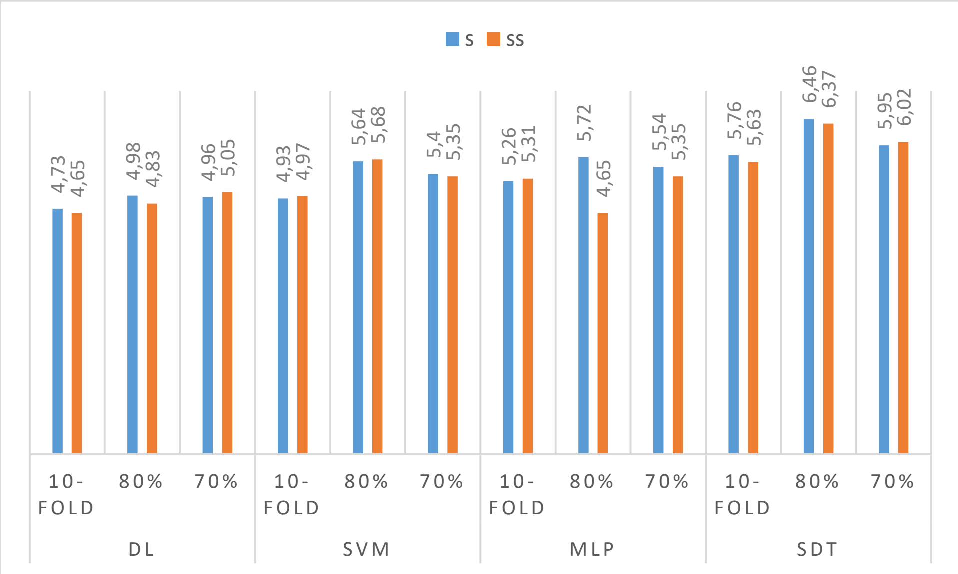


Figure 4.13. Arithmetical mean in RMSE's of Non-exercise predictor variable

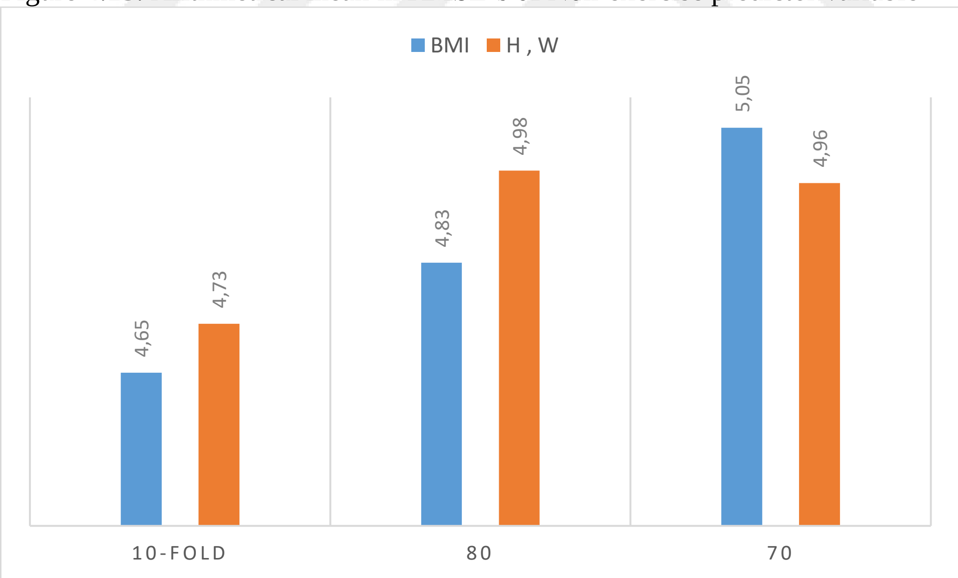
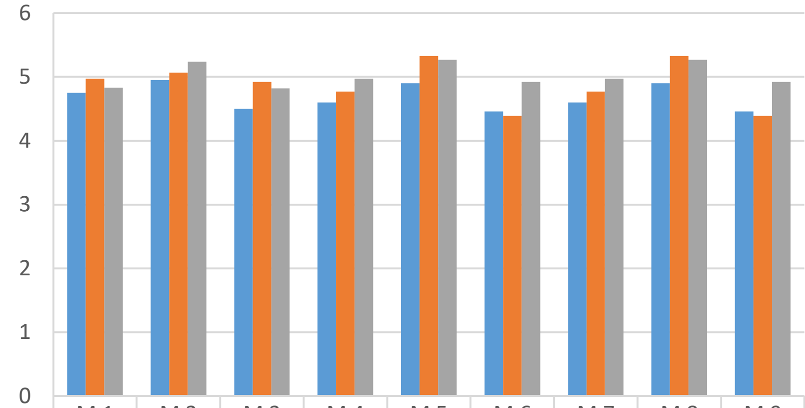


Figure 4.14. Arithmetical mean in RMSE's of DL with maximal predictor variables



	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9
DL 10-fold	4,75	4,95	4,5	4,6	4,9	4,46	4,6	4,9	4,46
DL 80%	4,97	5,07	4,92	4,77	5,33	4,39	4,77	5,33	4,39
DL 70%	4,83	5,24	4,82	4,97	5,27	4,92	4,97	5,27	4,92

Figure 4.15. Arithmetical mean in RMSE's of DL with all models

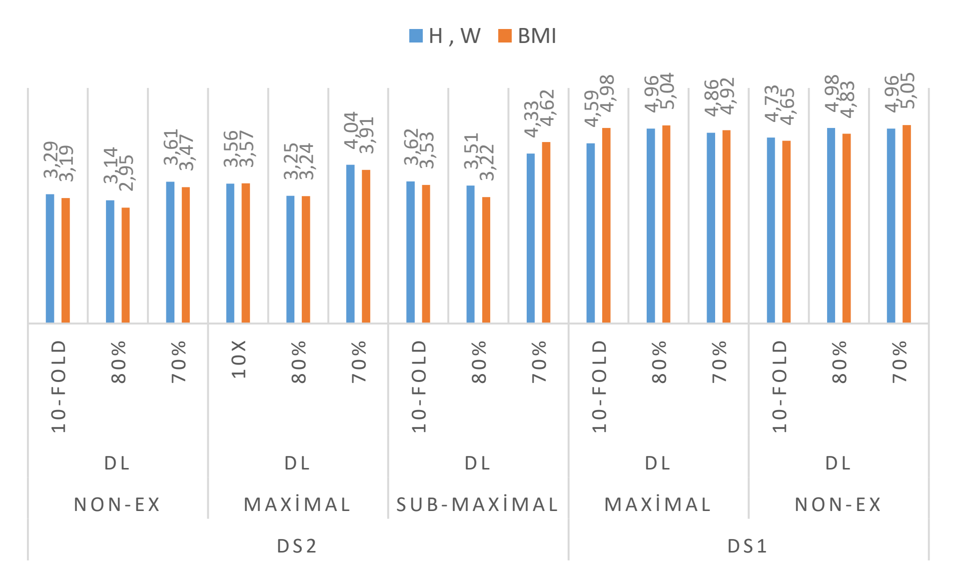


Figure 4.16. Arithmetical mean for DL RMSE's with all models and for each split ratio

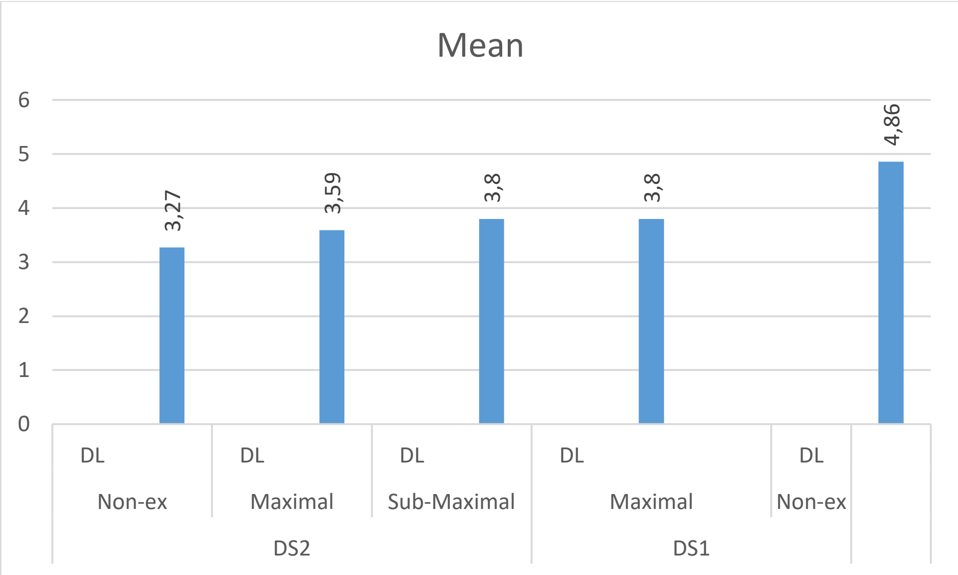


Figure 4.17. Arithmetical mean for DL RMSE's with all models

4.3 Discussion

Discussion based on the results

In this study, the promising results revealed that DL-based prediction models have the best results with the lowest error comparing with SVM-based models, MLP-based models, and SDT-based models respectively. Furthermore:

- DL-based models have an average of RMSE value of 3.86 *R*. On the other hand, the prediction models including BMI predictor variables have RMSE average value of 4.07*R*, while the prediction models including Height, Weight have RMSE average value of 4.09 *R*.
- The 80% - 20% split ratio has an average of RMSE value of 3.91 *R*, while 10-fold cross validation has an average of RMSE value of 4.37 *R*.
- MLP-based models have an average of RMSE value of 4.64 *R*. On the other hand, the prediction models including BMI predictor variables have RMSE average value of 4.56 *R*, while the prediction models including Height, Weight have RMSE average value of 4.72 *R*.
- SVM-based models have an average of RMSE value of 4.61 *R*. On the other hand, the prediction models including BMI predictor variables have RMSE average value of 4.59 *R*, while the prediction models including Height, Weight have RMSE average value of 4.67 *R*.
- Finally, SDT-based models have an average of RMSE value of 5.95 *R*. On the other hand, the prediction models including BMI predictor variables have RMSE average value of 5.92 *R*, while the prediction models including Height, Weight have RMSE average value of 5.97 *R*.
- The ranking of the methods in terms of their prediction performance (from the best to the worst) is given as DL, SVM, MLP, and SDT.

- The models including height and weight predictor variables show lower error than the one including BMI.

Discussions on the Different Predictor Variables of the Prediction Models Based on best and worst model:

- **The Maximal models**: The prediction models with MX-ES and MX-GRD including gender, height, weight, and age yields the lowest *RMSE* value of 2.07*R*, while the prediction model with MX-HR and MX-RPE including gender, age, height, weight shows the worst result in terms of *RMSE* with 5.16*R*.
- **The Non-Exercise Models**: prediction models with PFA-1, PFA-2, PAR including gender, BMI, and age yields the lowest *RMSE* value of 2.61*R*, while the prediction model with variable PFA-2 including age, gender, height, and weight shows the highest error in terms of *RMSE* with 4.09*R*.
- **The Sub-Maximal Models:** prediction model with SM-ES including gender, height, weight, and age yields the lowest *RMSE* value of 2.88*R*, while the prediction model with variable SM-S including age, gender, and BMI shows the highest error in terms of *RMSE* with 5.37*R*.
- **The Maximal models:** prediction model with MX-HR, MX-TIME predictors including age, gender, height, weight yields the lowest *RMSE* value of 3.93*R* which is the lowest *RMSE* through all the models. On the other hand, predictor MX-RPE has the highest *RMSE* with 6.15*R* value which is the highest error among all the models. Therefore, in this maximal group, we obtained the lowest and the highest error rate in terms of accuracy.

- **The Non-Exercise Models**: The prediction model with AC and BF including gender, BMI, and age yields the lowest *RMSE* value of 4.39*R* when the prediction model with predictor BF including physiological variables, and BMI gives the highest error in terms of *RMSE* with 5.33*R*.

Based on models with each split ratio:

- **The Sub-Maximal Models with different split ratio:** the best performance has been given from split ratio 80%-20% with 2.54*R* including: physiological variables with BMI, SM-ES, SM-HR, SM-S, while in split ratio 70%-30% it is 2.8*R* for: physiological variables with (Height, Weight), and SM-ES, in 10-fold it is 3.3*R* for: physiological variables with BMI, SM-ES, SM-S.
- **The Maximal Models with different split ratio:** the best performance has been given from split ratio 80%-20% with 2.07*R* including: physiological variables with (Height, Weight), MX-ES, while in split ratio 70%-30% it is 2.08 *R* also for: physiological variables with (Height, Weight), MX-ES, in 10-fold it is 2.3*R* for: physiological variables with (Height, Weight), MX-HR, MX-RPE, MX-GRD.
- **The Non-exercise Models with different split ratio:** the best performance has been given from split ratio 80%-20% with 2.41*R* including: physiological variables with BMI, PFA-1 while in split ratio 70%-30% it is 2.95*R* including; physiological variables with (Height, Weight), PFA-1, in 10-fold it is 2.95*R* including physiological variables with BMI, PFA-1, PFA-2, PAR.
- **The Maximal 2nd Models with different split ratio:** the best performance has been given from split ratio 70%-30% with 3.93*R*

including: physiological variables with (Height, Weight), MX-HR, MX-TIME, while in split ratio 80%-20% it is 3.99*R* with: physiological variables with (Height, Weight), MX-HE, MX-TIME, in 10-fold it is 3.99*R* including: physiological variables with (Height, Weight), MX-RPE, MX-TIME.

- **The Maximal Models with different split ratio:** the best performance has been given from split ratio 80%-20% with 4.39*R* including: physiological variables with BMI, AC, BF, while in 10-fold it is 4.67*R* including: physiological variables with BMI, AC, BF, in 70%-30% it is 4.82*R* with: physiological variables with (Height, Weight), AC.

Based on split ratio performance:

- **10-fold cross:** achieved the lowest RMSE with Non-exercise predictor variables and BMI physiological variable with a 3.19*R*, after that with Sub-maximal predictor variables for a 3.53*R* also with BMI physiological variable.
- **80%-20% split ratio:** achieved the lowest RMSE with Non-exercise predictor variables and BMI physiological variable with a 2.95*R*, after that with Maximal predictor variables for a 3.24*R* also with BMI physiological variable.
- **70%-30% split ratio:** achieved the lowest RMSE with Non-exercise predictor variables and BMI physiological variable with a 3.41*R*, after that with Maximal predictor variables for a 3.91*R* also with BMI physiological variable.

Based on Parameters:

As we mentioned that the number of hidden layers can’t be specified in advanced, it depends on the dataset size and the problem complexity.

In the first dataset: larger number of hidden layer have been used up to 30 hidden layers especially with maximal models.

In the second dataset: up to 9 hidden layers have been used with maximal model.

Increasing the hidden layers and neurons numbers in small dataset could increase the network complexity and causes undesired output and underfitting problem. Also using quasi network function in small dataset with large number of hidden layers causes a stuck point during the algorithm running time. For that reason, DL with large dataset gives an ability to use large number of hidden layers, otherwise it would not give optimized results.

5 CONCLUSION

A lot of studies have been conducted in the last years to predict VO_2max of various target audiences. However, DL-based VO_2max prediction models can be safely used in the prediction domain. In order to research and conduct the thesis, there were four main machine learning methods to use DL, SVM, MLP, and SDT. The dataset has been split into training and test sets using 70% - 30% and 80% - 20%split ratios, in addition to 10-fold cross-validation. The first dataset contains Non-Exercise variables group; (age, gender, height, weight, BMI, body fat, activity code, and lean body mass), and Maximal variables group; (age, gender, height, weight, BMI, heart rate, respiratory exchange ratio, self-reported rating of perceived exertion, and time to exhaustion). The second dataset contains Non-Exercise variables group; (age, gender, height, weight, BMI, PFA-1, PFA-2, and PAR), Maximal variables group; (age, gender, height, weight, BMI, ending speed, heart rate, self-reported rating of perceived exertion, and grade), and Sub-Maximal variables group; (age, gender, height, weight, BMI, ending speed, heart rate, stage). We had considered that [age, gender, (height, weight) || BMI] are fixed variables, and the test predictor variables had been distributed in many groups and different combinations. The first dataset includes data of 439 participants. Participants were required from Amherst and Worcester, Massachusetts, while the second dataset includes one hundred participants (50 males, 50 female). Participants were required from the Y-Be-Fit Wellness Program at Brigham Young University and employees from the LDS Hospital in Salt Lake City, Utah. For computing, the prediction root means squared errors (*RMSE*) has been used.

Multiple conclusions can be made based on this study results. First of all, DL-based prediction results show that deep learning can be safely used to predict VO_2max with acceptable accuracy. Secondly, the obtained *RMSE's* average rate by DL-based prediction models ranges between 3.27% and 4.86%.

Deep Learning shows the best performance while SDT was the worst for the VO_2max prediction. The prediction models of the Maximal (1^{st} dataset) yield lower *RMSE's* than the Non-Exercise (2^{nd} dataset) yield lower *RMSE's* than Maximal and Sub-Maximal models, regardless of which split ratio is considered. Finally, BMI predictor variable shows a higher error in the prediction models that have used Height and Weight instead.

REFERENCES

Abut, F., & Akay, M. F. (2015). Determination of the variables affecting the maximal oxygen uptake of cross-country skiers by using machine learning and feature selection algorithms. 2015 23nd Signal Processing and Communications Applications Conference (SIU).

Acikkar M, Akay MF, Akturk E, Gulec M. Intelligent regression techniques for non-exercise prediction of VO_2max. In: 21st Signal Processing and Communications Applications Conference (SIU).

Acikkar M, Akay MF, George JD, Delil M, Aktürk E. Artificial neural network models for predicting maximum oxygen uptake from submaximal exercise involving walking, jogging or running. In: International Symposium on Electrical & Electronics Engineering and Computer Systems. North Cyprus; 2012:17-21.

Akay MF, Abut F, Daneshvar S, Heil D. Prediction of Maximal Oxygen Uptake of Cross-Country Skiers Using Different Regression Methods. In: International Symposium on Engineering, Artificial Intelligence & Applications. North Cyprus; 2014:18-19.

Akay MF, Akgol D, George J. Support Vector Machines Combined with Feature Selection for Prediction of Maximal Oxygen Uptake. In: International Symposium on Engineering, Artificial Intelligence & Applications. North Cyprus; 2014:3-4.

Akay MF, Ghahremanlou NS, Aktürk E, George JD, Aktarla E. Performance comparison of different regression methods for VO_2max prediction. In: 7th International Symposium on Electrical & Electronics Engineering and Computer Systems. North Cyprus; 2012:7-11.

Akay MF, Inan C, Bradshaw DI, George JD. Support vector regression and multilayer feedforward neural networks for non-exercise prediction of

VO_2max. Expert Syst. Appl. 2009; 36(6):10112-10119. doi:10.1016/j.eswa.2009.01.009.

Akay MF, Özsert G, Genc E, George J. Development of New Non-Exercise Maximal Oxygen Uptake Models by Using Support Vector Machines. In: International Symposium on Engineering, Artificial Intelligence & Applications. North Cyprus; 2014:11-12.

Akay MF, Özsert G, George J. Prediction of Maximal Oxygen Uptake Using Support Vector Machines from Submaximal Data. DEÜ Eng. Fac. - Eng. Sci. J. 2014; 16(483):42-48.

Akay MF, Zayid EIM, Aktürk E, George JD. Artificial neural network-based model for predicting VO_2max from a submaximal exercise test. Expert Syst. Appl. 2011; 38(3):2007-2010. doi:10.1016/j.eswa.2010.07.135.

Akay, M. F., Cetin, E., Yarim, I., Bozkurt, O., & Ozciloglu, M. M. (2017). Development of novel maximal oxygen uptake prediction models for Turkish college students using machine learning and exercise data. 2017 9th International Conference on Computational Intelligence and Communication Networks (CICN).

Aktürk E, Akay MF. Prediction of Maximal Oxygen Uptake Using Multilayer Perceptron Combined with Feature Selection. In: 7th Engineering and Technology Symposium. Ankara; 2014:3-6.

Bandyopadhyay A. Validity of 20-meter multi-stage shuttle run test for estimation of maximum oxygen uptake in male university students. Indian J. Physiol. Pharmacol. 2014.

Billinger SA, Swearingen E, McClain M, Lentz AA, Good MB. Recumbent stepper submaximal exercise test to predict peak oxygen uptake. Med. Sci. Sports Exerc. 2012; 44(8):1539-1544. doi:10.1249/MSS.0b013e31824f5be4.

Bruce, R.A., Kusumi, F., & Hosmer, D. (1973). Maximal oxygen intake and monographic assessment of functional aerobic impairment in cardiovascular disease. American Heart Journal, 85, 546-562.

Cao Z-B, Miyatake N, Aoyama T, Higuchi M, Tabata I. Prediction of maximal oxygen uptake from a 3-minute walk based on gender, age, and body composition. 2014.

Cao Z-B, Miyatake N, Higuchi M, Miyachi M, Ishikawa-Takata K, Tabata I. Predicting VO_2max with an objectively measured physical activity in Japanese women.

Chatterjee P, Banerjee AK, Das P, Debnath P. A regression equation for the estimation of maximum oxygen uptake in Nepalese adult females. Asian J. Sports Med. 2010; 1(1):41-5.

Coquart JBJ, Eston RG, Grosbois J-M, et al. Prediction of peak oxygen uptake from age and power output at RPE 15 in obese women. Eur. J. Appl. Physiol. 2010; 110(3):645-9. Doi: 10.1007/s00421-010-1524-5.

Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge, UK: Cambridge University Press.

Daros LB, Osiecki R, Dourado AC, et al. Maximum Aerobic Power Test for Soccer Players. J. Exerc. Physiol. 2012; 15(2):80.

George, J.D., Stone, W.J., & Burkett, L.N. (1997). Non-exercise VO_2max estimation for physically active college students. Medicine and Science in Sports and Exercise, 29(3), 415-423.

Goyal, V. Deolia, V. K., & Sharma, T. N. (2016). Neural Network Based Sliding Mode Control for Uncertain Discrete-Time Nonlinear Systems with Time-Varying Delay. International Journal of Computational Intelligence Research,12 (2), 125-138.

Greenhalgh, H.A., George, J.D., & Hager, R.L. (2001). Cross-validation of a quarter-mile walk test using two VO_2max regression models. Measurement in Physical Education and Exercise Science, 5 (3): 139-151.

Hunn HM, Lapuma PT, Holt DT. The Influence of Pre-Test Anxiety, Personality, and Exercise on VO_2max Estimation. J. Exerc. Physiol. 2002; 5(1):5-14.

IVANOVICH, E. V., and HAMID, K., 2014. An Alternative Approach for Assessing Sediment Impact on Aquatic Ecosystems Using Single Decision Tree (SDT). Journal of Water Sustainability, 4(3), 181–204.

Jang T-W, Park S-G, Kim H-R, Kim J-M, Hong Y-S, Kim B-G. Estimation of maximal oxygen uptake without exercise testing in Korean healthy adult workers. 2014.

John McCarthy, J. (2017). Homage to John McCarthy, the father of Artificial Intelligence (AI).

Machado FA, Denadai BS. Predição da potência aeróbiaVO_2max de crianças e adolescentes em teste incremental na esteira rolante. Mot. Rev. Educ. Física 2013; 19(1):126-132. Doi: 10.1590/S1980-65742013000100012.

Mahar MT, Guerrieri AM, Hanna MS, Kemble CD. Estimation of aerobic fitness from 20-m multistage shuttle run test performance. Am. J. Prev. Med. 2011; 41(4 Suppl 2): S117-23. doi:10.1016/j.amepre.2011.07.008.

Nielson, D.E. (2009). Predicting VO2max in college-aged participants using cycle ergometry and nonexercise measures. MSc thesis, Department of Exercise Sciences, Brigham Young University.

Schembre SM, Riebe DA. Non-exercise estimation of VO_2max using the International Physical.

Shenoy S, Tyagi B, Sandhu J, Sengupta D. Development of non-exercise based VO_2max prediction equation in college-aged participants in India. J. Sports Med. Phys.2014.

Silva G, Oliveira NL, Aires L, Mota J, Oliveira J, Ribeiro JC. Calculation and validation of models for estimating VO_2max from the 20-m shuttle run test in children and adolescents. Arch. Exerc. Heal. Dis. 2012; 3(1-2):145-152. doi:10.5628/aehd.v3i1-2.20.

Tönis TM, Gorter K, Vollenbroek-Hutten MMR, Hermens H. Comparing VO_2max determined by using the relation between heart rate and accelerometry with submaximal estimated VO_2max. J.SportsMed. Phys. Fitness 2012; 52(4):337-43.

Veronese da Costa A, Manoel da Cunha C, Melo de Oliveira SF, Lima de Albuquerque F, José de Sá Pereira Guimarães F, Barbosa TM. Validation of an equation for estimating maximal oxygen consumption of nonexpert adult swimmers.

Zakariás, G., Petrekanits, M., & Laukkanen, R. (2003). The validity of a 2-km walk test in predicting the maximal oxygen uptake in moderately active Hungarian men. European Journal of Sports Science, 3 (1): 1-8.

CURRICULUM

Hala ABDULKADER was born ın Syrıa, Aleppo ın 1992. She has two brothers. She studıed computer engıneerıng ın Aleppo unıversıty 2011-2014. Then she moved to Turkey wıth Famıly ın 2014 and contınued her study ın Toros Unıversıty, Mersın computer engıneerıng department. She graduated ın 2016. In 2017 she was accepted ın Cukurova unıversıty, Adana, Computer Engıneerıng Master program and had Prof. Dr. Mehmet Fatıh AKAY as a supervısor.