ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

Msc THESIS

Boubacar Sow

PREDICTION OF HAMSTRING AND QUADRICEPS MUSCLE STRENGTH OF ATHLETES USING MACHINE LEARNING METHODS

DEPARTMENT OF COMPUTER ENGINEERING

ADANA, 2017

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PREDICTION OF HAMSTRING AND QUADRICEPS MUSCLE STRENGTH OF ATHLETES USING MACHINE LEARNING METHODS

Boubacar Sow

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

………………………..….	……………………………..…….	……………………...….….……
Assoc. Prof. Dr. Fatih AKAY	Asst. Prof. Dr. Melis ÖZYILDIRIM	Assoc. Prof. Dr. Serdar YILDIRIM
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-2015-3845.

ABSTRACT

MASTER THESIS

PREDICTION OF HAMSTRING AND QUADRICEPS MUSCLE STRENGTH OF ATHLETES USING MACHINE LEARNING METHODS

Boubacar Sow

ÇUKUROVAUNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF COMPUTER ENGINEERING

Supervisor : Assoc. Prof. Dr. Mehmet Fatih AKAY
Year: 2017, Pages 97
Jury : Assoc. Prof. Dr. Mehmet Fatih AKAY
: Asst. Prof. Dr. B. Melis ÖZYILDIRIM
: Assoc. Prof. Dr. Serdar YILDIRIM

The success of athletes is closely related to the muscles of the thigh, namely the hamstring and quadriceps. Among the different techniques used for measuring the hamstring and quadriceps muscle strength, the use of isokinetic equipments is the most accurate. However, their utilization is associated with several difficulties and limitations. The aim of this study is to build new and more comprehensive models for predicting the hamstring and quadriceps muscle strength of athletes using four main machine learning methods, namely Support Vector Machine (SVM), Multilayer Perceptron Neural Network (MLP), Radial Basis Function Neural Network (RBFNN) and Single Decision Tree (SDT). The root mean square errors (*RMSE*'s) and the multiple correlation coefficients (*R*'s) have been used for computing the prediction errors. On the basis of the results obtained, it has been proved that machine learning methods especially SVM can be used for the hamstring and quadriceps muscle strength prediction with an acceptable accuracy

Key words: Machine Learning, Prediction, Hamstring, Quadriceps

ÖZ

YÜKSEK LİSANS TEZİ

FARKLI MAKINE ÖĞRENME YÖNTEMLERI ILE HAMSTRING VE KUADRISEPS KAS GÜCÜ TAHMINI İÇIN EGZERSIZE DAYALI OLMAYAN YENI MODELLERIN GELIŞTIRILMESI

Boubacar Sow

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Doç. Dr. M. Fatih AKAY
Yıl: 2017, Sayfa: 97
Jüri : Doç. Dr. M. Fatih AKAY
: Yrd. Doç. Dr. B. Melis ÖZYILDIRIM
: Doç. Dr. Serdar YILDIRIM

Hamstring ve Kuadriseps isimli iki tür üst bacak kası, atletlerin spor performanslarıyla yakından ilgilidir. Hamstring ve Kuadriseps kaslarının üstün izokinetik test teknikleriyle ölçümünden gerçeğe yakın sonuçlar elde edilebilmesine rağmen bu yöntemlerin belirli zorlukları ve sınırlamaları vardır. Bu çalışmanın temel amacı, üniversite çağındaki atletlerin Hamstring ve Kuadriseps kas güçlerinin Destek Vektör Makineleri (SVM), Çok Katmanlı Algılayıcılar (MLP), Radyal Temel Fonksiyon Yapay Sinir Ağları (RBFNN), ve Tekli Karar Ağaçları (SDT) kullanılarak daha yeni ve kapsamlı tahmin modelleri oluşturulmasıdır. Tahmin hatalarının hesaplanması için Ortalama Hata Kareleri Toplamı Karekökü (*RMSE*) ve Çoklu Korelasyon Katsayısı *(R)* değerleri kullanılmıştır. Temel olarak elde edilen sonuçlara bakıldığında başta SVM olmak üzere makine öğrenme metodlarının kabul edilebilir ve doğru sonuçlar vermesi nedeniyle Hamstring ve Kuadriseps kas gücü tahmininde kullanılabileceği kanıtlanmıştır.

Anahtar Kelimeler: Destek Vektör Makineleri, Tahmin, Hamstring, Kuadriseps

GENİŞLETİLMİŞ ÖZET

Kas gücü, bir kasın belli bir dirence karşı tek seferde üretebildiği maksimum güç miktarıdır. Kaslar insan vücudunun hareket etmesinden ve vücudun dik duruşunun sağlanmasından sorumludur. Bacaklarda bulunan hamstring ve kuadriseps gibi güçlü kaslar bünyeye vücudun dik durması için gerekli güç ve dayanıklılığı sağlar. Kas gücünün günlük aktiviteler gereken performans için ve kronik rahatsızlıkların önlenmesi açısından önemli bir rol oynadığı görülmüştür. Özellikle atletlerin, yapmış oldukları spor dallarında başarılı sonuçlar elde edebilmeleri için yüksek seviyede bir kas gücüne sahip olmaları önemlidir. Hamstring ve Kuadriseps isimli iki tür üst bacak kası, atletlerin spor performanslarıyla yakından ilgilidir. Üst bacağın arka kısmında yer alan Hamstring ve ön kısmında yer alan Kuadriseps kaslarının günlük aktiviteler için de büyük önemi vardır. Kuadriseps kasları zıplama ve tekmelemede önemli rol oynarken, Hamstring kasları ise koşu aktivitelerini kontrol eder ve dönüşler ve darbeler esnasında diz bölgesinin stabilizasyonuna yardımcı olur.

Günümüzde Hamstring ve Kuadriseps kaslarının gücünün doğrudan ölçülebilmesi adına çeşitli teknikler uygulanmaktadır. İzokinetik testler bu alandaki en popüler ölçüm tekniği olmuştur. Hamstring ve Kuadriseps kaslarının üstün izokinetik test teknikleriyle ölçümünden gerçeğe yakın sonuçlar elde edilebilmesine rağmen bu yöntemlerin belirli zorlukları ve sınırlamaları vardır. Örneğin bu ölçümler için kullanılan aletler çok zor bulunabilmeleriyle birlikte oldukça maliyetlidirler. Ayrıca bu yolla yapılan ölçümlerin kalabalık test grupları üzerinde uygulanması, aynı zamanda sadece bir adet ölçüm yapılabildiğinden dolayı oldukça zordur. Bununla birlikte bu teknikler yardımıyla doğru sonuçlar elde edilebilmesi için profesyonel olarak eğitilmiş personeller ve uzun zaman dilimleri gerekmektedir. Verilen durumlar değerlendirildiğinde, araştırmacılar direkt ölçüm yöntemlerinin yerine makine öğrenmesi metodları kullanılarak Hamstring ve Kuadriseps kas gücünün tahmin edilebilmesini sağlamayı

önermişlerdir. Bu konuda bazı çalışmalar yapılmıştır. Ancak bu çalışmalar, kullanılan makine öğrenmesi metodu sayısı ile ilgili kısıtlamalar içermektedir. Bu kısıtlamalar, tahmin modellerinde kullanılan her bir tahmin değişkenine ait performansın ayrıntılı bir şekilde kavranmasını engellemiştir. Bu sebepten dolayı Hamstring ve Kuadriseps kas gücü ölçümünde kullanılan tahmin değişkenlerinin sonuca etkilerinin kapsamlı bir şekilde kavranabilmesi için daha yeni ve fazla sayıda tahmin modeli üreten çalışmaların yürütülmesine ihtiyaç duyulmuştur.

Bu çalışmanın temel amacı, üniversite çağındaki atletlerin Hamstring ve Kuadriseps kas güçlerinin Destek Vektör Makineleri (Support Vector Machines, SVM), Çok Katmanlı Algılayıcılar (Multilayer Perceptron, MLP), Radyal Temel Fonksiyon Yapay Sinir Ağları (Radial Basis Function Neural Network, RBFNN), ve Tekli Karar Ağaçları (Single Decision Tree, SDT) kullanılarak daha yeni ve kapsamlı tahmin modelleri oluşturulmasıdır. Klasik ve Statik Eğitim olarak bilinen iki adet popüler eğitim metodunun kullanılmasıyla, Hamstring ve Kuadriseps kas güçleri Cinsiyet (Sex), Yaş (Age), Boy (Height), Ağırlık (Weight), Vücut Kitle İndeksi (Body Mass Index, BMI) ve Spor Dalı (Sports Branch, SB) olmak üzere 6 adet bilimsel olarak ilgili tahmin değişkeninden yararlanılarak tahmin edilmiştir. Bu tez çalışmasında kullanılan veri seti, Gazi Üniversitesi Beden Eğitimi ve Spor Yüksekokulu bünyesindeki 70 adet atlete ait bilgileri içermektedir. Çalışma için seçilen atletler izokinetik dinamometre kullanılarak değerlendirilmiştir. 10-fold Cross Validation yönteminin kullanılmasıyla tahmin modellerine ait genel hata oranları hesaplanmıştır. Tahmin hatalarının hesaplanması için Ortalama Hata Kareleri Toplamı Karekökü (Root Mean Square Error, *RMSE*) ve Çoklu Korelasyon Katsayısı (Multiple Correlation Coefficient, R) değerleri kullanılmıştır.

Bu çalışmalardan elde edilen sonuçlara göre SVM tabanlı tahmin modelleri ile MLP, RBFNN ve SDT tabanlı tahmin modellerine kıyasla daha yüksek performans sağlanmıştır. Bir başka deyişle, Hamstring ve Kuadriseps kas güçlerinin tahmin edilmesinde SVM tabanlı modeller ile atletler üzerinde kullanılan eğitim metodlarından ve kullanılan tahmin modeli kategorisinden bağımsız olarak en düşük *RMSE* değerleri bulunmuştur. Daha ayrıntılı olarak, SVM tabanlı modeller ile Hamstring kas gücünün tahmini sonucunda *RMSE* değerleri ortalama olarak Klasik Eğitim yönteminde 20.25 Nm, Statik Eğitim yönteminde ise 20.19 Nm elde edimiştir. Bunun yanında *RMSE* değerleri Kuadriseps kas gücünün tahmininde ortalama olarak Klasik Eğitim metodu için 30.62 Nm, Statik Eğitim metodu için ise 29.40 Nm şeklindedir. Bunun yanı sıra, tahmin değişkeni olarak boy ve ağırlık yerine BMI kullanılarak oluşturulan SVM tabanlı tahmin modelleriyle Hamstring kas grubu için *RMSE* değerleri ortalama olarak Klasik eğitim metodunda 21.60 Nm ve Statik eğitim metodunda 21.89 Nm değerlerini alırken, aynı değişkenler için Kuadriseps kas grubunda *RMSE* değerleri Klasik eğitimde 32.98 Nm ve statik eğitimde 31.62 Nm şeklindedir.

Hamstring kas gücünün tahmin edilmesinde kullanılan makine öğrenme metodlarının performans açısından sıralanması, kullanılan kas eğitim metodu ve tahmin modeli türü hesaba katılmaksızın, en düşük *RMSE* değerinden en yüksek *RMSE* değerine doğru olacak şekilde SVM, MLP, RBFNN ve SDT biçimindedir. Bununla birlikte bu sıralama Kuadriseps kas gücü ölçümünde BMI tahmin değişkeninin kullanıldığı ve kullanılmadığı, ayrıca hem Klasik hem de Statik eğitim metodlarının kullanıldığı toplam durumlarda ise SVM, MLP, SDT ve RBFNN şeklindedir. SVM tabanlı ve diğer tahmin modellerine ait *RMSE* değerlerinin karşılaştırılması sonucu, SVM tabanlı modellerde *RMSE* değerinin azalma miktarının yüzdelik oranlar cinsinden değerleri ortalama olarak Klasik eğitimde %7.24, %14.32 ve %20.61, Statik eğitimde ise %7.89, %14.48 ve %20.14 olmaktadır. Aynı oranlar Kuadriseps kas gücü ölçümlerinde Klasik eğitimde %5.65, %11.19 ve %15.47, Statik eğitimde ise %6.21, %11.34 ve %16.11 değerlerini almaktadır. Bunun yanında Hamstring kas gücü tahmininde tahmin değişkeni olarak boy ve ağırlık yerine BMI kullanıldığında SVM tabanlı modellerle *RMSE* azalma oranları ortalama olarak Klasik eğitimde %6.44, %13.04 ve %19.51 iken Statik eğitimde %6.33, %12.29 ve %18.21 olmaktadır. Aynı değerler

Kuadriseps kas gücü ölçümünde Klasik eğitimde %6.29, %12.28 ve %16.15 iken Statik eğitimde ise %7.03, %13.16 ve %17.72 şeklindedir.

Temel olarak elde edilen sonuçlara bakıldığında başta SVM olmak üzere makine öğrenme metodlarının kabul edilebilir ve doğru sonuçlar vermesi nedeniyle Hamstring ve Kuadriseps kas gücü tahmininde kullanılabileceği kanıtlanmıştır.

DEDICATION

... to my late sister Yaye Béla.

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Assoc. Prof. Dr. M. Fatih AKAY, for his motivation, patience and enthusiasm. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor for my MSc thesis.

I would like to thank members of the MSc thesis jury, Asst. Prof. Dr. Buse Melis ÖZYILDIRIM and Assoc. Prof. Dr. Serdar YILDIRIM, for their suggestions and corrections.

I would also like to thank Cukurova University Scientific Research Projects Center for supporting this work (Project no: FYL-2015-3845).

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**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF ABBREVIATIONS

MLP	:	Multi-Layer Perceptrons
RBF	:	Radial Basis Function
RBFNN		Radial Basis Function Neural Network
SVM	:	Support Vector Machines
SDT	:	Single Decision Tree
1-RM		1-repetition maximum
RMSE	:	Root Mean Square Errors
R	:	Multiple Correlation Coefficient
CT	:	Classic Training
ST	:	Static Training
Nm	:	Newton-meters
W	:	Weight
G	:	Gender
H	:	Height
SB	:	Sport Branch
BMI	:	Body Mass Index

I. INTRODUCTION

1.1. Overview of Hamstring and Quadriceps Muscle Strength

Muscular strength is the maximum amount of force that a muscle can produce against some form of resistance in a single effort. Muscles are responsible for the movement of human body and maintaining posture. The creation of force, stabilizing and balancing the skeleton or generating movement are the main roles of skeletal muscles. Strong muscles in the legs such as hamstrings and quadriceps, but also muscles of back, chest, and abdomen give a person the necessary strength to stand up straight and maintain good posture. The vital role of muscular strength in the performance of daily activities of life and sport field, as well as in the prevention of chronic disease, is widely recognized (Ruiz et al, 2008).

In sport fields, in order to remain competitive, athletes must optimize their muscular strength. Therefore in order to achieve success in their respective sports, besides certain very important parameters such as endurance, flexibility and speed, athletes must also develop a high level of muscular strength that is considered to be in close relation with the mentioned parameters. As an example of the close relation between muscular strength and the above mentioned parameters, in addition to speed practices, performing training exercises in order to enhance the ability of fast-running include muscle strength exercises as well, because there is a close relation between the muscle strength of athletes and their speed (Di Michele et al, 2012). Also, a good follow-up of the muscular strength of an athlete has a capital role in setting up an adequate training program, attaining defined objective in term of performance, avoiding possible injuries that occur because of the vulnerability of athletes and identifying appropriate medical therapy to treat them. It is important to mention that certain well-known factors such as gender, age and level of physical conditioning can affect muscular strength (Faigenbaum et al, 2010).

The success of athletes is closely related to the muscles of the thigh, namely the hamstring and quadriceps (Aginsky et al (2014), Greco et al (2012) & Jordan et al (2015)). Taking place respectively on the back and the front of the upper leg, the hamstring and quadriceps muscles play a crucial role in many daily activities. Although working together, every muscle performs a specific function. The functions of quadriceps muscles are to flex the leg at the knee, to allow jumping and kicking ,whereas the functions of hamstring muscles are to extend the knee joint, straightening the leg, to control running activities. Figure 1.1 shows the hamstring and quadriceps muscles.

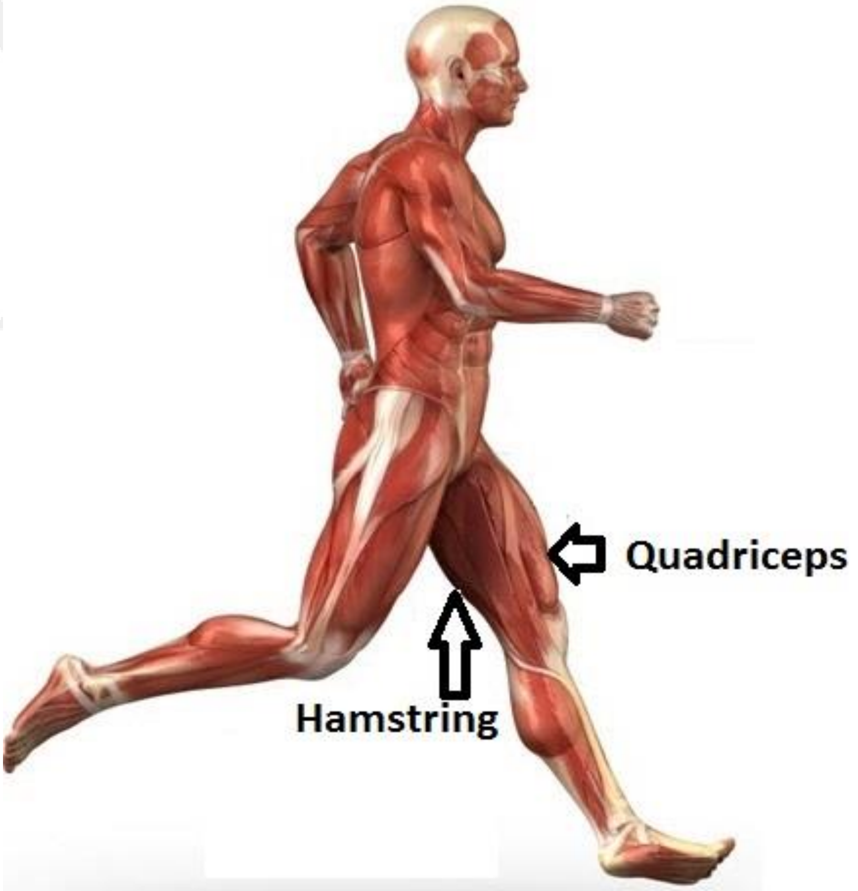

Figure 1.1. Hamstring and quadriceps muscles

Nowadays various techniques are proposed for directly measuring the hamstring and quadriceps muscle strength such as certain tests performed by the use of isokinetic technique (Ustun et al, 2013), dynamometer (Ford-Smith et al, 2001) and tensiometer (Clarke, 2013). Nevertheless, tests with isokinetic techniques have become the most widely known. The usual purpose of isokinetic testing is to assess muscle function during different intervals of exercise using a computer-driven device. Isokinetic exercise is usually conducted by using dynamometers which maintain a constant speed during movements. Before doing the isokinetic exercises, the participant is stationed in such a way that the body movement to be measured is isolated. The dynamometer is then adjusted at different speed and the force exerted by the participant can be measured over the entire range of movement (Kılınç et al, 2015). Particularly muscle performance testing is realized safely because muscles or group of muscles can be targeted at different angle options and speeds by isokinetic techniques. Due to numeric measurements they give for the performance of muscles, isokinetic techniques are widely used for monitoring of athletes' performance (Manca et al, 2015).

1.2. Motivation and Purpose of this Thesis

Although isokinetic equipments give the most accurate results, their utilization is associated with several difficulties and limitations. Some of these difficulties and limitations are listed below:

- The necessary equipment normally used to perform the measurements is costly and not easily available. More particularly, it is within the research projects at educational institutions or in certain sports research centers that this kind of measurement is carried out.
- The required equipment is voluminous, i.e. it is not easily movable. Therefore it is extremely difficult to use the equipment at different places for measurement purposes.

- Due to the fact that only one individual can be assessed at a time, measuring directly the muscle strength for a large population is almost impossible. On the other hand, the data obtained from a direct measurement via isokinetic devices demands qualified staff and time consuming practices in order to be interpreted.

Considering the above mentioned drawbacks, rather than directly measuring the hamstring and quadriceps muscle strength, researchers have recently used empirical equations for the hamstring and quadriceps muscle strength prediction. In that respect, some studies have been carried out. However, these previous studies have important limitations in terms of number of models used and they did not use the promising machine learning algorithms for the prediction of hamstring and quadriceps muscle strength. These limitations prevent having detailed comprehension of the performance of each predictor variable used for the hamstring and quadriceps muscle strength prediction. Therefore additional studies that develop new and more comprehensive prediction models are absolutely needed in order to determine the predictor variables that have greater impact on hamstring and quadriceps muscle strength prediction and to have more precise results for hamstring and quadriceps muscle strength prediction.

The aim of this study is to build new and more comprehensive models for the prediction of hamstring and quadriceps muscle strength of young athletes using four main machines learning methods, namely Support Vector Machine (SVM) Multilayer Perceptron Neural Network (MLP), Radial Basis Function Neural Network (RBFNN) and Single Decision Tree (SDT). Using two well-known training methods, namely classic muscle training and static muscle training, the hamstring and quadriceps muscle strength have been predicted by utilizing six scientifically relevant predictor variables which are: body mass index (BMI), sport branch (SB), age(A), height(H), weight(W), and gender(G). These predictor variables have been proved in literature to be in a correlation with hamstring and quadriceps muscle strength (Lemmer et al (2016) & Lindle et al (1997)). By carrying out 10-fold cross-validation, the generalization error of the prediction models has been calculated. The root mean square errors (*RMSE*'s) and multiple correlation coefficients (R's) have been used for computing the prediction errors.

1.3. Literature Review

In the related literature, to the best of our knowledge, there is no study that used machine learning methods to predict hamstring and quadriceps muscle strength. However, to the best of our knowledge, there exists only one study (McNair et al, 2011) that estimates the quadriceps muscle strength using several empirical equations (1-repetition maximum equations). Although directly related to this thesis because predicting quadriceps muscle strength, these equations are not derived from machine learning algorithms.

As mentioned above in (McNair et al, 2011), authors used prediction equations known as 1-repetition maximum (1-RM) equations which are listed in the table below:

Table 1.2. 1-RM prediction equations

Equations	**Formula**
Adams,1993	W/(1 − 0.02 × R)
Berger, 1970	W/(1.0261e−0.00262 × R)
Brown, 1996	(R × 0.0328 + 0.9849) × W
Brzycki, 1993	W/(1.0278 − 0.0278 × R)
Epley, 1985	(0.033 × R × W) + W
Kemmler et al, 2006	W × (0.988 − (0.0000584 × R3 + 0.00190 × R2 + 0.0104 × R))
Lander, 1985	W/(1.013 − 0.0267123 × R)
Lombardi, 1989	R0.1 × W
Mayhew et al, 2004	W/(0.522 + 0.419e−0.055 × R)
O'Conner et al, 1989	0.025 × (W × R) + W
Wathen, 1994	W/(0.488 + 0.538e−0.075 × R)

Where W = weight used for repetitions to failure; R = repetitions to failure.

Eighteen Individuals with osteoarthritis (OA) of the knee joint who attended a rehabilitation gymnasium participated to the tests. The participants performed two exercises, namely the leg press extension and the knee extension. Eleven prediction equations have been utilized for the prediction of the maximal strength of the quadriceps muscle. The Poliquin chart which is widely utilized in fitness and exercises centers has also been utilized. The results of their study have shown that the Brown, Brzycki, Epley, Lander, Mayhew et al, Poliquin, and Wathen prediction equations showed the highest prediction accuracy for the knee extension exercise with typical errors varying between 3% and 4%. Whereas for the knee press exercise, the Adams, Berger, Kemmler et al, and O'Conner et al equations showed the highest accuracy for predicting maximal strength of quadriceps muscle with the typical errors ranged from 5.9–6.3%. Therefore authors concluded that prediction equations can be used to assess quadriceps maximal strength in individuals with a knee joint with OA.

The following studies have been carried out and published as part of the thesis using the material included in this thesis:

- The first preliminary study that compares the performance of different machine learning algorithms for predicting hamstring and quadriceps muscle strength was in (Akay et al, 2015). In the Study, SVM, RBFNN and SDT have been applied to develop a single model for hamstring and quadriceps muscle strength prediction. The authors have concluded that SVM-based model yields lower RMSE's than the ones obtained using RBFNN-based and SDT-based prediction models. In more detail, the results they obtained show that the SVM-based quadriceps muscle strength prediction model yields an RMSE value of 24.16 W and an R value of 0.89 for classic training; and an RMSE value of 22.40 Nm and an R value of 0.90 for static training, respectively. Similarly, the hamstring muscle strength prediction model gives an RMSE value of 15.54 Nm and an R value of 0.90 for classic training; and finally an RMSE value of 15.74 Nm and an R value of 0.90 for static training, respectively.
- The second study was in (Akay et al, 2016). The authors used SVM, RBFNN and SDT to develop two categories of hamstring and quadriceps muscle strength prediction models. For the first category, predictor variables weight, height, gender, age and SB have been used. The second category included the same predictor variables except that in place of predictor variables height and weight, sport branch has been utilized. The authors also used four different muscle training methods including classic training(CT), static training(ST), static training followed by a rest of 5 minutes (ST-5MIN) and static training

with a rest of 15 minutes(ST-15MIN) directly after the static training for the hamstring and quadriceps muscle strength prediction. In summary their study shows that the SVM-based models show the best performance among the machine learning methods used, regardless of which muscle training method has been performed by the participants or whether quadriceps or hamstring muscle strength is predicted. The hamstring muscle strength prediction models give better result than the quadriceps muscle strength prediction models, independent of which machine learning method or muscle training method has been utilized. The performances of classic training and static training models for the prediction of the hamstring and quadriceps muscle strength are comparable. Whereas ST-5MIN and ST-15MIN models show lower performance than CT and ST models. All prediction models that have predictor variables height and weight in place of BMI exhibit relatively higher performance regardless of the regression method or training type utilized.

- The third study has been carried out in (Sow et al, 2016). In this study SVM, MLP and SDT have been applied and thirty prediction models have been developed for only the hamstring muscle strength using the classic training. The thirty prediction models have been divided in four categories depending on the number of predictor variables employed in the prediction model. The results have shown that SVM-based prediction models outperform MLP-based and SDT-based prediction models independent of the category of prediction models applied. Besides, the authors concluded that among the thirty prediction models developed, prediction models including predictor variables gender and weight show relatively the best performance whereas prediction models including age and sport branch give lower performance for the hamstring muscle strength prediction. The category of models with four predictor variables on the average leads to the best results whereas the category of prediction models with one predictor variable shows the worst results.

1.4. Contributions of the Thesis

The major differences between this thesis and the previous study can be listed as follow:

1. This study is the first that propose to use machine learning methods for hamstring and quadriceps muscle strength prediction.
2. In this thesis, new, and more comprehensive models for the hamstring and quadriceps muscle strength prediction have been developed.
3. In this thesis, several predictor variables including gender, weight, height, age, SB and BMI have been utilized. On the contrary, in (McNair et al, 2011) 1-repetition maximum prediction equation with only two predictor variables, namely repetitions to failure and weight have been used.
4. 70 young athletes participated in the creation of the dataset used in this thesis, whilst in (McNair et al, 2011) only 18 subjects attended the study.

1.5. Overview of the Dataset

The dataset used in this thesis included 70 athletes from the College of Physical Education and Sport at Gazi University. The athletes who participated in the experiments have been assessed by the use of an isokinetic dynamometer (Figure 1.2).

Figure 1.2. Isokinetic dynamometer

The right upper leg hamstring and quadriceps muscle strength of all athletes have been measured by the use of an isokinetic dynamometer. The classic training which is a light run for 5 minutes and the static training which involved a light run for 5 minutes followed by an active static stretching for 4 minutes, have been performed in order to measure the hamstring and quadriceps muscle strength. The variables age, gender, height, body mass index (BMI), weight and sport branch as predictors, but also the variables hamstring and quadriceps muscle strength as targets have been used for the dataset creation. In Table 1.3, the maximum, minimum, average and standard deviation values for each predictor and target variables are given.

Table 1.3. Statistics of the dataset

Variables	**Minimum**	**Maximum**	**Mean**	**Standard Deviation**
Gender	0	1.00	0.36	0.48
Age (Year)	19.00	38.00	21.79	3.06
Height (m)	1.57	2.02	1.71	0.08
Weight (kg)	45.00	93.00	62.04	11.27
BMI (Kg/m2)	16.46	26.31	21.06	2.57
SB	1.00	17.00	9.31	5.03
Hamstring Muscle Strength (Nm)(classic training)	50.10	195.90	111.84	36.10
Hamstring Muscle Strength (Nm)(static training)	72.2	285.2	154.77	54.80
Quadriceps Muscle Strength (Nm)(classic training)	61.2	197.7	111.62	36.45
Quadriceps Muscle Strength (Nm)(static training)	85.2	278.1	157.01	54.41

2. OVERVIEW OF METHODS

In this thesis as mentioned above, in order to predict the hamstring and quadriceps muscle strength, SVM, MLP, RBFNN and SDT have been used.

2.1. Support Vector Machines

SVM is a supervised machine learning algorithm invented by Vladimir N. Vapnik and Alexey Chervonenkis in 1963 and introduced by Boser, Guyon and Vapnik in 1992. Founded on the intuitive geometric concepts of large margin separation and regularized risk minimization, SVM's are well embedded into statistical learning theory (Dogan et al, 2016). Nowadays, among the machines learning methods, SVMs are among the most widely used methods. The popularity of SVM's is due to its remarkable accuracy in both prediction and classification. SVM, in general, has been reported to be superior to other machine learning methods especially in the field of sport physiology (Akay et al (2016), Abut & Akay (2015)). Because of these reasons, in this study SVM has been chosen as the main regression method for the prediction of hamstring and quadriceps muscles strength.

2.1.1. Linear SVM

Considering a training set of N data points, $S = \{x_k, y_k\}$ where $x_k \in R^n$ is an input vector and $y_k \in R$ is an output vector. SVM outputs an optimal hyperplane which categorizes new examples. A vector w and a bias b define the equations of hyperplanes. Decision function for the hyperplanes is $w^t - x + b = 0$. In order to maximize the margin of separation ρ , an optimal hyperplane is created. SVM constructs a classifier

$$f(x) = sign(w^t.x + b) \quad (2.1.)$$

The limit of parameters *w* and *b* are given in (2.2):

$$\min_i |w.x_i + b| \geq 1. \tag{2.2.}$$

If it does not result any problem after the vectors dividing process, then as a result of this process, if the distance separating the nearest vector and the hyperplane is maximum, it is again divided by a hyperplane. Therefore, a dividing hyperplane must satisfy the constraints given in (2.3) in order to be in a standard form.

$$y_i\left(w.x_i + b\right) \geq 1,\ i = 1, 2, ..., n. \tag{2.3.}$$

An input point x_i that has the distance *d* from the hyperplane $\left(\mathbf{w}, b\right)$ is,

$$d((w,b), xi) = \frac{y_i(x_i.w+b)}{\|w\|} \geq \frac{1}{\|w\|} \tag{2.4.}$$

ρ can be calculated as

$$\rho = \frac{2}{\|w\|} \tag{2.5.}$$

Hence, SVM searches for a separating hyperplane by minimizing

$$\Phi\left(w\right) = \frac{1}{2}\left(w.w\right) \tag{2.6.}$$

$\Phi(w)$ in (2.6) can be minimized by performing the structural risk minimization principle,

$$\|w\|^2 \le c. \tag{2.7.}$$

h is the series of standard hyperplanes in space that has n-dimension and is limited by,

$$h \le \min\left[\left(R^2 c\right), d\right] + 1 \tag{2.8.}$$

where R is a radius of a hypersphere that surrounds all training vectors. Consequently, minimizing (2.6) is equal to minimization of the upper bound.

The limitations of (2.3) can be reduced by presenting slack variables $\xi_i \ge 0,\ i = 1, 2, ..., n$; therefore (2.3) can be rewritten as

$$y_i.\left(w.x_i + b\right) \ge 1 - \xi_i,\ i = 1, 2, ..., n. \tag{2.9.}$$

Under these circumstances, the problem of optimization becomes

$$\Phi(w, \xi) = \frac{1}{2}(w.w) + C\sum_{i=1}^{n}\xi_i\ . \tag{2.10.}$$

In (2.10.) C is a user specified positive fixed constant. The saddle point of Lagrangian function is utilized in the solution of the problem given in (2.10).

$$L(w,b,\alpha,\xi,\gamma)=\frac{1}{2}(w.w)+C\sum_{i=1}^{n}\xi_i-\sum_{i=1}^{n}\alpha_i\left|y_i\left(w.x_i+b\right)-1+\xi_i\right|-\sum_{i=1}^{n}\gamma_i\xi_i \tag{2.11.}$$

In (2.11), $\alpha_i \geq 0,\ \xi_i \geq 0,\ i=1,2,...,n$ are Lagrange multipliers. (2.11) must be solved in terms of $\mathbf{w},\ b,$ and ξ_i. Classical Lagrangian duality empowers the first issue, turning (2.11) into a dual problem of it and this makes the solution easier. (2.12) shows the dual problem to be solved

$$\max_{\alpha}\left[\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j\left(x_i.x_j\right)\right] \tag{2.12.}$$

with constraints

$$\sum_{i=1}^{n}\alpha_i y_i=0,\ 0\leq\alpha_i\leq C,\ i=1,2,...,n. \tag{2.13.}$$

There is a unique solution for this classic quadratic optimization problem. In respect of the Kuhn-Tucker theorem of optimization theory (2.16.),

$$\alpha_i\left[y_i\left(w.x_i+b\right)-1\right]=0,\ i=1,2,...,n. \tag{2.14.}$$

if $\mathbf{x}_i$ satisfy (2.15), then (2.14) will have non-zero Lagrange multipliers

$$y_i\left(\mathbf{w}.\mathbf{x}_i+b\right)=1. \tag{2.15.}$$

The points in (2.15.) are called support vectors (SV's). Small subset of the training vectors of the SV's determines the hyperplane. Therefore, if the optimal solution α_i^* does not take a value of zero, the classifier function can be represented as

$$f\left(x\right)=\operatorname{sgn}\left\{\sum_{i=1}^{n}\alpha_i^* y_i\left(x_i.x\right)+b^*\right\} \tag{2.16.}$$

In (2.16.) b^* is the solution of (14) for any non-zero α_i^*.

2.1.2. Non-linear SVM

In general it is not possible to divide the datasets accurately using a linear separating hyperplane. However, they can be divided linearly if mapped into a higher dimensional field using a nonlinear mapping. Therefore, $z=\phi(x)$ that converts the input vector x that has a dimension d into a vector z that has a dimension d is defined and $\phi()$ is selected so that $\{\phi(x_i, y_i)\}$ (new training data) is divisible with a hyperplane.

The data points from the input space into some space of higher dimension are mapped by using the function

$$\varphi(.):R^n \to R^{nh}. \tag{2.17.}$$

Optimal function (2.9) transforms (2.18) using the same constraints,

$$\max_{\alpha}\left[\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K\left(\mathbf{x}_i,\mathbf{x}_j\right)\right] \tag{2.18.}$$

where

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\left\{\varphi\left(\mathbf{x}_i\right).\varphi\left(\mathbf{x}_j\right)\right\} \tag{2.19.}$$

is the kernel function.

The exponential kernel function is given by

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\exp\left\{-\frac{\left|\mathbf{x}_i-\mathbf{x}_j\right|^2}{2\sigma^2}\right\} \tag{2.20.}$$

whereas the polynomial kernel function is given by

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\left(\mathbf{x}_i.\mathbf{x}_j+1\right)^q,\ q=1,2,... \tag{2.21.}$$

where the parameters σ and q in (2.20) and (2.21) have to be prearranged. Then, the classifier function is specified as

$$f\left(\mathbf{x}\right)=\operatorname{sgn}\left\{\sum_{i=1}^{n}\alpha_i^* y_i K\left(\mathbf{x}_i,\mathbf{x}\right)+b^*\right\} \tag{2.22.}$$

and b^* is the solution of (2.22) for any non-zero α_i^*.

2.2. Multi-Layer Perceptron

An MLP is a feed-forward neural network with one or more layers between input and output layer. It is called Feedforward because data moves from input to output layer in one direction.

An MLP consists of a system of simple interconnected neurons, or nodes, which is a model representing a nonlinear mapping between an input vector and an output vector. In MLP the computation is carried out using a set of simple units with weighted connections between them. An MLP includes multiple layers of nodes in a coordinated graph and each layer is completely connected to the following one. Only the input neurons do not have an activation function. Each of all other nodes node is a neuron with a nonlinear activation. In order to train the network, a technique called back propagation which is also known as a handled learning technique is used. It has been shown that an MLP with a hidden layer can approximate nonlinear functions accurately (Goyal et al, 2016). The complexity of the MLP network can be changed by varying the number of layers and the number of units in each layer. An MLP can be represented as given in Figure 2.1.

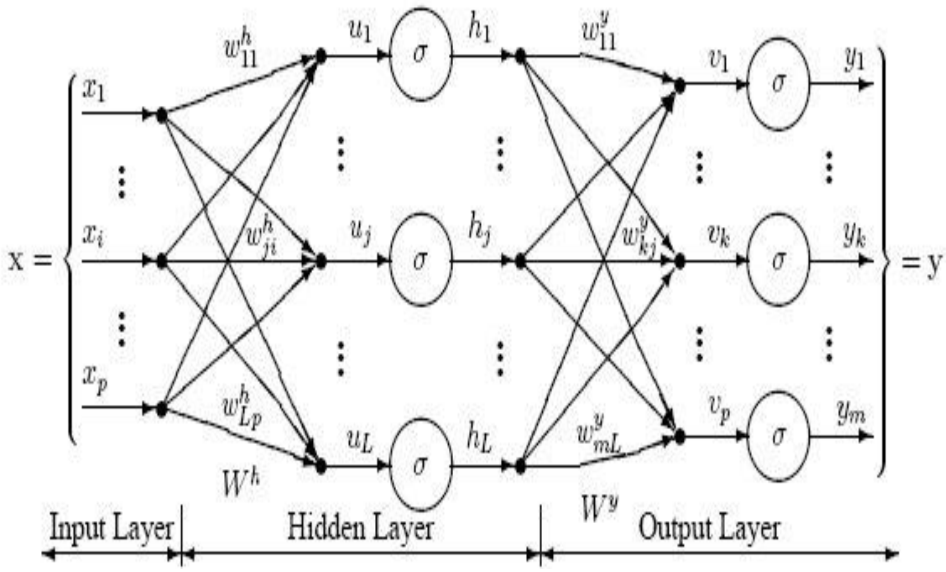

Figure 2.1. A typical MLP Structure

In the figure above, the network has an input layer (on the left, one hidden layer (in the middle) and an output layer (on the right) with three neurons for each layer respectively.

Input Layer: A vector of predictor variable values ($x_1...x_p$) arrives at the input layer. The input layer standardizes these values so that the range of each variable is -1 to 1. The input layer forwards the values to each of the neurons in the hidden layer.

Hidden Layer: Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight (w_{ji}), and the resulting weighted values are added together producing a combined value u_j.

$$U_j=\sum_{i=0}^{p} w_{ji}\, x_i \qquad (2.23.)$$

The weighted sum (u_j) is fed into a transfer function, σ, which outputs a value h_j. The outputs from the hidden layer are distributed to the output layer.
The transfer function may be a logistic function or a hyperbolic tangent.

$\sigma = (1 + e^{-x})^{-1}$ (logistic function)

$\sigma = = tangh(x)$ (hyperbolic tangent)

$$h_j= f(u_j)= (1 + e^{-uj})^{-1} \qquad (2.24.)$$

Output Layer: Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight (w_{kj}), and the resulting weighted values are added together producing a combined value v_j.

A. Non-Linear Activation/Transfer Function

The values of the logistic sigmoid function range from 0 to 1. The standard sigmoid is basically used with the hyperbolic tangent.

It has the feature,

$$f(x) = \tanh(x) = 2Sigmoid(2x) - 1 \;\; f(-x) = -f(x) \tag{2.25.}$$

and its derivative is given by

$$f'(x) = 1 - f(x)^2 . \tag{2.26.}$$

B. Learning

For training *N-Layer* neural networks as well as the networks having a single layer, the same steps are used. The network weights $w_{ij}^{(n)}$ are set to make the output cost function given in (2.27) minimum.

$$E_{SSE} = \frac{1}{2}\sum_p \sum_j (t\arg_j^p - out_j^{(N)p})^2 \tag{2.27.}$$

or

$$E_{CE} = -\sum_p \sum_j [t\arg_j^p .\log(out_j^{(n)p}) + (1 - t\arg_j^p).\log(1 - out_j^{(N)p})] \tag{2.28.}$$

and once again this can be done by a series of gradient descent weight changes

$$\Delta w_{kl}^{(m)} = -\eta \frac{\partial E(\{w_{ij}^{(n)}\})}{\partial w_{kl}^{(m)}} \quad . \tag{2.29.}$$

$out_j^{(N)}$ is the only output of the last layer. This becomes apparent in the error function E of the output. However, outputs of the final layer are related with all the weights of the previous layers. This learning algorithm automatically sets $out_j^{(n)}$ of the previous layers.

C. Training

The training for MLP is realized following the steps below:

1. Let's suppose that we have the series of training patterns given below in (2.30)

$$\{in_i^p, out_j^p : i = 1...ninputs, j = 1...noutputs, p = 1...npatterns\} \qquad (2.30.)$$

2. Network is set up with N inputs, $N-1$ hidden layers of $nhidden^{(n)}$ nonlinear units, and N outputs in layer N. Each layer is completely connected with the previous layer by weight $w_{ij}^{(n)}$.
3. First weights are generated randomly in the range $[-smwt, +smwt]$.
4. The learning rate η and a convenient error function $E(w_{jk}^{(n)})$ are selected.
5. For each training pattern p, the weight update equation $\Delta w_{jk}^{(n)} = -\eta \partial E(w_{jk}^{(n)}) / \partial w_{jk}^{(n)}$ is applied to each weight $w_{jk}^{(n)}$.
6. Step 5 is repeated until the network's error function gives negligible error rates.

2.3. Radial Basis Function Neural Network

The RBFNN is a feed-forward neural network model with good performance, global approximation, and is free from the local minima problems. It is a multi-input, single-output system consisting of an input layer, a hidden layer, and an output layer. During the data processing, the hidden layer performs nonlinear transforms for the feature extraction and the output layer gives a linear combination of output weights. The advantage of RBFNN is its fast learning algorithms (Korürek & Doğan, 2010). As mentioned above, RBFNN has the feed-forward structure. It has one layer that consists of hidden units. These units are completely adjusted with the output units. (2.31) indicates the output units (ψ_j) from a linear combination of the main functions (Ghosh-dastidar et al, 2008).

$$\psi_j(x) = K\left(\frac{\left\| x - c_j \right\|}{\sigma_j^2}\right) \quad (2.31.)$$

$f : R^n \to R^1$ (a function) is estimated by using an RBFNN. $x \in R^n$ (a vector) is an input, $\psi(x, c_j, \sigma_j)$ is the j[th] function with centre $c_j \in R^n$, width σ_j, and $w = (w_1, w_2, ..., w_m) \in R^M$ is the vector of linear output weights. M represents the main number function. The M centers $c_j \in R^n$ are connected to obtain $c = (c_1, c_2, ..., c_m) \in R^{nM}$. Lastly, the widths are connected to obtain $\sigma = (\sigma_1, \sigma_2, ..., \sigma_M) \in R^M$. The output of the network for $x \in R^n$ and $\sigma \in R^M$ is shown by (2.32).

$$F(x, c, \sigma, \omega) = \sum_{j=1}^{M} (\omega_j \psi(x, c_j, \sigma_j)) \, . \quad (2.32.)$$

Suppose that $y = (y_1, y_2, ..., y_n)$ is the weighted output vector and $(x_i, y_i) : i = 1, 2, ..., N$ is a series of training pairs. For each $c \in R^{nM}$,

$\omega \in R^M, \sigma \in R^M$ and for random weights $\lambda_i, i = 1, 2, \ldots, N$, which are taken as positive numbers to point out importance of definite domains of the input space, set (2.33) (Wright et al, 2013).

$$E(c,\sigma,\omega) = \frac{1}{2}\sum_{i=1}^{N}\left[\lambda_i(y_i - F(x_i, c, \sigma, \omega))\right]^2 \tag{2.33}$$

2.4. Single Decision Trees

A decision tree can be defined as a graphical representation of possible ways to make a decision that depends on certain conditions. It's called a decision tree because it is a tree-like graph that consists of nodes with a starting node called root. The root node does not have any incoming edges. All other nodes have one incoming edge. A node with outgoing edges is called an internal node. All other nodes which are also known as decision nodes are called leaves. In order to split the instance space into two or more sub-spaces, a discrete function of the input attributes values is used.

In most of the cases, the instance space is divided considering a single predictor variable and according to the value of the predictor variable. If the predictor variables are numeric, a range is utilized. The most relevant target variable represented by one class is linked to each leaf. An alternative way is that the leaf holds a probability vector showing the probability of the target variable having a certain value. Instances are classified according to the results of the tests along the path (from the root to a leaf). Figure 2.2 from (Rokach & Maimon, 2005) is an example of a decision tree, in which the response of a customer to direct mailing is predicted. In this example the decision tree includes nominal and numeric predictor variables. The response of a customer can be predicted using this classifier. It is also possible by the use of this classifier to understand the behavior of the entire population of customers regarding direct mailing. Each node is labeled with the attribute it tests, and its branches are labeled with its corresponding values. If predictor variables are numeric, decision trees can be seen as a collection of hyperplanes. Since decision trees with less complexity are easier to comprehend, they are preferred in most of cases. Furthermore, as noted in (Breiman et al, 1984) the accuracy of a decision tree is highly associated with its complexity. It is clear that the tree complexity is in turn controlled by the stopping criteria employed and the pruning method applied.

The tree complexity is generally determined by certain parameters which include the total number of nodes, total number of leaves, tree depth and number of attributes used. The induction of a decision tree is related to rule induction. A rule may be created by putting together the tests in different paths from the root of a decision tree to one of its leaves and using the leaf's estimation as the target value. In Figure 2.2 mentioned above, one of the paths can give the rule: If the age of a customer is below 30 and the customer is male – then the customer will respond to the mail". The comprehensibility of the resulting rule set to a human user can be improved by making the resulting rule set as simple as possible and if possible by improving also its accuracy (Quinlan, 1987).

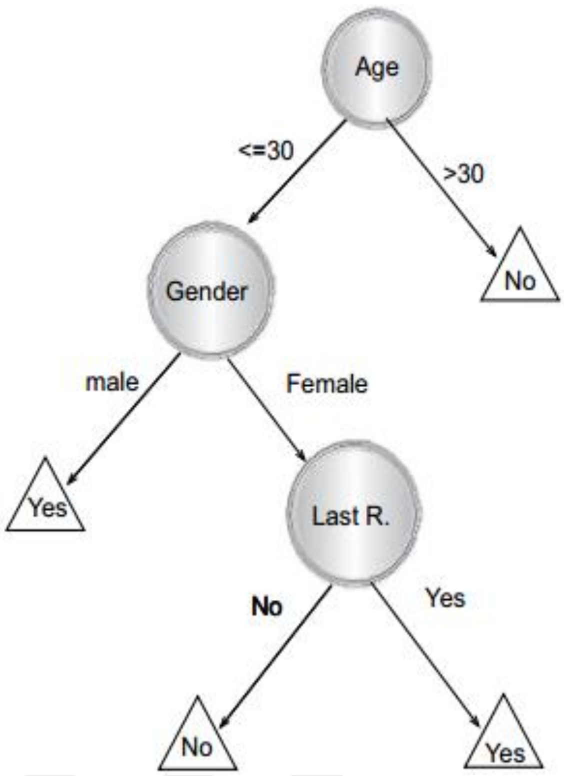

Figure 2.2. Decision tree example

3. DEVELOPMENT OF PREDICTION MODELS

3.1. Support Vector Machine Prediction Models

In order to have an effective SVM in terms of accuracy, certain principal parameters such as the cost C, the ε-insensitive loss function and gamma must have their optimal values. The parameter C affects the tradeoff between proportion and complexity of nonseparable samples. In other words, the parameter C controls the tradeoff between errors of the SVM on training data and margin maximization. A low value of C tends to make decision surface smooth (increment of the training error quantity). On the other hand if the value of C is large, a high misclassification cost is observed. For measuring the quality of estimation and quantifying the empirical risk, the loss function of Vapnik (Vapnik, 2000) called also ε-insensitive is generally the most utilized function. In order to control the width of the ε-insensitive zone used to fit the training data, the ε parameter is utilized. Since the number of selected support vectors is affected by the value of ε, therefore a large value of ε implies that the number of selected support vectors is small.

The type of the kernel including linear, polynomial, radial basis function (RBF) and sigmoid has also an important impact on the accuracy of SVM. However in this thesis RBF that has been found to do best job in the majority of cases has been utilized. The regularization parameter gamma is necessary for the optimization of the RBF kernel function. The gamma parameter is defined as the inverse of the standard deviation of the RBF kernel function, which is used as similarity measure between two points. If gamma is small, the variance of the kernel function is large. Consequently, even if the distance between two points is large, they can be seen as similar. On the other hand, if gamma is too large, the variance is small, thus two points are similar only when they are close to each other.

In order to minimize the error, it is necessary to have the optimal values of the above mentioned parameters. Several techniques such as swarm optimization and grid search are employed to find these optimal values. However the grid search technique has been used in this thesis. The grid search has been proved to be very efficient for the optimization of values of C, ε, and gamma on medium dataset. In grid search the values of the parameters are changed within a specific interval and the three parameters showing the lowest error rate are retained. Figure 3.1 shows the flow chart of the SVM-based prediction models. As a first step, the dataset is standardized in order to have zero mean and unity variance of the predictor variables. Applying the standardization process on the dataset gives new test and train sets. For the validation of models and the enhancement of the reliability of the results, 10-fold cross validation has been applied on the dataset. Therefore, the training data comprised 63 elements and the test data had 7 elements for each fold. Table 3.1 shows the list of the intervals for values of the parameters for SVM models.

Table 3.1. List of the intervals for values of the parameters for the SVM model

Parameter	**Value**
Cost (*C*)	[2^{-6} - 2^{16}]
Epsilon(**ε**)	[0.01-150]
Gamma (*γ*)	[2^{-10}- 2^{8}]
Kernel Function	RBF

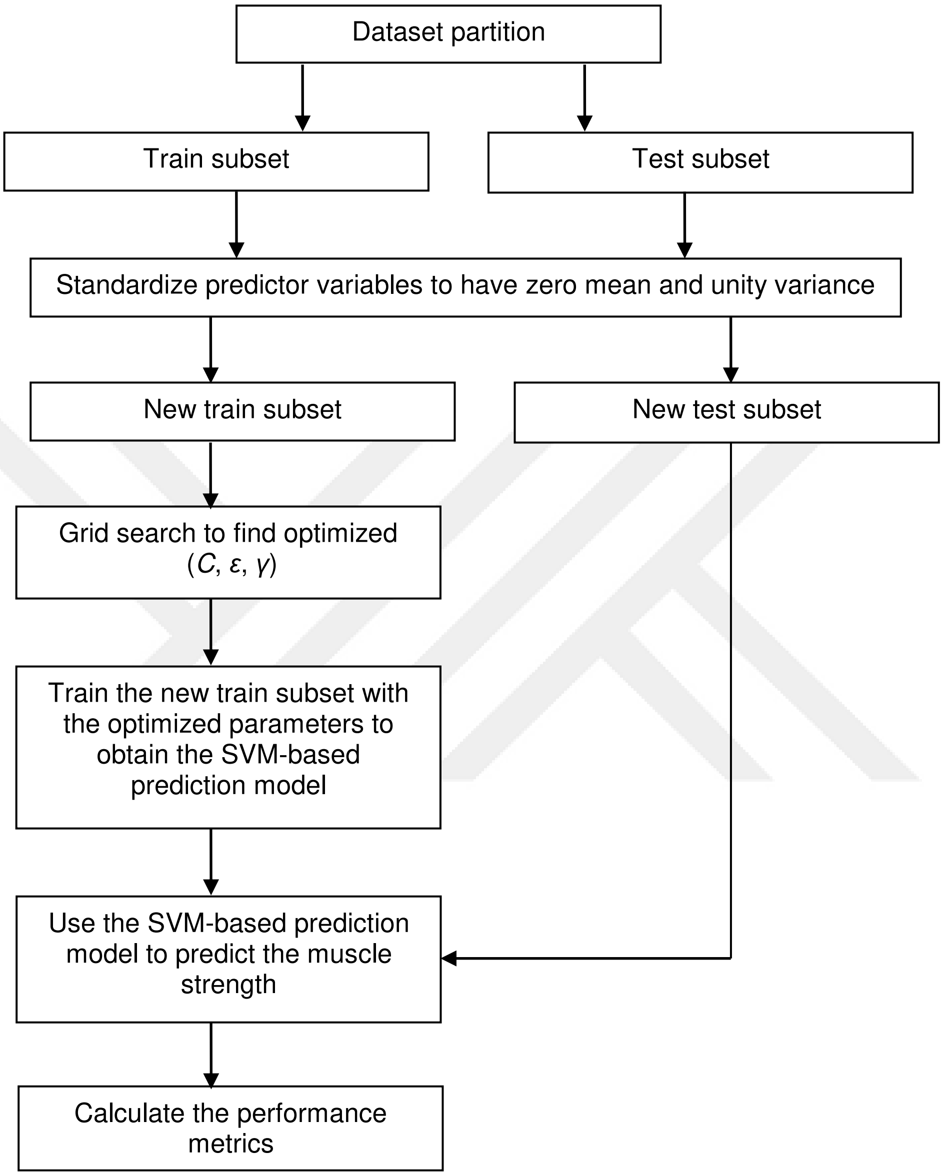

Figure 3.1. Flow chart of SVM-based models

3.2. Multilayer Perceptron Prediction Models

The MLP based hamstring and quadriceps prediction models have been created using the so-called back-propagation. The accuracy of an MLP based prediction model is improved by varying the number of layers and the number of units in each layer. There is a close relationship between the number of neurons in the hidden layer and the performance of an MLP based prediction model. Using a small number of neuron may cause a loss of information, whereas using a large number of neurons increases the complexity of the model. However, the number of neurons in a hidden layer cannot be determined in advance. Generally, the optimal number of neurons is obtained with try and error process depending on the complexity of the problem. The tansigmoid function and the pure-linear function have been utilized in the hidden layer and output layer respectively. The flow chart of SVM-based models is given in Figure 3.2. Table 3.2 shows the list of the values of the parameters for MLP models

Table 3.2. List of intervals for the values of the parameters for MLP models

Parameter	**Value**
Number of neurons in the hidden layer	[2-20]
Learning rate	(0-1)
Momentum	(0-1)
Maximum iteration	10000
Number of epochs	500
Number of convergence tries	[2-20]

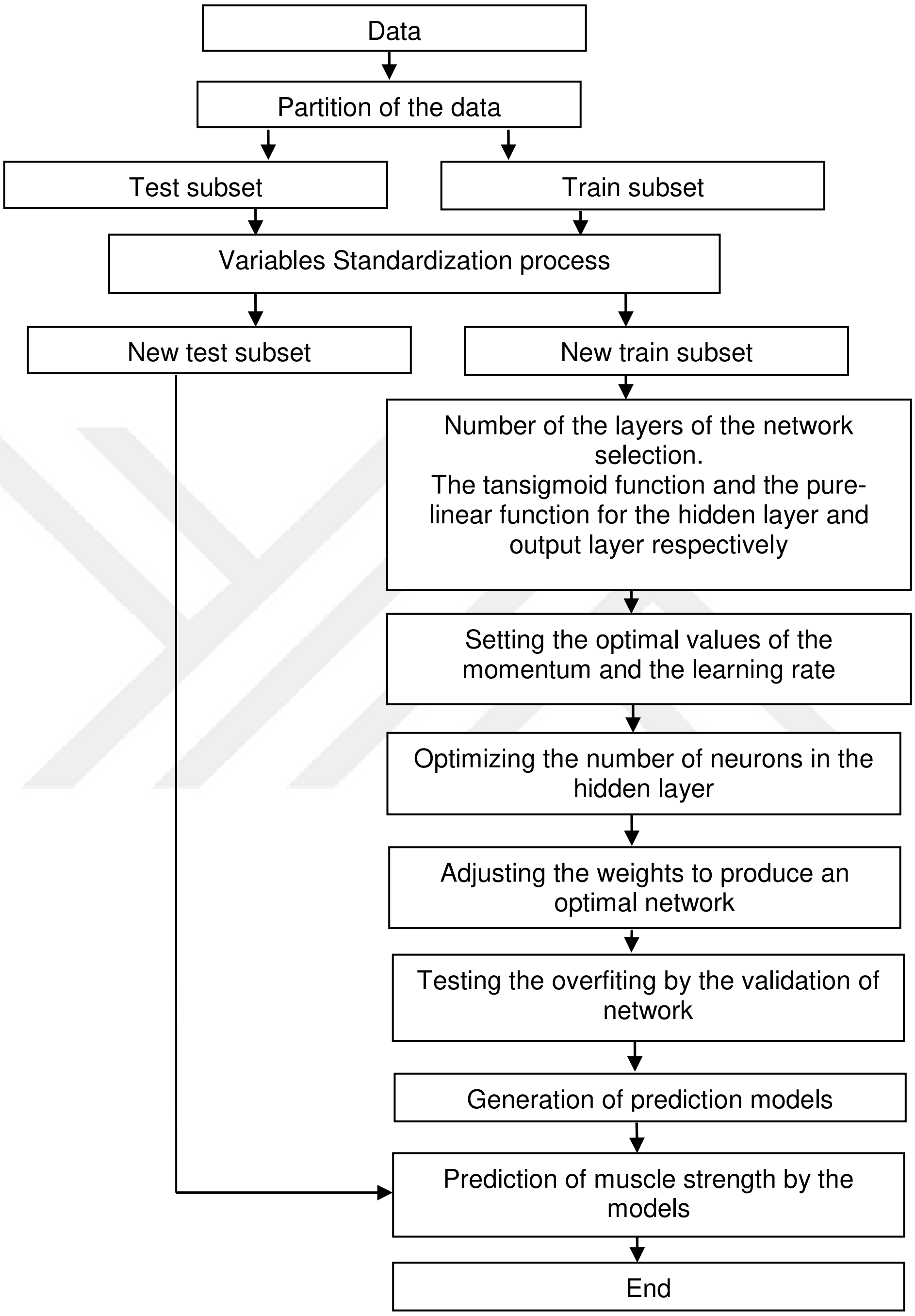

Figure 3.2. Flow chart of MLP based prediction models

3.3. RBFNN Prediction Models

An RBFNN consists of an input layer, a hidden layer and an output layer. In the hidden layer, the neurons utilize the Gaussian transfer function whose outputs are inversely proportional to the distance from the center of the neuron. In order to create an efficient RBFNN-based prediction model, several steps have to be followed. Reading the content of the dataset is the first step, afterwards the variables are standardized. The addition of a new neuron to the network and the adjustment of weight for the output layer is the following step. Then calculation of the output error of the network is performed. If the error is too large, the addition of a second neuron and the calculation process of the error are reproduced. The performance of the RBFNN for the training and test data is determined when an acceptable error occurs. The processes of addition a neuron and the error evaluation are performed again if a low performance of the network is observed. The main parameters that are related to the performance of an RBFNN-based prediction model include the population size, radius of the RBFNN, the maximum number of neurons and the regularization parameter (λ). Table 3.3 sows the list of the values of the parameters for RBFNN prediction models.

Table 3.3. List of the values of the parameters for RBFNN prediction models

Parameter	**Value**
Regularization parameter (λ)	[0.001 - 25]
Population size	[100 - 400]
Radius of RBFNN	[0.001- 450]
Maximum nomber of neurons	[95-100]

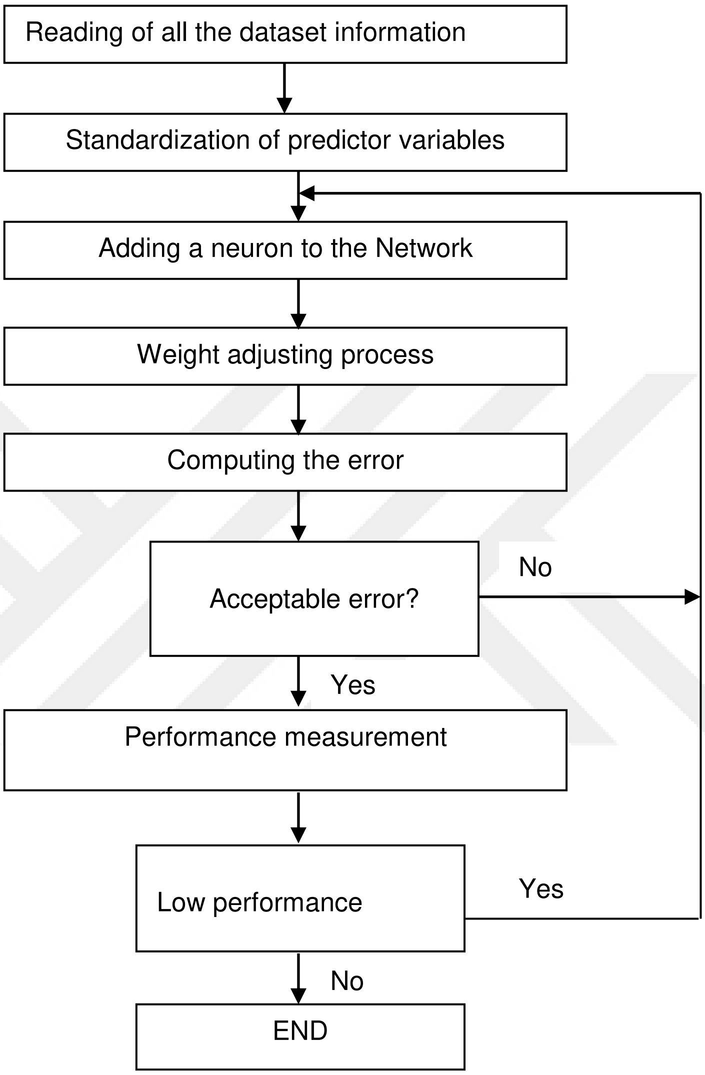

Figure 3.3. Flow chart of RBFNN based prediction models

3.4. Single Decision Tree Prediction Models

The performance of SDT-based prediction models is related to three main parameters, namely the minimum rows in a node, the minimum size node to split and the maximum tree levels. The minimum rows in a node specify the minimum number of rows that may fall in a node after splitting. The minimum size node to split indicates that a node or a group of nodes should never be split if it contains fewer rows than the specified value. The maximum tree levels specify the maximum number of levels in the tree that are created when the tree is being built. In order to have an efficient SDT-based prediction models, the above mentioned parameters must have their optimal values. Table 3.4 shows the list off intervals values of the parameters for SDT prediction models.

Table 3.4. List of intervals values of the parameters for SDT-based prediction

Parameter	**Value**
Minimum size node to split	[2 - 15]
Minimum rows in a node	[2 - 25]
Maximum tree levels	[5- 10]

4. RESULTS AND DISCUSSIONS

4.1 Approach of this Thesis

Using the combinations of predictor variables height, weight, gender, age, sport branch and BMI, two categories of hamstring and quadriceps muscle strength models have been created. The first category includes predictor variables height, weight, gender, age and SB. In the second category, predictor variable BMI has been used instead of predictor variables weight and height. Thus the second category includes predictor variables BMI, gender, age and sport branch. In order to evaluate the performance and the accuracy of the different models, the values of *RMSE* and *R* have been computed using 10-fold cross validation method. The equations of *RMSE* and *R* are given in (3.1) and (3.2) respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y - Y')^2} \tag{3.1}$$

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(Y-Y')^2}{\sum_{i=1}^{n}(Y-\bar{Y})^2}} \tag{3.2}$$

In (3.1) and (3.2), Y is the measured value, Y' is the predicted value, $\bar{Y}$ is the mean of the measured values and n is the number of samples in a test subset.

Table 4.1 through Table 4.5 show the prediction models with one, two, three, four and five predictor variables respectively for the hamstring and quadriceps muscle strength prediction.

Table 4.1. Overview of prediction models with one predictor variable

Model Name	**Abbreviation**	**Predictor variables**
Model 1	M1	Gender
Model 2	M2	Weight
Model 3	M3	Height
Model 4	M4	Sport Branch
Model 5	M5	Age
Model 32	M32	BMI

Table 4.2. Overview of prediction models with two predictor variables

Model Name	**Abbreviation**	**Predictor variables**
Model 6	M6	Gender, Weight
Model 7	M7	Gender, Height
Model 8	M8	Gender, Sport Branch
Model 9	M9	Gender, Age
Model 10	M10	Weight, Height
Model 11	M11	Weight, Sport Branch
Model 12	M12	Weight, Age
Model 13	M13	Height, Sport Branch
Model 14	M14	Height, Age
Model 15	M15	Sport Branch, Age
Model 33	M33	Gender, BMI
Model 34	M34	Sport Branch, BMI
Model 35	M35	Age, BMI

Table 4.3. Overview of prediction models with three predictor variables

Model Name	**Abbreviation**	**Predictor variables**
Model 16	M16	Gender, Weight, Height
Model 17	M17	Gender, Weight, Age
Model 18	M18	Gender, Weight, Sport Branch
Model 19	M19	Gender, Height, Age
Model 20	M20	Gender, Height, Sport Branch
Model 21	M21	Gender, Sport Branch, Age
Model 22	M22	Weight, Height, Age
Model 23	M23	Weight, Height, Sport Branch
Model 24	M24	Weight, Sport Branch, Age
Model 25	M25	Sport Branch, Height, Age
Model 36	M36	Gender, Sport Branch, BMI
Model 37	M37	Gender, Age, BMI
Model 38	M38	Sport Branch, Age, BMI

Table 4.4. Overview of prediction models with four predictor variables

Model Name	**Abbreviation**	**Predictor variables**
Model 26	M26	Gender, Height, Weight, Age
Model 27	M27	Gender, Sport Branch, Height, Weight
Model 28	M28	Gender, Sport Branch, Weight, Age
Model 29	M29	Gender, Sport Branch, Height, Age
Model 30	M30	Sport Branch, Height, Weight, Age
Model 39	M39	Gender, Sport Branch, Age, BMI

Table 4.5. Overview of prediction models with five predictor variables

Model Name	**Abbreviation**	**Predictor variables**
Model 31	M31	Gender, Height, Weight, Sport Branch, Age

4.2. Results

Table 4.6 through Table 4.25 show the test results for models without the predictor variable BMI, whilst Table 4.26 through Table 4.41 show the test results for prediction models with the predictor variable BMI.

Table 4.6. Test results for hamstring-CT prediction models with one predictor variable

Models	**SVM**		**MLP**		**RBFNN**		**SDT**	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 1	19.04	0.72	21.58	0.70	23.29	0.64	24.51	0.53
Model 2	19.63	0.70	22.15	0.65	23.53	0.59	24.80	0.55
Model 3	25.91	0.48	28.54	0.45	30.19	0.32	31.79	0.29
Model 4	31.15	0.24	33.99	0.10	35.78	0.08	36.95	0
Model 5	35.87	0	37.74	0	39.46	0	40.81	0

Table 4.7. Test results for hamstring-CT prediction models with two predictor variables

Models	**SVM**		**MLP**		**RBFNN**		**SDT**	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 6	15.31	0.82	16.29	0.79	17.37	0.77	18.88	0.72
Model 7	17.26	0.77	18.49	0.73	20.31	0.68	23.89	0.56
Model 8	19.05	0.72	20.81	0.66	23.22	0.64	24.75	0.52
Model 9	18.69	0.73	19.79	0.70	22.03	0.62	24.95	0.52
Model 11	18.70	0.73	19.78	0.70	20.98	0.66	22.42	0.61
Model 10	19.32	0.71	20.94	0.66	22.08	0.62	25.40	0.50
Model 12	20.05	0.69	21.34	0.65	22.53	0.60	23.70	0.56
Model 13	22.93	0.59	24.84	0.52	26.20	0.47	28.53	0.37
Model 14	25.68	0.49	27.44	0.41	28.97	0.35	30.95	0.25
Model 15	30.62	0.27	32.64	0.17	35.11	0.04	37.14	0

Table 4.8. Test results for hamstring-CT prediction models with three predictor variables

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 17	15.48	0.81	16.81	0.78	18.11	0.74	20.15	0.68
Model 16	15.36	0.82	16.53	0.79	17.57	0.76	18.93	0.72
Model 18	15.68	0.81	17.39	0.76	19.23	0.71	22.41	0.61
Model 20	17.33	0.77	19.03	0.72	23.93	0.55	25.70	0.49
Model 19	16.78	0.78	17.97	0.75	19.02	0.72	20.03	0.69
Model 21	18.99	0.72	20.47	0.67	22.43	0.61	23.93	0.55
Model 23	17.89	0.75	19.50	0.70	21.84	0.63	24.07	0.55
Model 24	19.39	0.71	21.52	0.64	24.31	0.54	25.44	0.50
Model 22	19.78	0.70	20.90	0.66	23.88	0.56	25.57	0.49
Model 25	22.88	0.59	24.26	0.54	25.53	0.49	27.22	0.42

Table 4.9. Test results for hamstring-CT prediction models with four predictor variables

Models	SVM		MLP		RBFNN		SDT	
	RMSE	**R**	***RMSE***	**R**	***RMSE***	**R**	***RMSE***	**R**
Model 26	15.19	0.82	16.29	0.79	17.86	0.75	18.93	0.72
Model 27	15.84	0.80	17.09	0.77	18.32	0.74	19.59	0.70
Model 28	15.89	0.80	17.66	0.76	19.12	0.72	20.62	0.67
Model 29	17.62	0.76	18.52	0.73	20.67	0.67	23.93	0.55
Model 30	17.97	0.75	19.83	0.69	22.48	0.61	24.96	0.52

Table 4.10. Test results for hamstring-CT prediction models with five predictor variables

Models	**SVM**		**MLP**		**RBFNN**		**SDT**	
	RMSE	**R**	***RMSE***	**R**	***RMSE***	**R**	***RMSE***	**R**
Model 31	15.55	0.81	16.46	0.79	17.60	0.76	19.05	0.72

Table 4.11. Test results for hamstring-ST prediction models with one predictor variable

Models	**SVM**		**MLP**		**RBFNN**		**SDT**	
	RMSE	**R**	***RMSE***	**R**	***RMSE***	**R**	***RMSE***	**R**
Model 1	18.99	0.72	22.71	0.68	24.07	0.66	25.44	0.67
Model 2	19.55	0.71	20.67	0.67	22.12	0.63	23.61	0.57
Model 3	25.68	0.50	27.81	0.52	28.67	0.37	30.26	0.30
Model 4	29.91	0.32	32.76	0.18	34.15	0.26	35.21	0.37
Model 5	36.32	0	38.56	0	40.39	0	41.46	0

Table 4.12. Test results for hamstring-ST prediction models with two predictor variables

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 6	15.59	0.81	16.52	0.79	17.49	0.77	19.73	0.70
Model 7	16.77	0.79	17.57	0.76	18.20	0.75	19.28	0.69
Model 8	17.39	0.77	19.63	0.71	20.78	0.73	22.71	0.61
Model 9	18.44	0.74	19.80	0.70	21.09	0.66	23.89	0.56
Model 11	20.01	0.69	21.42	0.65	24.39	0.55	25.35	0.51
Model 10	19.37	0.71	20.37	0.68	21.90	0.63	23.61	0.57
Model 12	20.08	0.69	21.08	0.66	22.44	0.62	23.59	0.57
Model 13	23.41	0.58	24.45	0.54	25.91	0.49	27.36	0.43
Model 14	25.71	0.49	27.23	0.43	29.11	0.35	30.76	0.28
Model 15	31.03	0.26	32.58	0.19	33.72	0.13	34.50	0.09

Table 4.13. Test results for hamstring-ST prediction models with three predictor variables

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 17	15.57	0.81	16.93	0.78	19.85	0.70	22.34	0.62
Model 16	15.44	0.82	16.48	0.79	18.38	0.74	20.53	0.68
Model 18	15.64	0.81	18.01	0.75	20.35	0.68	23.43	0.58
Model 20	16.15	0.80	18.54	0.74	19.61	0.71	21.46	0.65
Model 19	16.80	0.78	18.70	0.73	20.20	0.69	21.21	0.66
Model 21	17.88	0.76	19.51	0.71	21.34	0.65	22.71	0.61
Model 23	17.34	0.77	20.09	0.69	22.84	0.60	25.10	0.52
Model 24	20.53	0.68	22.17	0.62	24.57	0.54	25.35	0.51
Model 22	19.64	0.71	20.87	0.67	22.67	0.61	24.54	0.54
Model 25	22.36	0.62	24.85	0.53	26.54	0.46	27.67	0.42

Table 4.14. Test results for hamstring-ST prediction models with four predictor variables

Models		**SVM**		**MLP**		**RBFNN**		**SDT**	
		RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 26		15.60	0.81	17.08	0.78	18.60	0.74	19.83	0.70
Model 27		15.60	0.81	17.39	0.77	19.15	0.72	21.60	0.64
Model 28		15.81	0.81	18.18	0.75	20.39	0.68	22.34	0.62
Model 29		16.22	0.80	18.46	0.74	20.92	0.67	22.74	0.61
Model 30		17.67	0.76	20.74	0.67	23.85	0.57	25.10	0.52

Table 4.15. Test results for hamstring-ST prediction models with five predictor variables

Models	**SVM**		**MLP**		**RBFNN**		**SDT**	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 31	15.75	0.81	16.89	0.78	17.54	0.76	19.31	0.72

Table 4.16. Test results for quadriceps-CT prediction models with one predictor variable

Models	**SVM**		**MLP**		**SDT**		**RBFNN**	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 1	28.04	0.73	29.58	0.70	32.11	0.65	33.35	0.63
Model 2	30.72	0.68	32.52	0.67	34.38	0.60	36.00	0.56
Model 3	39.13	0.48	41.30	0.46	43.50	0.36	45.35	0.34
Model 4	47.78	0.23	50.14	0.15	52.13	0.11	55.21	0
Model 5	54.44	0	56.74	0	58.97	0	61.36	0

Table 4.17. Test results for quadriceps-CT prediction models with two predictor variables

Models		SVM		MLP		SDT		RBFNN	
		RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 6		21.99	0.84	23.92	0.81	24.94	0.79	26.47	0.76
Model 7		25.31	0.78	28.83	0.72	30.22	0.69	32.67	0.64
Model 8		27.74	0.74	29.83	0.70	31.71	0.66	34.75	0.61
Model 9		28.23	0.73	29.66	0.70	32.15	0.65	33.57	0.62
Model 11		29.40	0.71	30.79	0.68	33.52	0.62	34.71	0.59
Model 10		30.18	0.69	32.28	0.65	33.98	0.61	34.48	0.60
Model 12		31.79	0.66	33.43	0.63	34.97	0.59	36.61	0.55
Model 13		33.05	0.63	35.76	0.57	37.65	0.52	38.75	0.49
Model 14		40.01	0.46	42.64	0.39	44.09	0.34	45.11	0.31
Model 15		47.17	0.25	49.35	0.18	51.31	0.17	52.74	0.06

Table 4.18. Test results for quadriceps-CT prediction models with three predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 17	22.06	0.84	23.08	0.82	24.34	0.80	26.02	0.77
Model 16	21.19	0.85	22.35	0.83	24.96	0.79	26.40	0.76
Model 18	23.98	0.81	25.08	0.79	26.90	0.76	27.99	0.74
Model 20	25.58	0.78	27.20	0.75	28.57	0.75	31.24	0.67
Model 19	24.84	0.79	26.60	0.76	28.48	0.73	30.00	0.70
Model 21	27.70	0.74	29.25	0.71	31.71	0.66	32.99	0.63
Model 23	27.91	0.74	29.66	0.70	30.81	0.61	32.05	0.65
Model 24	30.19	0.69	31.61	0.66	34.69	0.59	37.68	0.52
Model 22	31.44	0.67	33.61	0.62	35.81	0.57	37.92	0.51
Model 25	33.28	0.63	35.87	0.57	37.05	0.54	38.36	0.50

Table 4.19. Test results for quadriceps-CT prediction models with four predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 26	21.48	0.84	23.01	0.82	24.96	0.79	26.19	0.77
Model 27	23.31	0.82	25.43	0.78	26.81	0.76	28.11	0.73
Model 28	24.28	0.80	25.47	0.77	27.56	0.74	28.59	0.72
Model 29	25.42	0.78	26.92	0.76	28.03	0.73	31.32	0.67
Model 30	28.40	0.73	30.48	0.69	34.07	0.61	37.90	0.51

Table 4.20. Test results for quadriceps-CT prediction models with five predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 31	24.17	0.80	25.53	0.78	26.81	0.78	28.75	0.72

Table 4.21. Test results for quadriceps-ST prediction models with one predictor variable

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 1	27.12	0.75	28.66	0.72	30.70	0.69	33.63	0.65
Model 2	27.99	0.73	29.31	0.71	31.32	0.66	33.88	0.63
Model 3	38.12	0.50	39.95	0.45	42.64	0.38	43.82	0.37
Model 4	43.94	0.38	46.75	0.30	49.60	0.16	51.70	0.08
Model 5	55.22	0	57.91	0	59.13	0	61.52	0

Table 4.22. Test results for quadriceps-ST prediction models with two predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 6	19.88	0.86	21.43	0.84	22.67	0.82	24.62	0.79
Model 7	23.71	0.81	25.18	0.78	26.34	0.76	27.86	0.73
Model 8	25.63	0.78	28.21	0.73	29.95	0.69	31.39	0.66
Model 9	27.12	0.75	28.60	0.74	30.08	0.69	31.61	0.66
Model 11	27.86	0.73	29.41	0.70	31.16	0.67	33.26	0.62
Model 10	27.54	0.74	29.60	0.70	31.42	0.66	33.27	0.62
Model 12	28.15	0.73	29.44	0.70	30.54	0.68	31.80	0.65
Model 13	34.66	0.59	36.99	0.53	39.46	0.47	40.97	0.42
Model 14	39.64	0.46	42.54	0.38	44.51	0.32	46.42	0.26
Model 15	49.91	0.15	52.14	0.07	53.44	0.03	54.67	0

Table 4.23. Test results for quadriceps-ST prediction models with three predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 17	20.62	0.85	22.69	0.82	24.75	0.79	27.10	0.75
Model 16	19.72	0.87	21.16	0.85	22.26	0.83	23.97	0.80
Model 18	22.90	0.82	24.11	0.80	25.97	0.77	27.93	0.73
Model 20	24.44	0.80	25.48	0.78	27.27	0.75	28.32	0.73
Model 19	24.14	0.80	26.82	0.75	28.78	0.72	30.92	0.67
Model 21	25.76	0.77	28.31	0.73	29.95	0.69	32.41	0.64
Model 23	26.35	0.76	28.66	0.72	30.76	0.68	31.96	0.65
Model 24	28.74	0.72	30.13	0.69	31.16	0.67	33.09	0.62
Model 22	28.53	0.72	29.96	0.69	31.54	0.66	33.17	0.62
Model 25	35.49	0.57	37.86	0.51	39.27	0.47	42.55	0.38

Table 4.24. Test results for quadriceps-ST prediction models with four predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 26	20.00	0.86	21.27	0.85	22.26	0.83	24.66	0.79
Model 27	22.35	0.83	23.64	0.81	24.79	0.79	26.08	0.77
Model 28	22.93	0.82	24.96	0.79	25.97	0.77	26.98	0.75
Model 29	24.82	0.79	26.05	0.77	27.27	0.75	28.87	0.71
Model 30	26.52	0.76	28.33	0.73	30.88	0.67	33.60	0.61

Table 4.25. Test results for quadriceps-ST prediction models with five predictor variables

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 31	23.76	0.81	24.86	0.79	25.97	0.79	27.16	0.75

Table 4.26. Test results for hamstring-CT prediction models with one predictor variable (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 32	25.33	0.50	26.76	0.49	28.09	0.39	29.24	0.33

Table 4.27. Test results for hamstring-CT prediction models with two predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 33	17.69	0.76	18.71	0.73	20.60	0.67	22.17	0.62
Model 34	24.84	0.52	26.34	0.46	28.43	0.37	30.71	0.27
Model 35	25.63	0.49	27.04	0.43	29.45	0.32	31.89	0.21

Table 4.28. Test results for hamstring-CT prediction models with three predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 36	17.89	0.75	19.89	0.69	21.62	0.64	23.93	0.55
Model 37	17.80	0.75	19.11	0.72	20.80	0.66	22.17	0.62
Model 38	26.00	0.47	27.24	0.42	28.32	0.38	30.67	0.27

Table 4.29. Test results for hamstring-CT prediction models with four predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 39	17.63	0.76	19.60	0.70	21.41	0.64	23.93	0.55

Table 4.30. Test results for hamstring-ST prediction models with one predictor variable (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 32	26.69	0.46	27.72	0.41	29.04	0.36	30.36	0.30

Table 4.31. Test results for hamstring-ST prediction models with two predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 33	18.40	0.75	19.95	0.70	21.53	0.65	23.56	0.58
Model 34	25.02	0.52	26.61	0.46	27.64	0.42	30.20	0.30
Model 35	25.95	0.49	27.14	0.44	28.17	0.39	29.28	0.35

Table 4.32. Test results for hamstring-ST prediction models with three predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 36	17.47	0.77	19.18	0.72	20.98	0.66	22.71	0.61
Model 37	18.51	0.74	20.05	0.69	21.71	0.64	23.56	0.58
Model 38	25.71	0.50	27.66	0.42	29.89	0.32	31.72	0.23

Table 4.33. Test results for hamstring-ST prediction models with four predictor variables (BMI case)

Models	SVM		MLP		RBFNN		SDT	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 39	17.35	0.77	18.63	0.73	20.68	0.67	22.71	0.61

Table 4.34. Test results for quadriceps-CT prediction models with one predictor variable (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 32	39.10	0.48	41.42	0.42	43.00	0.38	44.38	0.33

Table 4.35. Test results for quadriceps-CT prediction models with two predictor variables (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 33	25.89	0.77	27.68	0.74	28.78	0.72	30.27	0.69
Model 34	39.69	0.47	41.91	0.41	43.55	0.36	45.37	0.30
Model 35	39.66	0.47	42.41	0.39	45.25	0.31	46.79	0.26

Table 4.36. Test results for quadriceps-CT prediction models with three predictor variables (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 36	27.19	0.75	28.78	0.72	31.71	0.66	33.62	0.62
Model 37	25.62	0.78	26.92	0.76	28.48	0.73	31.34	0.67
Model 38	39.53	0.47	43.16	0.37	47.56	0.24	49.18	0.18

Table 4.37. Test results for quadriceps-CT prediction models with four predictor variables (BMI case)

Models		SVM		MLP		SDT		RBFNN	
		RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 39		27.20	0.75	29.31	0.71	32.46	0.64	33.75	0.62

Table 4.38. Test results for quadriceps-ST prediction models with one predictor variable (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 32	37.59	0.52	39.45	0.47	42.73	0.37	44.64	0.32

Table 4.39. Test results for quadriceps-ST prediction models with two predictor variables (BMI case)

Models		SVM		MLP		SDT		RBFNN	
		RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 33		25.96	0.77	27.32	0.74	28.77	0.72	31.31	0.66
Model 34		37.08	0.53	40.32	0.44	42.97	0.37	45.72	0.28
Model 35		37.52	0.52	39.54	0.46	42.85	0.37	45.28	0.30

Table 4.40. Test results for quadriceps-ST prediction models with three predictor variables (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 36	25.47	0.78	27.90	0.73	29.95	0.69	31.76	0.65
Model 37	25.48	0.78	28.14	0.73	30.34	0.68	32.58	0.64
Model 38	38.44	0.49	41.03	0.42	43.70	0.35	44.82	0.31

Table 4.41. Test results for quadriceps-ST prediction models with four predictor variables (BMI case)

Models	SVM		MLP		SDT		RBFNN	
	RMSE	***R***	***RMSE***	***R***	***RMSE***	***R***	***RMSE***	***R***
Model 39	25.38	0.78	28.36	0.72	29.95	0.69	31.28	0.66

Figure 4.1 through Figure 4.20 show the decrement rates in percentage of hamstring and quadriceps muscle strength prediction obtained by comparing the *RMSE*'s of SVM to *RMSE*'s of MLP, RBFNN and SDT for both classic and static training without the predictor variable BMI.

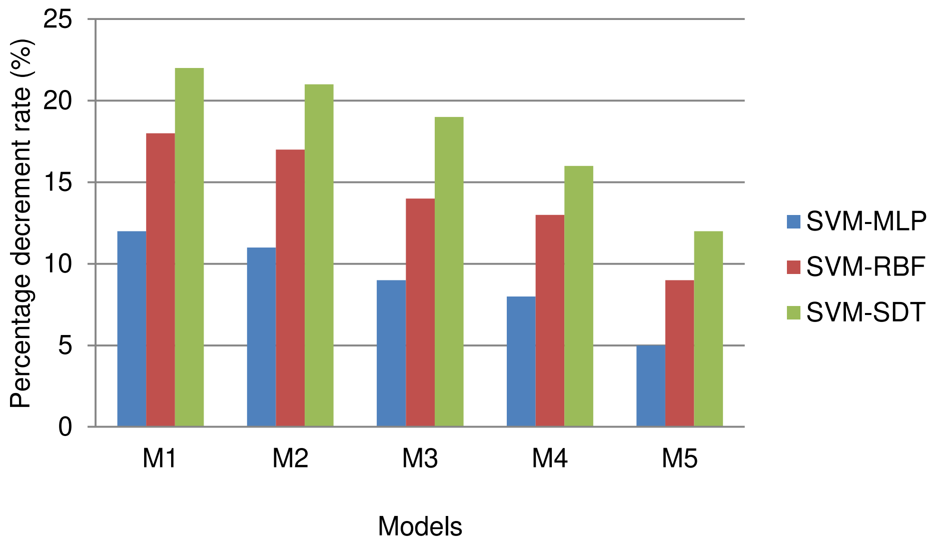

Figure 4.1. Decrement rates in *RMSE*'s of hamstring-CT with one predictor variable

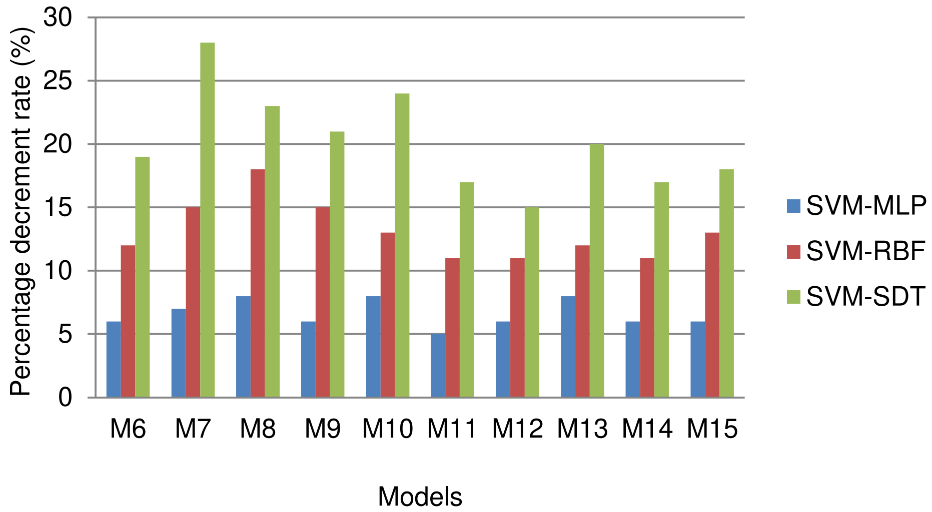

Figure 4.2. Decrement rates in *RMSE*'s of hamstring-CT with two predictor variables.

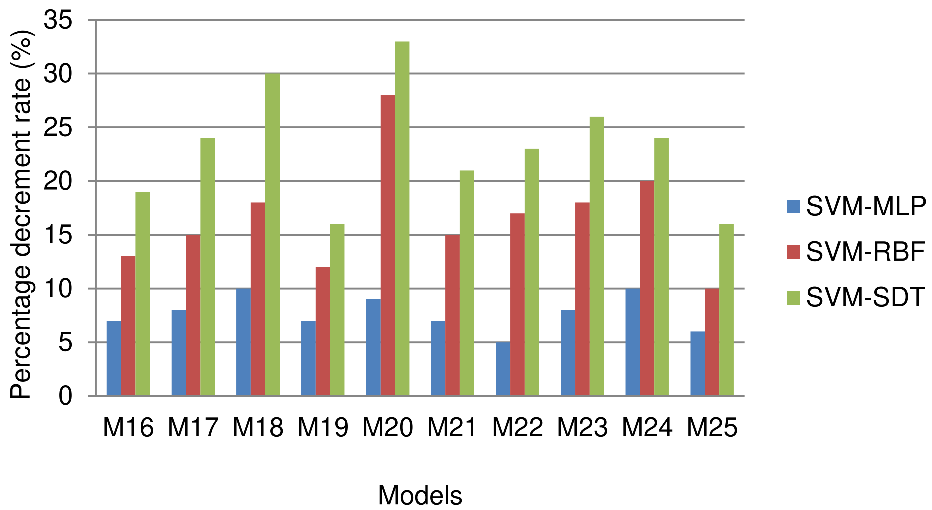

Figure 4.3. Decrement rates in *RMSE*'s of hamstring-CT with three predictor variables.

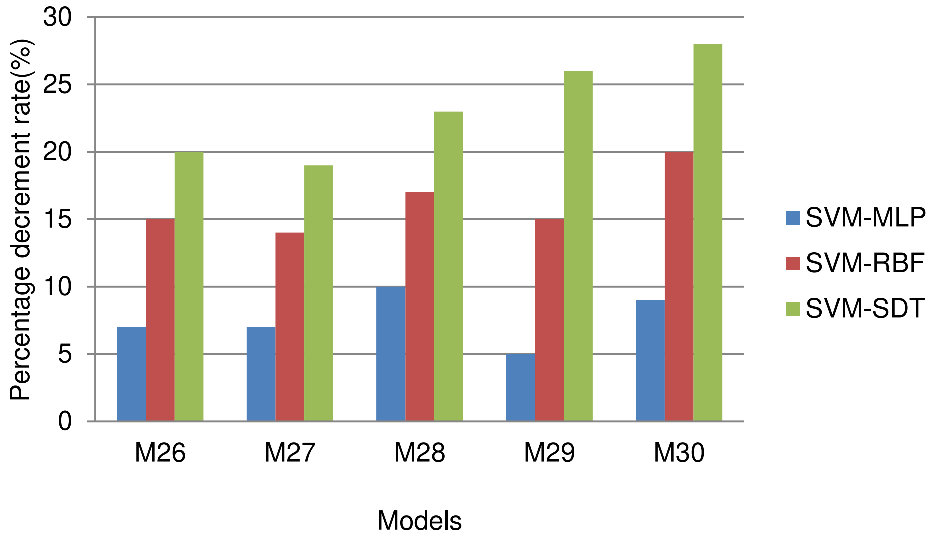

Figure 4.4. Decrement rates in *RMSE*'s of hamstring-CT with four predictor variables.

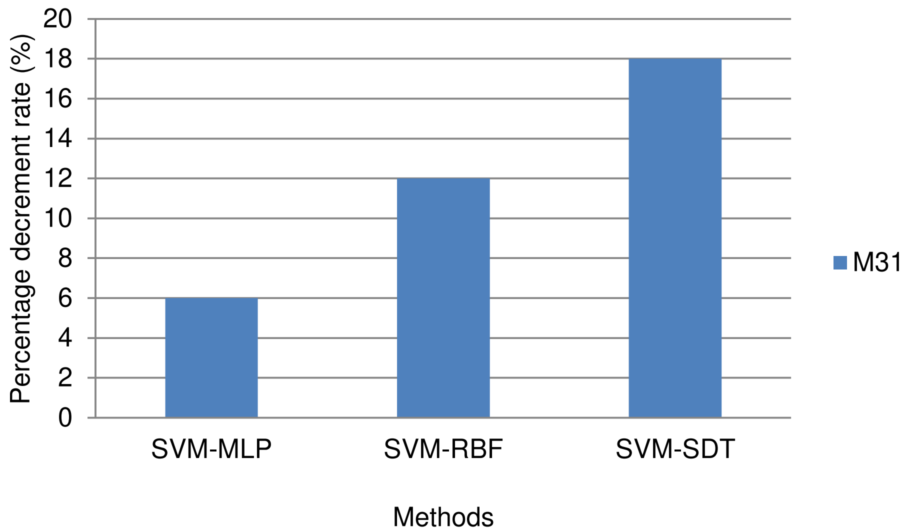

Figure 4.5. Decrement rates in *RMSE*'s of hamstring-CT with five predictor variables.

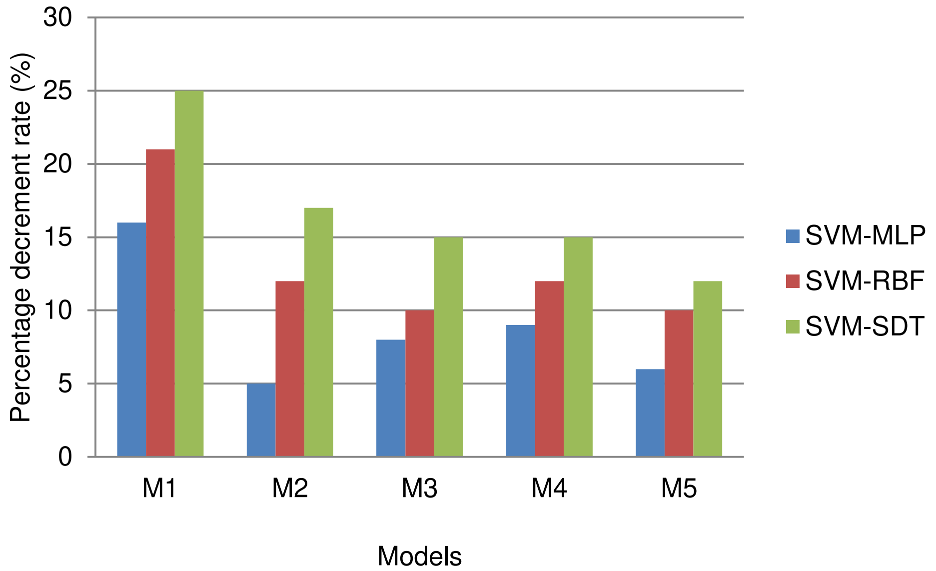

Figure 4.6. Decrement rates in *RMSE*'s of hamstring-ST with one predictor variable.

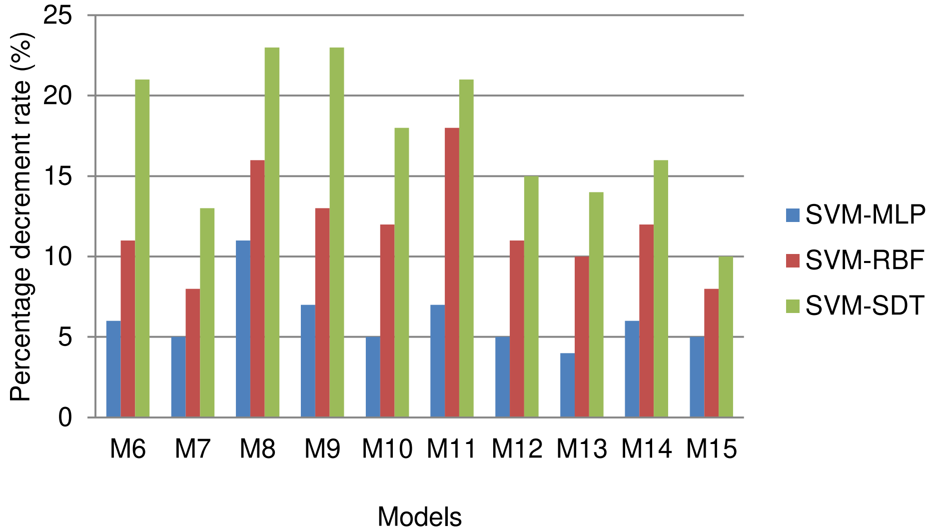

Figure 4.7. Decrement rates in *RMSE*'s of hamstring-ST with two predictor variables

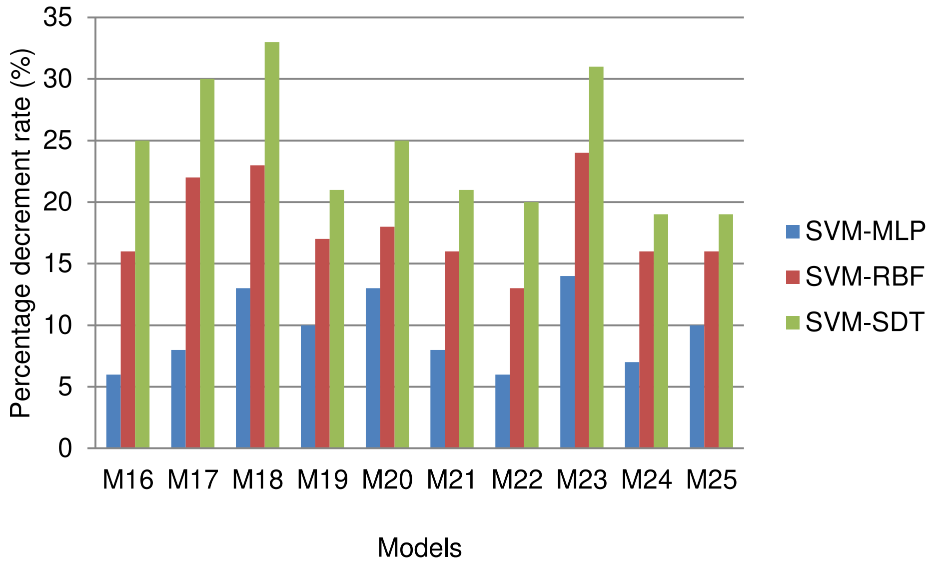

Figure 4.8. Decrement rates in *RMSE*'s of hamstring-ST with three predictor variables

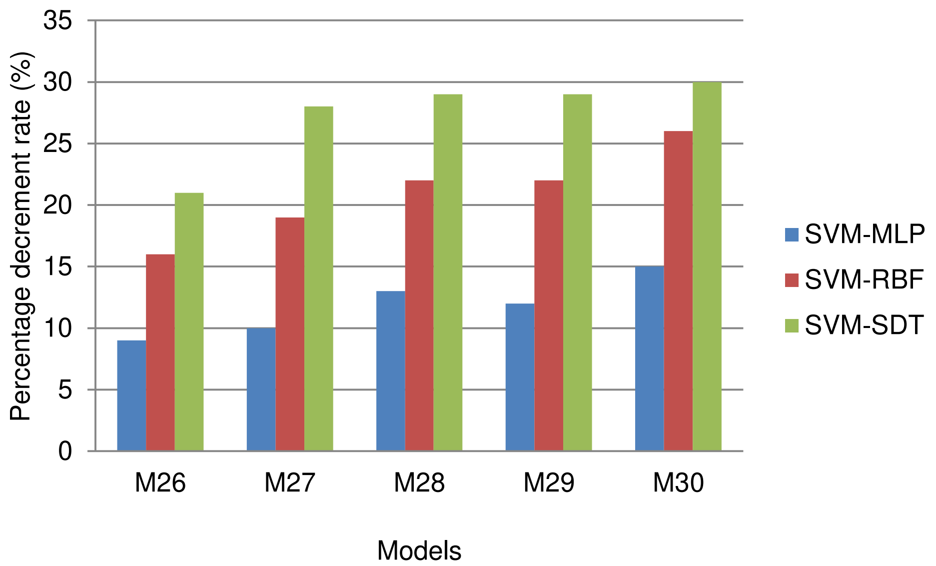

Figure 4.9. Decrement rates in *RMSE*'s of hamstring-ST with four predictor variables

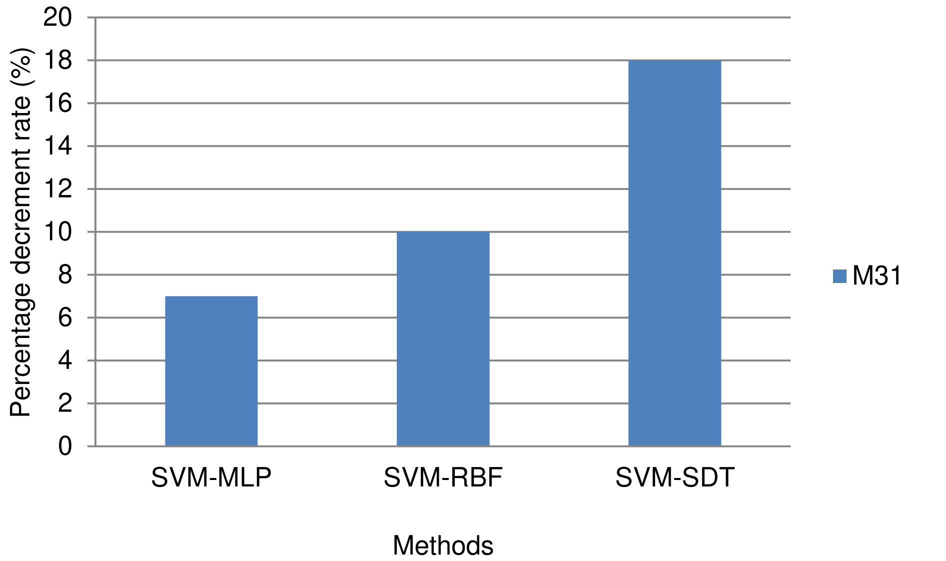

Figure 4.10. Decrement rates in *RMSE*'s of hamstring-ST with five predictor variables

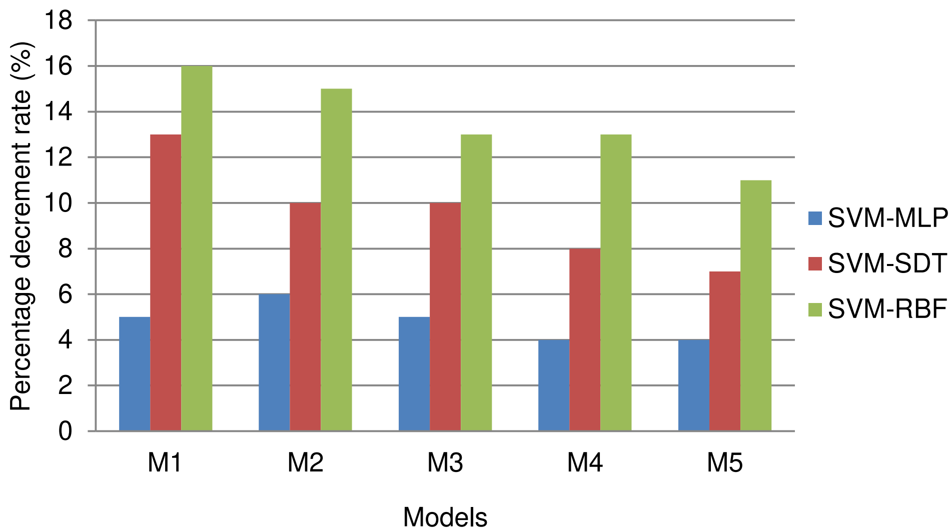

Figure 4.11. Decrement rates in *RMSE*'s of quadriceps-CT with one predictor variables

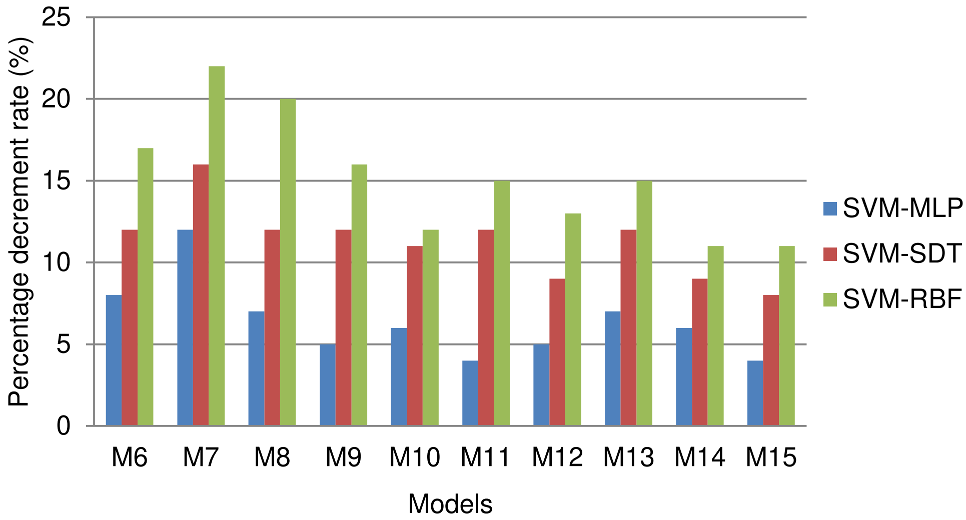

Figure 4.12. Decrement rates in *RMSE*'s of quadriceps-CT with two predictor variables

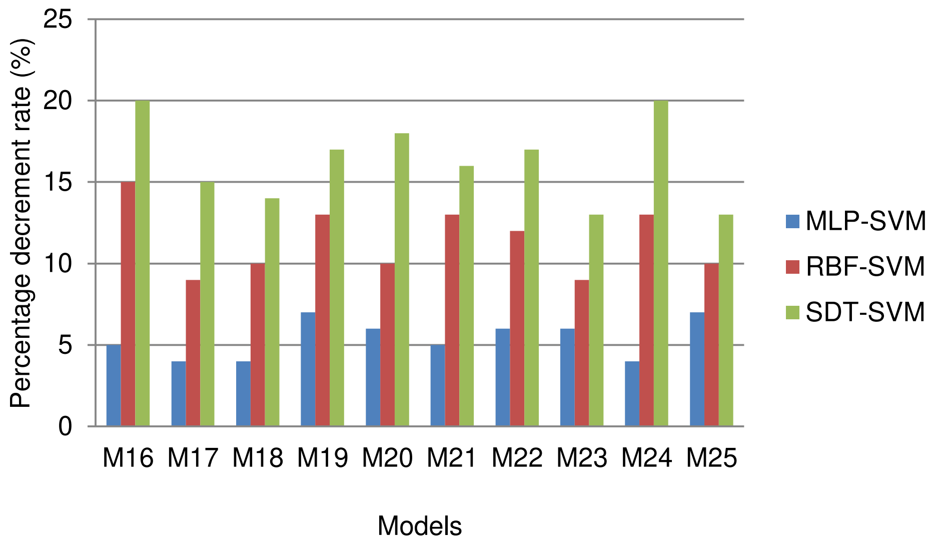

Figure 4.13. Decrement rates in *RMSE*'s of quadriceps-CT with three predictor variables

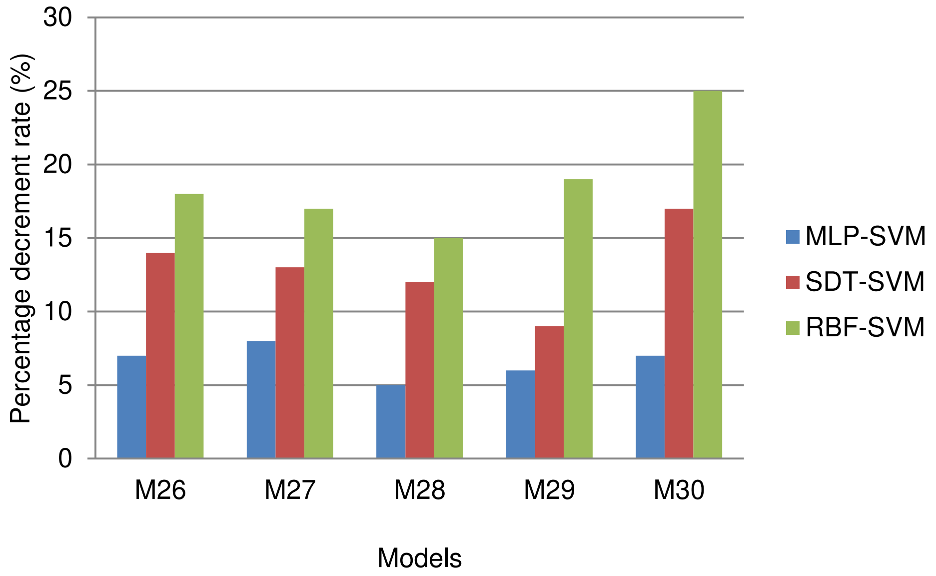

Figure 4.14. Decrement rates in *RMSE*'s of quadriceps-CT with four predictor variables

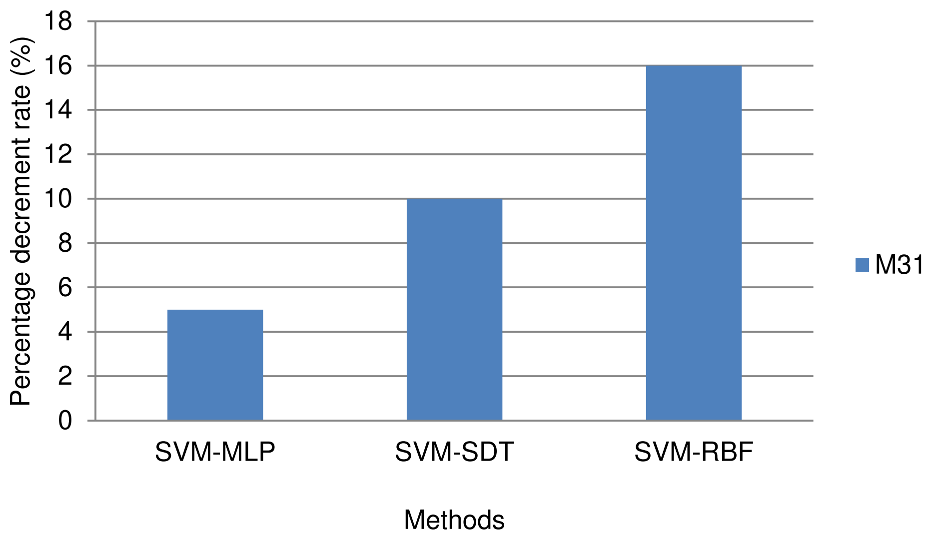

Figure 4.15. Decrement rates in *RMSE*'s of quadriceps-CT with five predictor variables

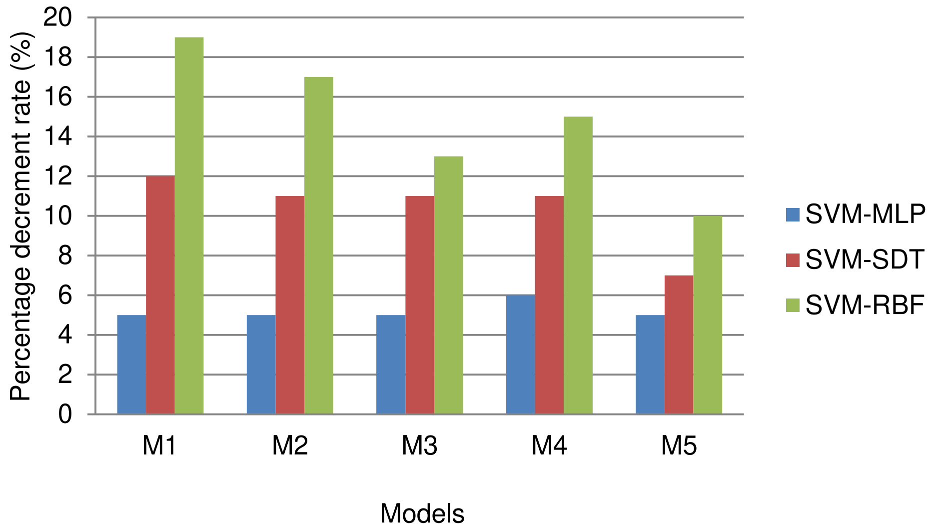

Figure 4.16. Decrement rates in *RMSE*'s of quadriceps-ST with one predictor variable

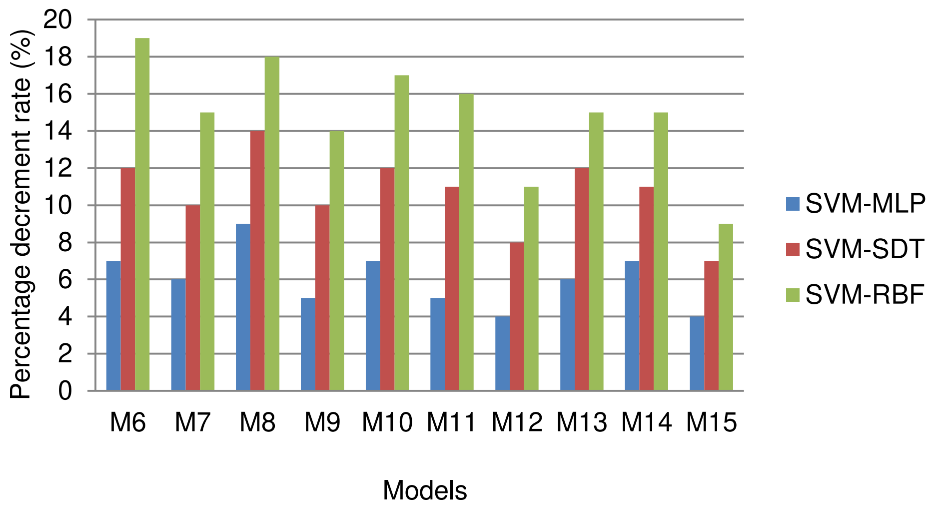

Figure 4.17. Decrement rates in *RMSE*'s of quadriceps-ST with two predictor variables

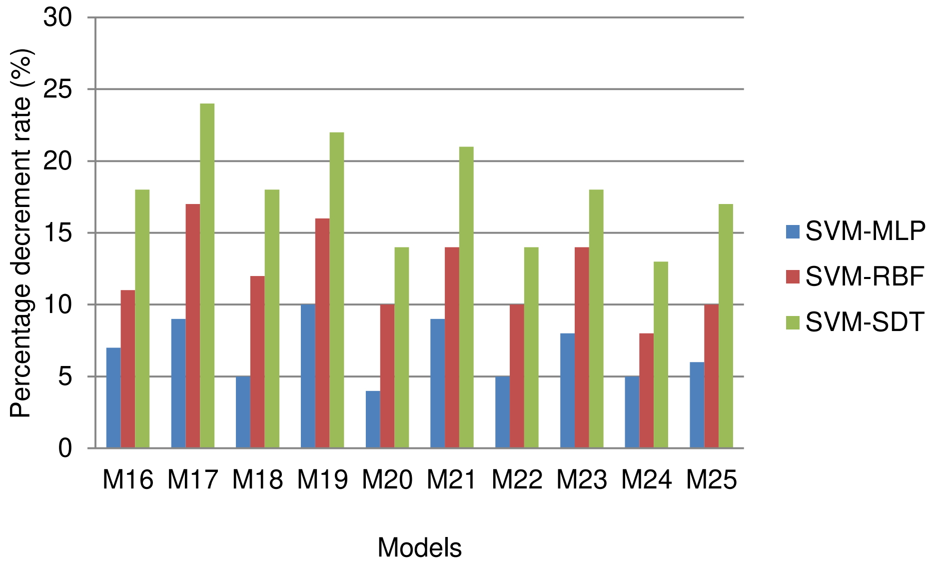

Figure 4.18. Decrement rates in *RMSE*'s of quadriceps-ST with three predictor variables

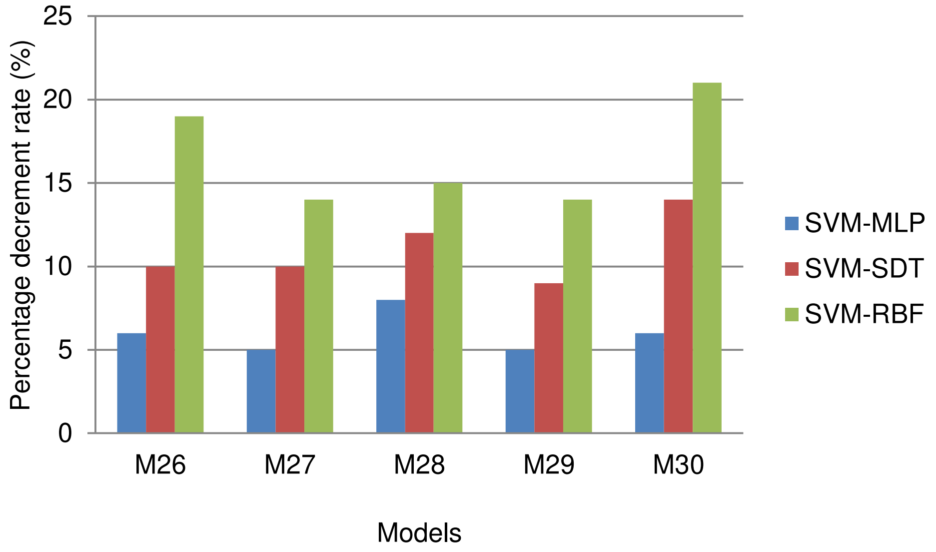

Figure 4.19. Decrement rates in *RMSE*'s of quadriceps-ST with four predictor variables

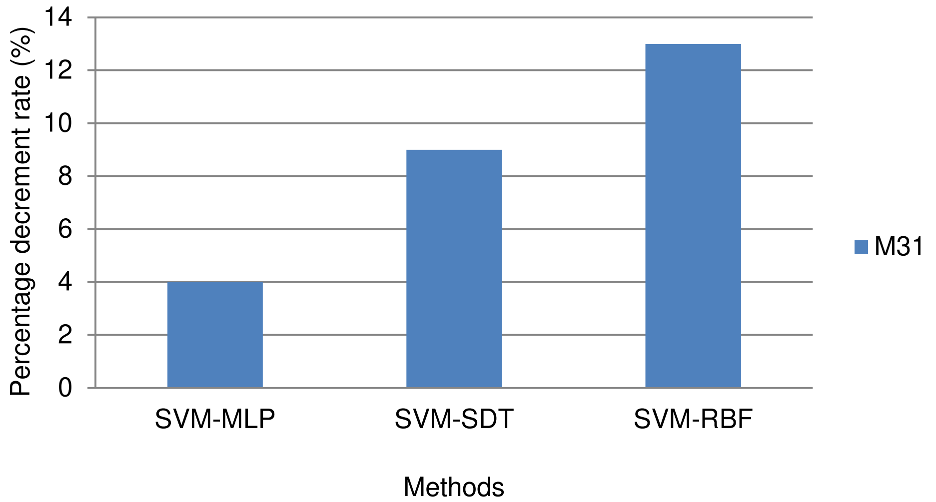

Figure 4.20. Decrement rates in *RMSE*'s of quadriceps-ST with five predictor variables

Figure 4.21 through Figure 4.36 show the decrement rates in percentage of hamstring and quadriceps muscle strength prediction for both classic and static training with the predictor variable BMI.

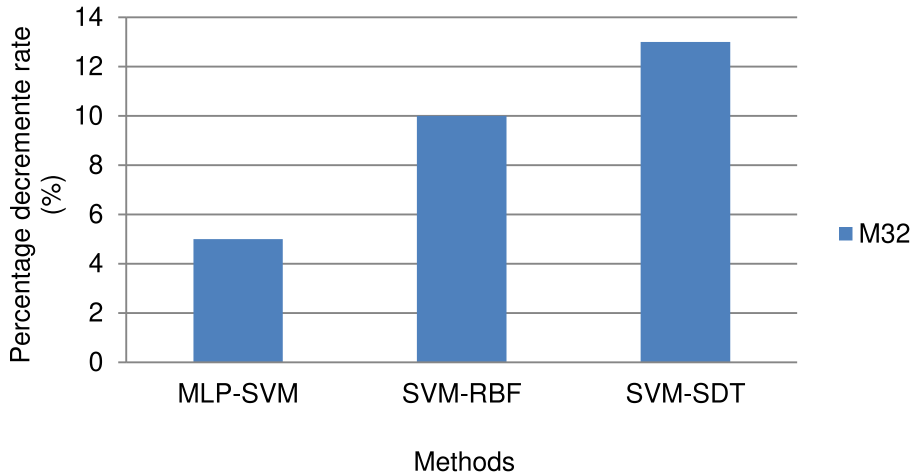

Figure 4.21. Decrement rates in *RMSE*'s of hamstring-CT with one predictor variable (BMI case)

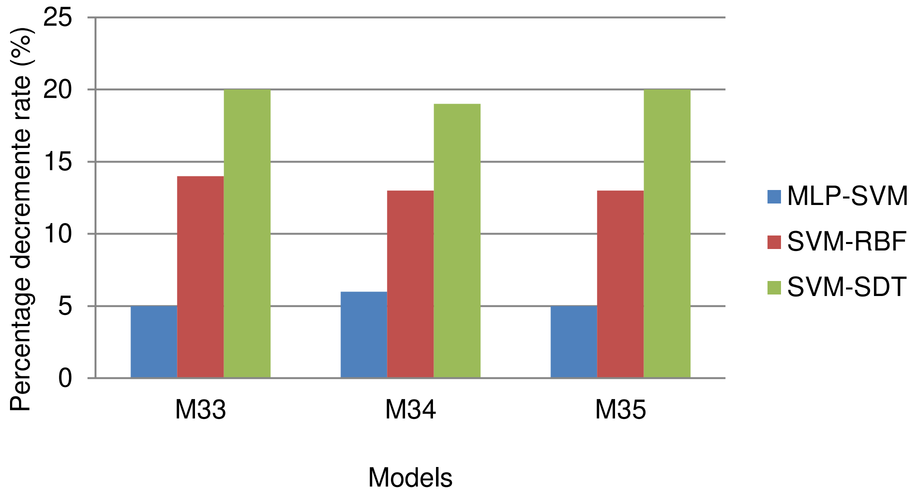

Figure 4.22. Decrement rates in *RMSE*'s of hamstring-CT with two predictor variables (BMI case)

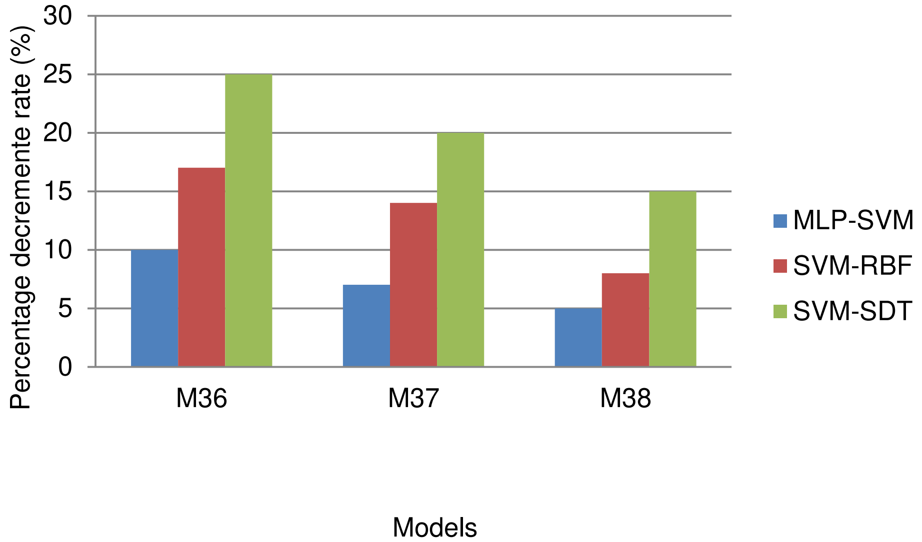

Figure 4.23. Decrement rates in *RMSE*'s of hamstring-CT with three predictor variables (BMI case)

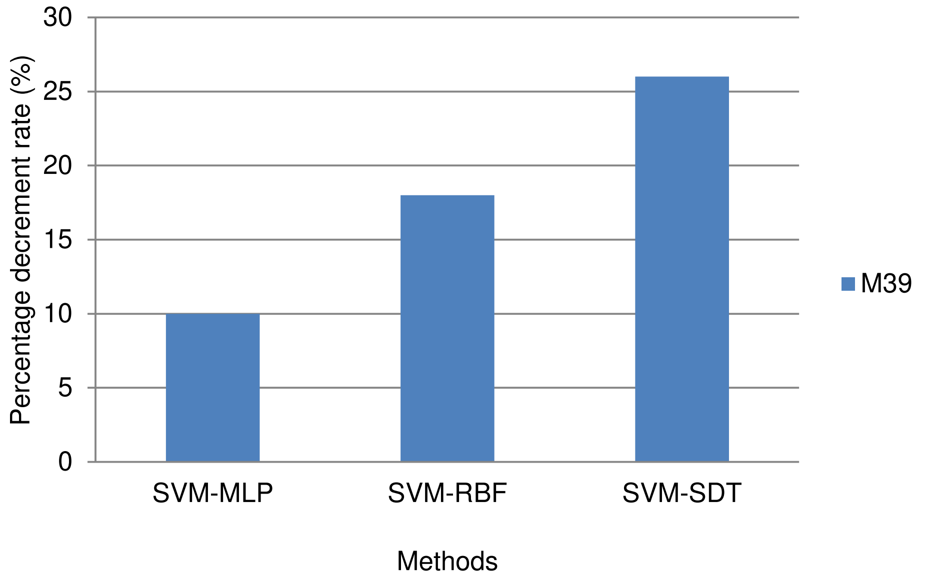

Figure 4.24. Decrement rates in *RMSE*'s of hamstring-CT with four predictor variables (BMI case)

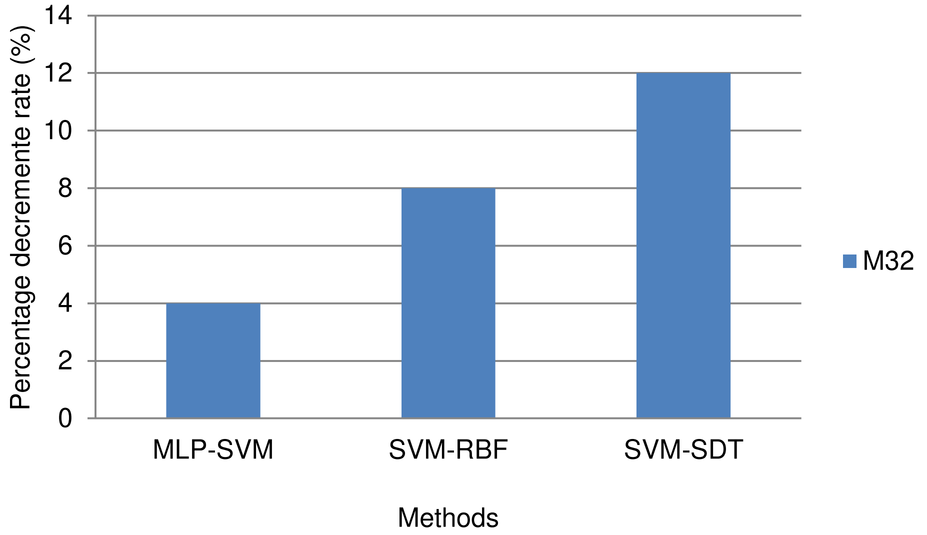

Figure 4.25. Decrement rates in *RMSE*'s of hamstring-ST with one predictor variable (BMI case)

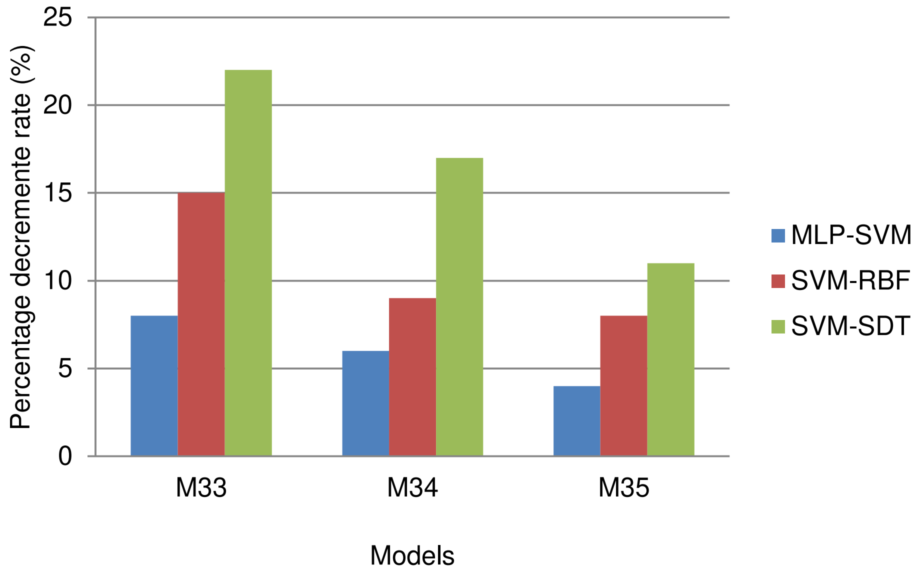

Figure 4.26. Decrement rates in *RMSE*'s of hamstring-ST with two predictor variables (BMI case)

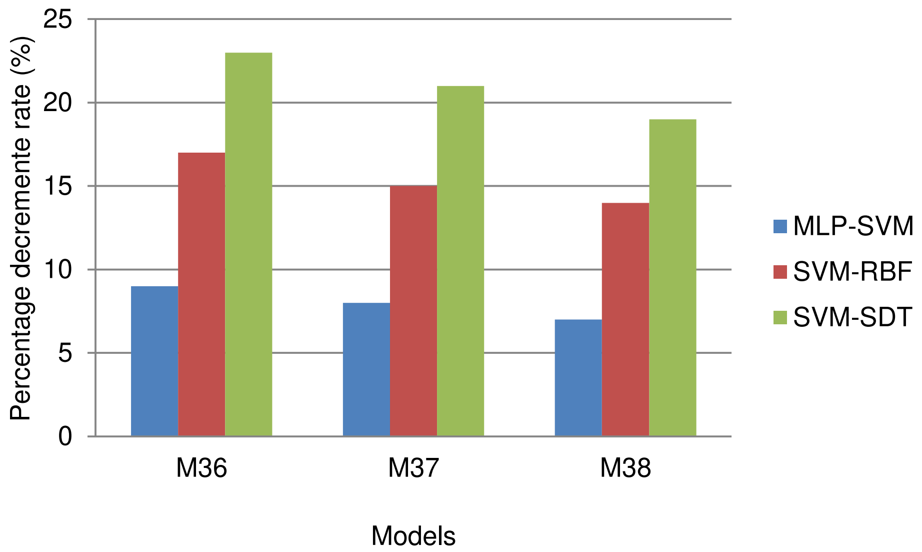

Figure 4.27. Decrement rates in *RMSE*'s of hamstring-ST with three predictor variables (BMI case)

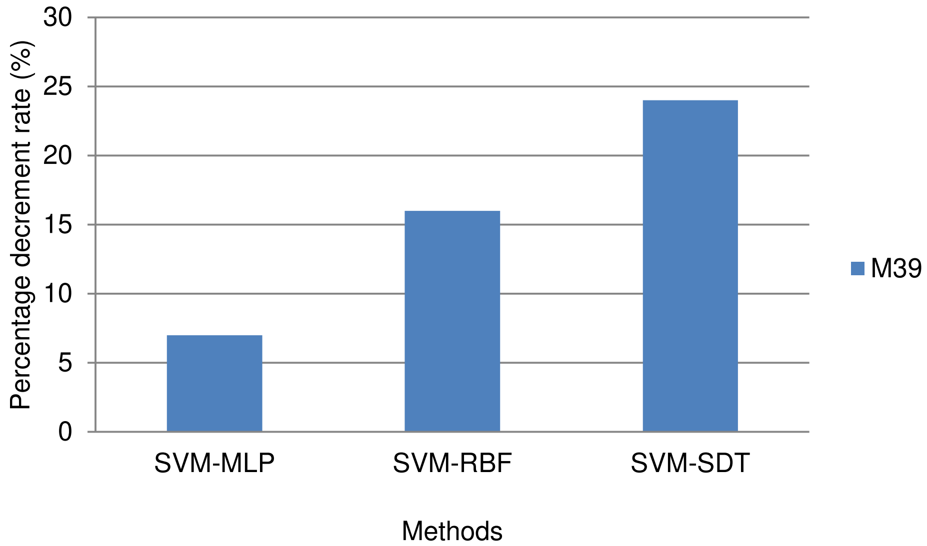

Figure 4.28. Decrement rates in *RMSE*'s of hamstring-ST with four predictor variables (BMI case)

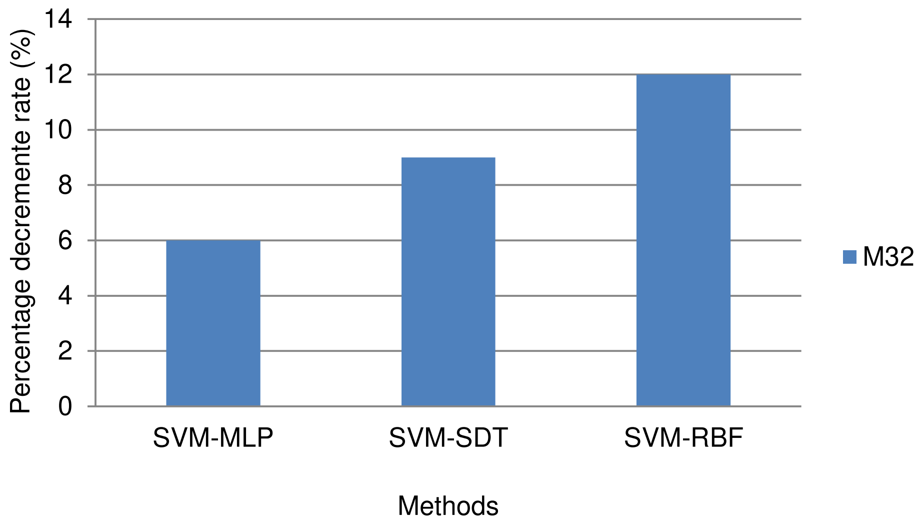

Figure 4.29. Decrement rates in *RMSE*'s of Quadriceps-CT with one predictor variable (BMI case)

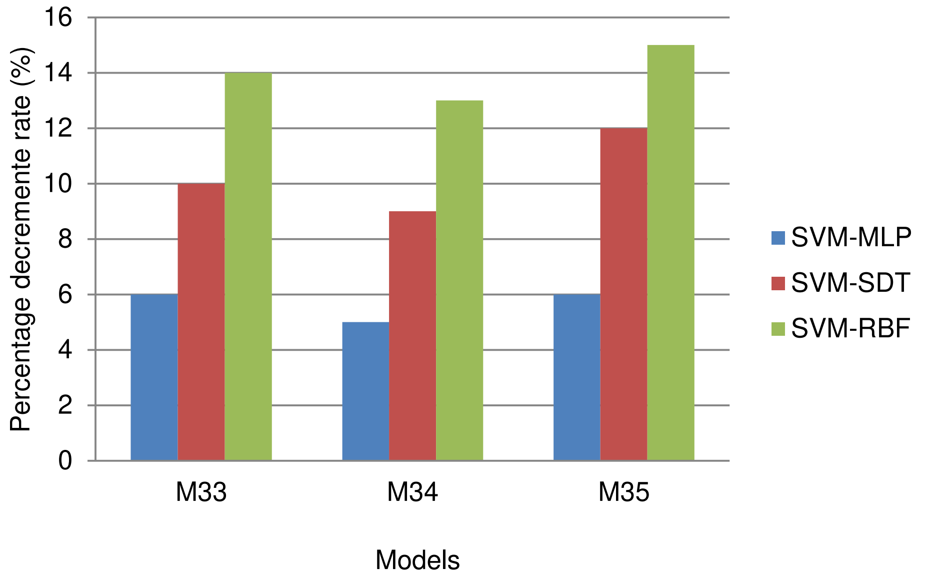

Figure 4.30. Decrement rates in *RMSE*'s of Quadriceps -CT with two predictor variables (BMI case)

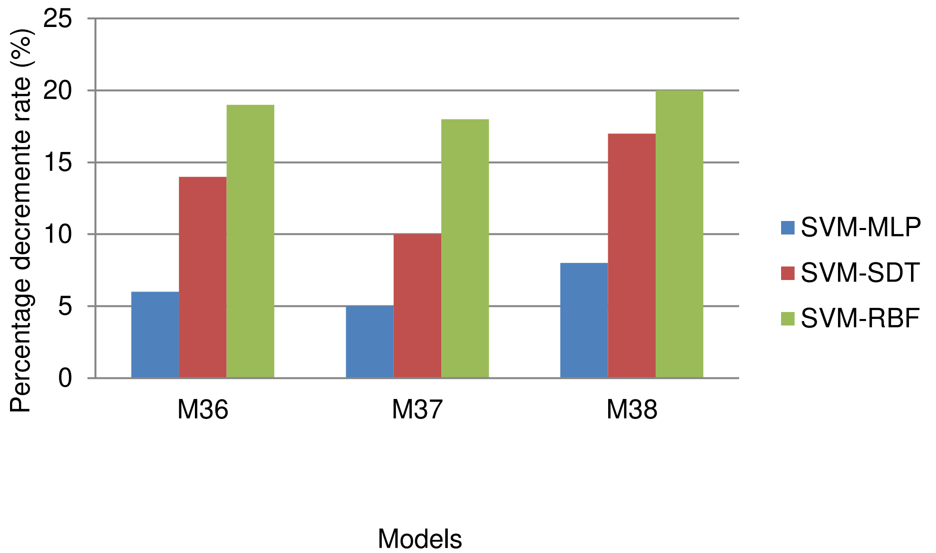

Figure 4.31. Decrement rates in *RMSE*'s of Quadriceps-CT with three predictor variables (BMI case)

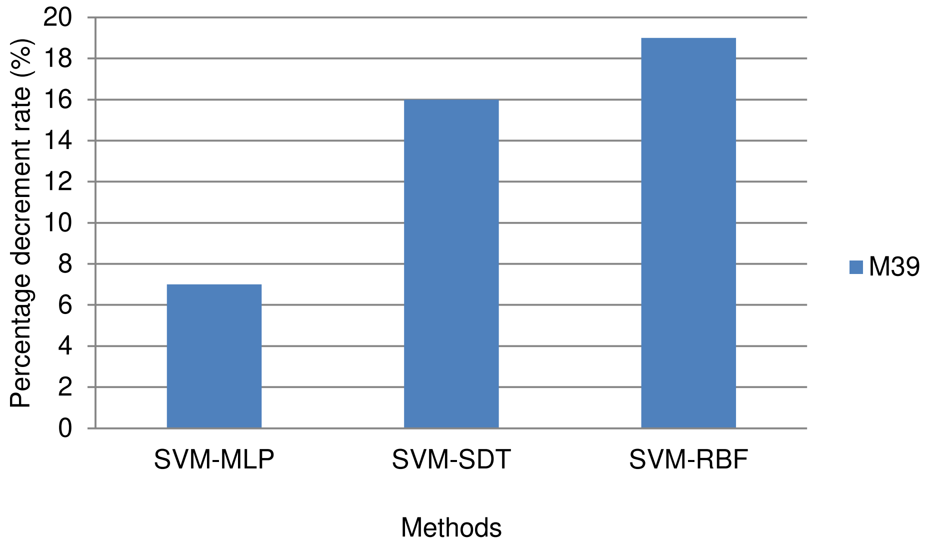

Figure 4.32. Decrement rates in *RMSE*'s of Quadriceps -CT with four predictor variables (BMI case)

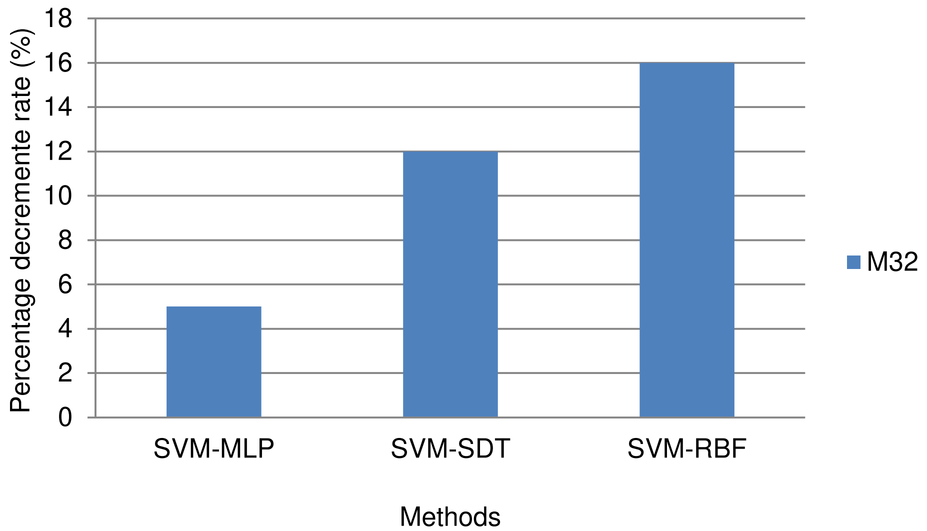

Figure 4.33. Decrement rates in *RMSE*'s of Quadriceps-ST with one predictor variable (BMI case)

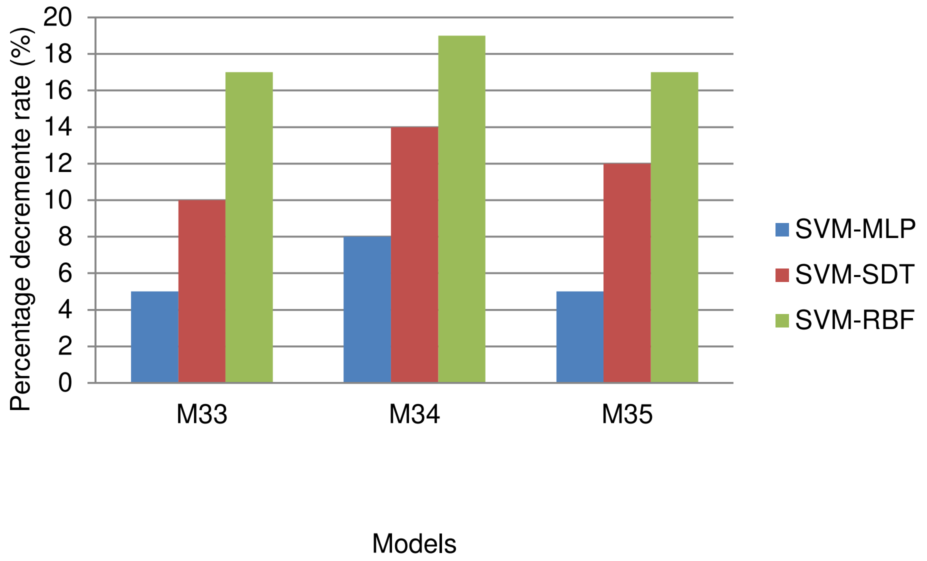

Figure 4.34. Decrement rates in *RMSE*'s of Quadriceps -ST with two predictor variables (BMI case)

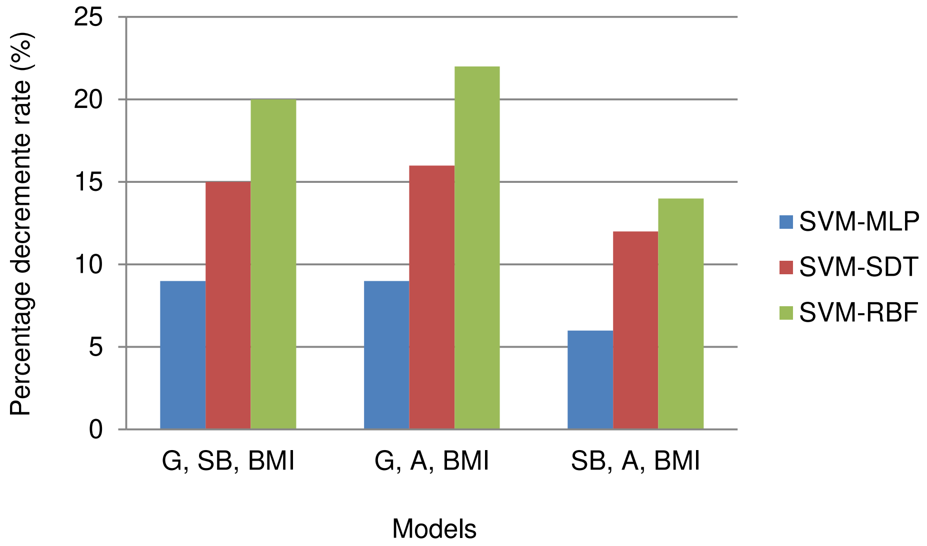

Figure 4.35. Decrement rates in *RMSE*'s of Quadriceps-ST with three predictor variables (BMI case)

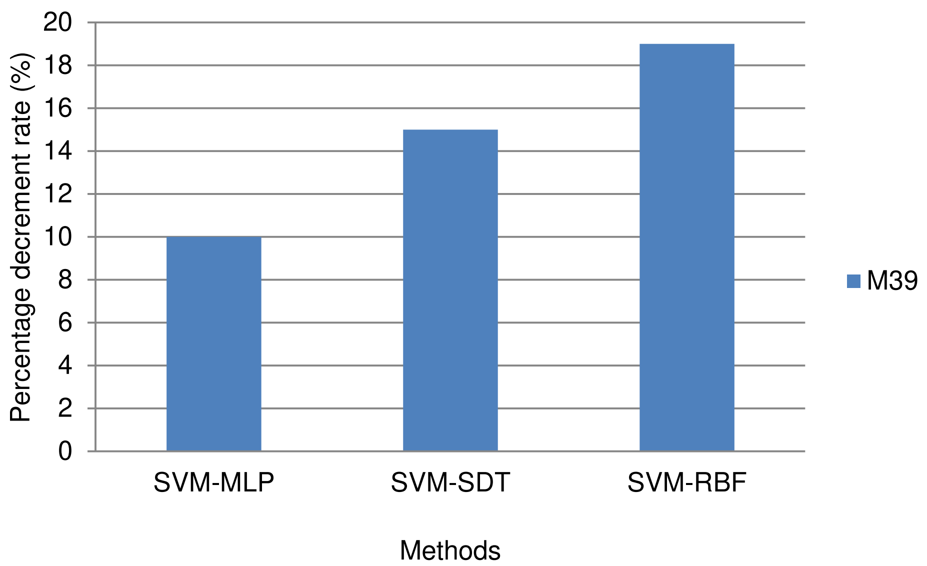

Figure 4.36. Decrement rates in *RMSE*'s of Quadriceps -ST with four predictor variables (BMI case)

Figure 4.37 through Figure 4.44 show the average of the decrement rates in *RMSE*'s of hamstring and quadriceps muscle strength prediction with SVM compared to *RMSE*'s obtained by MLP, RBFNN and SDT for both classic and static training with and without the predictor variable BMI.

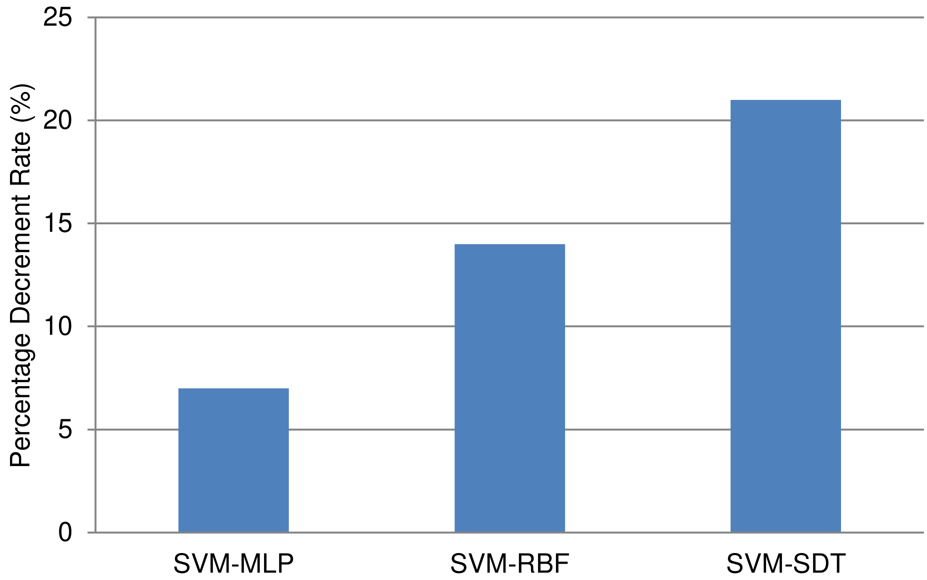

Figure 4.37. Average decrement rates in *RMSE*'s of hamstring-CT muscle strength prediction

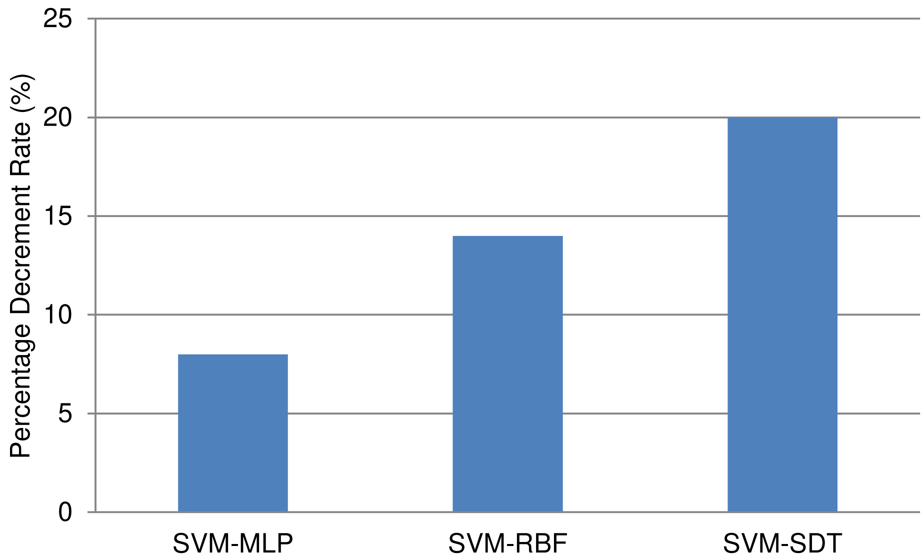

Figure 4.38. Average decrement rates in *RMSE*'s of hamstring-ST muscle strength prediction

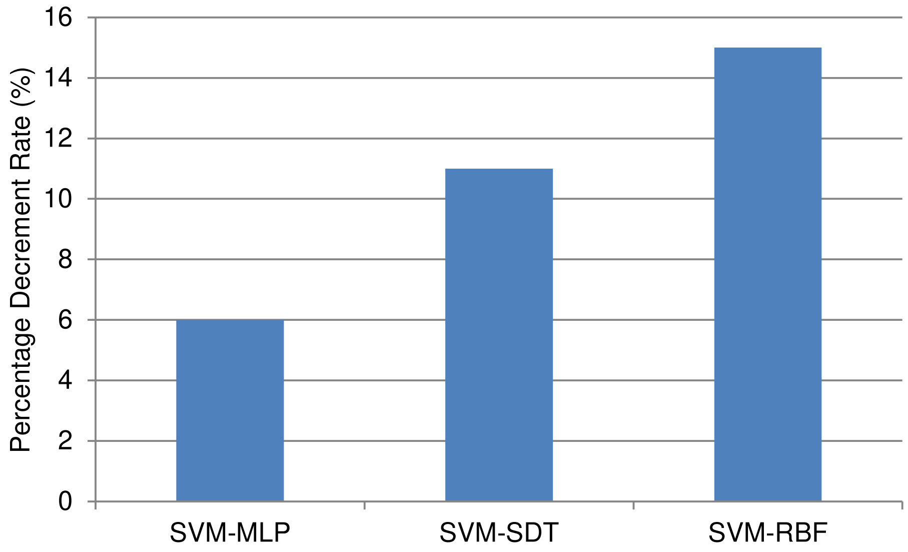

Figure 4.39. Average decrement rates in *RMSE*'s of quadriceps-CT muscle strength prediction

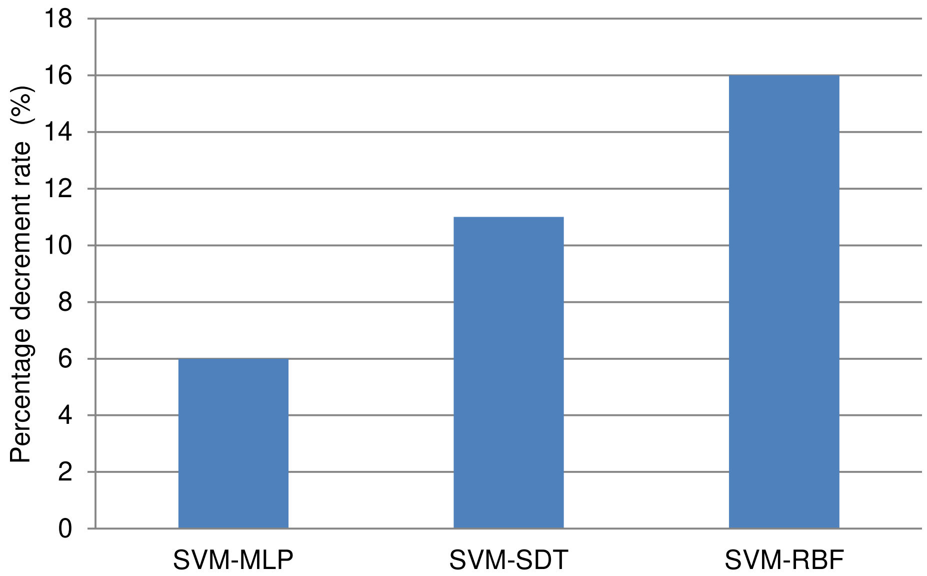

Figure 4.40. Average decrement rates in *RMSE*'s of quadriceps-ST muscle strength prediction

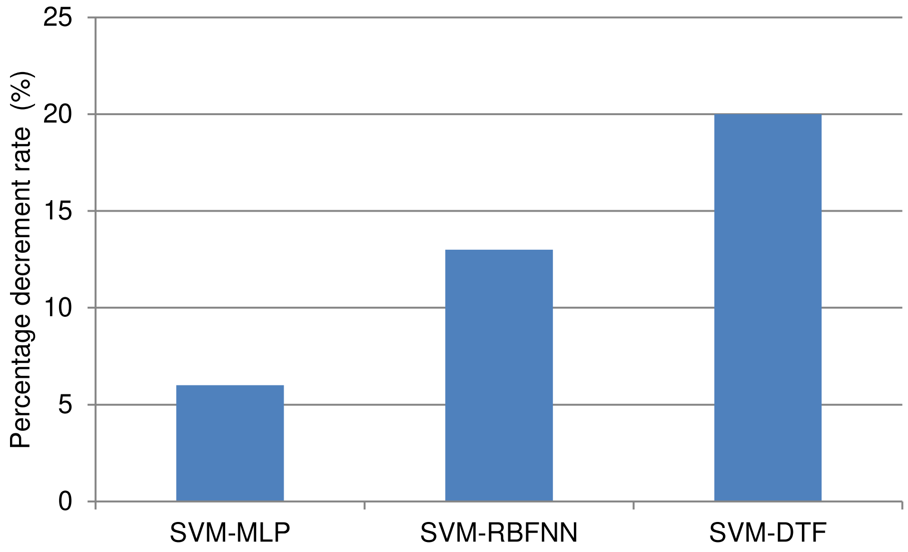

Figure 4.41. Average decrement rates in *RMSE*'s of hamstring-CT muscle strength prediction with BMI

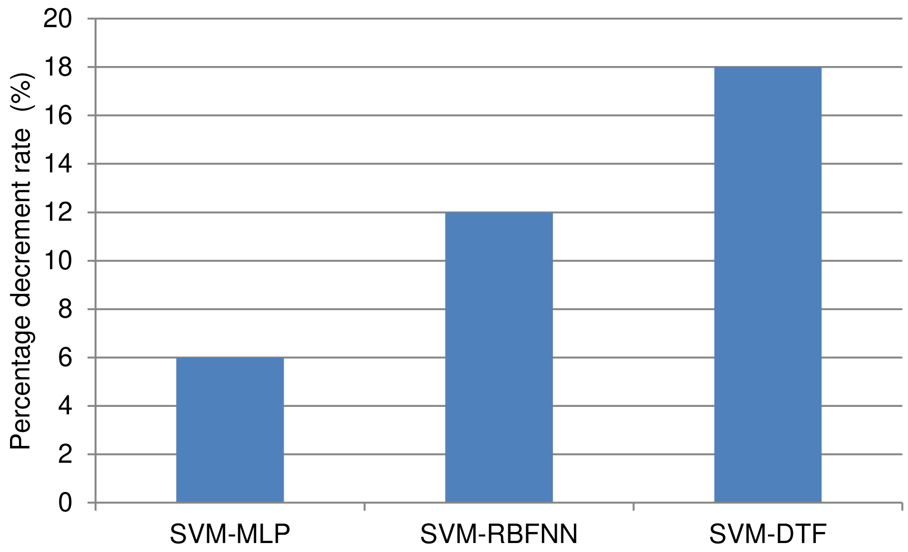

Figure 4.42. Average decrement rates in *RMSE*'s of hamstring-ST muscle strength prediction with BMI

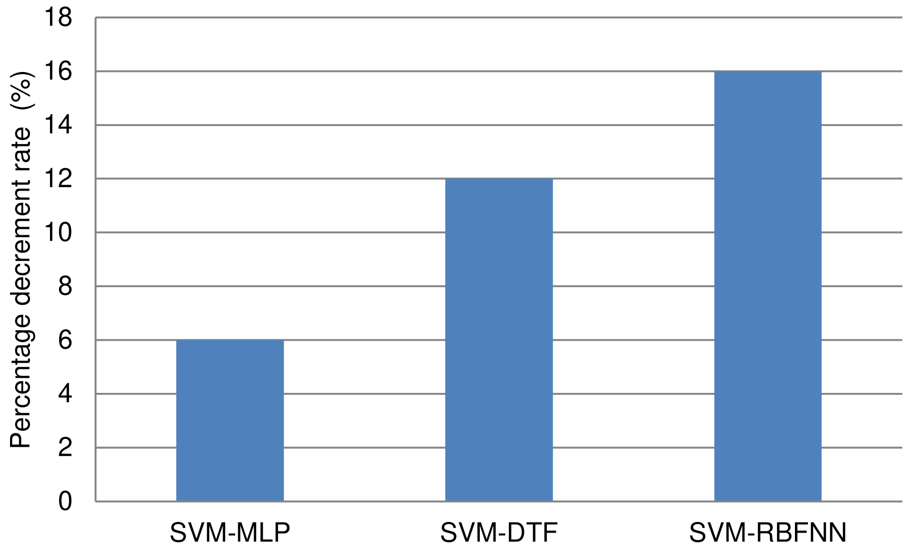

Figure 4.43. Average decrement rates in *RMSE*'s of quadriceps-CT muscle strength prediction with BMI

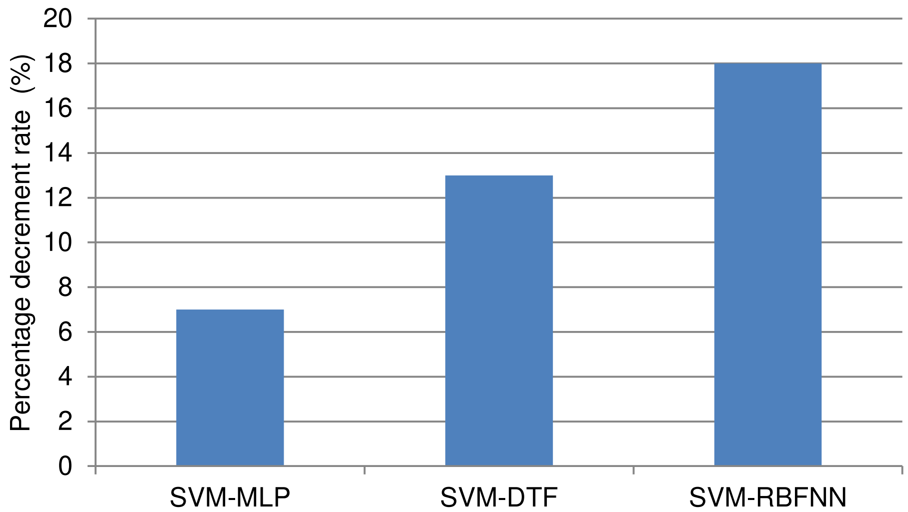

Figure 4.44. Average decrement rates in *RMSE*'s of quadriceps-ST muscle strength prediction with BMI

Figure 4.45 through Figure 4.448 show the average *RMSE*'s for hamstring prediction models with classic training.

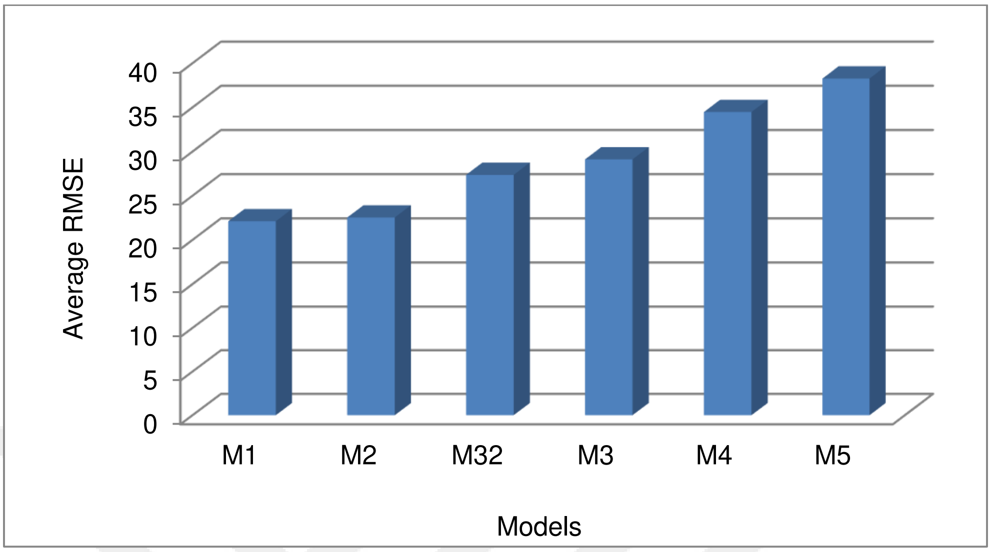

Figure 4.45. Average *RMSE*'s for hamstring-CT prediction models with one predictor variable

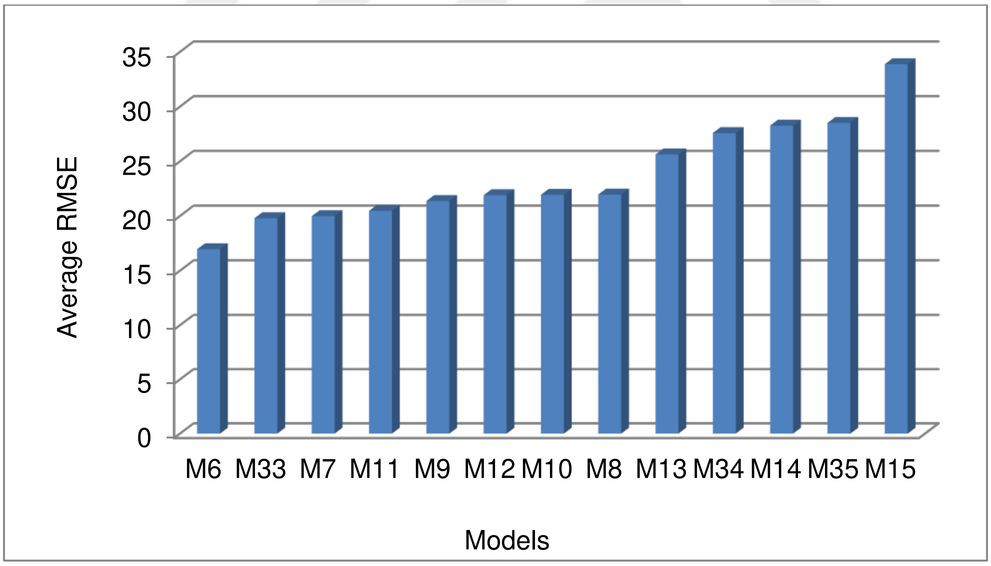

Figure 4.46. Average *RMSE*'s for hamstring-CT prediction models with two predictor variables

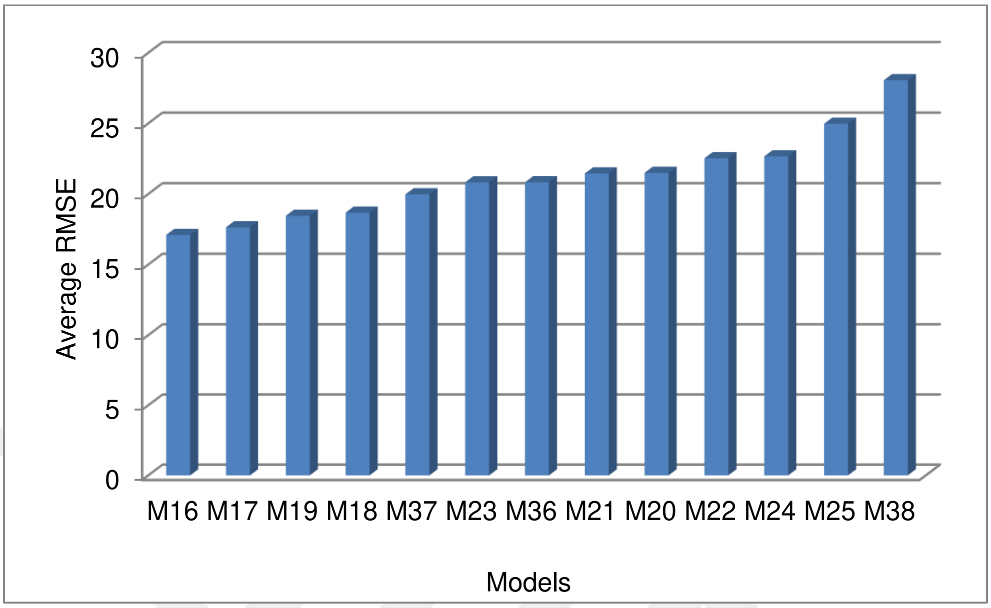

Figure 4.47. Average *RMSE*'s for hamstring-CT prediction models with three predictor variables

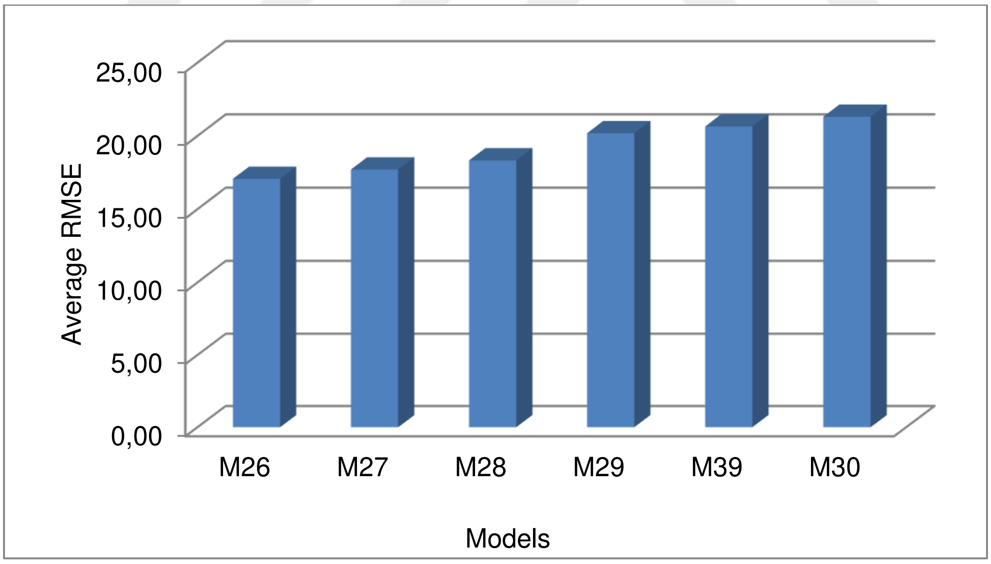

Figure 4.48. Average *RMSE*'s for hamstring-CT prediction models with four predictor variables.

Figure 4.49 through Figure 4.52 show the average *RMSE*'s for quadriceps prediction models with classic training.

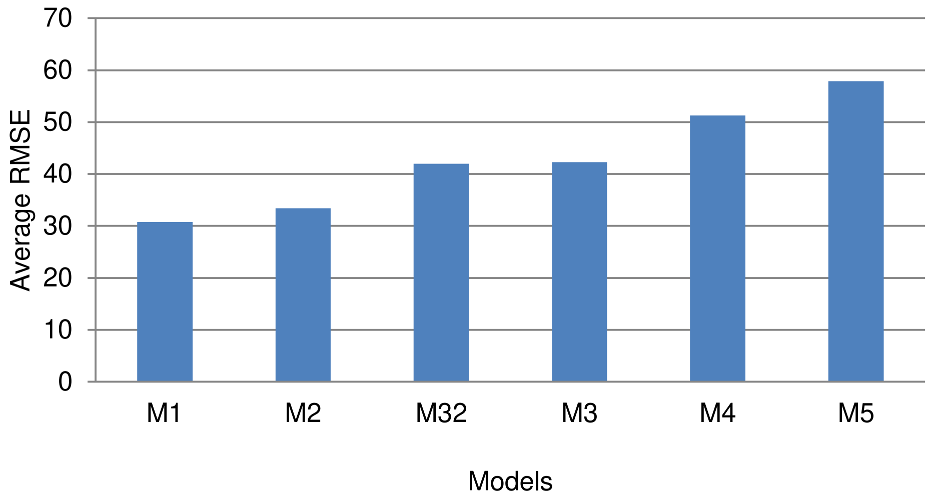

Figure 4.49. Average *RMSE's* for quadriceps-CT prediction models with one predictor variable.

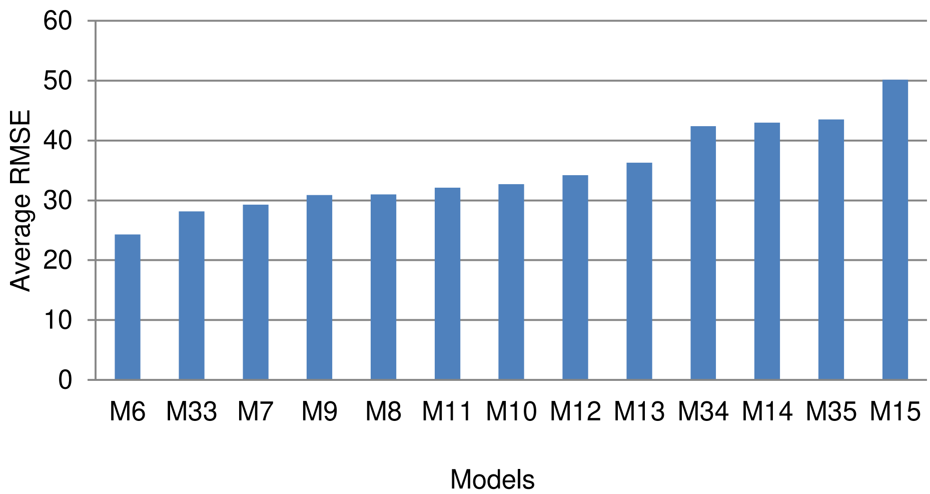

Figure 4.50. Average *RMSE*'s for quadriceps-CT prediction models with two predictor variables.

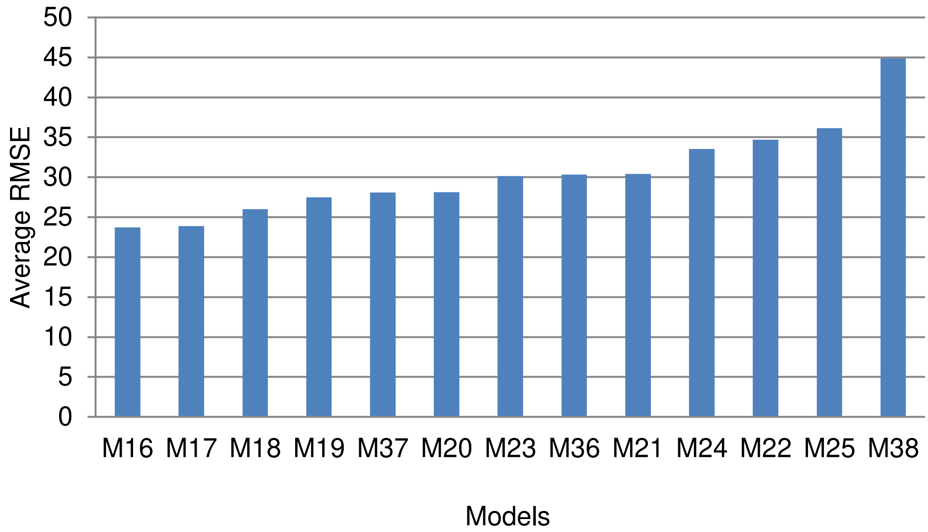

Figure 4.51. Average *RMSE*'s for quadriceps-CT prediction models with three predictor variables.

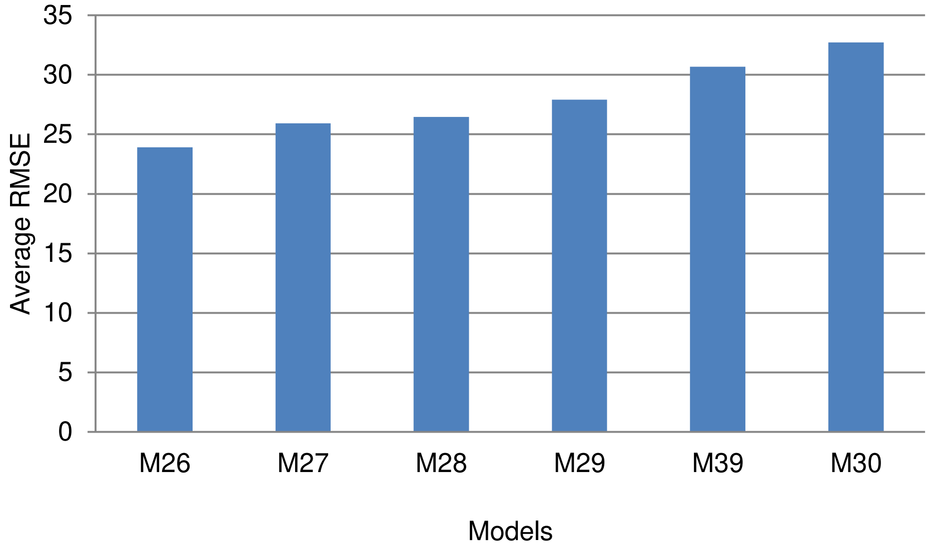

Figure 4.52. Average *RMSE*'s for quadriceps-CT prediction models with four predictor variables.

4.3 Discussions

4.3.1 Overall Discussion on the Results

The results obtained in this study reveal that SVM-based prediction models outperform MLP-based, RBFNN-based and DTF-based prediction models. In other words SVM models show the highest accuracy for the prediction of hamstring and quadriceps muscle strength, regardless of which category of prediction models has been utilized or which muscle training method has been applied to the participants. In more details:

- SVM-based prediction models have in average an *RMSE* value of 20.25 Nm, 20.19 Nm for the hamstring muscle strength prediction with the classic and static training respectively. They yield an average *RMSE* value of 30.62 Nm and 29.40 Nm for the quadriceps muscle strength prediction with the classic and static training respectively. On the other hand, SVM-based prediction models that have predictor variable BMI instead of predictor variables height and weight have in average an *RMSE* value of 21.60 Nm, 21.89 Nm for the hamstring muscle strength prediction with the classic and static training respectively. Similarly, SVM-based quadriceps muscle strength prediction models yield an average *RMSE* value of 32.98 Nm and 31.62 Nm with the classic and static training respectively.
- MLP-based prediction models have in average an *RMSE* value of 21.83 Nm, 21.92 Nm for prediction of hamstring muscle strength with the classic and static training respectively. They yield an average *RMSE* value of 32.45 Nm and 31.35 Nm for the quadriceps muscle strength prediction with the classic and static training respectively. On the other hand, MLP-based prediction models that have predictor

variable BMI instead of predictor variables height and weight have in average an *RMSE* value of 23.09 Nm, 23.37 Nm for the hamstring muscle strength prediction with the classic and static training respectively. Similarly, MLP-based quadriceps muscle strength prediction models yield an average *RMSE* value of 35.20 Nm and 34.01 Nm with the classic and static training respectively.

- RBFNN-based models have in average an *RMSE* value of 23.63 Nm, 23.61 Nm for the hamstring muscle strength prediction with classic and static training respectively. RBFNN-based models yield an average *RMSE* value of 36.22 Nm and 35.04 Nm for the quadriceps muscle strength prediction with the classic and static training respectively. On the other hand, RBFNN-based prediction models that have predictor variable BMI instead of predictor variables height and weight have in average an *RMSE* value of 24.84 Nm, 24.96 Nm for the hamstring muscle strength prediction with the classic and static training respectively. Similarly, RBFNN-based quadriceps muscle strength prediction models yield an average *RMSE* value of 39.34 Nm and 38.42 Nm for the classic and static training respectively.
- Finally, SDT-based models have in average an *RMSE* value of 25.51 Nm, 25.29 Nm for the hamstring muscle strength with the classic and static training respectively. SDT-based models yield an average *RMSE* value of 34.48 Nm and 33.16 Nm for the quadriceps muscle strength prediction with the classic and static training respectively. On the other hand, SDT-based prediction models that have predictor variable BMI instead of predictor variables height and weight have in average an *RMSE* value of 26.84 Nm, 26.76 Nm for the hamstring muscle strength prediction with the classic and static training respectively. Similarly, SDT-based quadriceps muscle strength prediction models

yield an average *RMSE* value of 37.60 Nm and 36.41 Nm for the classic and static training respectively.

- The ranking of machine learning methods in terms of performance from lowest *RMSE*'s to highest ones can be listed as SVM, MLP, RBFNN and SDT for the hamstring muscle strength prediction irrespective of which muscle training type is used or which prediction model is considered. On the other hand this ranking is listed as SVM, MLP, SDT and RBFNN for the quadriceps muscle strength prediction with both classic and static training and with or without the predictor variable BMI. Therefore DTF-based and RBFNN-based prediction models show relatively the worst performance in terms of prediction accuracy for hamstring and quadriceps muscle strength prediction respectively.
- By comparing the *RMSE*'s obtained by SVM-based prediction models to the *RMSE*'s obtained by the other methods (MLP, RBFNN, DTF respectively for the hamstring muscle strength prediction and MLP, DTF, RBFNN respectively for quadriceps muscle strength prediction), the percentage decrement rates in *RMSE*'s obtained by SVM-based prediction models are in average:
  - 7.24%, 14.32% and 20.61% with the classic training, and 7.89%, 14.48% and 20.14% with the static training for the hamstring muscle strength prediction.
  - 5.65%, 11.19% and 15.47% with the classic training, and 6.21%, 11.34% and 16.11% with the static training for the quadriceps muscle strength prediction.
- On the other hand when predictor variable BMI is used instead of predictor variables height and weight, the percentage decrement rates in *RMSE*'s obtained by SVM-based prediction models are in average:

    - 6.44%, 13.04% and 19.51% with the classic training, and 6.33%, 12.29% and 18.21% with the static training for the hamstring muscle strength prediction.
    - 6.29%, 12.28% and 16.15% with the classic training, and 7.03%, 13.16% and 17.72% with the static training for the quadriceps muscle strength prediction.
- In general, prediction models that use predictor variables height and weight yield lower RMSE’s than the ones that use predictor variable BMI instead of height and weight.
- The results reveal that the hamstring muscle strength models yield lower *RMSE*'s than the quadriceps muscle strength prediction models, regardless of which machine learning method has been used or which muscle training method has been utilized. In more detail, as compared to *RMSE*’s of quadriceps muscle strength prediction models, the percentage decrement rates in *RMSE*’s for hamstring muscle strength prediction models are in average (MLP, SDT and RBFNN respectively):
    - 33.87%, 32.74%, 26.02% and 34.75% with the classic training.
    - 31.31%, 30.06%, 28.78% and 27.84% with the static training.
    - 34.51%, 34.41%, 28.62% and 36.86% with the classic training (BMI case).
    - 30.77%, 31.29%, 26.76% and 35.05% with the static training (BMI case).
- Considering the results in terms of muscle training type, the *RMSE*’s of the classic and static training have been found to be very close to each other, independent of which machine learning method has been used or which prediction model has been utilized. Therefore they are comparable.

4.3.2 Discussions on the Different Categories of the Prediction Models

In general, among all the prediction models developed, prediction models including predictor variables gender and weight lead to the lowest *RMSE*s and highest *R* whereas prediction models including sport branch and age lead to the highest *RMSE'*s and lowest *R'*s. On the other hand, the category of prediction models with four and three predictor variables exhibit in general the best performance, whilst the category of prediction models with one predictor variables exhibits the worst performance in terms of *RMSE*s. In more detail, the results obtained in this study reveal that :

- The SVM-based prediction model with four predictor variables including gender, height, weight and age for the hamstring muscle strength prediction with the classic training yields an *RMSE* value of 15.19 Nm which is the lowest *RMSE* among all the prediction models. On the Other hand, the RBFNN-based prediction model with one predictor variable age for the quadriceps muscle strength prediction with the static training has an *RMSE* value of 61.52 Nm which is the highest *RMSE* among all the prediction models. Therefore, these two prediction models are respectively the best and worst among all the prediction models in terms of prediction accuracy.
- Among the category with one predictor variable, prediction model with variable gender yields in average the lowest *RMSE* and the highest *R*, whilst the prediction model with variable age shows the worst result in terms of *RMSE*. This high performance of the prediction model that include predictor variable gender in this category is due to the fact the performance of the muscular strength of athetes is closely related to their physiological characteristics.
- The prediction model with variables gender and weight and that one with predictor variables sport branch and age occupy the first and last

place respectively in terms of the performance regarding the average of *RMSE* for the category with two predictor variables.

- Among the category with three predictor variables, prediction model with variables gender, weight and height outperforms the other prediction models, whereas prediction model sport branch, height and age yields the worst result in terms of the average of RMSE.
- The prediction models with variables gender, weight, height and age yields in average the lowest *RMSE* and highest *R* whilst that one with variables sport branch, height, weight and age yields in average the highest *RMSE* and the lowest *R* for the category with four predictor variables.

4.3.3. Discussions on the Hamstring Muscle Strength

The hamstring prediction models yields the lowest *RMSE*'s and the highest *R*'s. Particularly, the SVM-based and MLP-based prediction model that includes gender, height, weight and age yield an *RMSE* value of 15.19 Nm and 16.29 Nm respectively, for the hamstring muscle strength prediction with the classic training. On the other hand, the RBFNN-based and SDT-based hamstring prediction model with the classic training that includes predictor variables gender and weight has an *RMSE* value of 17.37 Nm and 18.88 Nm respectively. Among all the prediction models developed in this study, these two prediction models are the most accurate models for SVM, MLP, SDT and RBFNN. In more detail, regardless which regression method is used, the results of hamstring muscle strength prediction (the classic training results that are found to be comparable to the ones with static training are used) reveal that:

- Among prediction models with one predictor variable, the prediction model including gender show better performance than the others regardless which regression method has been used. Particularly, this prediction model, in average yields an *RMSE* value of 22.11 Nm. On the contrary, the prediction model including age shows the lowest performance with an *RMSE* value of 38.26 Nm in average.

- For the prediction models with two predictor variable, prediction model including gender and weight has the lowest *RMSE*, whilst prediction model including sport branch and age has the highest *RMSE*. The best prediction model and the worst one have in average an *RMSE* value of 19.96 Nm and 33.88 Nm respectively.
- The prediction model with three predictor variables including gender, height and weight is the best prediction model in terms of accuracy with in average an *RMSE* value of 17.10 Nm, whereas the prediction model including sport branch, age and BMI with in average an *RMSE* value of 28.06 Nm shows the worst performance among the prediction model with three predictor variables.
- Within prediction models with four predictor variables, the prediction model including gender, height, weight, age with in average an *RMSE* value of 17.07 Nm yields the best performance, whereas prediction model including sport branch, height, weight and age with in average an *RMSE* value of 21.31 Nm yields the worst performance.

4.3.4. Discussions on the Quadriceps Muscle Strength

The ranking of prediction models within the two categories of the quadriceps muscle strength is the same as seen in the discussion part of the hamstring muscle strength. The following discussions are based only on the result of the classic training results that are found to be comparable to the results obtained with the static training. In more details, regardless which regression method is used, the results of quadriceps muscle strength prediction reveal that:

- Firstly, prediction model with variable gender yields the lowest *RMSE* with in average 30.77 Nm, whilst the one with variable age shows the worst performance in terms of *RMSE* with 57.88 Nm within the category including a single predictor variable.
- Secondly, among the prediction model with two predictor variables, prediction model including gender and weight, and the prediction model including sport branch and age with respectively 24.33 Nm and 50.14 Nm occupy the first and last place, respectively, in terms of the performance regarding the average of *RMSE*.
- Thirdly, within the category with three predictor variables, prediction model with variables gender, weight and height with in average an *RMSE* value of 23.72 Nm outperforms the others whereas the prediction model including sport branch, age and BMI with in average an *RMSE* value of 44.86 Nm yields the worst performance.
- Finally, within the category with four predictor variables, the prediction model including predictor variables gender, weight, height and age with in average an *RMSE* value of 23.91 yields the best performance Nm whilst that one including predictor variables sport branch, height, weight and age with in average an *RMSE* value of 32.71 yields the worst performance.

5. CONCLUSION

The use of machine learning methods for the hamstring and quadriceps muscle strength prediction is relatively recent. However the results obtained especially in this thesis are clearly promising. In order to carry out this study, four main machine learning methods namely SVM, MLP, RBFNN and SDT have been used to build the hamstring and quadriceps muscle strength prediction models. These prediction models are divided in five categories according to the number of predictor variables. Gender, age, weight, height and sport branch are the predictor variable used in this study. Predictor variable BMI has been used in place of weight and height for a certain number of prediction models. The dataset used for the hamstring and quadriceps muscle strength prediction included 70 young students from the College of Physical Education and Sport at Gazi University. The classic and static training methods have been used for the hamstring and quadriceps muscle strength measurement.10-fold cross-validation has been used for the calculation of the generalization error of the prediction models. The root mean square errors (*RMSE*'s) and multiple correlation coefficients (*R*'s) have been used for computing the prediction errors.

On the basis of the results obtained, several conclusions can be made on this study. Firstly, the results reveal that machine learning methods especially SVM can be used for estimating the hamstring and quadriceps muscle strength with an acceptable accuracy. Secondly, SVM-based prediction models outperform the other regression methods utilized in this study, regardless of whether hamstring or quadriceps muscle strength is predicted; or whether classic training or static training is employed; or which category of prediction models is used. The Average decrement rate in *RMSE*'s obtained by SVM-based prediction models varies between 5.65% and 20.61%. On the other hand, SDT exhibits the worst performance for the hamstring muscle strength prediction, whereas RBFNN yields the highest *RMSE*'s for the quadriceps muscle strength prediction. Thirdly, the prediction models of hamstring muscle strength yield lower *RMSE*'s than the quadriceps muscle strength prediction models, regardless which category of prediction models is considered or which muscle training type is used. Fourthly, the results reveal that the classic training and the static training show comparable performance for hamstring and quadriceps muscle strength prediction models. Fifthly, the prediction models that include variables gender and weight yield the lowest errors, whilst the ones that include variable sport branch and age exhibit the highest errors. Finally, the prediction models that use predictor variable BMI in place of variables weight and height show the worst performance in terms of accuracy independent of which regression method is utilized or which training type is used.

While this thesis has revealed the potential of using machine learning methods to predict with accuracy the hamstring and quadriceps muscle strength, many opportunities for improving and extending the result of this thesis remain. For instance, in order to determine discriminative features, feature selection algorithms such as F-scores can be employed. The combination of feature selection algorithms and machines learning methods can improve the future results. Others machines learning methods such as Tree Boost, Logistic Regression and Gene expression programming can be used for estimating the hamstring and quadriceps muscle strength. Different muscle training methods and additional datasets that include more predictor variable such as the length, the thickness and the inner diameter of the bone of the thigh, can be explored to create more accurate model.

REFERENCES

Abut, F., & Akay, M. F. , 2015. Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. Medical Devices (Auckland, NZ), 8:369.

Adams, G. M., 1998. Exercise physiology laboratory manual. Boston:McGraw-Hill

Aginsky, K. D., Neophytou, N., & Charalambous, T. (2014). Isokinetic hamstring and quadriceps muscle strength profiles of elite South African football players: sport science. African Journal for Physical Health Education, Recreation and Dance, 20(3):1225-1236.

Akay, M. F., Abut, F., Çetin, E., Yarım, İ., & Sow, B., 2015. Data-Driven Modeling of Quadriceps and Hamstring Muscle Strength Using Support Vector Machines. Third International Symposium on Engineering, Artificial Intelligence & Applications, North Cyprus, Nov 4-6, 2-4.

Akay, M. F., Abut, F., Çetin, E., Yarım, İ., & Sow, B., 2016. Support vector machines for predicting the hamstring and quadriceps muscle strength of college-aged athletes. Turkish Journal of Electrical Engineering and Computer Science, DOI:10.3906/elk-1603-304 (accepted for publication)

Akay, M. F., Abut, F., Özçiloğlu, M., & Heil, D., 2016. Identifying the discriminative predictors of upper body power of cross-country skiers using support vector machines combined with feature selection. Neural Computing and Applications, 27(6):1785-1796.

Berger, R., 1961. Determination of resistance load for 1-RM and 10-RM. J Assoc Physical Mental Rehabil, 15: 100–10.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A., 1984. Classification and regression trees. CRC press.

Brown, H., 1992. Lifetime fitness. 3rd ed. Scottsdale (AZ): Gorscuch Scarisbrick

Brzycki, M., 1993. Strength testing—predicting a one-rep max from reps-to-fatigue. Journal of Physical Education, Recreation & Dance, 64(1):88-90.

Clarke H. H.,2013. Comparison of Instruments for Recording Muscle Strength. American Association for Health, Physical Education and Recreation, 25:398-411.

Di Michele, R., Del Curto, L., & Merni, F., 2012. Mechanical and metabolic responses during a high-intensity circuit training workout in competitive runners. The Journal of sports medicine and physical fitness, 52(1): 33-39.

Dogan, U., Glasmachers, T., & Igel, C.,2016. A unified view on multi-class support vector classification. Journal of Machine Learning Research, 17:1-32.

Epley, B., 1985. Poundage chart. Lincoln, NE: Boyd Epley Workout.

Faigenbaum, A. D., & Myer, G. D. (2010). Resistance training among young athletes: safety, efficacy and injury prevention effects. British journal of sports medicine, 44(1): 56-63.

Ford-Smith, C. D., Wyman, J. F., Elswick, R. K., & Fernandez, T. (2001). Reliability of stationary dynamometer muscle strength testing in community-dwelling older adults. Archives of physical medicine and rehabilitation, 82(8):1128-1132.

Ghosh-dastidar, S., Adelı, H., Dadhmer, N., 2008. Principal Component Analysis Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection. Biomedical Engineering, 55(2):512 – 518.

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

Greco, C. C., Da Silva, W. L., Camarda, S. R., & Denadai, B. S. (2012). Rapid hamstrings/quadriceps strength capacity in professional soccer players with different conventional isokinetic muscle strength ratios. Journal of sports science & medicine, 11(3): 418.

Jordan, M. J., Aagaard, P., & Herzog, W. (2015). Rapid hamstrings/quadriceps strength in ACL-reconstructed elite alpine ski racers. Medicine & Science in Sports & Exercise, 47(1):109-119.

Kemmler, W. K., Lauber, D., Wassermann, A., & Mayhew, J. L., 2006. Predicting maximal strength in trained postmenopausal woman. Journal of strength and conditioning research, 20(4): 838-842.

Kılınç, B. E., Kara, A., Savas Camur, Y. O., & Celik, H. (2015). Isokinetic dynamometer evaluation of the effects of early thigh diameter difference on thigh muscle strength in patients undergoing anterior cruciate ligament reconstruction with hamstring tendon graft. Journal of exercise rehabilitation, 11(2):95.

Korürek, M., & Doğan, B., 2010. ECG beat classification using particle swarm optimization and radial basis function neural network. Expert systems with Applications, 37(12):7563-7569.

Lander, J., 1985. Maximum based on reps. NSCA journal, 6(6): 60-61.

Lemmer, J. T., Hurlbut, D. E., Martel, G. F., Tracy, B. L., EY IV, F. M., Metter, E. J., ... & Hurley, B. F., 2000. Age and gender responses to strength training and detraining. Medicine & Science in Sports & Exercise, 32(8):1505-1512.

Lindle, R. S., Metter, E. J., Lynch, N. A., Fleg, J. L., Fozard, J. L., Tobin, J., ... & Hurley, B. F.,1997. Age and gender comparisons of muscle strength in 654 women and men aged 20–93 yr. Journal of Applied Physiology, 83(5):1581-1587.

Lombardi, V. P., 1989. Beginning weight training: the safe and effective way. William C Brown Pub.

Maeda, T., Oowatashi, A., Kiyama, R., Yoshida, Y., & Sakae, K., 2001. Prediction of muscle strength using length and width of the bone. Journal of Physical Therapy Science, 13(1):27-30.

Manca, A., Solinas, G., Dragone, D., & Deriu, F. (2015). Isokinetic testing of muscle performance: new concepts for strength assessment. Isokinetics and Exercise Science, 23(2):69-75.

Mayhew, J. L., Kerksick, C. D., Lentz, D., Ware, J. S., & Mayhew, D. L., 2004. Using repetitions to fatigue to predict one-repetition maximum bench press in male high school athletes. Pediatric Exercise Science, 16(3): 265-276.

McNair, P. J., Colvin, M., & Reid, D., 2011. Predicting maximal strength of quadriceps from submaximal performance in individuals with knee joint osteoarthritis. Arthritis care & research, 63(2):216-222.

Muraki, S., Fukumoto, K., & Fukuda, O., 2013. Prediction of the muscle strength by the muscle thickness and hardness using ultrasound muscle hardness meter. Springerplus, 2(1):457.

O'Connor, R., O'Connor, B., Simmons, J., & O'Shea, P., 1989. Weight training today. Thomson Learning.

Quinlan, J. R. , 1987.Generating Production Rules from Decision Trees. In IJCAI, 87:304-307.

Rokach, L., & Maimon, O. , 2005. Decision trees. Data mining and knowledge discovery handbook . Springer US, 165-192.

Ruiz, J. R., Sui, X., Lobelo, F., Morrow, J. R., Jackson, A. W., Sjöström, M., & Blair, S. N., 2008. Association between muscular strength and mortality in men: prospective cohort study. Bmj, 337, a439.

Sow, B., Akay, M. F., Abut, F., Çetin, E., Yarım, İ., & Alak, H., 2016. New Regression Models for Predicting the Hamstring Muscle Strength Using Support Vector Machines. Çukurova University Journal of the Faculty of Engineering and Architecture, 31(2):153-159.

Tan, S., Wang, J., & Liu, S. , 2015. Establishment of the prediction equations of 1RM skeletal muscle strength in 60-to 75-year-old Chinese men and women. Journal of aging and physical activity, 23(4):640-646.

Ustun, N., Erol, O., Ozcakar, L., Ceceli, E., Ciner, O. A., & Yorgancioglu, Z. R. (2013). Association with isokinetic ankle strength measurements and normal clinical muscle testing in sciatica patients. Journal of back and musculoskeletal rehabilitation, 26(4):361-365.

Vapnik V., 2000. The Nature of Statistical Learning Theory. 2nd ed. New York: Springer.

Wathen, D., 1994. Muscle balance. Essentials of strength training and conditioning, 424-430.

Wright, G. B., Flyer, N., and Fornberg, B., 2013. Radial Basis Functiongenerated Finite Differences: A Mesh-free Method for Computational Geosciences, Handbook of Geomathematics. Springer, Berlin, 1-30.

CURRICULUM VITAE

Boubacar SOW was born in Guinea, in 1985. He completed the elementary education at Conakry, Guinea. He went to Lambadji and Yacine Diallo High Schools in Conakry, Guinea. In 2010, he graduated from the Department of Computer Science, University of Gamal Abdel Nasser, Conakry, Guinea. He worked as a web developer from 2011 to 2014 at Guinea Webdev, Conakry. He started MSc program at the department of Computer Engineering, Çukurova University, Adana, in 2015 under the supervision of Assoc. Prof. Dr. Fatih Akay. He completed MSc program at the Department of Computer Engineering in 2017.