ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS

Eser YÜCEL

PREDICTION OF MAXIMUM OXYGEN UPTAKE (VO_2max) WITH SUBMAXIMAL AND QUESTIONNAIRE VARIABLES USING DIFFERENT REGRESSION METHODS

DEPARTMENT OF COMPUTER ENGINEERING

ADANA, 2014

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PREDICTION OF MAXIMUM OXYGEN UPTAKE (VO_2max) WITH SUBMAXIMAL AND QUESTIONNAIRE VARIABLES USING DIFFERENT REGRESSION METHODS

Eser YÜCEL

MSc THESIS

DEPARTMENT OF COMPUTER ENGINEERING

We certify that the thesis titled above was reviewed and approved for the award of degree of the Master of Science by the board of jury on 11/11/2014

.....	.....	.....
Assoc. Prof. Dr. M. Fatih AKAY	Asst. Prof. Dr. Çiğdem ACI	Asst. Prof. Dr. Zehan KESİLMİŞ
SUPERVISOR	MEMBER	MEMBER

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

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

This study was supported by Research Projects Unit Ç.Ü. Project Number: FYL-2014-1961

Not: The usage of the presented specific declerations, tables, figures, and photographs either in this thesis or in any other reference without citiation is subject to "The law of Arts and Intellectual Products" number of 5846 of Turkish Republic

ABSTRACT

MSc THESIS

PREDICTION OF MAXIMUM OXYGEN UPTAKE (VO_2max) WITH SUBMAXIMAL AND QUESTIONNAIRE VARIABLES USING DIFFERENT REGRESSION METHODS

Eser YÜCEL

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

Supervisor : Assoc. Prof. Dr. M. Fatih AKAY
Year: 2014, Pages: 73
Jury : Assoc. Prof. Dr. M. Fatih AKAY
: Asst. Prof. Dr. Çiğdem ACI
: Asst. Prof. Dr. Zehan KESİLMİŞ

Maximal oxygen uptake (VO_2max) refers to the maximum amount of oxygen that an individual can utilize during intense or maximal exercise. The purpose of this thesis is to develop accurate VO_2max prediction models using submaximal and questionnaire variables. Regression methods such as Support Vector Machines (SVM), Multilayer Perceptron (MLP) and Multiple Linear Regression (MLR) have been used for developing VO_2max prediction models. The performance of prediction models has been evaluated by calculating their multiple correlation coefficients (*R*'s) and standard error of estimates (*SEE*'s). The results show that the accuracy of VO_2max prediction models based on submaximal and standard non-exercise variables could be significantly improved by including questionnaire variables in prediction models. The results of SVM models have been also compared with the ones obtained by MLP and MLR and it turned out that SVM-based VO_2max prediction models perform better (i.e. yield lower *SEE*'s and higher *R*'s) than the prediction models developed by other regression methods.

Key Words: Support Vector Machines, Multilayer Perceptron, Multiple Linear Regression, Maximal Oxygen Uptake

ÖZ

YÜKSEK LİSANS TEZİ

SUBMAKSİMAL VE ANKET DEĞİŞKENLERİ KULLANARAK MAKSİMUM OKSİJEN TÜKETİMİNİN FARKLI REGRESYON YÖNTEMLERİYLE TAHMİN EDİLMESİ

Eser YÜCEL

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Doç. Dr. M. Fatih AKAY
Yıl: 2014, Sayfa: 73
Jüri : Doç. Dr. M. Fatih AKAY
: Yrd. Doç. Dr. Çiğdem ACI
: Yrd. Doç. Dr. Zehan KESİLMİŞ

Maksimal oksijen tüketimi (VO_2max), kişinin yoğun ve maksimal egzersiz sırasında kullandığı maksimum oksijen miktarını ifade eder. Bu tezin amacı, submaksimal ve anket değişkenlerini kullanarak doğru Maksimal Oksijen Tüketimi (VO_2max) tahmin modelleri geliştirmektir. VO_2max tahmin modellerini geliştirmek için Destek Vektör Makineleri (SVM), Çoklu Doğrusal Regresyon (MLR) ve Çok Katmanlı Algılayıcı (MLP) gibi regresyon yöntemleri kullanılmıştır. Tahmin modellerin performansı çoklu korelasyon katsayısı (R) ve standart tahmin hatası (SEE) hesaplanarak belirlenmiştir. Sonuçlar göstermiştir ki, submaksimal ve standart egzersize dayalı olmayan değişkenlere anket değişkenlerinin eklenmesi VO_2max tahmin modellerinin doğruluğunu belirgin şekilde arttırmıştır. SVM modellerinin sonuçları MLP ve MLR modellerinden elde edilenler ile karşılaştırılmış ve SVM bazlı VO_2max tahmin modellerinin diğer regresyon modellerinden üretilen tahmin modellerinden daha iyi performans gösterdiği görülmüştür (Daha düşük SEE ve daha düşük R).

AnahtarKelimeler: Destek Vektör Makineleri, Çok Katmanlı Algılayıcı, Çoklu Doğrusal Regresyon, Maksimal Oksijen Tüketimi

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Assoc. Prof. Dr. M. Fatih AKAY, for his supervision guidance, encouragements, patience, motivation, useful suggestions and his valuable time for this work.

I would like to thank Assoc. Prof. Dr. James GEORGE for providing us with the datasets.

I would like to thank members of the MSc thesis jury, Asst. Prof. Dr. Çiğdem ACI and Asst. Prof. Dr. Zehan KESİLMİŞ, for their suggestions and corrections.

I would also like to thank Cukurova University Scientific Research Projects Center for supporting this work (Project no: FYL-2014-1961)**.**

Last but not the least, I would like to thank my family for their endless support and encouragements for my life and career.

CONTENTS **PAGE**

ABSTRACT I
ÖZ II
ACKNOWLEDGMENTS III
CONTENTS IV
LIST OF TABLES VI
LIST OF FIGURES VIII
LIST OF ABBREVIATONS X
l. INTRODUCTION 1
1.1. Maximal Oxygen Uptake (VO2max) 1
1.2. Assessing Maximal Oxygen Uptake 3
1.3. How is VO2max Measured? 5
1.4. Advantages and Disadvantages of Maximal Tests 5
1.5. Motivation of This Thesis 6
2. PREVIOUS WORKS 9
2.1. Submaximal prediction models based studies 9
2.2. Submaximal and questionnaire prediction models based studies 15
3. OVERVIEW OF METHODS 19
3.1. Support Vector Machines (SVMs) 19
3.1.1. Classification 20
3.1.2. Regression 22
3.1.3. Kernel Functions 23
3.2. Artifical Neural Networks (ANN) 25
3.2.1. Neural Networks Architectures 27
3.2.2. Hebbian Learning 29
3.2.3. Perceptron Learning 29
3.2.4. Backpropagation Learning 30
3.3. Multilayer Perceptron 31
3.4. Multiple Linear Regression (MLR) 35
3.4.1. Uses of Multiple Regression 36

3.5. Overview of Dataset 37
3.5.1. The Dataset 37
4. RESULTS AND DISCUSSION 41
5. CONCLUSION 65
6. REFERENCES 67
CURRICULUM VITAE 73

LIST OF TABLES **PAGE**

Table 1.1. Female and male normative data for VO2max (values in ml/kg/min) (Achten, and Jeukendrup, 2003) 2

Table 2.1. Summary of submaximal regression models to estimate VO2max 15

Table 2.2. Summary of submaximal and questionnaire variables models to estimate VO2max..... 18

Table 3.1. Descriptive statistics of the dataset 39

Table 4.1. Overview of VO2max prediction models along with the predictor variables for the data set (Model 1- Model 8) (Category 1)..... 42

Table 4.2. Overview of VO2max prediction models along with the predictor variables for the data set (Model 9- Model 16) (Category 2)..... 42

Table 4.3. Overview of VO2max prediction models along with the predictor variables for the data set (Model 17- Model 24) (Category 3)..... 42

Table 4.4. Overview of VO2max prediction models along with the predictor variables for the data set (Model 25- Model 32) (Category 4)..... 43

Table 4.5. Overview of VO2max prediction models along with the predictor variables for the data set (Model 33- Model 39) (Category 5)..... 43

Table 4.6. Overview of VO2max prediction models along with the predictor variables for the data set (Model 40- Model 48) (Category 6)..... 43

Table 4.7. Overview of VO2max prediction models along with the predictor variables for the data set (Model 49- Model 56) (Category 7)..... 44

Table 4.8. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 1- Model 8) 45

Table 4.9. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 9 - Model 16) 46

Table 4.10. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 17 – Model 24).... 47

Table 4.11. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 25 - Model 32) 48

Table 4.12. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 33 – Model 39).... 49
Table 4.13. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 40 - Model 48) 50
Table 4.14. R and SEE values for VO2max prediction models using different regression methods on the VO2 max dataset (Model 49 – Model 56).... 51
Table 4.15. Category and Models 54

LIST OF FIGURES **PAGE**

Figure 3.1. Architecture of an artificial neuron and a multilayered neural network 27
Figure 3.2. Consept of an artifical neural network..... 32
Figure 3.3. Geometry of a multilayer feed-forward neural network..... 34
Figure 4.1. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR (Model 1-Model 8)..... 45
Figure 4.2. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR (Model 9-Model 17)..... 46
Figure 4.3. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR(Model 18-Model 24)..... 47
Figure 4.4. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR (Model 25-Model 32)..... 48
Figure 4.5. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR (Model 33-Model 39)..... 49
Figure 4.6. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR (Model 40-Model 48)..... 50
Figure 4.7. Percentage decrease rates in SEE's of maximal VO2max for machine learning methods compared to SEE's obtained by MLR(Model 49-Model 56)..... 51
Figure 4.8. Scatter plot of actual VO2max vs. predicted VO2max for Model 6 using SVM 58
Figure 4.9. Scatter plot of actual VO2max vs. predicted VO2max for Model 8 using SVM 58

Figure 4.10. Scatter plot of actual VO2max vs. predicted VO2max for Model 14 using SVM 59
Figure 4.11. Scatter plot of actual VO2max vs. predicted VO2max for Model 16 using SVM 59
Figure 4.12. Scatter plot of actual VO2max vs. predicted VO2max for Model 22 using SVM 60
Figure 4.13. Scatter plot of actual VO2max vs. predicted VO2max for Model 24 using SVM 60
Figure 4.14. Scatter plot of actual VO2max vs. predicted VO2max for Model 30 using SVM 61
Figure 4.15. Scatter plot of actual VO2max vs. predicted VO2max for Model 32 using SVM 61
Figure 4.16. Scatter plot of actual VO2max vs. predicted VO2max for Model 38 using SVM 62
Figure 4.17. Scatter plot of actual VO2max vs. predicted VO2max for Model 40 using SVM 62
Figure 4.18. Scatter plot of actual VO2max vs. predicted VO2max for Model 46 using SVM 63
Figure 4.19. Scatter plot of actual VO2max vs. predicted VO2max for Model 48 using SVM 63
Figure 4.20. Scatter plot of actual VO2max vs. predicted VO2max for Model 54 using SVM 64
Figure 4.21. Scatter plot of actual VO2max vs. predicted VO2max for Model 56 using SVM 64

LIST OF ABBREVIATIONS

3MWD	:	3-Minute Walk Distance
ACSM	:	American College of Sports Medicine
AD	:	Accelerometer Data
AHR	:	After Heart Rate
ANN	:	Artificial Neural Network
AOD	:	Accumulated Oxygen Deficit
ARC	:	Anaerobic Rowing Capacity
BF	:	Body Fat
BF%	:	Percent Body Fat
BIA	:	Bioelectrical Impedance Analysis
BM	:	Body Mass
BMI	:	Body Mass Index
BSA	:	Body Surface Area
BYU	:	Brigham Young University
CFS	:	Correlation-Based Feature Selection
CRF	:	Cardiorespiratory Fitness
CV	:	Critical Velocity
EX	:	Exercise
FEV1	:	Forced Expiratory Volume
GRNN	:	Generalized Regression Neural Networks
GXT	:	Graded Exercise Test
HR	:	Heart Rate
HR3	:	Exercise Heart Rate At The End Of 1.5 Mile
LoPAR	:	Self-Report Physical Activity Assessment
LPA	:	Leisure-time Physical Activity

L-PA	:	Leisure-time Physical Activity
MAX-GRD	:	Maximal Grade
MAX-HR	:	Maximal Heart Rate
MAX-RPE	:	Maximal Rating of Perceived Exertion
MAX-Time	:	Maximal Exercise Time
MFANN	:	Multilayer Feed-Forward Artificial Neural Networks
MIN3	:	Elapsed Exercise Time At The End Of 1.5 Mile
MLP	:	Multilayer Perceptron
MLR	:	Multi Linear Regression
MPH	:	Speed
MVPA	:	Moderate to Vigorous Physical Activity
N-Ex	:	Non-Exercise
PA	:	Physical Activity
PACER	:	Progressive Aerobic Cardiovascular Endurance Run
PA-R	:	Physical Activity Rating
PAR-Q	:	Physical Activity Readiness Questionnaire
PARS	:	Physical Activity Rating Score
PASS	:	NASA Physical Activity Status Scale
PFA	:	Perceived Functional Ability
PFI	:	Physical Fitness Index
PHR15	:	15-Second Post-Exercise Heart Rate
PHR60	:	60-Second Post-Exercise Heart Rate
PRPE 15	:	Power At RPE 15
PO	:	Power Output
R	:	Multiple Correlation Coefficient
RER	:	Respiratory Exchange Ratio
RPC	:	Rating of Perceived Capacity

RPE	:	Rating of Perceived Exertion
SC	:	Step Count
SD	:	Standard Deviation
SEE	:	Standard Error of Estimation
SM	:	Skeletal Muscle
SMAT	:	Skating Multistage Aerobic Test
SM-GRD	:	Submaximal Grade
SM-HR	:	Submaximal Heart Rate
SM-MPH	:	Submaximal Speed
SM-RPE	:	Submaximal Rating of Perceived Exertion
SM-Stage	:	Submaximal Treadmill Stage
SRT	:	Shuttle Run Test
SVM	:	Support Vector Machine
SVR	:	Support Vector Regression
VO_2	:	Oxygen Consumption
VO_2max	:	Maximum Oxygen Uptake
VPA	:	Vigorous Physical Activity
VPA	:	Vigorous Physical Activity
VSAQ	:	Veterans Specific Activity Questionnaire
WC	:	Waist Circumference
W-PA	:	Work-related Physical Activity
WR	:	Work Rate

1. INTRODUCTION

1.1. Maximal Oxygen Uptake (VO_2max)

VO_2max (also maximal oxygen consumption, maximal oxygen uptake, peak oxygen uptake or maximal aerobic capacity) is the maximum rate of oxygen consumption as measured during incremental exercise, most typically on a motorized treadmill (Clemente, 2009)(Dlugosz, 2013). Maximal oxygen consumption reflects the aerobic physical fitness of the individual, and is an important determinant of their endurance capacity during prolonged, sub-maximal exercise. The name is derived from V - volume, O_2 - oxygen, max - maximum.

VO_2max is expressed either as an absolute rate in (for example) litres of oxygen per minute (L/min) or as a relative rate in (for example) millilitres of oxygen per kilogram of body mass per minute (e.g., mL/(kg*min)). The latter expression is often used to compare the performance of endurance sports athletes. However, VO_2max generally does not vary linearly with body mass, either among individuals within a species or among species, so comparisons of the performance capacities of individuals or species that differ in body size must be done with appropriate statistical procedures, such as analysis of covariance.

Oxygen consumption increases with exercise intensity. However, a point is reached where exercise intensity can continue to increase without the associated rise in oxygen consumption. The point at which oxygen consumption plateaus is the VO_2max or a person's maximal aerobic capacity. VO_2max is more useful as an indicator of a person's aerobic potential or upper limit than as a predictor of success in endurance events. This is because there are other factors besides VO_2max that influence a competitor's success, such as mental attitude (ability to tolerate pain), running (or other activity) economy (how efficiently one runs), and lactate threshold (fastest pace that can be maintained without accumulating large amounts of lactic acid in the blood)(David Hampson, 2008). A higher maximal oxygen consumption translates into a greater submaximal work capacity, which in turn means greater caloric expenditure in the same time while exercising or doing physical labor.

The average male aged 20-29 has a VO_2max of 44 to 51 and the average female aged 20-29 has a VO_2max of 35 to 43. (Lance Armstrong's VO_2max is 85 milliliters per kg.) The highest ever recorded VO_2max is 96 ml/kg/min, attributed to Bjørn Dæhlie, and 77 ml/kg/min in women. Both were cross-country skiers. The world-record test was purportedly taken when Bjørn was out of season; his physiologist, Erlend Hem, has stated it's very likely that Dæhlie could pass the 100 ml/kg/min score if the test was taken mid-season.

A person who is fit, would have a higher VO_2max than someone who is less fit. However, it is important to understand that if a person has a larger VO_2max than others, it does not necessarily mean that person could beat them in, for example, a marathon race. What it means is that the body is more able to absorb and use oxygen to generate energy for the muscles, and this will certainly give an edge, but what a person is actually capable of doing with that energy depends on many other factors (for instance, the mechanical efficiency with which a person runs at a given speed) (Shamsi et al., 2011).

Table 1.1. Female and male normative data for VO_2max (values in ml/kg/min) (Achten, and Jeukendrup, 2003)

Gender	Age	Very Poor	Poor	Fair	Good	Excellent	Superior
Female	13-19	<25.0	25.0 - 30.9	31.0 - 34.9	35.0 - 38.9	39.0 - 41.9	>41.9
	20-29	<23.6	23.6 - 28.9	29.0 - 32.9	33.0 - 36.9	37.0 - 41.0	>41.0
	30-39	<22.8	22.8 - 26.9	27.0 - 31.4	31.5 - 35.6	35.7 - 40.0	>40.0
	40-49	<21.0	21.0 - 24.4	24.5 - 28.9	29.0 - 32.8	32.9 - 36.9	>36.9
	50-59	<20.2	20.2 - 22.7	22.8 - 26.9	27.0 - 31.4	31.5 - 35.7	>35.7
	60+	<17.5	17.5 - 20.1	20.2 - 24.4	24.5 - 30.2	30.3 - 31.4	>31.4
Male	13-19	<35.0	35.0 - 38.3	38.4 - 45.1	45.2 - 50.9	51.0 - 55.9	>55.9
	20-29	<33.0	33.0 - 36.4	36.5 - 42.4	42.5 - 46.4	46.5 - 52.4	>52.4
	30-39	<31.5	31.5 - 35.4	35.5 - 40.9	41.0 - 44.9	45.0 - 49.4	>49.4
	40-49	<30.2	30.2 - 33.5	33.6 - 38.9	39.0 - 43.7	43.8 - 48.0	>48.0
	50-59	<26.1	26.1 - 30.9	31.0 - 35.7	35.8 - 40.9	41.0 - 45.3	>45.3
	60+	<20.5	20.5 - 26.0	26.1 - 32.2	32.3 - 36.4	36.5 - 44.2	>44.2

Table 1.1 shows normative data for the VO_2max assessment and comprises of the following grades: Very Poor, Poor, Fair, Good, Excellent and Superior.

1.2. Assessing Maximal Oxygen Uptake

Cardiorespiratory fitness is known to be an objective, reproducible measure that reflects the functional consequences of physical training and recent physical activity habits, and is a powerful predictor of chronic disease morbidity and mortality. Prospective observational studies have shown that low cardiorespiratory fitness is strongly associated with the risk for developing coronary heart disease, hypertension, type 2 diabetes mellitus, and metabolic syndrome as well as mortality from cardiovascular disease, cancer, and all causes of mortality. Hence, cardiorespiratory fitness has been suggested to be included in the European Health Monitoring System for the adult population, National Health and Nutrition Examination Survey in the United States, and the Exercise and Physical Activity Reference for Health Promotion 2006 in Japan, in which the recommended reference value for VO_2max to prevent lifestyle-related disease was reported. Although cardiorespiratory fitness is an important health indicator, cardiorespiratory fitness assessment is usually not performed in many health care settings because of the absence of feasible and practical assessment methods (Cao et al., 2009).

There are a variety of choices when it comes to performing a VO_2max test and the type of test that will be chosen often depends on who is being tested and where. When testing an individual, exercise or Monarc bikes, treadmills, and steps are the most common methods used, while the beep test is a popular choice when training sports teams.

Two treadmill tests that are commonly used are the Balke Protocol and the Bruce Protocol. Both tests involve a subject on the treadmill with increasing speed, grade or both. These tests are effective in estimating VO_2max without using too much time, effort or equipment. Using these tests, the subject walks or runs, and factors including grade and speed are changed. With the Balke test the grade is changed every minute with a constant speed until the subject can no longer perform.

The final time is recorded and then put into an equation to estimate VO_2max (Chavda et al., 2013).

The most common option for fitness testing in a gym is bike tests. Most gyms will offer the choice of an exercise bike with built in fitness test such as the LifeCycle bike (the best option for older adults, or for someone who is obese, or very unfit) or a Monarc bike. The Monarc bike is designed specifically for fitness testing and has a weight attached to the wheel. This weight is adjusted every 3 minutes to increase the work load while the volunteer is exercising and the way in which heart rate responds helps to determine fitness level (Cao et al., 2010).

Personal trainers who don't work in a gym, prefer walking test. In this test, heart rate is recorded as volunteer walks for a set distance (normally 1.6 kms) and the time it takes is also noted. Heart rate and the time it takes to walk the distance is then used to determine fitness level. One of the advantages of this test is that volunteer can instantly see that his (her) fitness has improved at the next test by the time it takes to walk the distance (Nes et al., 2011).

A popular home testing method and often used by schools, the step test combines the height of the step with the number of steps that volunteer completes in one minute, to calculate fitness level. As with the walking tests, volunteer can instantly tell if his (her) fitness has improved at his (her) next test by how many steps he (she) is able to do (Chatterjee et al., 2005).

Popular with sporting clubs and other organizations, the beep test requires participants to run between two markers set 20 meters apart and is known as a multi-stage test. The athlete or client is required to run between the two markers, starting when the beep sounds and arrive at the other end before the beep is heard again. As the test progresses, the beeps sound closer together, making it necessary to run faster as an athlete is getting tired, just to keep up. Once the athlete misses two beeps in a row (i.e. cannot reach the marker before the beep sounds) they need to stop and the last staged noted is the equivalent of their fitness level (Chatterjee et al., 2008).

1.3. How is VO_2max Measured?

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 specific increases in the speed and intensity of the exercise and collection and measurement of the volume and oxygen concentration of inhaled and exhaled air. This determines how much oxygen the athlete is using.

An athlete's oxygen consumption rises in a linear relationship with exercise intensity - up to a point. There is a specific point at which oxygen consumption plateaus even if the exercise intensity increases. This plateau marks the VO_2max. It's a painful point in VO_2max testing where the athlete moves from aerobic metabolism to anaerobic metabolism. From here, it's not long before muscle fatigue forces the athlete to stop exercising. The test usually takes between 10 and 15 minutes and requires an athlete to be completely rested and motivated to endure the pain long enough to find the true VO_2max.

VO_2max also can be estimated. There are a variety of protocols used to estimate VO_2max, but none are as accurate as direct testing (Elizabeth Quinn, 2008).

1.4. Advantages and Disadvantages of Maximal Tests

Maximal tests are expensive to administer and can be risky because the subject reaches his maximal heart rate near complete exhaustion. The requirement of expensive gas analysis and ventilation equipment, and the need for medical attendance are the other major disadvantages of maximal tests. Because of these drawbacks of maximal tests, other methods that do not require maximal effort have been developed to determine VO_2max. These methods use either non-exercise tests or submaximal exercise tests to predict VO_2max. The main idea is to develop a VO_2max prediction equation by using non-exercise variables (perceived functional ability and physical ability rating questionnaires) or submaximal exercise variables (heart rate, speed, exercise time, rating of perceived exertion, etc.) along with physical features (age, gender, body mass, height, body mass index, percentage of

body fat, etc.). The advantage of VO_2max prediction equations is that they provide a simpler means of determining a complex measurement by using variables that are easily measured.

Non-exercise prediction equations provide a convenient estimate of VO_2max without needing to perform a maximal or submaximal exercise test. Researchers may favor non-exercise models over maximal and submaximal exercise tests because of their ease of administration to large populations and their independence from expensive instrumentation and exercise testing. The main drawback of non-exercise tests is that they depend upon the truthful self-report of activity and cannot be used when the individuals have an interest in the results, giving them a reason to falsify the self-reported activity levels.

Submaximal tests are faster, safer, and cheaper to administer than maximal tests. In addition, submaximal exercise testing provides the administrator an opportunity to observe responses to exercise and to teach participants about selection of an appropriate intensity of exercise. However, submaximal tests have been found to be less accurate than maximal tests yet more accurate than the non-exercise tests.

1.5. Motivation of This Thesis

As discussed before, although direct measurement of VO_2max is the most reliable method, it has also several limitations. Because of these drawbacks, numerous VO_2max prediction models have been developed by using exercise and non-exercise variables. Most of these models have been developed by MLR. In recent years, Akay et. al (Akay et al., 2013) have shown that machine learning methods such as SVM and MLP perform better than MLR-based VO_2max prediction models.

The purpose of this thesis is to develop accurate VO_2max prediction models using submaximal and questionnaire variables.

Two types of questionnaire variables are considered. Perceived functional ability (PFA) questions are designed to determine how fast participants believe they could walk, jog, or run one mile distance (PFA- 1) and a three mile distance (PFA-

2) without becoming overly breathless or tired. The responses to the PFA-1 and PFA-2 questions are the PFA-1 and PFA-2 scores (range = 1–13), respectively. The physical activity rating (PA-R) question allows participants to judge their level of physical activity during the past six months. The response to the PFA-3 question is the PFA-3 score (range = 0–10).

Regression methods such as SVM, MLP and MLR are used for developing VO_2max prediction models. The performance of prediction models has been evaluated by calculating their multiple correlation coefficients (*R*'s) and standard error of estimates (*SEE*'s). The results show that the accuracy of VO_2max prediction models based on submaximal and standard non-exercise variables can be significantly improved by including questionnaire variables.

The main contributions of this thesis are as follows:

- This is the first study in literature that develops VO_2max prediction models using machine learning methods (SVM, MLP and MLR) combined with questionnaire data.
- Use of different regression methods allows to determine their relative order for VO_2max prediction.
- In related literature for predicting VO_2max, the two questionnaire variables PFA-1 and PFA-2 have been always used together. As compared to such previous works, this study uses each questionnaire variable separately to investigate and calculate the effect of adding each questionnaire to the models.
- In this thesis 56 different models have been developed for VO_2max prediction. By comparing all of the models, best and worst models have been determined. It turns out that working on a higher amount of models increases the accuracy of the study.

2. PREVIOUS WORKS

2.1. Submaximal Prediction Models Based Studies

Swank et al. (Swank et. Al., 2001) determined the reliability and validity of a branching treadmill protocol in predicting VO_2max. Thirty-seven, apparently healthy individuals (19 women and 18 men); volunteered to participate. On 2 separate testing days, each subject underwent maximal exercise testing using the protocol developed. Stepwise regression analysis indicated that the percentage of age-predicted maximum heart rate (APMHR) achieved at stage 3, speed and grade at stage 3, and APMHR accounted for 89% of the variance in VO_2max. The 4 predictor variables had been statistically significant ($p<0.01$), and the standard error of the estimate had been $4.56 ml \cdot kg^{-1}.min^{-1}$ Their results indicated that health and fitness professionals could incorporate that protocol into their practices for the purpose of predicting VO_2max for their clients outside the laboratory environment. Furthermore, their results indicated that using the proposed regression model is reliable.

Larsen et al. (Larsen et. al., 2002) developed a submaximal, 1.5-mile endurance test for college-aged students using walking, jogging, or running exercise. College students (N = 101: 52 men, 47 women), ages 18-26 years, successfully completed the 1.5-mile test twice, and a maximal graded exercise test. Participants were instructed to achieve a "somewhat hard" exercise intensity (rating of perceived exertion = 13) and maintain a steady pace throughout each 1.5-mile test. This submaximal 1.5-mile test accurately predicted maximal oxygen uptake (VO_2max) without measuring heart rate and was similar to the 1.5-mile run in that it allowsfor mass testing and requires only a flat, measured distance and a stopwatch. Further, it could accommodate a wide range of fitness levels (from walkers to runners). They reported *SEE* values as 2.87 respectively.

Swain et al (2004) validates a new method that uses a final 6-min stage and that estimates VO_2max from the relationship between heart rate reserve (HRR) and VO_2 reserve. A cycle ergometry protocol was designed to elicit 65-75 % HRR in the fifth and sixth minutes of the final stage. Maximal workload was estimated by

dividing the workload of the final stage by % HRR. VO_2max was then estimated using the ACSM metabolic equation for cycling. After the 6-min stage was completed, an incremental test to maximal effort was used to measure actual VO_2max. Forty-nine subjects completed a pilot study using one protocol to reach the 6-min stage, and 50 additional subjects completed a modified protocol. A new method for estimating VO_2max from submaximal cycling based on VO_2 reserve has been found to be valid and more accurate than previous methods. They reported ***SEE*** values as 4.00.

Roy et al (2004) evaluated the SM Predicted VO_2max protocol by comparing it to results from a VO_2max treadmill test in 20 young healthy women aged 20-25 years. Subjects were 10 step-trained (ST) women who had performed aerobic activities and exercised on a step ergometer for 20-30 minutes at least 3 times per week for the past 3 months, and 10 nonstep-trained (NST) women who had performed aerobic activities no more than twice a week during the past 3 months and had no previous experience on a step ergometer. The SM Predicted VO_2max protocol used 2 steady state heart rates between approximately 115-150 b min^{-1} to estimate VO_2max. The Bruce maximal treadmill protocol (Actual VO_2max) was used to measure VO_2max by open circuit spirometry. Each subject performed both tests within a 7-day period. Their best results for ***SEE*** value reported as 5.32.

Dalleck et al. (Dalleck, Kravitz, and Robergs,2006) developed an equation to predict VO_2max from a submaximal elliptical cross-trainer test. Fifty-four apparently healthy subjects (25 men and 29 women) participated in the study and were randomly assigned to an original sample group (n = 40) and a cross-validation group (n = 14). Each subject completed an elliptical cross-trainer submaximal (35-minute submaximal stages) and a VO_2max test on the same day, with a 15-minute rest period in between. Stepwise multiple regression analyses were used to develop an equation for estimating elliptical cross-trainer VO_2max from the data of the original sample group. The accuracy of the equation was tested by using data from the cross-validation group. Results indicated that the protocol and equation developed in their study could be used by exercise professionals to provide acceptably accurate

estimates of VO_2max in non-laboratory-based settings. They reported ***SEE*** and ***R*** values as , 3.91 and 0.86, respectively.

McComb et al. (McComb, Roh, and Williams, 2006) had a study to develop a prediction equation that could be used to estimate VO_2max from a submaximal water running protocol. Thirty-two volunteers (n =19 males, n = 13 females), ages 18 - 24 years, underwent the following testing procedures: (a) a 7-site skin fold assessment; (b) a land VO_2max running treadmill test, and (c) a 6 min water runing test. For the water running submaximal protocol, the participants were fitted with an Aqua Jogger Classic Uni-Sex Belt and a Polar Heart Rate Monitor, the participants' head, shoulders hips and feet were vertically aligned, using a modified running/bicycle motion. They used a regression model to predict VO_2max. They reported ***SEE*** value as 3.31 and 0.89

Vehrs et al. (Vehrs et. al., 2007) tried to develop a single-stage submaximal treadmill jogging (TMJ) test to predict VO_2max in fit adults. Participants (*N*=400; men=250 and women=150), ages 18 to 40 years, successfully completed a maximal graded exercise test (GXT) at 1 of 3 laboratories to determine VO_2max. The TMJ test was completed during the first 2 stages of the GXT. Following 3 min of walking (Stage 1), participants achieved a steady-state heart rate (HR) while exercising at a comfortable self-selected submaximal jogging speed at level grade (Stage 2). Gender, age, body mass, steady-state HR, and jogging speed (mph) were included as independent variables. They used multiple linear regression to predict VO_2max. They reported ***SEE*** value as 2.52.

Waddoups et al. (Waddoups et. al., 2008) determined if single-stage treadmill walking test correctly predicts VO_2max at the low (50% of maximum heart rate) and high (70% of maximum heart rate) ends of the specified heart rate range for males and females aged 18 - 55 years. In their study each of the 34 participants completed one low-intensity and one high-intensity trial. The two trials resulted in significantly different estimates of VO_2max (low-intensity trial and high-intensity trial). Their best results for ***SEE*** value reported as 5.7.

Akalan et al. (Akalan, Robergs, and Kravitz, 2008) hypothesized that a large proportion of the error of VO_2max prediction comes from individual differences in heart rate responses to submaximal exercise, and that if these differences could be decreased the accuracy of VO_2max prediction would increase. Eighty (43 male, 37 female) sedentary to highly trained, healthy volunteers first completed a self-report physical activity assessment (Lo-Par), and then performed a modified YMCA protocol with 4-minute stages, a second submaximal test involving an individualized ramp submaximal protocol that was terminated at 80% of their cycle ergometer age-predicted maximum heart rate. Exercise and five-minute recovery heart rate data were collected. A ramp cycle ergometer protocol with expired gas analysis was used to measure actual VO_2max. Multiple regression analysis was used to predict VO_2max. They reported *SEE* value as 4.23.

Akay et al (2009) developed non-exercise (N-Ex) VO2max prediction models by using support vector regression (SVR) and multilayer feed forward neural networks (MFFNN). VO2max values of 100 subjects (50 males and 50 females) are measured using a maximal graded exercise test. The variables; gender, age, body mass index (BMI), perceived functional ability (PFA) to walk, jog or run given distances and current physical activity rating (PA-R) are used to build two N-Ex prediction models. Using 10-fold cross validation on the dataset, standard error of estimates (*SEE*) and multiple correlation coefficients (*R*) of both models are calculated. The MFFNN-based model yields lower *SEE* (3.23 ml kg−1 min−1) whereas the SVR-based model yields higher *R* (0.93). Compared with the results of the other N-Ex prediction models in literature that are developed using multiple linear regression analysis, the reported values of*SEE* and *R* in this study are considerably more accurate. Therefore, the results suggest that SVR-based and MFFNN-based N-Ex prediction models can be valid predictors of VO2max for heterogeneous samples.

Akay et al. (Akay et. al., 2010) developed an accurate ANN-based model to predict maximal VO_2max of fit adults from a single stage submaximal treadmill jogging test. Participants (81 males and 45 females), aged from 17 to 40 years, successfully completed a maximal graded exercise test (GXT) to determine VO_2max.

The variables; gender, age, body mass, steady-state heart rate and jogging speed are used to build the ANN prediction model. Compared with the results of the other prediction models in literature that were developed using Multiple Linear Regression Analysis, the reported values of ***SEE*** and ***R*** in their study were considerably more accurate. They reported ***SEE*** value as 1.8.

Lee J et al (2011), The main purpose of this study was to assess the validity of the Cosmed Fitmate (FM) for the prediction of maximal oxygen consumption (VO_2max). In addition, this study examined whether measuring submaximal VO_2, rather than predicting it, can improve upon the prediction of VO_2max. Participants for the study were 48 young to middle-age adults (32 men, 16 women), with a mean age of 31 yr. Each participant completed a submaximal and maximal treadmill test on 2 separate occasions. During the submaximal test, VO_2max was predicted using the FM. This device extrapolates the linear regression relating heart rate (HR) and measured VO_2 at submaximal work rates to age-predicted maximum HR (HR = 220 - age). The criterion measure was obtained using a graded, maximal treadmill test, with VO_2 measured by the Douglas bag (DB) method. There was no significant difference between VO_2max predicted by the FM and VO_2max measured by the DB method. The results of this study showed that a strong positive correlation (r = 0.897) existed between VO_2max predicted by the FM and VO_2max measured by the DB method, with a standard error of the estimate (***SEE***) = 3.97 ml·kg(-1)·min(-1). There was a significant difference in VO_2max predicted by the American College of Sports Medicine (ACSM) metabolic equations. These findings indicate that a small, portable, and easy-to-use metabolic system provides valid estimates of VO_2max, and improves upon predictive accuracy, compared to using generalized ACSM metabolic equations.

Schembre et all (2011) aimed to develop and test a VO_2max estimation equation derived from the IPAQ-Short Form (IPAQ-S). College-aged males and females (n = 80) completed the IPAQ-S and performed a maximal exercise test. The estimation equation was created with multivariate regression in a gender-balanced subsample of participants, equally representing five levels of fitness (n = 50) and validated in the remaining participants (n = 30). The resulting equation explained

43% of the variance in measured VO_2max (*SEE* = 5.45 ml·kg(-1)·min(-1)). Estimated VO_2max for 87% of individuals fell within acceptable limits of error observed with submaximal exercise testing (20% error). The IPAQ-S can be used to successfully estimate VO_2max as well as submaximal exercise tests. Development of other population-specific estimation equations is warranted.

Clare E.Marsh et all (2012) used submaximal treadmill running test in predicting VO_2Max. Twenty-one moderately well-trained men aged 18-34 years performed 1 maximal treadmill test to determine maximal oxygen uptake (M VO_2max) and 2 submaximal treadmill tests using 4 stages of continuous submaximal exercise. Estimated VO_2max was predicted by extrapolation to age-predicted maximal heart rate (HRmax) and calculated in 2 ways: using data from all submaximal stages between 110 b·min(-1) and 85% HRmax (P VO_2max-All), and using data from the last 2 stages only (P VO_2max-2).In conclusion ACSM equation provides a reasonably good estimation of VO_2max with no difference in predictive accuracy between P VO_2max-2 and P VO_2max-All, and hence, either approach may be equally useful in tracking an individual's aerobic fitness over time. However, if a precise knowledge of VO_2max is required, then it is recommended that this be measured directly. They reported *SEE* value as 4.35

The purpose of this thesis is to develop accurate VO_2max prediction models using submaximal and questionnaire variables. Regression methods such as Support Vector Machines (SVM), Multilayer Perceptron (MLP) and Multiple Linear Regression (MLR) have been used for developing VO_2max prediction models

Table 2.1. Summary of submaximal regression models to estimate VO_2max

Study	**Ages**	**Predictor Variables**	***SEE***
Swank et al. (2001)	8–43	Age, BM, BMI, HR, RER	4.56
Larsen et al. (2002)	8–26	G, BM, Exercise Time	3.37
Swain et al. (2004)	18–44	A, Height, BM,HR	4.00
Roy et al (2004)	20-25	Height, Weight, BM, HR	5.32
Dalleck et al. (2006)	0–49	G, Age, HR, BM	3.91
McComb et al. (2006)	8–56	G, Height, BM, BF, HR	3.31
Vehrs et al. (2007)	18-40	G, Age, BM, MPH, HR	2.52
Waddoups et al.(2008)	8–55	G, Age, Height, BM, BMI, HR, RPE	5.70
Akalan et al. (2008)	8–49	Age, G, Height, Weight, BF, BM,W/min, Delta Recovery HR, Recovery HR, LoPAR	4.23
Akay et al (2009)	18-65	Age, G, BM, PFA, PA-R	3.23
Akay et al.(2010)	17-40	G,Age, BM, MPH,HR	1.80
Lee at al (2011)	18-59	A,Height, Weight, BMI,HR	3,97
Schembre (2011)	18-25	Age,Weight, Height, BMI,PA-R	5.45
Clare E.Marsh (2012)	8-34	HR,MPH	4.35

G, gender; **BM**, body mass; **BMI**, body mass index; **BF,** body fat;**HR**, heart rate; **RER**, respiratory exchange ratio; **RPE**, rating of perceived exertion; **LoPAR**, self-report physical activity assessment; **W/min** ,minutes that watt increase; **PFA**, perceived functional ability; **PA-R**, physical activity rating; **MPH**, speed.

The performance of prediction models has been evaluated by calculating their multiple correlation coefficients (R's) and standard error of estimates (SEE's). The results show that the accuracy of VO_2max prediction models based on submaximal and standard non-exercise variables could be significantly improved by including questionnaire variables in prediction models. The results of SVM models have been also compared with the ones obtained by MLP and MLR and it turned out that SVM-based VO_2max prediction models perform better (i.e. yield lower SEE's and higher R's) than the prediction models developed by other regression methods.

2.2. Submaximal and Questionnaire Prediction Models Based Studies

George H (2009), sought to develop a regression model to predict maximal oxygen uptake (VO_2max) based on submaximal treadmill exercise (EX) and non-exercise (N-EX) data involving 116 participants, ages 18-65 years. The EX data

included the participants' self-selected treadmill speed (at a level grade) when exercise heart rate first reached at least 70% of predicted maximum heart rate (HRmax; 220 - age) by the end of any one of three 4-min consecutive stages involving walking (3.0-4.0 mph; Stage 1), jogging (4.1-6.0 mph; Stage 2), and running (> 6.0 mph; Stage 3). The N-EX data included various demographic (age, gender), biometric (body mass), and questionnaire (participants' perceived functional ability [PFA] to walk, jog, or run given distances, and their self-reported level of physical activity [PA-R]) information. All participants (n = 100) who completed the study requirements and successfully achieved a maximal level of exertion during a graded exercise test (GXT) to assess VO_2max (mean ± SD; 41.39 ± 9.15 ml · kg-1 · min-1) were included in the data analysis. Stepwise regression was used to generate the following prediction equation (*R* = .94, *SEE* = 3.09 ml · kg-1 · min-1): VO_2max (ml · kg-1 · min-1) = 30.04 + (6.37 × gender; females = 0, males = 1) - (0.243 × age) - (0.122 × body mass) + (3.263 × ending self-selected treadmill speed; mph) + (0.391 × PFA) + (0.669 × PA-R). Each of the predictor variables were statistically significant ($p < .001$) and cross-validation procedures using PRESS (predicted residual sum of squares) statistics revealed minimal shrinkage (*Rp* = .92 and *SEEp* = 3.29 ml · kg-1 · min-1). In summary, this submaximal treadmill test and accompanying regression model yields relatively accurate VO_2max estimates in healthy men and women (ages 18-65 years) using both EX and N-EX data

Nielson et al (2010) developed a multiple linear regression model to predict treadmill VO_2max scores using both exercise and non-exercise data. One hundred five college-aged participants (53 male, 52 female) successfully completed a submaximal cycle ergometer test and a maximal graded exercise test on a motorized treadmill. The submaximal cycle protocol required participants to achieve a steady-state heart rate equal to at least 70% of age-predicted maximum heart rate (220-age), while the maximal treadmill graded exercise test required participants to exercise to volitional fatigue. Relevant submaximal cycle ergometer test data included a mean (±SD) ending steady-state heart rate and ending workrate equal to 164.2 ± 13.0 bpm and 115.3 ± 27.0 watts, respectively. Relevant non-exercise data included a mean (±SD) body mass (kg), perceived functional ability score, and physical activity rating

score of 74.2 ± 15.1, 15.7 ± 4.3, and 4.7 ± 2.1, respectively. Multiple linear regression was used to generate the following prediction of (R = .91, standard error of estimates (*SEE*) = 3.36 ml·kg−1·min−1): VO_2max = 54.513 + 9.752 (gender, 1 = male, 0 = female) – .297 (body mass, kg) + .739 (perceived functional ability, 2–26) + .077 (work rate, watts) – .072 (steady-state heart rate). Each predictor variable was statistically significant ($p < .05$) with beta weights for gender, body mass, perceived functional ability, exercise workrate, and steady-state heart rate equal to .594, –.544, .388, .305, and –.116, respectively. The predicted residual sums of squares (PRESS) statistics reflected minimal shrinkage (*RPRESS* = .90, *SEE*PRESS = 3.56 ml·kg−1·min−1) for the multiple linear regression model. In summary, the submaximal cycle ergometer protocol and accompanying prediction model yield relatively accurate VO_2max estimates in healthy college-aged participants using both exercise and non-exercise data.

Akay et al (2012), developed accurate VO_2max prediction models using maximal and N-Ex variables and show that the accuracy of maximal prediction models can be improved by adding N-Ex questionnaire predictor variables. Machine learning methods such as SVM, Artificial Neural Network (ANN), MLR and General Reggresion Neural Network (GRNN) are used for developing VO_2max prediction models. Two different datasets are used for VO_2max prediction. Using 10-fold cross validation, the performance of the models are evaluated by calculating R's and *SEE*'s. It is shown that adding questionnaire to ANN-based prediction models yields 4.25% to 10.81% increments in *R*'s and 27.53% to 35.59% decrements in *SEE*'s for the first dataset and 1.08% to 6.97% increments in *R*'s and 0.53% to 17.97% decrements in *SEE*'s for the second dataset. Also it is concluded that ANN-based prediction models are more accurate than GRNN-based, SVM-based and MLR-based prediction models.

Table 2.2. Summary of submaximal and questionnaire models to estimate VO_2max

Study	**Ages**	**Predictor Variables**	***SEE***
George H et al (2009)	18-65	Age,G,Height, BM, BMI,PFA,PA-R, HR,RPE	3,09
Nielson et al (2010)	9-29	G,BM,PFA,HR,	3,36
Akay et al (2012)	18-65	G,Age,BMI,HR,RPE,PFA,PA-R	3,73

G, gender; **BM**, body mass; **BMI**, body mass index; **HR**, heart rate; **RER**, respiratory exchange ratio; **RPE**, rating of perceived exertion; **LoPAR**, self-report physical activity assessment; **W/min** ,minutes that watt increase; **PFA**, perceived functional ability; **PA-R**, physical activity rating; **MPH**, speed.

3. OVERVIEW OF METHODS

3.1. Support Vector Machines (SVMs)

Support Vector learning is based on simple ideas which originated in statistical learning theory (Vapnik, 1998). The simplicity comes from the fact that Support Vector Machines (SVMs) apply a simple linear method to the data but in a high-dimensional feature space non-linearly related to the input space. Moreover, even though we can think of SVMs as a linear algorithm in a high-dimensional space, in practice, it does not involve any computations in that high-dimensional space. This simplicity combined with state of the art performance on many learning problems (classification and regression) has contributed to the popularity of the SVM.

SVMs use an implicit mapping $\Phi$of the input data into a high-dimensional feature space defined by a kernel function, i.e., a function returning the inner product $(\Phi(x), \Phi(x'))$ between the images of two data points x, x'in the feature space. The learning then takes place in the feature space, and the data points only appear inside dot products with other points. This is often referred to as the "kernel trick" (Schölkopf, and Smola, 2002). More precisely, if a projection $\Phi: X \rightarrow H$ is used, the dot product $(\Phi(x), \Phi(x'))$ can be represented by a kernel function k

$$k(x, x') = \langle \Phi(x), \Phi(x') \rangle \tag{3.1}$$

which is computationally simpler than explicitly projecting x and x'into the feature space H.

One interesting property of support vector machines and other kernel-based systems is that, once a valid kernel function has been selected, one can practically work in spaces of any dimension without any significant additional computational cost, since feature mapping is never effectively performed. In fact, one does not even need to know which features are being used.

Another advantage of SVMs and kernel methods is that one can design and use a kernel for a particular problem that could be applied directly to the data without the need for a feature extraction process. This is particularly important in problems where a lot of structure of the data is lost by the feature extraction process (e.g., text processing).

3.1.1. Classification

In classification, support vector machines separate the different classes of data by a hyperplane

$$\langle \mathbf{w}, \Phi(x) \rangle + b = 0 \tag{3.2}$$

corresponding to the decision function

$$f(x) = \text{sign}(\langle \mathbf{w}, \Phi(x) \rangle + b) \tag{3.3}$$

It can be shown that the optimal, in terms of classification performance, hyper-plane (Vapnik, 1998) is the one with the maximal margin of separation between the two classes. It can be constructed by solving a constrained quadratic optimization problem whose solution w has an expansion $\mathbf{w} = \sum_i \alpha_i \Phi(x_i)$ in terms of a subset of training patterns that lie on the margin. These training patterns, called support vectors, carry all relevant information about the classification problem. Omitting the details of the calculation, there is just one crucial property of the algorithm that we need to emphasize: both the quadratic programming problem and the final decision function depend only on dot products between patterns. This allows the use of the “kernel trick” and the generalization of this linear algorithm to the nonlinear case.

In the case of the L2-norm soft margin classification the primal optimization problem takes the form:

$$\begin{aligned} &\text{minimize} t(\mathbf{w}, \xi) = \frac{1}{2} \| \mathbf{w} \|^2 + \frac{C}{m} \sum_{i=1}^{m} \xi_i \\ &\text{subject to} \quad y_i(\langle \Phi(x_i), \mathbf{w} \rangle + b) \geq 1 - \xi_i (i = 1, \dots, m) \\ &\qquad \xi_i \geq 0 \quad (i = 1, \dots, m) \end{aligned} \tag{3.4}$$

Where m is the number of training patterns, and $y_i = \pm 1$. As in most kernel methods, the SVM solution **w** can be shown to have an expansion

$$\mathbf{w} = \sum_{i=1}^{m} \alpha_i \, y_i \Phi(x_i) \tag{3.5}$$

where non-zero coefficients (support vectors) occur when a point (x_i, y_i) meets the constraint. The coefficients α_i are found by solving the following (dual) quadratic programming problem:

$$\begin{aligned} &\text{minimize} \quad W(\alpha) = \sum_{i=1}^{m} \alpha_i + \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j \, k(x_i, x_j) \\ &\qquad \text{subject to} \quad 0 \leq \alpha_i \leq \frac{C}{m} \quad (i = 1, \dots, m) \\ &\qquad\qquad \sum_{i=1}^{m} \alpha_i y_i = 0. \end{aligned} \tag{3.6}$$

This is a typical quadratic problem of the form:

$$\begin{aligned} &\text{minimize} \quad c^T x + \frac{1}{2} x^T H x \\ &\text{subject to} \quad b \leq Ax \leq b + r \\ &\qquad\qquad l \leq x \leq u \end{aligned} \tag{3.7}$$

where$H \in \mathbb{R}^{m \times m}$ with entries $H_{ij} = y_i y_j k(x_i, x_j)$, $C = (1, \dots, 1) \in \mathbb{R}^m$, $u = (C, \dots, C) \in \mathbb{R}^m$, $l = (0, \dots, 0) \in \mathbb{R}^m$, $A = (y_1, \dots, y_m) \in \mathbb{R}^m$, $b = 0$, $r = 0$. The cost parameter *C* of the SVM formulation in Equation 3.11 controls the penalty paid by the SVM for miss-classifying a training point and thus the complexity of the prediction function. A high cost value *C* will force the SVM to create a complex enough prediction function to miss-classify as few training points as possible, while a

lower cost parameter will lead to a simpler prediction function. Therefore, this type of SVM is usually called C-SVM.

Another formulation of the classification with a more intuitive hyperparameter than C is the ϑ-SVM (Schölkopf, Smola, Williamson, and Bartlett, 2000). The ϑ parameter has the interesting property of being an upper bound on the training error and a lower bound on the fraction of support vectors found in the data set, thus controlling the complexity of the classification function build by the SVM (see Appendix for details). For multi-class classification, mostly voting schemes such as one-against-one and one-against-all are used. In the one-against-all method k binary SVM classifiers are trained, where k is the number of classes, each trained to separate one class from the rest. The classifiers are then combined by comparing their decision values on a test data instance and labeling it according to the classifier with the highest decision value.

In the one-against-one classification method (also called pair wise classification (Knerr, Personnaz, and Dreyfus, 1990; Kreßel 1999), $\binom{k}{2}$ classifiers are constructed where each one is trained on data from two classes. Prediction is done by voting where each classifier gives a prediction and the class which is most frequently predicted wins ("Max Wins"). This method has been shown to produce robust results when used with SVMs (Hsu, and Lin, 2002a). Although this suggests a higher number of support vector machines to train the overall CPU time used is less compared to the one-against-all method since the problems are smaller and the SVM optimization problem scales super-linearly.

3.1.2. Regression

By using a different loss function called the _-insensitive loss function$\| y - f(x) \|_\epsilon = max\{0, \| y - f(x) \| - \epsilon\}$, SVMs can also perform regression. This loss function ignores errors that are smaller than a certain threshold $\epsilon > 0$ thus creating a tube around the true output. The primal becomes:

$$\text{minimize} t(\boldsymbol{w}, \xi) = \frac{1}{2} \| w \|^2 + \frac{C}{m} \sum_{i=1}^{m} (\xi_i + \xi_i^*)$$
$$\text{subject to} \quad (\langle \Phi(x_i), \mathbf{w} \rangle + b) - y_i \leq \epsilon - \xi_i \tag{3.8}$$

$$y_i - (\langle \Phi(x_i), \mathbf{w} \rangle + b) \leq \epsilon - \xi_i^* \tag{3.9}$$
$$\xi_i^* \geq 0 \qquad (i = 1, \ldots, m)$$

We can estimate the accuracy of SVM regression by computing the scale parameter of a

Laplacian distribution on the residuals $\zeta = y - f(x)$, where $f(x)$ is the estimated decision function.

3.1.3. Kernel functions

As seen before, the kernel functions return the inner product between two points in a suitable feature space, thus defining a notion of similarity, with little computational cost even in very high-dimensional spaces. Kernels commonly used with kernel methods and SVMs in particular include the following:

- the linear kernel implementing the simplest of all kernel functions

$$k(\mathrm{x}, \mathrm{x}') = \langle \mathrm{x}, \mathrm{x}' \rangle \tag{3.10}$$

- the Gaussian Radial Basis Function (RBF) kernel

$$k(\mathrm{x}, \mathrm{x}') = \exp(-\sigma \| \mathrm{x} - \mathrm{x}' \|^2) \tag{3.11}$$

- the polynomial kernel

$$k(\mathrm{x}, \mathrm{x}') = (scale. \langle \mathrm{x}, \mathrm{x}' \rangle + offset)^{degree} \tag{3.12}$$

· the hyperbolic tangent kernel

$$k(\mathrm{x},\mathrm{x}') = \tanh(scale.\langle \mathrm{x},\mathrm{x}'\rangle + offset) \tag{3.13}$$

· the Bessel function of the first kind kernel

$$k(\mathrm{x},\mathrm{x}') = \frac{Bessel^{n}_{(\vartheta+1)}(\sigma\|\mathrm{x}-\mathrm{x}'\|)}{(\|\mathrm{x}-\mathrm{x}'\|)^{-n(\vartheta+1)}} \tag{3.14}$$

· the Laplace Radial Basis Function (RBF) kernel

$$k(\mathrm{x},\mathrm{x}') = \exp(-\sigma \parallel \mathrm{x}-\mathrm{x}' \parallel) \tag{3.15}$$

· the ANOVA radial basis kernel

$$k(\mathrm{x},\mathrm{x}') = (\sum_{k=1}^{n} \exp(-\sigma(x^k - x'^k)^2))^d \tag{3.16}$$

· the linear splines kernel in one dimension

$$k(x,x') = 1 + xx' \min(x,x') - \frac{x+x'}{2}(\min(x,x'))^2 + \frac{(\min(x,x'))^3}{3} \tag{3.17}$$

and for the multidimensional case $k(\mathrm{x},\mathrm{x}') = \prod_{k=1}^{n} k(x^k, x'^k)$.

The Gaussian and Laplace RBF and Bessel kernels are general-purpose kernels used when there is no prior knowledge about the data. The linear kernel is useful when dealing with large sparse data vectors as is usually the case in text categorization. The polynomial kernel is popular in image processing and the sigmoid kernel is mainly used as a proxy for neural networks. The splines and ANOVA RBF kernels typically perform well in regression problems.

3.2. Artificial Neural Networks (ANN)

The human brain provides proof of the existence of massive neural networks that can succeed at those cognitive, perceptual, and control tasks in which humans are successful. The brain is capable of computationally demanding perceptual acts (e.g. recognition of faces, speech) and control activities (e.g. body movements and body functions). The advantage of the brain is its effective use of massive parallelism, the highly parallel computing structure, and the imprecise information-processing capability. The human brain is a collection of more than 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive, process, and transmit information. Treelike networks of nerve fibers called dendrites are connected to the cell body or soma, where the cell nucleus is located. Extending from the cell body is a single long fiber called the axon, which eventually branches into strands and substrands and are connected to other neurons through synaptic terminals or synapses. The transmission of signals from one neuron to another at synapses is a complex chemical process in which specific transmitter substances are released from the sending end of the junction. The effect is to raise or lower the electrical potential inside the body of the receiving cell. If the potential reaches a threshold, a pulse is sent down the axon and the cell is ‘fired’. ANN has been developed as generalizations of mathematical models of biological nervous systems. A first wave of interest in neural networks (also known as connectionist models or parallel distributed processing) emerged after the introduction of simplified neurons (McCulloch, and Pitts, 1943). The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. In a simplified mathematical model of the neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons is represented by a transfer function. The neuron impulse is then computed as the weighted sum of the input signals, transformed by the transfer function. The learning capability of an artificial neuron is achieved by adjusting the weights in accordance to the chosen learning algorithm.

A typical artificial neuron and the modeling of a multilayered neural network are illustrated in Figure 3.1. Referring to Figure 3.1, the signal flow from inputs $x_1, \dots, x_n$is considered to be unidirectional, which are indicated by arrows, as is a neuron's output signal flow (*O*). The neuron output signal *O* is given by the following relationship:

$$O = f(net) = f(\sum_{j=1}^{n} w_j \ x_j) \tag{3.18}$$

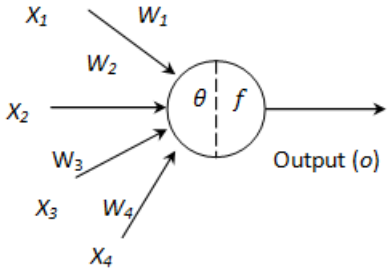
Where *wj* is the weight vector, and the function *f(net)* is referred to as an activation (transfer) function. The variable *net* is defined as a scalar product of the weight and input vectors,

$$net = w^T x = w_1 x_1 + \cdots + w_n x_n \tag{3.19}$$

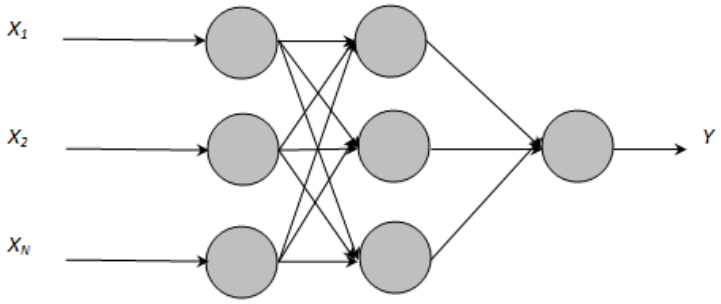
Where *T* is the transpose of a matrix, and, in the simplest case, the output value *O* is computed as

$$O = f(net) = \begin{cases} 1 & if\, w^T x \geq \theta \\ 0 & otherwise \end{cases} \tag{3.20}$$

where θ is called the threshold level; and this type of node is called a linear threshold unit.



(a) Artificial neuron



(b) Multilayered artificial neural network

Figure 3.1. Architecture of an artificial neuron and a multilayered neural network.

3.2.1. Neural Networks Architectures

The basic architecture consists of three types of neuron layers: input, hidden, and output layers. In feed-forward networks, the signal flow is from input to output units, strictly in a feed-forward direction. The data processing can extend over multiple (layers of) units, but no feedback connections are present. Recurrent networks contain feedback connections. Contrary to feed-forward networks, the dynamical properties of the network are important. In some cases, the activation values of the units undergo a relaxation process such that the network will evolve to a stable state in which these activations do not change anymore. In other applications, the changes of the activation values of the output neurons are significant, such that the dynamical behavior constitutes the output of the network. There are several other neural network architectures (Elman network, adaptive resonance theory maps, competitive networks, etc.), depending on the properties and requirement of the application.

A neural network has to be configured such that the application of a set of inputs produces the desired set of outputs. Various methods to set the strengths of the connections exist. One way is to set the weights explicitly, using a priori knowledge. Another way is to train the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule. The learning situations in neural networks may be classified into three distinct sorts. These are supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, an input vector is presented at the inputs together with a set of desired responses, one for each node, at the output layer. A forward pass is done, and the errors or discrepancies between the desired and actual response for each node in the output layer are found. These are then used to determine weight changes in the net according to the prevailing learning rule. The term supervised originates from the fact that the desired signals on individual output nodes are provided by an external teacher.

The best-known examples of this technique occur in the back propagation algorithm, the delta rule, and the perceptron rule. In unsupervised learning (or self-organization), a (output) unit is trained to respond to clusters of pattern within the input. In this paradigm, the system is supposed to discover statistically salient features of the input population. Unlike the supervised learning paradigm, there is no a priori set of categories into which the patterns are to be classified; rather, the system must develop its own representation of the input stimuli. Reinforcement learning is learning what to do – how to map situations to actions – so as to maximize a numerical reward signal. The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them. In the most interesting and challenging cases, actions may affect not only the immediate reward, but also the next situation and, through that, all subsequent rewards. These two characteristics, trial-and-error search and delayed reward are the two most important distinguishing features of reinforcement learning.

3.2.2. Hebbian Learning

The learning paradigms discussed above result in an adjustment of the weights of the connections between units, according to some modification rule. Perhaps the most influential work in connectionism's history is the contribution of Hebb (Hebb, 1949), where he presented a theory of behavior based, as much as possible, on the physiology of the nervous system. The most important concept to emerge from Hebb's work was his formal statement (known as Hebb's postulate) of how learning could occur. Learning was based on the modification of synaptic connections between neurons. Specifically, when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased. The principles underlying this statement have become known as Hebbian Learning. Virtually, most of the neural network learning techniques can be considered as a variant of the Hebbian learning rule. The basic idea is that if two neurons are active simultaneously, their interconnection must be strengthened. If we consider a single layer net, one of the interconnected neurons will be an input unit and one an output unit. If the data are represented in bipolar form, it is easy to express the desired weight update as

$$w_i(new) = w_i(old) + x_i o \tag{3.21}$$

Where *o* is the desired output for *i*= 1 to *n(*inputs*).*

Unfortunately, plain Hebbian learning continually strengthens its weights without bound (unless the input data is properly normalized).

3.2.3. Perceptron Learning

The perceptron is a single layer neural network whose weights and biases could be trained to produce a correct target vector when presented with the

corresponding input vector. The training technique used is called the Perceptron learning rule. Perceptrons are especially suited for simple problems in pattern classification. Suppose we have a set of learning samples consisting of an input vector *x* and a desired output *d(k)*. For a classification task, the *d(k)* is usually +1 or −1. The perceptron-learning rule is very simple and can be stated as follows:

1. Start with random weights for the connections.
2. Select an input vector *x* from the set of training samples.
3. If output $y_k \neq d(k)$(the perceptron gives an incorrect response), modify all connections *wi* according to: $\delta w_i = \eta(d_k - y_k)x_i$; ($\eta$ = learning rate).
4. Go back to step 2.

Note that the procedure is very similar to the Hebb rule; the only difference is that when the network responds correctly, no connection weights are modified.

3.2.4. Back propagation Learning

The simple perceptron is just able to handle linearly separable or linearly independent problems. By taking the partial derivative of the error of the network with respect to each weight, we will learn a little about the direction the error of the network is moving. In fact, if we take the negative of this derivative (i.e. the rate change of the error as the value of the weight increases) and then proceed to add it to the weight, the error will decrease until it reaches a local minima. This makes sense because if the derivative is positive, this tells us that the error is increasing when the weight is increasing. The obvious thing to do then is to add a negative value to the weight and vice versa if the derivative is negative. Because the taking of these partial derivatives and then applying them to each of the weights takes place, starting from the output layer to hidden layer weights, then the hidden layer to input layer weights (as it turns out, this is necessary since changing these set of weights requires that we know the partial derivatives calculated in the layer downstream), this algorithm has been called the back propagation algorithm.

A neural network can be trained in two different modes: online and batch modes. The number of weight updates of the two methods for the same number of data presentations is very different. The online method weight updates are computed for each input data sample, and the weights are modified after each sample. An alternative solution is to compute the weight update for each input sample, but store these values during one pass through the training set which is called an epoch. At the end of the epoch, all the contributions are added, and only then the weights will be updated with the composite value. This method adapts the weights with a cumulative weight update, so it will follow the gradient more closely. It is called the batch-training mode.

Training basically involves feeding training samples as input vectors through a neural network, calculating the error of the output layer, and then adjusting the weights of the network to minimize the error.

3.3. Multilayer Perceptron

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined by the connections between elements. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between the elements. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation is shown in Figure 3.2. Here, the network is adjusted, based on a comparison of the output and the target, up to the network output matches the target. Typically many such input/target output pairs are used to train a network. Batch training of a network proceeds by making weight and bias changes based on an entire set (batch) of input vectors. Incremental training changes the weights and biases of a network as needed after presentation of each individual input vector. Incremental training is sometimes referred to as ‘‘online’’ or ‘‘adaptive’’ training. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, and vision and control systems.

Today neural networks can be trained to solve problems that are difficult for conventional computer programs or human beings

Considering the geometry of the network, there are several types of neural networks: Hopfield, Hamming, Campenter and Grossberg, Kohonen, multi-layer feed forward neural network and others. Neural networks with different geometry are used for solving various problems.

The fist three types of neural networks are usually used for binary input data and with problems of classification into classes. The last two types of neural networks are appropriate for the approximation of an unknown function.

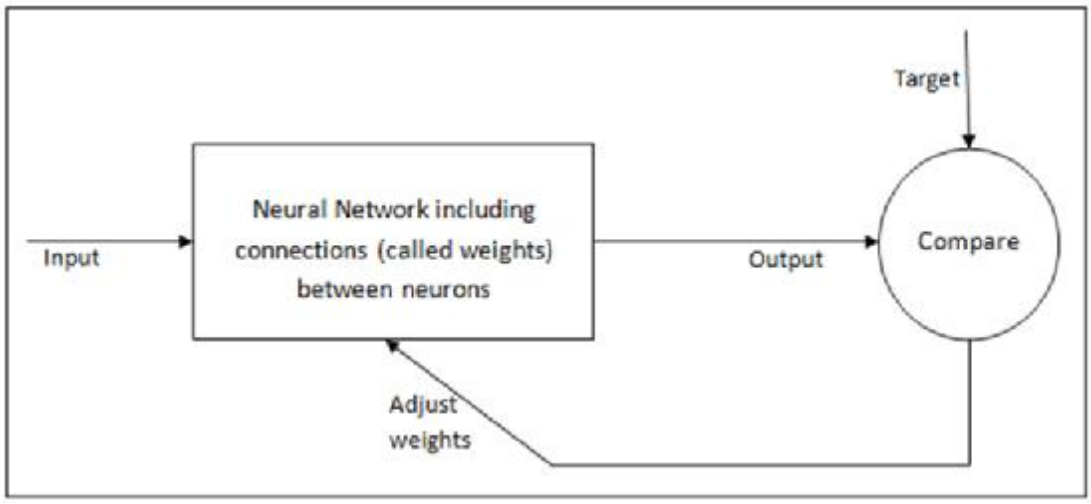


Figure 3.2. Concept of an artificial neural network

The geometry of a multilayer feed-forward neural network is shown in Figure 3.3. Input units are connected to the first layer of hidden units which are further connected to the units of the second hidden layer. The units of the last hidden layer are connected to the output units. The input units represent the input data, and the output units represent the output data. The hidden layers may be considered as a black box which performs the necessary transformations of the input data so that the target output data are obtained.

Each unit is represented by its value y_i^k : Each connection between the units is represented by its weight w_{ij}^k ; where index i corresponds to the unit number of the k^{th} layer, while index j corresponds to the unit number of the $(k-1)^{th}$ layer. The input

layer is denoted by 0, whereas the output layer is denoted by n_1: The signals travel in one direction only, i.e. from the input layer towards the output layer.

The value of a unit is multiplied by the corresponding weight and added to the value of the signal in the unit of the next layer. In addition, the value of bias neuron or threshold θ_i^k is added to the equation, which is given in (3.22)

$$y_i^k = f\left(\sum_{i=1}^{n_{k-1}} w_{ij}^k y_i^{k-1} + \theta_i^k\right) \tag{3.22}$$

Activation function $f(.)$ enables the modeling of an arbitrary nonlinear relation between input and output variables. Different functions could be used as an activation function. The usual choices of activation function are a sigmoid function, Gaussian and radial basis function.

The behavior of the neural network depends on the values of weights w_{ij}^k and thresholds θ_i^k which have to be determined by the training procedure. The supervised training is in fact a general optimization problem in which the minimum of error E_p is sought

$$E_p = \frac{1}{2}\sum_{i=1}^{n_0}\left(t_{pi} - y_{pi}^{nl}\right)^2, \tag{3.23}$$

where t_{pi} are the target output values, y_{pi}^{nl} are the values of neurons in the output layer, i.e. the output values evaluated by neural network, n_0 is the number of neurons in output layer, i.e. the number of output variables.

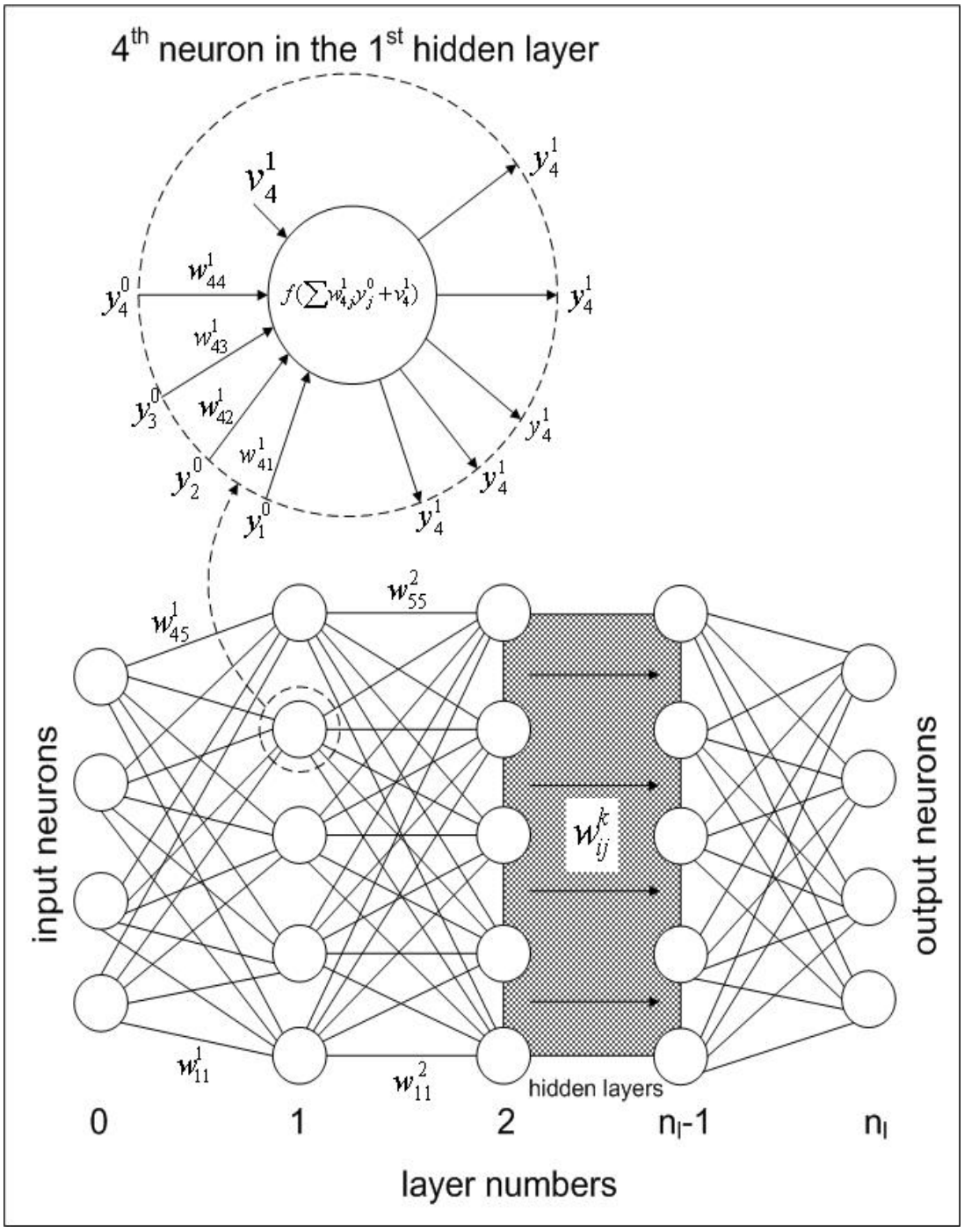


Figure 3.3. Geometry of a multilayer feed-forward neural network

3.4. Multiple Linear Regression (MLR)

The multiple linear regression model is an extension of a simple linear regression model to incorporate two or more explanatory variable in a prediction equation for a response variable. Multiple regression modeling is now a mainstay of statistical analysis in most fields because of it's power and flexibility. It requires very little effort (and sometimes even less thought) to estimate very complicated models with large numbers of variables.

When we wish to model a continuous outcome variable, then an appropriate analysis is often multiple linear regression. For simple linear regression we have one continuous input variable. In multiple regression we generalize the method to more than one input variable and we will allow them to be continuous or categorical. Multiple regression is a generalization of the analysis of variance and analysis of covariance.

In multiple regression the basic model is the following:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ip} + \varepsilon_i \tag{3.24}$$

We assume that the error term $\varepsilon_i$is normally distributed, with mean 0 and standard deviation σ.

Here $y_i$is the output for unit or subject i and there are k input variables $X_{i1}, X_{i2}, \ldots, X_{ip}$. Often $y_i$is termed the dependentvariable and the input variables $X_{i1}, X_{i2}, \ldots, X_{ip}$are termed the independent variables. The latter can be continuous or nominal. However the term "independent" is a misnomer since the *X*'s need not be independent of each other. Sometimes they are called the explanatoryor predictor variables. Each of the input variables is associated with a regression coefficient$\beta_1, \beta_2, \ldots, \beta_k$. There is also an additive constant term β_0. These are the modelparameters.

We can write the first section on the right-hand side of equation (3.24.) as

$$LP_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ip}$$

Where LP_i is known as the *linear predictor* and is the value of y_i predicted by the input variables. The difference $y_i - LP_i = \varepsilon_i$is the error term.

The models are fitted by choosing estimates $b_0, b_1, \ldots, b_p$, which minimize the sum of squares (SS) of the predicted error. These estimates are termed ordinaryleastsquaresestimates. Using these estimates we can calculate the fitted values y_i^{fit}, and the observed residuals $e_i = y_i - y_i^{fit}$. Here it is clear that the residuals estimate the error term (Draper, and Smith,1998).

3.4.1. Uses of Multiple Regression

To adjust the effects of an input variable on a continuous output variable for the effects of confounders. For example, to investigate the effect of diet on weight allowing for smoking habits. Here the dependent variable is the outcome from a clinical trial. The independent variables could be the two treatment groups (as a 0/1 binary variable), smoking (as a continuous variable in numbers of packs per week) and baseline weight. The multiple regression model allows one to compare the outcome between groups, having adjusted for differences in baseline weight and smoking habit. This is also known as analysis of covariance.

To analyze the simultaneous effects of a number of categorical variables on an output variable. An alternative technique is the analysis of variance but the same results can be achieved using multiple regression.

To predict a value of an outcome, for given inputs. For example, an investigator might wish to predict the forced expiratory volume (*FEV*1) of a subject given age and height, so as to be able to calculate the observed *FEV*1 as a percentage of predicted, and to decide if the observed *FEV*1 is below, say, 80% of the predicted one.

3.5. Overview of Dataset

One dataset was used in this thesis. The dataset contains some input parameters and one output parameter (VO_2max). As will be described in the following sections, using different input parameters of the dataset, different models are developed to predict VO_2max.

3.5.1. The Dataset

One hundred participants, 50 females and 50 males, aged 18-65 years will volunteer for this study. This study will recruit employees from the LDS Hospital and participants of the Brigham Young University Y-Be-Fit Program. Participants will be classified as either low risk or moderate risk according to ACSM (2000) guidelines. Low risk participants will be tested at BYU while moderate risk participants will be tested at the LDS Hospital. Participants will read and sign an informed consent document which contains written instructions explaining testing procedures, benefits, and risks of the test. Participants will also complete a PAR-Q and a non-exercise questionnaire. All participants will receive $25 compensation for their time and efforts. The study will be approved by the BYU Institutional Review Board for Human Subjects.

All participants will complete a maximal treadmill GXT. Prior to testing, participants will be instructed to wear comfortable clothing, acquire an adequate amount of sleep the night before the test (6-8 hours), and avoid exercise or strenuous physical activity the day of the test (ASCM, 2000). Participants will be instructed to drink plenty of fluids 24-hours prior to testing, and avoid food, tobacco, alcohol, and caffeine 3 hours before the test (ACSM).

Participants' height and body mass will be measured and recorded while wearing lightweight clothing and no shoes using a stadiometer and balance beam scale, respectively. Information pertaining to the maximal exercise test such as the electronic heart rate monitor, head-gear, mouth piece, nose-clip, and RPE scale (Borg, 1982) will be explained prior to testing.

Participants will perform a maximal GXT using a slightly modified version of the ASU maximal protocol (George, 1996). The maximal exercise test will begin with a 4-12 minute warm up on a treadmill at level ground. The warm up will consist of participants' completing up to three, four-minute stages until a steady state heart rate of 75% of age predicted heart rate max is achieved. During the first four-minute stage participants will walk at a brisk pace, 3-4 mph. If 75% of heart rate max is not achieved in this stage, participants will progress to the second stage, and speed will increase to a self-selected jogging pace between 4 and 6 mph. A third stage will be performed if heart rate criteria have not been met by the end of the second stage, and speed will increase to speeds of 6 mph or greater. Following the warm-up, participants will rest for approximately 5 minutes, while the test administrator explains the test procedures for the maximal GXT and fits participants with a mouthpiece and nose clip. The ending speed of the warm-up will serve as the necessary pace for the maximal GXT. Thereafter, the grade will increase 1.5% each minute, while pace remains the same, and continue until participants reach volitional fatigue and are no longer able to continue.

During the maximal GXT, metabolic gases will be collected using the TrueMax 2400 metabolic measurement system (Consentius Technologies, Sandy, UT). Participants' heart rate, VO_2, and RER will be averaged and analyzed by the computer every 15 seconds. The participants will give an RPE at the end of each stage to estimate the exercise intensity. VO_2max is considered valid when at least two of the following three criteria are met (ACSM, 2000; Howley et al., 1995):

1. Maximal heart rate within 15 beats of age predicted maximal heart rate.
2. Respiratory exchange ratio equal to or greater than 1.10.
3. Plateau in VO_2 despite an increase in work load.

Participants who failed to meet at least two of these criteria were dropped from the study (Howley et al., 1995).

The dataset includes non-exercise variables (gender, age, PFA, BMI and PA-R), maximal exercise test variables (MAX-HR, MAX-GRD and MAX-RPE) and submaximal exercise test variables (SM-HR, SM-GRD and SM-RPE).

Table 3.1. Descriptive statistics of the dataset

Variables	**Mean ± Standard Deviation**
Gender	0.5 ± 0.503
Age (year)	36.23 ± 13.223
BM (kg)	74.86 ± 15.291
Height (m)	1.717 ± 0.089
BMI (kg/m^2)	25.246 ± 4.005
SM-MPH(mph)	5.248 ± 1.013
SM-HR (bpm)	147.73 ± 12.357
SM-RPE	12.663 ± 1.699
SM-Stage	2.11 ± 0.49
MAX-MPH (mph)	5.313 ± 1.042
MAX-HR (bpm)	185.16 ± 13.19
MAX-RPE	19.01 ± 0.785
MAX-GRD	9.385 ± 2.712
MAX-RER	1.168 ± 0.052
PFA	14.75 ± 4.831
PAR	4.83 ± 2.252
VO_2max (ml/kg/min)	41.387 ± 9.145

4. RESULTS AND DISCUSSION

As described in Section 3, one dataset is used in this study. Using each of them, several different VO_2max prediction models have been developed. These models contain various combinations of input parameters. For each model VO_2max is predicted using MLR, MLP, SVM. Weka 3.6 is used for prediction

The performance of MLP, MLR and SVM models are evaluated by using 10-fold cross validation and calculating the ***SEE*** and ***R***, whose formulas are given in (4.1) and (4.2), respectively,

$$SEE = \sqrt{\frac{\sum(Y-Y')^2}{N}} \tag{4.1}$$

$$R = \sqrt{1 - \frac{\sum(Y-Y')^2}{\sum(Y-\bar{Y})^2}} \tag{4.2}$$

In (4.1) and (4.2), Y is the measured VO_2max value, $Y'$is the predicted VO_2max value, $\overline{Y}$ is the mean of the measured values of VO_2max and N is the number of samples in a test subset. Using 10-fold cross validation, 10 training and 10 test sets are developed for each model

56 VO_2max prediction models (for MLP, MLR and SVM) have been developed for the dataset, respectively.

Table 4.1 through Table 4.7 categorically shows 56 models and variables that each model contains (for the dataset). For MLR-based, MLP-based and SVM-based VO_2max prediction models, the individual ***SEE*** and ***R*** values for each fold have been calculated for each model are shown in Table 4.8 through Table 4.14 shows the performance (in terms of ***SEE*** and ***R***) of all VO_2max prediction models.

Table 4.1. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 1- Model 8) (Category 1)

Models	**Predictor Variables**
Model 1	Gender, Age, Weight, Height, Speed
Model 2	Gender, Age, Weight, Height, Speed, PFA-1
Model 3	Gender, Age, Weight, Height, Speed, PFA-2
Model 4	Gender, Age, Weight, Height, Speed, PFA-3
Model 5	Gender, Age, Weight, Height, Speed, PFA-1, PFA-2
Model 6	Gender, Age, Weight, Height, Speed, PFA-1, PFA-3
Model 7	Gender, Age, Weight, Height, Speed, PFA-2, PFA-3
Model 8	Gender, Age, Weight, Height, Speed, PFA-1, PFA-2, PFA-3

Table 4.2. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 9- Model 16) (Category 2)

Models	**Predictor Variables**
Model 9	Gender, Age, Weight, Height, Hr
Model 10	Gender, Age, Weight, Height, Hr, PFA-1
Model 11	Gender, Age, Weight, Height, Hr, PFA-2
Model 12	Gender, Age, Weight, Height, Hr, PFA-3
Model 13	Gender, Age, Weight, Height, Hr, PFA-1, PFA-2
Model 14	Gender, Age, Weight, Height, Hr, PFA-1, PFA-3
Model 15	Gender, Age, Weight, Height, Hr, PFA-2, PFA-3
Model 16	Gender, Age, Weight, Height, Hr, PFA-1, PFA-2, PFA-3

Table 4.3. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 17 - Model 24) (Category 3)

Models	**Predictor Variables**
Model 17	Gender, Age, Weight, Height, Stage
Model 18	Gender, Age, Weight, Height, Stage, PFA-1
Model 19	Gender, Age, Weight, Height, Stage, PFA-2
Model 20	Gender, Age, Weight, Height, Stage, PFA-3
Model 21	Gender, Age, Weight, Height, Stage, PFA-1, PFA-2
Model 22	Gender, Age, Weight, Height, Stage, PFA-1, PFA-3
Model 23	Gender, Age, Weight, Height, Stage, PFA-2, PFA-3
Model 24	Gender, Age, Weight, Height, Stage, PFA-1, PFA-2, PFA-3

Table 4.4. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 25 - Model 32) (Category 4)

Models	**Predictor Variables**
Model 25	Gender, Age, Weight, Height, Speed, Hr
Model 26	Gender, Age, Weight, Height, Speed, Hr, PFA-1
Model 27	Gender, Age, Weight, Height, Speed, Hr, PFA-2
Model 28	Gender, Age, Weight, Height, Speed, Hr, PFA-3
Model 29	Gender, Age, Weight, Height, Speed, Hr, PFA-1, PFA-2
Model 30	Gender, Age, Weight, Height, Speed, Hr, PFA-1, PFA-3
Model 31	Gender, Age, Weight, Height, Speed, Hr, PFA-2, PFA-3
Model 32	Gender, Age, Weight, Height, Speed, Hr, PFA-1, PFA-2, PFA-3

Table 4.5. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 33 - Model 40) (Category 5)

Models	**Predictor Variables**
Model 33	Gender, Age, Weight, Height, Speed, Stage
Model 34	Gender, Age, Weight, Height, Speed, Stage, PFA-1
Model 35	Gender, Age, Weight, Height, Speed, Stage, PFA-2
Model 36	Gender, Age, Weight, Height, Speed, Stage, PFA-3
Model 37	Gender, Age, Weight, Height, Speed, Stage, PFA-1, PFA-2
Model 38	Gender, Age, Weight, Height, Speed, Stage, PFA-1, PFA-3
Model 39	Gender, Age, Weight, Height, Speed, Stage, PFA-2, PFA-3
Model 40	Gender, Age, Weight, Height, Speed, Stage, PFA-1, PFA-2, PFA-3

Table 4.6. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 41 - Model 48) (Category 6)

Models	**Predictor Variables**
Model 41	Gender, Age, Weight, Height, Hr, Stage
Model 42	Gender, Age, Weight, Height, Hr, Stage, PFA-1
Model 43	Gender, Age, Weight, Height, Hr, Stage, PFA-2
Model 44	Gender, Age, Weight, Height, Hr, Stage, PFA-3
Model 45	Gender, Age, Weight, Height, Hr, Stage, PFA-1, PFA-2
Model 46	Gender, Age, Weight, Height, Hr, Stage, PFA-1, PFA-3
Model 47	Gender, Age, Weight, Height, Hr, Stage, PFA-2, PFA-3
Model 48	Gender, Age, Weight, Height, Hr, Stage, PFA-1, PFA-2, PFA-3

Table 4.7. Overview of VO_2max prediction models along with the predictor variables for the data set (Model 49 - Model 56) (Category 7)

Models	**Predictor Variables**
Model 49	Gender, Age, Weight, Height, Speed, Hr, Stage
Model 50	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-1
Model 51	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-2
Model 52	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-3
Model 53	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-1, PFA-2
Model 54	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-1, PFA-3
Model 55	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-2, PFA-3
Model 56	Gender, Age, Weight, Height, Speed, Hr, Stage, PFA-1, PFA-2, PFA-3

Table 4.8. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max dataset (Model 1 - Model8).

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 1	0,667	8,694	0,7831	6,0133	0,8499	5,4757
Model 2	0,8128	5,864	0,8723	4,6092	0,9054	4,4191
Model 3	0,6467	8,4376	0,7888	5,9237	0,8708	5,2274
Model 4	0,6816	7,7186	0,8261	5,637	0,8798	4,9773
Model 5	0,8135	5,7856	0,8836	4,4969	0,9147	4,0145
Model 6	0,8962	4,2004	0,9225	3,5568	0,9348	3,256
Model 7	0,748	7,5893	0,8307	5,5036	0,8874	4,8741
Model 8	0,8445	5,2624	0,901	4,1783	0,9238	3,7951

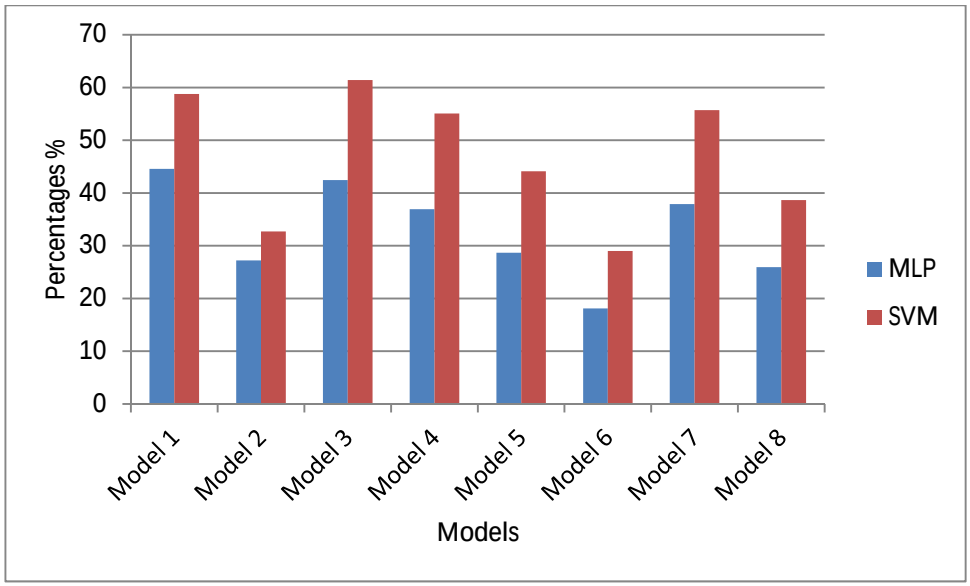


Figure 4.1. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 1-Model 8)

Table 4.9. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model 9-Model 16).

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 9	0,5175	11,3844	0,6879	7,1294	0,7274	6,66
Model 10	0,6699	7,7941	0,8192	5,5397	0,8651	5,1614
Model 11	0,586	11,0387	0,7201	7,0468	0,771	6,5169
Model 12	0,5944	9,8739	0,7709	6,5269	0,7868	6,0469
Model 13	0,7004	7,1461	0,8316	5,3123	0,8698	4,9601
Model 14	0,8384	5,1849	0,8929	4,1719	0,9056	3,8806
Model 15	0,6224	8,7689	0,7817	6,0226	0,8194	5,628
Model 16	0,7415	7,127	0,8357	5,2241	0,8729	4,8357

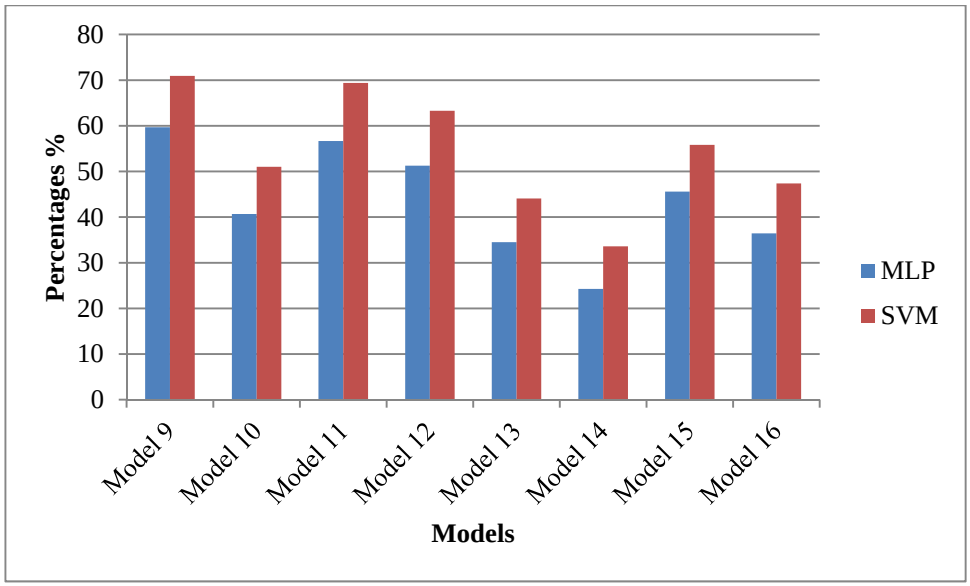


Figure 4.2. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 9-Model 16)

Table 4.10. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model 17-Model 24)

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 17	0,2692	17,1071	0,4822	9,7335	0,6699	8,229
Model 18	0,4146	11,7595	0,7341	6,9086	0,8464	6,2084
Model 19	0,4624	15,1877	0,6207	9,0687	0,7205	8,0638
Model 20	0,4465	14,7874	0,6499	8,7626	0,724	7,8063
Model 21	0,6081	9,9192	0,782	6,2377	0,8694	5,7095
Model 22	0,7343	7,4993	0,858	4,975	0,9038	4,5475
Model 23	0,5362	13,5934	0,7203	7,6692	0,8026	6,8879
Model 24	0,6152	9,8181	0,8113	5,9248	0,881	5,4627

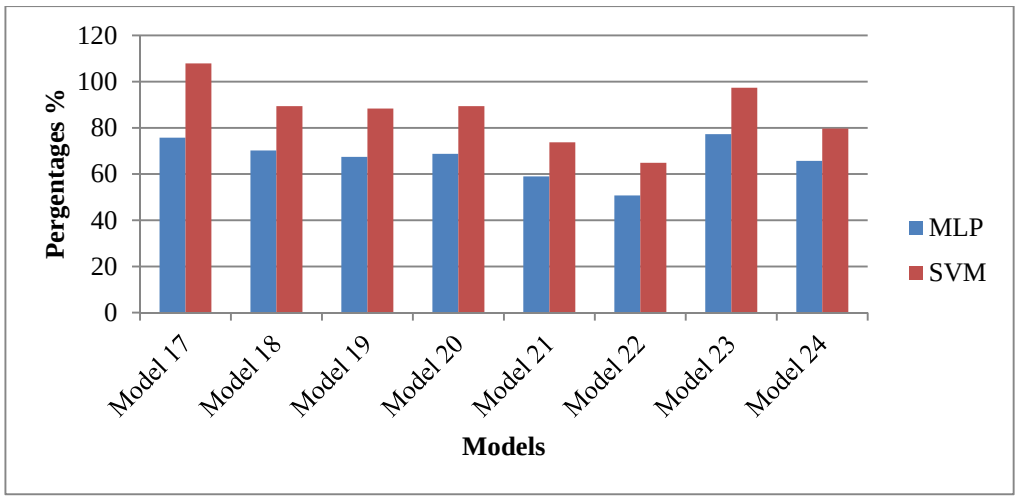


Figure 4.3. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 17- Model 24)

Table 4.11. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model 25- Model 32).

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 25	0,6816	12,1461	0,7592	6,7894	0,8384	6,2309
Model 26	0,7432	7,9208	0,8449	5,0601	0,8972	4,6372
Model 27	0,6249	11,8443	0,7528	6,662	0,8321	6,1054
Model 28	0,5783	10,8124	0,7983	6,2701	0,8423	5,7388
Model 29	0,6854	7,5327	0,8781	4,8801	0,9022	4,4766
Model 30	0,861	4,9361	0,9186	3,6506	0,9314	3,3496
Model 31	0,7018	8,9467	0,8308	5,499	0,8712	5,0965
Model 32	0,797	6,4859	0,8834	4,4271	0,9145	4,0767

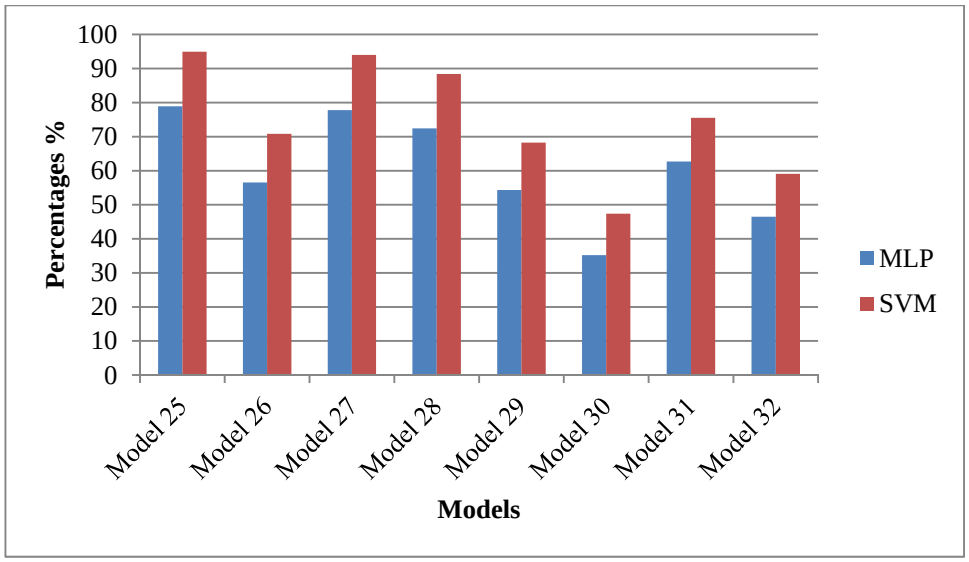


Figure 4.4. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 25- Model 32)

Table 4.12. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model 33-Model 40)

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 33	0,4712	13,4354	0,6816	8,7812	0,7603	7,6408
Model 34	0,7976	8,925	0,7752	6,5441	0,8728	5,8648
Model 35	0,4881	13,0624	0,5974	8,3155	0,7855	7,4717
Model 36	0,5235	11,453	0,6896	7,2434	0,8265	6,4338
Model 37	0,758	7,4691	0,7978	5,8153	0,8878	5,4245
Model 38	0,8184	5,7213	0,8742	4,6261	0,9175	4,3298
Model 39	0,6844	9,7356	0,7741	6,9632	0,8428	6,2205
Model 40	0,8346	6,009	0,8667	4,7788	0,8967	5,1542

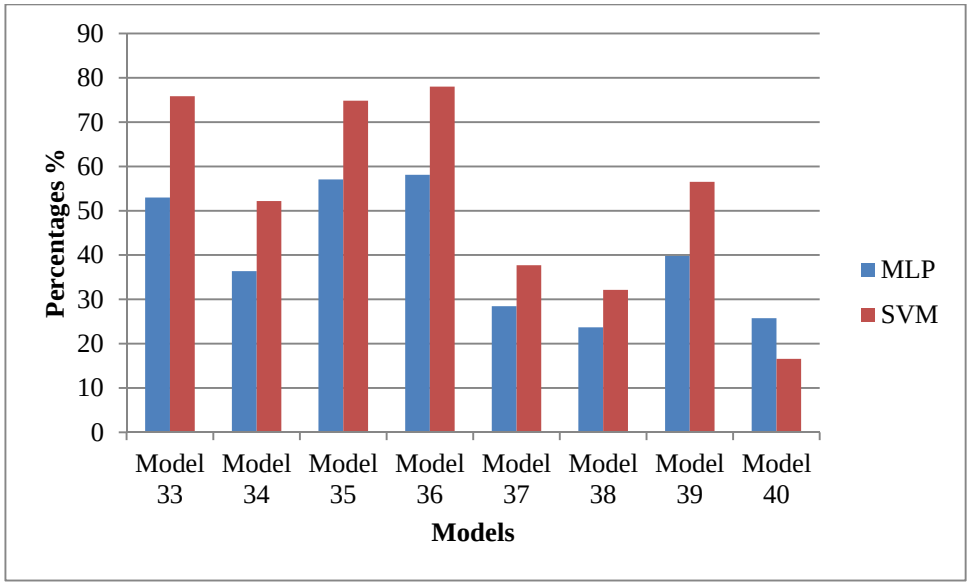


Figure 4.5. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 33-Model40)

Table 4.13. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model41-Model 48)

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 41	0,4325	14,2432	0,4826	9,708	0,7024	7,7255
Model 42	0,5832	9,9936	0,7877	6,8773	0,8545	5,8332
Model 43	0,4481	14,724	0,5938	8,8532	0,7494	7,5007
Model 44	0,4673	12,9244	0,6771	8,2354	0,7764	6,93
Model 45	0,7325	8,2325	0,7916	5,9501	0,8705	5,5412
Model 46	0,7561	6,626	0,8643	4,7442	0,9041	4,3772
Model 47	0,4749	11,1639	0,7368	7,4465	0,8129	6,253
Model 48	0,7078	7,7067	0,8463	5,2674	0,8783	5,3384

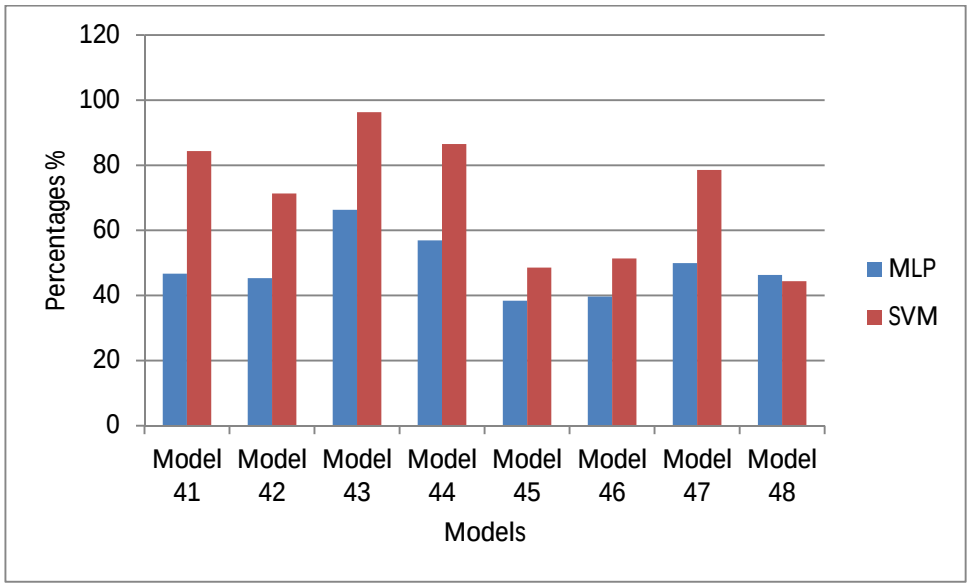


Figure 4.6. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 41- Model 48)

Table 4.14. *R* and *SEE* values for VO_2max prediction models using different regression methods on the VO_2 max data set (Model 49-Model 56)

Models	**MLR**		**MLP**		**SVM**	
	R	***SEE***	**R**	***SEE***	**R**	***SEE***
Model 49	0,6993	12,9673	0,7126	8,0636	0,7677	7,4954
Model 50	0,6992	8,5546	0,796	6,0395	0,8771	5,5724
Model 51	0,6107	12,7844	0,6908	7,754	0,7939	7,3808
Model 52	0,5576	10,3896	0,8032	6,8406	0,8264	6,3663
Model 53	0,6663	7,9771	0,8092	5,7532	0,8867	5,3465
Model 54	0,8366	5,3996	0,8856	4,3894	0,9184	4,076
Model 55	0,7079	8,7094	0,7859	6,1995	0,8492	5,7791
Model 56	0,763	6,9916	0,8431	5,3162	0,8983	4,9566

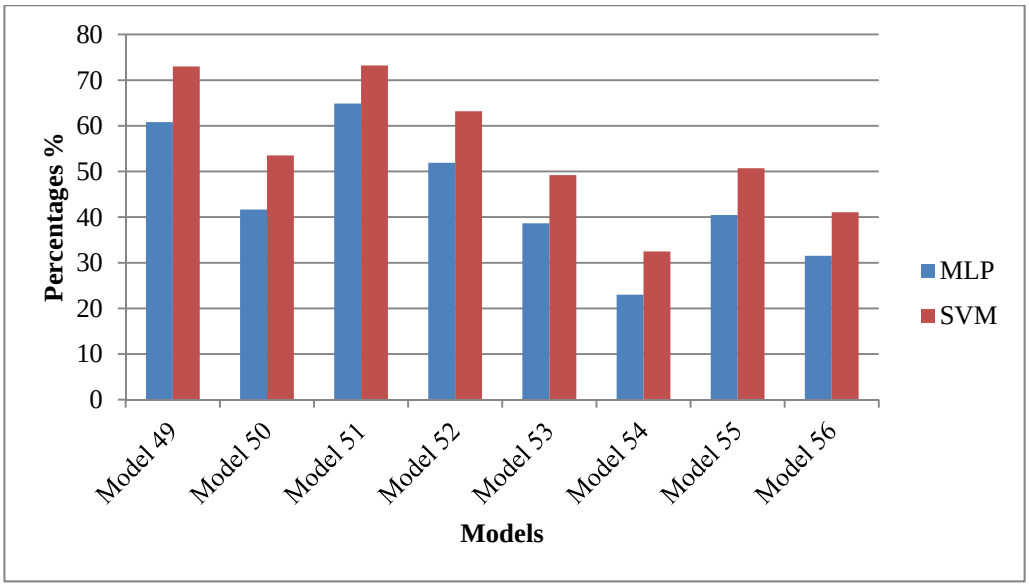


Figure 4.7. Percentage decrease rates in *SEE*'s of VO_2max for machine learning methods compared to *SEE*'s obtained by MLR (Model 49-Model 56)

- The results obtained from the Table 4.8, Table 4.9, Table 4.10, Table 4.11, Table 4.12, Table 4.13 and Table 4.14 presents that SVM-based VO_2max prediction models have lower *SEE*'s than MLP-based and MLR-based prediction models. MLP-based VO_2max prediction models comparatively give more accurate results than MLR-based prediction models.
- From Table 4.8, it can be seen that the best score is obtained by Model 6 including the predictor variables age, gender, weight, height, speed and PFA-1 and PFA-3 has the lowest *SEE*'s (3.25, 3.55, and 4.20) and highest *R*'s (0.89, 0.92, and 0.93) for the SVM-based, MLP-based and MLR-based models, in the given order.
- Model 17 (From Table 4.10) including the predictor variables age, gender, weight, height and stage has the highest *SEE*'s (8.22, 9.73 and 17,10) and the lowest *R*'s (0.66, 0.48, and 0.26) for the SVM-based, MLP-based and MLR-based models, in the given order.
- The accuracy of VO_2max prediction models based on submaximal and standard non-exercise variables can be improved by including questionnaire variables in the models.
- Inclusion of PFA-1, PFA-2, PFA-3 in all prediction models yields lower *SEE*'s and higher *R*'s, independent of whether SVM, MLP or MLR have been used.

The study shows that adding

- PFA-1 to SVM-based prediction models yields 23.90% to 34.50% increments in *SEE*'s performance.
- PFA-2 to SVM-based prediction models yields 1.17% to 9.65% increments in *SEE*'s performance.
- PFA-3 to SVM-based prediction models yields 5.41% to 18.76% increments in *SEE*'s performance.
- PFA-1 and PFA-2 to SVM-based prediction models yields 34.27% to 44.12% increments in *SEE*'s performance.

- PFA-1 and PFA-3 to SVM-based prediction models yields 68.17% to 86.01% increments in *SEE*'s performance.
- PFA-2 and PFA-3 to SVM-based prediction models yields 12.34% to 29.69% increments in *SEE*'s performance.
- PFA-1, PFA-2 and PFA-3 to SVM-based prediction models yields 37.72% to 52.84% increments in *SEE*'s performance.
- PFA-1 to MLP-based prediction models yields 28.69% to 41.16% increments in *SEE*'s performance.
- PFA-2 to MLP-based prediction models yields 1.17% to 9.65% increments in *SEE*'s performance.
- PFA-3 to MLP-based prediction models yields 6.67% to 21.23% increments in *SEE*'s performance.
- PFA-1 and PFA-2 to MLP-based prediction models yields 33.72% to 63.15% increments in *SEE*'s performance.
- PFA-1 and PFA-3 to MLP-based prediction models yields 9.98% to 34.91% increments in *SEE*'s performance.
- PFA-2 and PFA-3 to MLP-based prediction models yields 9.26% to 30.36% increments in *SEE*'s performance.
- PFA-1, PFA-2 and PFA-3 to MLP-based prediction models yields 36.47% to 84.30% increments in *SEE*'s performance.
- PFA-1 to MLR-based prediction models yields 42.52% to 53.34% increments in *SEE*'s performance.
- PFA-2 to MLR-based prediction models yields -3.26% to 12.63% increments in *SEE*'s performance.
- PFA-3 and MLR-based prediction models yields 10.20% to 24.81% increments in *SEE*'s performance.
- PFA-1 and PFA-2 to MLR-based prediction models yields 50.26% to 79.87% increments in *SEE*'s performance.
- PFA-1 and PFA-3 to MLR-based prediction models yields 106.98% to 146,06% increments in *SEE*'s performance.

- PFA-2 and PFA-3 to MLR-based prediction models yields 14.55% to 48.88% increments in *SEE*'s performance
- PFA-1,PFA-2 and PFA-3 with MLR-based prediction models yields 59.73% to 123.58% increments in *SEE*'s performance.

· Inclusion of PFA-1 and PFA-3 yields the lowest *SEE*'s in all models, independent of whether they are evaluated with SVM, MLP and MLR.

· On the other hand, inclusion of only PFA-2 leads the highest *SEE*'s.

Table 4.15. shows the seven categories each of which includes eight models. For each category, the following comments can be made for SVM, MLP and MLR:

Category	**Models**
Category 1	Model 1,2,3,4,5,6,7,8
Category 2	Model 9,10,11,12,13,14,15,16
Category 3	Model 17,18,19,20,21,22,23,24
Category 4	Model 25,26,27,28,29,30,31,32
Category 5	Model 33,34,35,36,37,38,39,40
Category 6	Model 41,42,43,44,45,46,47,48
Category 7	Model 49,50,51,52,53,54,55,56

Table 4.15 Category and Models

· In Category 1 including the models 1 through 8:

- For SVM, the best *SEE* score is achieved by Model 6 (*SEE* = 3,256 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 1 (*SEE* = 5,4757 mL kg^{-1} min^{-1})
- For MLP, the best *SEE* score was achieved by Model 6 (*SEE*= 3,5568 mL kg^{-1} min^{-1}) and worst *SEE* score is obtained for Model 1 (*SEE* = 6,0133)
- For MLR, the best *SEE* score is achieved by Model 6 (*SEE* = 4,2004 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained for Model 1 (*SEE*= 8,694 mL kg^{-1} min^{-1})

- In Category 2 including the models 9 through 16:

  - For SVM, the best *SEE* score is achieved by Model 14 (*SEE* = 3,8806 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 9 (*SEE* = 6,66 mL kg^{-1} min^{-1})
  - For MLP, the best *SEE* score is achieved by Model 14 (*SEE* = 4,1719 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 9 (*SEE* = 7,1294 mL kg^{-1} min^{-1})
  - For MLR, the best *SEE* score is achieved by Model 14 (*SEE* = 5,1849 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 9 (*SEE* = 11,3844 mL kg^{-1} min^{-1})

- In Category 3 including the models 17 through 24:

  - For SVM, the best *SEE* score is achieved by Model 22 (*SEE* = 4,5475 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 17 (*SEE* = 8,229 mL kg^{-1} min^{-1})
  - For MLP, the best *SEE* score is achieved by Model 22 (*SEE* = 4,975 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 17 (*SEE* = 9,7335 mL kg^{-1} min^{-1})
  - For MLR, the best *SEE* score is achieved by Model 22 (*SEE* = 7,4993 mL kg^{-1} min^{-1}) ant the worst *SEE* score is obtained by Model 17 (17,1071 mL kg^{-1} min^{-1})

- In Category 4 including the models 25 through 32:

  - For SVM, the best *SEE* score is achieved by Model 30 (*SEE* = 3,3496 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 25 (*SEE* = 6,2309 mL kg^{-1} min^{-1})

- For MLP, the best *SEE* score is achieved by Model 30 (*SEE* = 3,6506 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 25 (*SEE* = 6,7894 mL kg^{-1} min^{-1})
- For MLR, the best *SEE* score is achieved by Model 30 (*SEE* = 4,9361 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 25 (*SEE* = 12,1461 mL kg^{-1} min^{-1})

· In Category 5 models including the models 33 through 40:

- For SVM, the best *SEE* score is achieved by Model 38 (*SEE* = 4,3298 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 33 (*SEE* = 7,6408 mL kg^{-1} min^{-1})
- For MLP, the best *SEE* score is achieved by Model 38 (*SEE* = 4,6261 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 33 (*SEE* = 8,7812 mL kg^{-1} min^{-1})
- For MLR, the best *SEE* score is achieved by Model 38 (*SEE* = 5,7213 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 33 (*SEE* = 13,4354 mL kg^{-1} min^{-1})

· In Category 6 models including the models 41 through 48:

- For SVM, the best *SEE* score is achieved by Model 46 (*SEE* = 4,3772 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 41 (*SEE* = 7,7255 mL kg^{-1} min^{-1})
- For MLP, the best *SEE* score is achieved by Model 46 (*SEE* = 4,7442 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 41 (*SEE* = 9,708 mL kg^{-1} min^{-1})
- For MLR, the best *SEE* score is achieved by Model 46 (*SEE* = 6,626 mL kg^{-1} min^{-1}) ant the worst *SEE* score is obtained by Model 41 (*SEE* = 14,2432 mL kg^{-1} min^{-1})

- In Category 7 models including the models 49 through 56:

  - For SVM, the best *SEE* score is achieved by Model 54 (*SEE* = 4,076 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 49 (*SEE* = 7,4954 mL kg^{-1} min^{-1})
  - For MLP, the best *SEE* score is achieved by Model 54 (*SEE* = 4,3894 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 49 (*SEE* = 8,0636 mL kg^{-1} min^{-1})
  - For MLR, the best *SEE* score is achieved by Model 54 (*SEE* = 5,3996 mL kg^{-1} min^{-1}) and the worst *SEE* score is obtained by Model 49 (*SEE* = 12,9673 mL kg^{-1} min^{-1})

Finally, below are given the scatter plots of actual vs. predicted *SEE* scores for each best two models from the seven categories using SVM.

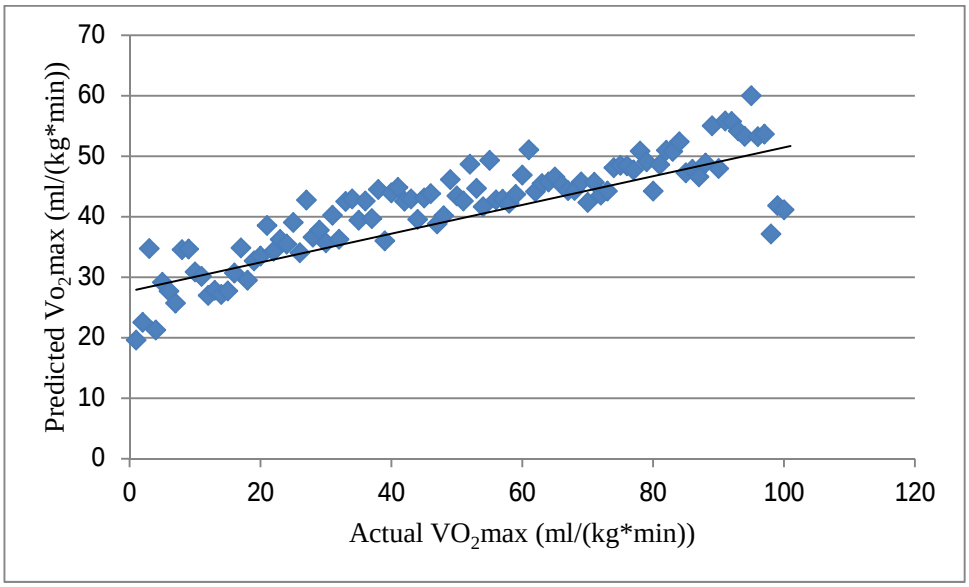


Figure 4.8. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 6 using SVM

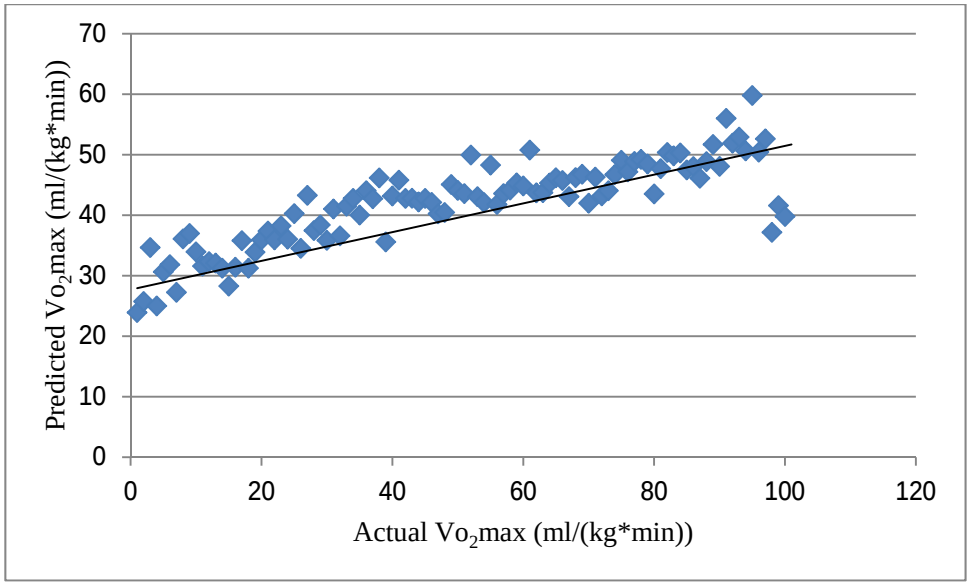


Figure 4.9. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 8 using SVM

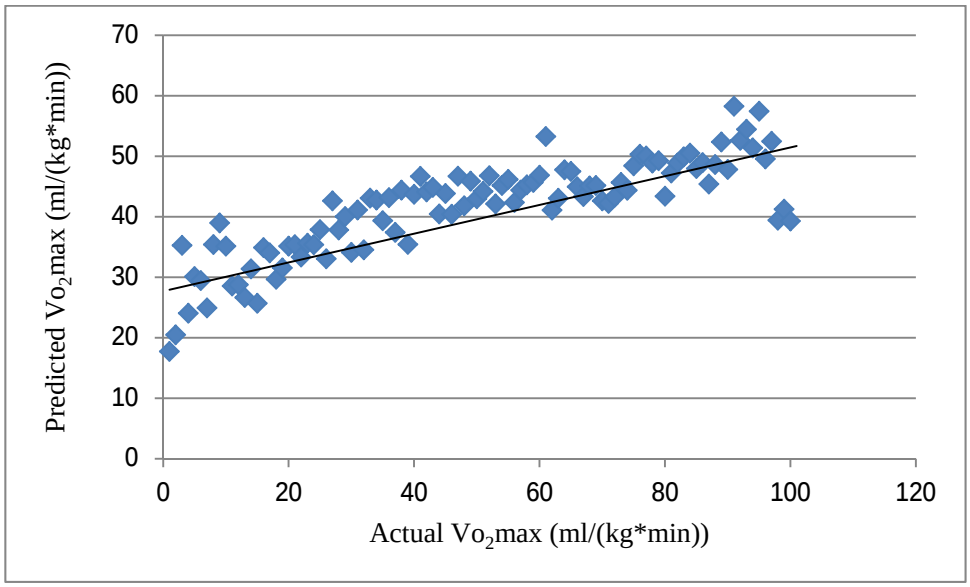


Figure 4.10. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 14 using SVM

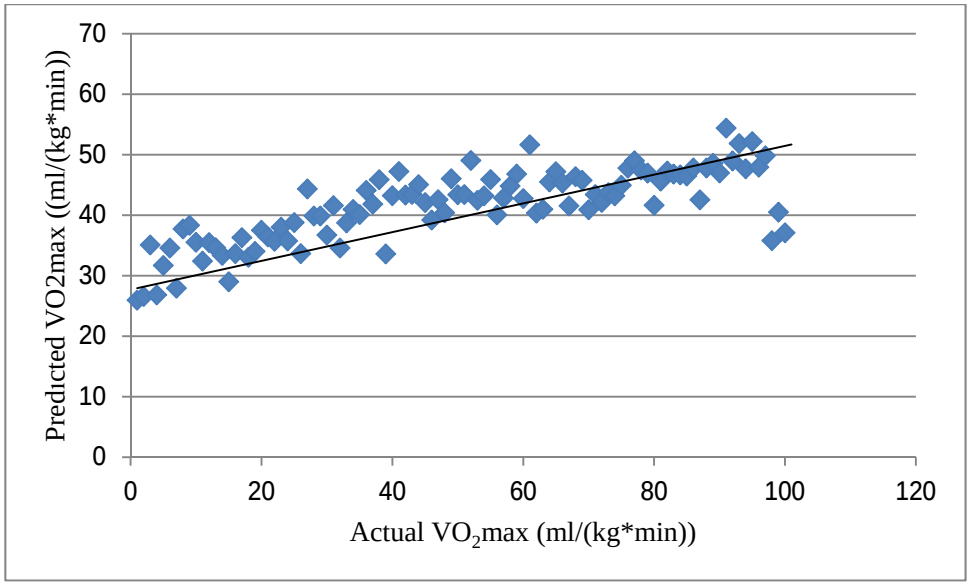


Figure 4.11. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 16 using SVM

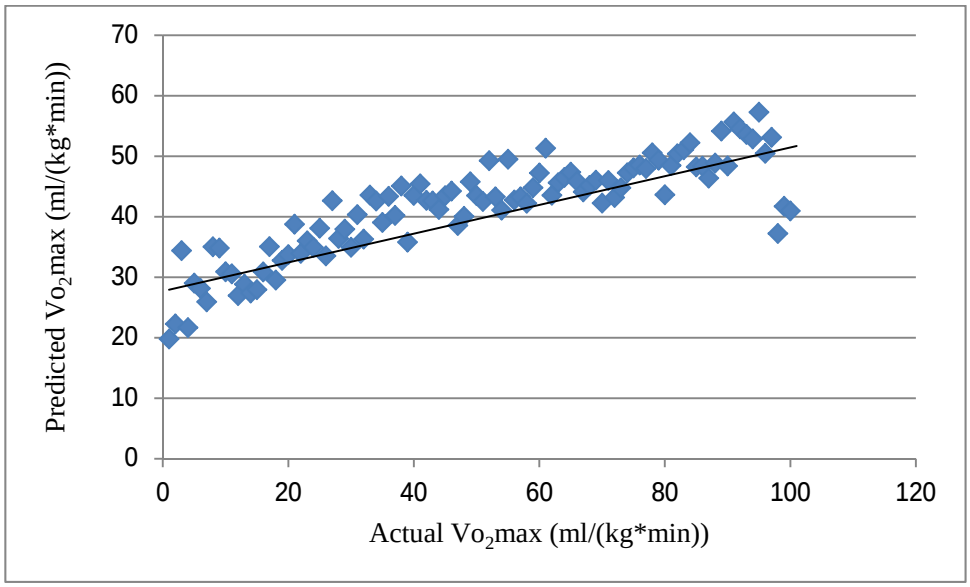


Figure 4.12. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 22 using SVM

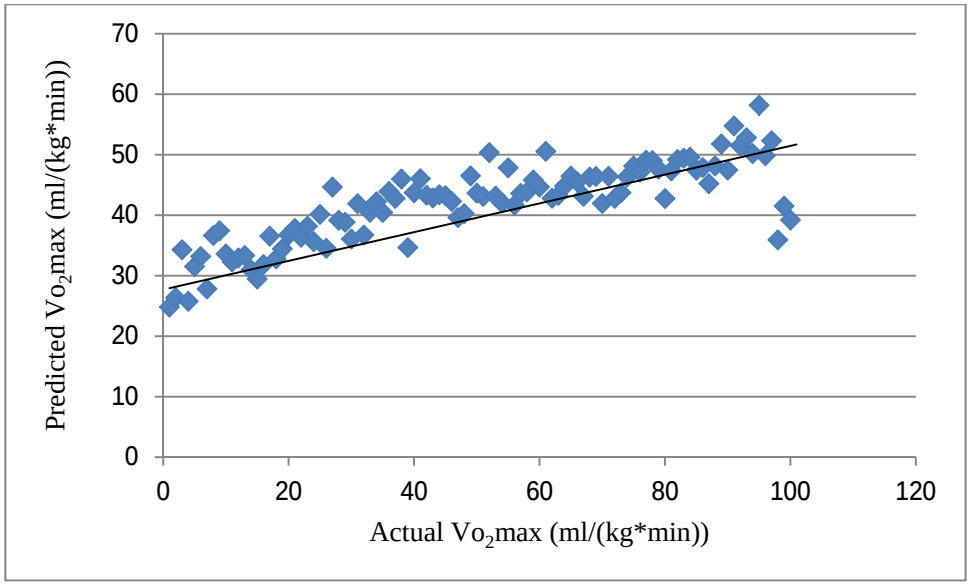


Figure 4.13. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 24 using SVM

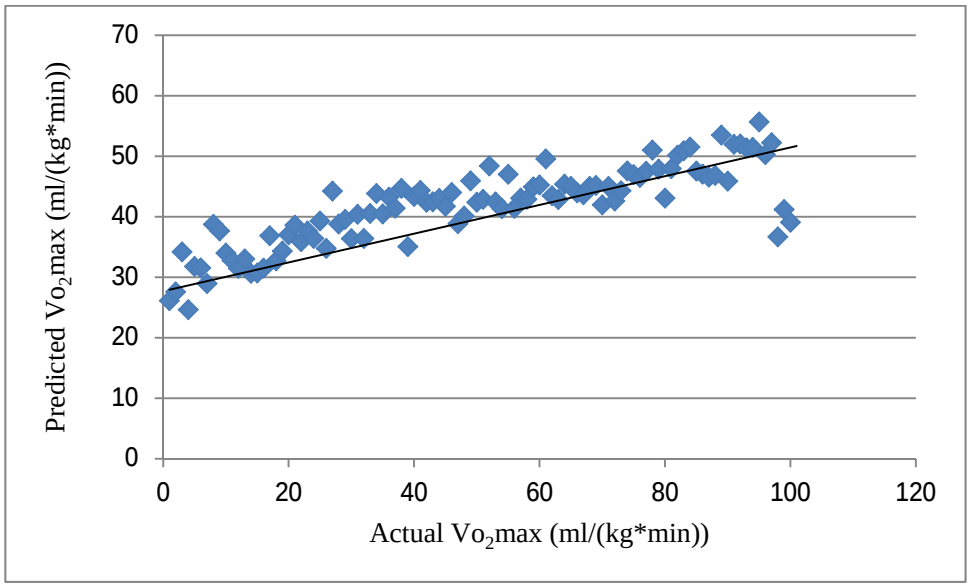


Figure 4.14. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 30 using SVM

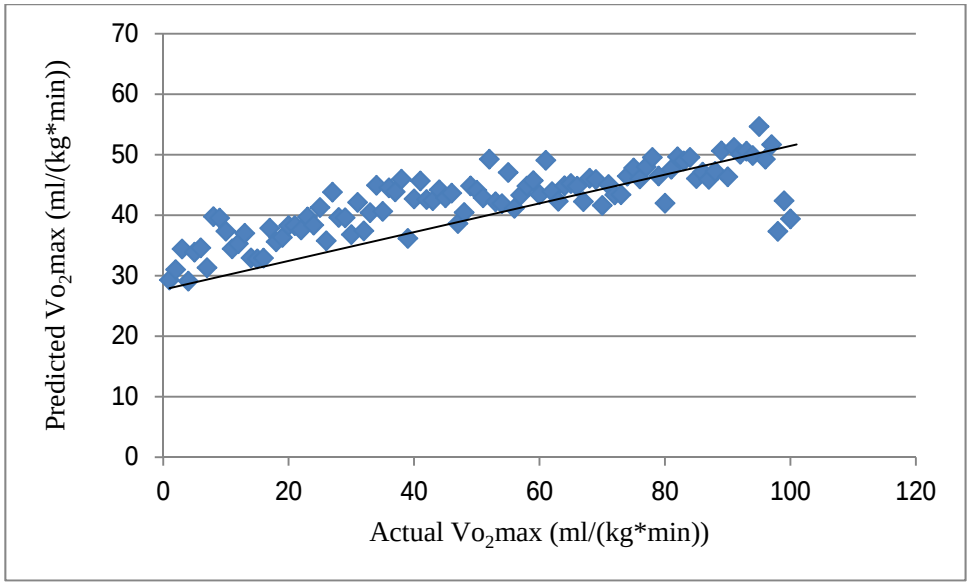


Figure 4.15. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 32 using SVM

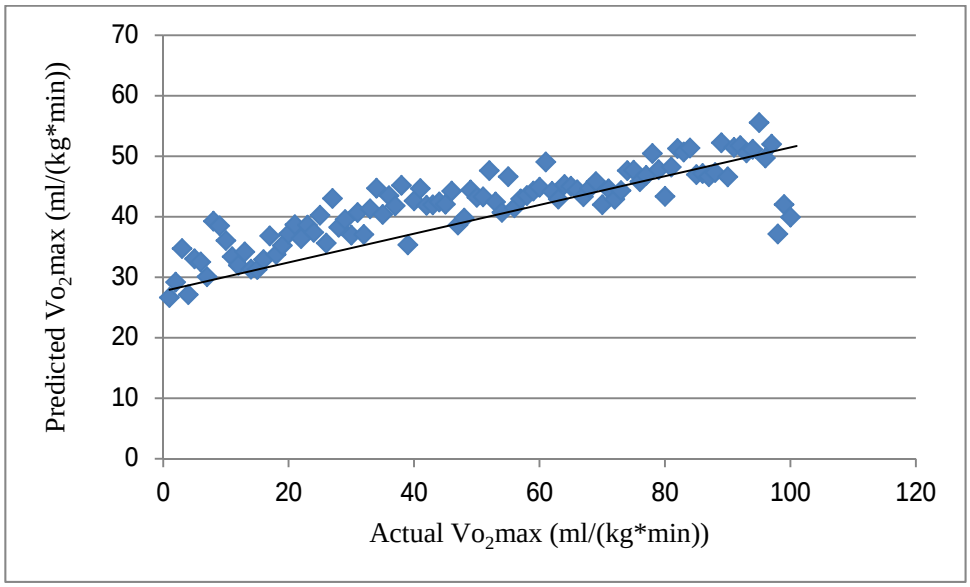


Figure 4.16. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 38 using SVM

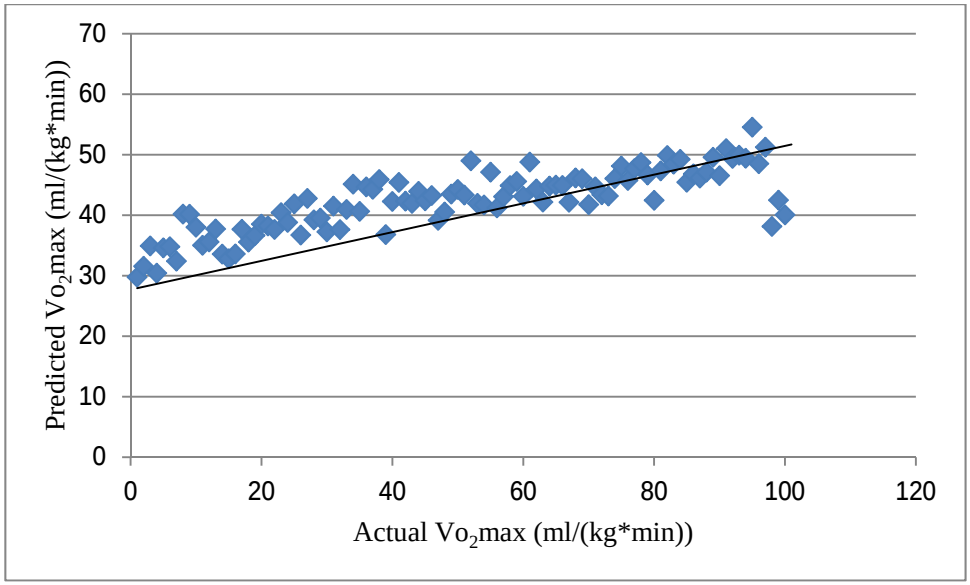


Figure 4.17. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 40 using SVM

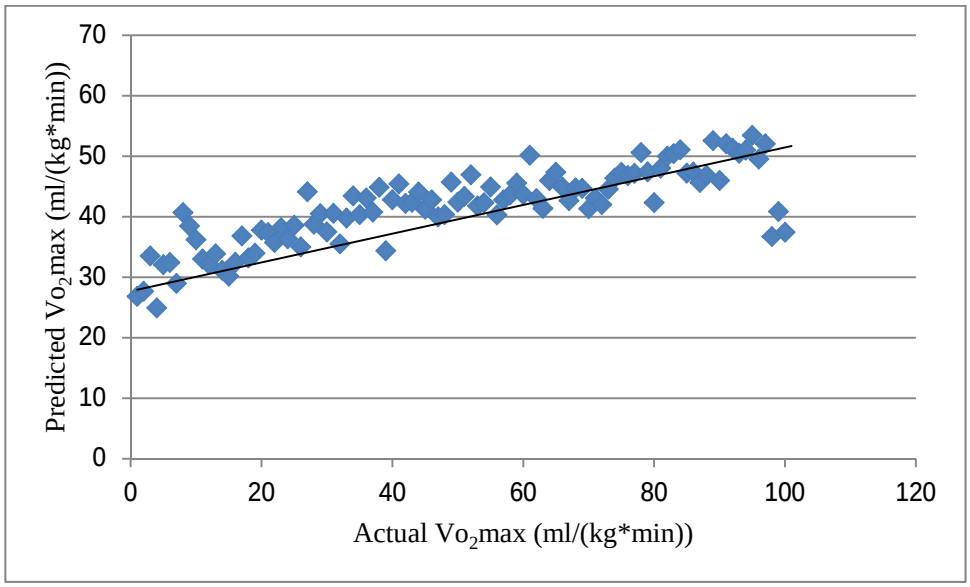


Figure 4.18. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 46 using SVM

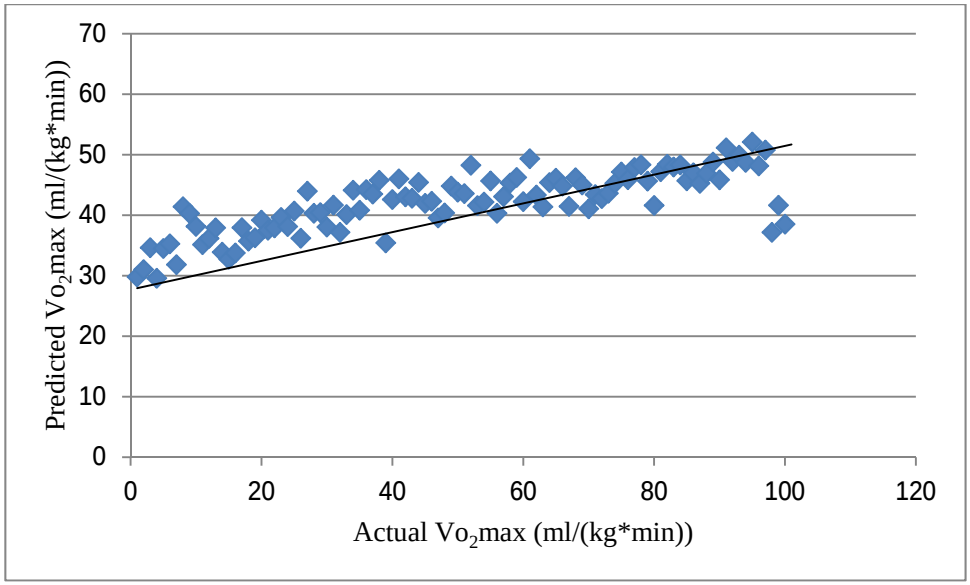


Figure 4.19. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 48 using SVM

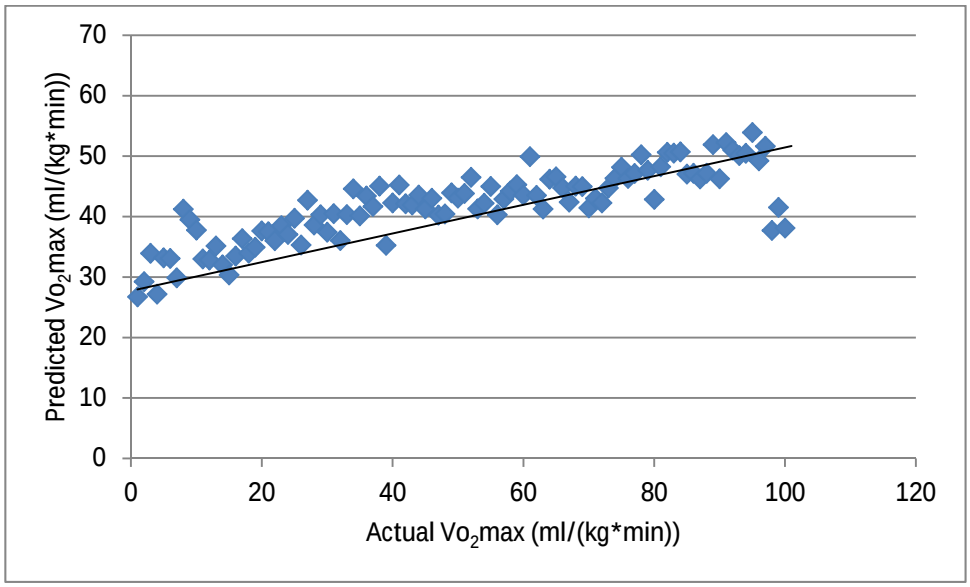


Figure 4.20. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 54 using SVM

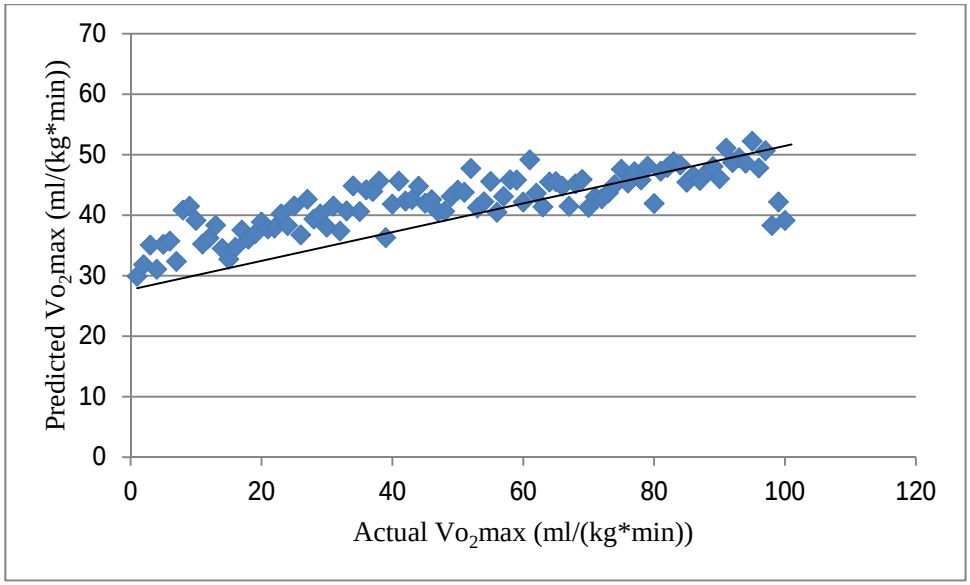


Figure 4.21. Scatter plot of actual VO_2max vs. predicted VO_2max for Model 56 using SVM

5. CONCLUSSION

In this study, SVM, MLP and MLR have been used to develop VO_2max prediction models using submaximal and questionnaire variables. A single dataset with different input variables have been utilized. The dataset includes submaximal variables speed, heart rate, grade self-reported from treadmill test, gender, age, weight, height and the non-exercise variables PFA-1, PFA-2 and PFA-3. By using 10-fold cross validation on the dataset, the performance of the models has been evaluated by calculating ***R***'s and ***SEE***'s. The results show that VO_2max can be predicted with reasonable error rates.

The results have shown that the accuracy of the VO_2max prediction models based on submaximal and standard non-exercise variables can be significantly improved by including questionnaire variables in the models. Inclusion of PFA-1, PFA-2 and PFA-3 in prediction models yields up to 146.06% increments in ***SEE***'s performance. Inclusion of PFA-1 and PFA-3 in predicton models yields the the lowest ***SEE***'s in all models, regardless of whether they are evaluated with SVM, MLP or MLR. On the other hand, inclusion of only PFA-2 leads to the highest ***SEE***'s.

The Model 6 including the predictor variables age, gender, weight, height, speed, PFA-1 and PFA-3 yields the lowest ***SEE***'s (3.25, 3.55, and 4.20) and highest ***R***'s (0.89, 0.92, and 0.93) for the SVM-based, MLP-based and MLR-based models, respectively. On the other hand, the Model 49 including the predictor variables age, gender, weight, height and stage leads the to highest ***SEE***'s (8.22, 9.73 and 17,10) and the lowest ***R***'s (0.66, 0.48, and 0.26) for the SVM-based, MLP-based and MLR-based models, in the order given.

Finally, it could be observed that SVM performs better (i.e. yields lower ***SEE*** and higher ***R***) than other regression methods. The order of the regression methods for prediction of VO_2max, from the best to the worst, is SVM, MLP and MLR.

REFERENCES

AKALAN,C., ROBERGS, R. A., and KRAVİTZ, L., 2008. Prediction of VO_2max from an individualized submaximal cycle ergometer protocol. JEPonline, 11(2):1-17.

AKAY, M. F., AKTÜRK, E., and BALIKÇI, A., 2013. VO_2max Prediction From Submaximal Exercise Test Using Artificial Neural Network. 21. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SİU2013, KKTC.

AKAY M. F., GHAHREMANLOU N.S., 2012. Prediction of Maximum Oxygen Uptake from Maximal and Non-Exercise Variables Using Machine Learning Methods

AKAY, M. F., MOHOMMOUD ZAYİD, E. I., AKTÜRK, E., and GEORGE, J. D., 2010. Artificial neural network-based model for predicting VO_2max from a submaximal exercise test.Expert Systems with Applications, 38(3):2007-2010.

AKAY, M.F.,, C.INAN, D.I. BRADSHAW and J.D. GEORGE, "Support vector regression and multilayer feed forward neural networks for non-exercise prediction of VO2max", Expert Systems with Applications, vol. 36, no. 6, pp. 10112 - 10119, 2009.

CAO, Z-B., MIYATAKE, N., HIGUCHI, M., MIYACHI, M., and TABATA, I., 2010. Predicting VO_2max with an objectively measured physical activity in Japanese men. European Journal of Applied Physiology, vol. 109, no. 3, pp. 465-472.

CAO, Z-B., MIYATAKE, N., HIGUCHI, M., MIYACHI, M., ISHIKAWA-TAKATA, K., and TABATA, I., 2009b. Predicting VO_2max with an objectively measured physical activity in Japanese women. Med Sci Sports Exerc., vol. 42, no. 1, pp. 179-186.

CHATTERJEE, P., BANERJEE, A.K., DAS, P., DEBNATH, P., and CHATTERJEE, P., 2008. Validity of 20 meter multi stage shuttle run test for prediction of maximum oxygen uptake in Indian female university students. Kathmandu Univ Med J, vol. 6, no. 2, pp. 176-180.

CHATTERJEE, S., CHATTERJEE, P., and BANDYOPADHYAY, A., 2005. Validity of Queen's College Step Test for estimation of maximum oxygen uptake in female students, Indian J Med Res., vol. 121, no. 1, pp. 32-35.

CHAVDA, V.V., DR. RAJPUT, M.H., HPARMAR, C., GOKHALE, P.A., MEHTA, H.B., and SHAH, J., 2013. Predicted Maximal Oxygen Consumption (VO_2max) Values Obtained During The Maximal Treadmill Test Using Different Protocols. National Journal of Integrated Research in Medicine, vol. 4, no. 2, pp. 149-155.

CLAIRE E. MARSH,2012. Evaluation of the American College of Sports Medicine submaximal treadmill running test for predicting VO_2max. School of Health Sport and Rehabilitation Sciences, Directorate of Sport, Exercise and Physiotherapy, University of Salford, Salford, United Kingdom. The Journal of Strength and Conditioning Research (Impact Factor: 1.8). 02/2012; 26(2):548-54. DOI: 10.1519/JSC.0b013e3181bac56e

CLEMENTE, C. J., P. C. WITHERS, AND G. G. THOMPSON. 2009. Metabolic rate and endurance capacity in Australian varanid lizards (Squamata; Varanidae; Varanus). Biological Journal of the Linnean Society 97:664-676

DALLECK, L. C., KRAVİTZ, L., and ROBERGS, R. A., 2006. Development of a submaximal test to predict elliptical cross-trainer VO_2max. National Strength & Conditioning Association, 20(2):278-283.

DAVID P. SWAIN, JAMES A. PARROTT, ANNA R. BENNETT, J. DAVID BRANCH, and ELIZABETH A. DOWLING, 2004. Validation of a new method for estimating VO_2max based on VO_2 reserve. Exercise Science Department, Old Dominion University, Norfolk, VA 23529-0196, USA. Medicine &amp Science in Sports &amp Exercise (Impact Factor: 4.48). 09/2004; 36(8):1421-6. Source: PubMed,

DLUGOSZ, E. M., M. A. CHAPPELL, T. H. MEEK, P. SZAFRAÑSKA, K. ZUB, M. KONARZEWSKI, J. H. JONES, J. E. P. W. BICUDO, V. CAREAU, AND T. GARLAND, JR, 2013. Phylogenetic analysis of mammalian maximal oxygen consumption during exercise. Journal of Experimental Biology 216:4712-4721.

DRAPER, N. R., SMITH, H., 1998. Applied Regression Analysis, 3rd edn. New York: John Wiley.

GEORGE J.D., SAMANTHA L. PAULA, ANNETTE HYDEA, DANIELLE I. BRADSHAWA, PAT R. VEHRSA, RONALD L. HAGERA & FRANK G. YANOWITZB,2009.Prediction of Maximum Oxygen Uptake Using Both Exercise and Non-Exercise Data.Measurement in Physical Education and Exercise Science Volume 13, Issue 1, 2009

HAMPSON, DAVID (N.D.). VO_2max: What is it, Why is it so important, and how do you improve it?. Cool Running. Retrieved on 2008-10-14.

HEBB, D. O., 1949. The Organization of Behavior. John Wiley, New York .

HSU, C. W., LIN, C. J., 2002a. A Comparison of Methods for Multi-class Support Vector Machines. IEEE Transactions on Neural Networks, 13:1045–1052

JANE L.P. ROY, I JOE F. SMITH, PHILLIP A. BISHOP, CHRISTOPHER HALLINAN, MINQI WANG, AND GARY R. HUNTER,2004. Prediction of maximal v02 from a submaximal stairmaster test in young women. Department of Human Studies, University of Alabama at Birmingham, Birmingham, AL 35294; Department of Kinesiology, The University of Alabama, Tuscaloosa, AL 35487; School of Human Movement, Recreation and Performance, Victoria University, Melbourne, Australia; Department of Public and Community Health, University of Maryland, College Park, MD 20742

KNERR, S., PERSONNAZ, L., DREYFUS, G., 1990. Single-layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network. In J Fogelman (ed.), Neurocomputing: Algorithms, Architectures and Applications, Springer-Verlag.

LARSEN, G. E., GEORGE, J. D., ALEXANDER, J. L., FELLINGHAM, G. W., ALDANA, S. G., and PARCELL, A. C., 2002. Prediction of maximum oxygen consumption from walking, jogging, or running. Research Quarterly for Exercise and Sport, 73(1):66-72.

LEE, J-M, BASSETT JR, DR, THOMPSON, DL, AND FİTZHUGH, EC, 2011. Validation of the Cosmed Fitmate for prediction of maximal oxygen consumption. J Strength Cond Res 25(9): 2573–2579, 2011

MCCOMB, J. J. R., ROH, D., and WILLIAMS, J. S., 2006. Explanatory variance in maximal oxygen uptake. Journal of Sports Science and Medicine, 5(2):296-303.

MCCULLOCH, W. S. and PITTS, W. H., 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 5:115–133.

NES, B.M., JANSZKY B.M.N.I., VATTEN L.J., NILSEN, T.I.L., ASPENE, S.T., and WISLOFF, U., 2011. Estimating VO_2peak from a non-exercise prediction model, The HUNT Study, Norway, vol. 43, no. 11, pp. 2024-2030.

NIELSON DAVID E., GEORGHE J.D., VEHRS PAT.R., HAGER RON L., WEBB CARRIE V.,2010. Predicting VO_2max in College-Aged Participants Using Cycle Ergometry and Perceived Functional Ability,Measurement in Physical Education and Exercise Science 01/2010; 14(4):252-264. DOI: 10.1080/1091367X.2010.520244

QUINN, ELIZABETH (June 9, 2008). What Is VO_2 Max?. About.com. Retrieved on 2008-10-14.

SCHEMBRE SUSAN M. and RIEBE DEBORAH A., 2011. Non-exercise estimation of VO_2max using the International Physical Activity Questionnaire Meas Phys Educ Exerc Sci. 2011 Jan 1;15(3):168-181.

SCHÖLKOPF, B., SMOLA, A., 2002. Learning with Kernels. MIT Press.

SCHÖLKOPF, B., SMOLA, A. J., WILLIAMSON, R. C., and BARTLETT, P. L., 2000. New Support Vector Algorithms. Neural Computation, 12:1207–1245.

SHAMSI, M.M., ALİNEJAD, H.A., GHADERİ, M., TALEBİ K., BADRABADİ, 2011. Queen's Collage Step Test Predicted VO_2max :The Effect of Stature, Annals of Biological Research, vol. 2, no. 6, pp. 371-377.

SWANK, A. M., SERAPIGLIA, L., FUNK, D. J, ADAMS, K. J., DURHAM, M., and BERNİNG, J. M., 2001. Development of a Branching Submaximal Treadmill Test for Predicting VO_2max. Journal of Strength and Conditioning Research, 15(3):302-308.

VAPNIK, V., 1998. Statistical Learning Theory. Wiley, New York.

VEHRS, P. R., GEORGE, J. D., FELLİNGHAM, G. W., PLOWMAN, S. A., and Dustman-Allen, K., 2007. Submaximal treadmill exercise test to predict VO_2max in fit adults. Measurement in Physical Education and Exercise Science, 11(2):61-72.

WADDOUPS, L., WAGNER, D., FALLON, J., and HEATH, E., 2008. Validation of a single-stage submaximal treadmill walking test. Journal of Sports Sciences, 26(5):491-497.

CURRICULUM VITAE

Eser YÜCEL was born in Adana-Turkey, in 1981. He completed the high school education at Adana. He graduated from the department of Computer Engineering, Mersin University, Mersin, at 2003. He worked several companies such as Kardelen Software and Bossa as a Network and Security Admin. He has been working at Temsa Global as an IT Administrator since 2011. He started MSc program at the department of Computer Engineering with Assoc. Prof. Dr. Fatih Akay. He is married and he has a son.