



## DESTEK VEKTÖR MAKİNELERİ KULLANILARAK SUBMAKSİMAL VERİLERDEN MAKSİMUM OKSİJEN TÜKETİMİNİN TAHMİN EDİLMESİ

### (*PREDICTION OF MAXIMAL OXYGEN UPTAKE USING SUPPORT VECTOR MACHINES FROM SUBMAXIMAL DATA*)

M. Fatih AKAY<sup>1</sup>, Gözde ÖZSERT<sup>1</sup>, James GEORGE<sup>2</sup>

#### ÖZET/ABSTRACT

Maksimum oksijen tüketimi ( $VO_2max$ ), kardiyο respiratuar uygunluğu belirleyen en önemli bileşendir. Bu çalışmada, submaksimal koşu bandı egzersizi uygulanan sağlıklı yetişkinlerin  $VO_2max$ 'ını tahmin etmek için Destek Vektör Makinesi (DVM) tabanlı modeller geliştirilmiştir.  $VO_2max$  regresyon denklemini oluşturabilmek için 185 denek içeren veri kümesi kullanılmıştır. 10 katlı çapraz doğrulama kullanılarak, modellerin standart tahmin hatası (STH) ve çoklu korelasyon katsayısı (KK) hesaplanmıştır. Karşılaştırma yapabilmek amacı ile Çoklu Doğrusal Regresyon (ÇDR) yöntemi ve Çok Katmanlı Algılayıcı (ÇKA) kullanılarak  $VO_2max$  tahmin modelleri de geliştirilmiştir. Sonuç olarak, DVM tabanlı modellere ait STH değerlerinin ÇDR ve ÇKA tabanlı modellerin STH değerlerinden daha düşük olduğu gözlemlenmiştir.

*Maximal Oxygen Uptake ( $VO_2max$ ) is the most significant indicator for cardiorespiratory fitness. In this study, Support Vector Machines (SVM) based prediction models have been developed to predict the  $VO_2max$  of 185 healthy subjects to which a submaximal treadmill exercise test has been applied. To form the  $VO_2max$  regression equation, a dataset including 185 test subjects have been utilized. Using 10-fold cross validation on the dataset, standard error of estimates (SEE's) and multiple correlation coefficients (R's) of the models have been calculated. For comparison purposes,  $VO_2max$  prediction models using Multiple Linear Regression (MLR) and Multilayer Perceptron (MLP) have been also developed. In conclusion, it is observed that SVM-based  $VO_2max$  prediction models yield lower SEE's than the ones obtained by using MLR-based and MLP-based prediction models.*

#### ANAHTAR KELİMELER/KEYWORDS

Destek vektör makineleri, Maksimum oksijen tüketimi, Submaksimal egzersiz testi  
Support vector machines, Maximal oxygen uptake, Submaximal exercise test

<sup>1</sup> Bilgisayar Mühendisliği Bölümü Çukurova Üniversitesi ADANA

<sup>2</sup> Dept. of Exercise Sciences, Brigham Young University, Provo, UT 84602, USA

## 1. GİRİŞ

$VO_2max$ , kardiyο respiratuar uygunluęu belirleyen en önemli bileşendir.  $VO_2max$ , egzersiz için çalışma süresini belirlemek, aerobik egzersiz programlarının etkilerini değerlendirmek ve bireylerin saęlık açısından risklerini sınıflandırmak için kullanılır.  $VO_2max$ 'ı belirlemenin standart yolu, koşu bandı veya siklet ergometresi üzerinde bireyin yorulana kadar egzersiz yaptığı süre sonunda  $VO_2max$ 'ın doğrudan ölçülmesidir (maksimal test). Bireye uygulanan egzersiz derecesinde artış olmasına rağmen, oksijen alımında artış gözlenmiyorsa,  $VO_2max$  değerine ulaşılmıştır (George vd., 2009).

Maksimal testler, yönetilmek için riskli olabilir çünkü kişi aşırı yorulduğunda kalp hızı maksimum seviyeye ulaşır. Maksimal testlerin dezavantajlarından dolayı,  $VO_2max$ 'ı belirlemek için maksimum çaba gerektirmeyen diğer yöntemler geliştirilmiştir. Bu yöntemler,  $VO_2max$ 'ı tahmin etmek için egzersize dayalı olmayan ve submaksimal egzersize dayalı verileri içerir. Egzersize dayalı olmayan modeller, maksimal ve submaksimal egzersiz testlerine ihtiyaç duyulmadan  $VO_2max$  tahmini sağlar (William ve Wilkins, 2000).

Submaksimal test,  $VO_2max$ 'ı dolaylı olarak belirler ve genelde koşu bandı, ergometre veya pist üzerinde gerçekleştirilir. Submaksimal egzersiz testini gerçekleştirmek için koşu bandı kullanımı daha yaygındır, çünkü a) Koşu bantları fitness salonları ve laboratuarlarda hazır bir şekilde mevcuttur. b) Yavaş bir tempoda koşu, egzersizin popüler bir şekli ve koşu bantları bir eğitim şeklidir. c) Koşu bandı protokolleri, yönetmede ve kontrol etmede kolaydır. Submaksimal testler maksimal testleri yönetmekten daha ucuz, hızlı ve güvenilirdir. Çizelge 1'de  $VO_2max$ 'ı belirlemek için submaksimal modelleri kullanan bazı önemli çalışmalar özetlenmiştir.

$VO_2max$  tahmin modelinin doğruluęunu etkileyen en önemli faktör uygun araştırma tekniklerinin kullanılmasıdır. Literatürde verilen birçok submaksimal  $VO_2max$  tahmin modeli için akıllı veri analizleri için önemli araçlar sunan makine öğrenme yöntemlerinin kullanılmadığı görülmektedir. Ayrıca, yine bir çok çalışmada çapraz doğrulama yapılmadığı için verilen sonuçların güvenilir olması tartışmalıdır.

Bu çalışmada, submaksimal koşu bandı egzersizi uygulanan saęlıklı yetişkinlerin  $VO_2max$ 'ını tahmin etmek için cinsiyet, yaş, boy, vücut aęırlığı, egzersiz süreleri ve kalp atım hızları deęişkenlerinin kombinasyonlarından oluşan DVM tabanlı 21 farklı model geliştirilmiştir. Veri kümesi, 185 deneęe ait bilgileri içermektedir. 10 katlı çapraz doğrulama kullanılarak, modellerin STH ve KK deęerleri hesaplanmıştır. Karşılaştırma yapabilmek amacı ile ÇDR ve ÇKA yöntemleri kullanılarak  $VO_2max$  tahmin modelleri de geliştirilmiştir.

Çizelge 1. Literatürde  $VO_2max$  tahmini için kullanılan bazı submaksimal regresyon modelleri

Çalışma	Tahmin Deęişkenleri	KK	STH
(Larsen vd., 2002)	G, BM, Egzersiz Süresi, HR	0.90	2.87
(Dalleck vd., 2006)	G, Yaş, HR, BM	0.86	3.91
(McComb vd., 2006)	G, Boy, BM, BF, HR	0.89	3.31
(Vehrs. vd., 2007)	G, Yaş, BM, MPH, HR	0.91	2.52
(Nielson, 2009)	G, BM, PFA, HR	0.91	3.36
(Kale vd., 2013)	Yaş, BM, MPH, HR	0.95	1.80

G, Cinsiyet; BM, vücut aęırlığı; BF, vücut yağ oranı; HR, kalp atış hızı; PFA, algılanan fonksiyonel yetenek; MPH, hız

## 2. EGZERSİZ PROTOKOLÜ VE VERİ SETİ OLUŞTURULMASI

18-26 yaş arası 185 (115 erkek ve 70 kadın) sağlıklı üniversite öğrencisi bu çalışmada yer almaktadır. Egzersiz testinden önce bütün katılımcılara bu araştırmanın klinik önemi ve amacıyla ilgili detaylı bir açıklama verilmiştir.

1.5 mil submaksimal testi kapalı alan pistinde 229.7 metrede tamamlanmıştır. 1.5 mil testi 10.5 (onbuçuk) turun tamamlanmasını gerektirir. Her katılımcının kalp atım hızı sürekli ölçülmüş ve elektronik bir ölçüm cihazı kullanılarak 0.5 mil, 1 mil ve 1.5 mil için değerler kaydedilmiştir. En az 24 saat sonra (fakat 7 günü geçmemek kaydı ile) 1.5 mil test katılımcıları, laboratuardaki maksimal egzersizlerini tamamlamışlardır.

Bütün katılımcılara maksimal testin yapılacağı gecedan önce yeterli uyku (6-8 saat) almaları, testten 3 saat önce ise yemek, kafein, tütün ürünleri ve alkol almamaları konusunda bilgi verilmiştir. Katılımcılardan maksimal test sırasındaki metabolik gaz ölçümünün yapılabilmesi için başlık, ağızlık ve burun klipsi giyinmeleri istenmiştir. Maksimal teste ısınmaları için katılımcılara 2-3 dakikalık bir süre kadar koşu bandı üzerinde yürümeleri ve daha sonra tempolu ve yavaşça koşmaları söylenmiştir. Isınma süresini takiben ise koşu bandının eğimi dakikada % 1.5 artırılmıştır. Bu işlem, katılımcılar maksimum zorlama seviyesine ulaşana ve sözlü cesaretle bile egzersize devam edemeyecek seviyeye gelene kadar devam etmiştir.

VO<sub>2</sub> ve solunum değişim oranı her 15 saniyede hesaplanmış, ortalaması bulunmuş ve online bir bilgisayar sisteminden çıktı alınmıştır. Maksimal test sırasında katılımcıların kalp atışları ve maksimum zorlama seviyeleri her bir bölüm sonunda kayıt edilmiş ve metabolik gaz TrueMax 2400 gaz ölçüm sistemi kullanılarak toplanmıştır. VO<sub>2</sub>max en yüksek ardışık 15 saniye skorlarının ortalaması alınarak hesaplanmıştır.

VO<sub>2</sub>max'ın geçerli olması için, aşağıdakilerden en az ikisinin sağlanması gerekmektedir:

- HRmax'ın (220 – yaş formülü ile verilen tahmini kalp atım hızı) 15 yakınında olması.
- Maksimum solunum oranının (Verilen karbondioksit / Alınan oksijen oranı) 1.1 veya daha fazla olması.
- İş yükündeki artışa rağmen VO<sub>2</sub>'de artış olmaması.

Çalışmada kullanılan; yaş, cinsiyet, vücut ağırlığı, boy, MIN1, MIN2, MIN3 (0.5 mil, 1 mil ve 1.5 milde egzersiz süreleri) ve HR1, HR2, HR3 (0.5 mil, 1 mil ve 1.5 milde kalp atım hızları) bağımsız değişkenlerini içermektedir.

## 3. MATERYAL VE YÖNTEM

### 3.1. DVM Modeli

DVM modelinin kalitesini ve performansını etkileyen en önemli bileşenler çekirdek fonksiyonunun tipi, parametreleri, C değeri ve ε değeridir. C parametresi hata ve karmaşıklık arasındaki ilişkiyi belirler. C'nin büyük değerleri için DVM oldukça az hata yapmaya çalışırken, karmaşıklık artar. C'nin küçük değerlerinde DVM daha çok hata yapabilirken daha basit bir model ortaya çıkarır.

ε parametresi, ε-duyarsız bölgenin büyüklüğünü kontrol eder. Buna ek olarak destek vektörlerinin sayısını belirler. ε'nin küçük değerleri daha fazla destek vektörlerine neden olur. Hangi C, ε ve gamma (γ) değerlerinin verilen problem için en iyi olacağı önceden bilinemez. Temel amaç bu parametrelerin değerlerini optimize ederek regresyon modelinin en az hatayı bulmasını sağlamaktır. Bu optimal parametrelerin seçimi için bir çok yöntem önerilmiştir. Örneğin; çapraz doğrulama, ızgara arama, genetik algoritmalar vb. parametrelerin optimal

değerlerini bulmaya çalışır (Hsu vd., 2003; Friedrichs ve Igel, 2005). k sayıdaki bir çapraz-doğrulamada orijinal veri seti k tane alt kümeye bölünür. İşlem k defa tekrarlanır ve daha sonra elde edilen performans değerlerinin ortalaması alınır.

Bu çalışmada, DVM tabanlı modellerin geliştirilmesinde kernel olarak radial tabanlı fonksiyon seçilmiştir. Modellerin C,  $\epsilon$  and  $\gamma$  değerlerini belirlemek için 5 katlı çapraz doğrulama kullanılmıştır.

### 3.2. ÇKA Modeli

ÇKA tabanlı modellerin tamamında gizli katmanda tansigmoid tanjant aktivasyon fonksiyonu ve çıkış katmanında ise doğrusal aktivasyon fonksiyonu kullanılmıştır. Ağ, Levenberg-Marquart algoritması kullanılarak eğitilmiştir. Giriş ve çıkışların değerleri normalize edilmiştir. ÇKA'nin diğer önemli parametreleri olan iterasyon sayısı 500 ve öğrenme oranı 0.5 olarak seçilmiştir. Bu değerler deneme yanılma ile bulunmuştur.

## 4. BULGULAR VE TARTIŞMA

DVM, ÇDR ve ÇKA modellerinin performansı, 10-katlı çapraz doğrulama kullanılarak, STH ve KK hesaplanarak değerlendirilmiştir. STH ve KK'nin formülleri sırasıyla Eşitlik 1 ve Eşitlik 2'de verilmiştir.

$$STH = \sigma_{Y'} \sqrt{1 - r^2} \quad (1)$$

$$KK = \sqrt{\frac{\sum(Y-Y')^2}{\sum(Y-\bar{Y})}} \quad (2)$$

Eşitlik 1 ve Eşitlik 2'de, ölçülen  $VO_2max$ 'ın değeri, tahmin edilen  $VO_2max$ 'ın değeri, ölçülen  $VO_2max$ 'ın ortalaması, tahmin edilen  $VO_2max$ 'ın ortalama değeri,  $VO_2max$ 'ın tahmin edilen değerlerinin sapması ve r Pearson korelasyon katsayıdır.

DVM, ÇDR ve ÇKA tabanlı  $VO_2max$  tahmin modellerinin her bir katı için STH ve KK değeri hesaplanmıştır ve her bir model için ortalama değerler Çizelge 2, Çizelge 3 ve Çizelge 4' te gösterilmiştir.

Elde edilen sonuçlara göre ulaşılan noktalar aşağıda maddeler halinde listelenmiştir.

1. DVM tabanlı modeller ÇKA tabanlı modellere göre, ÇKA tabanlı modeller de ÇDR tabanlı modellere göre daha düşük STH değerleri üretmiştir.

2. MIN değişkenlerini içeren  $VO_2max$  tahmin modellerinin HR değişkenlerini içeren modellere göre daha düşük STH ürettiği gözlemlenmiştir.

3. Yalnızca MIN değişkenlerini içeren modeller arasında, MIN3 değişkenini içeren modelin (model no: 3) en düşük STH değerini ürettiği gözlemlenmiştir.

4. Yalnızca HR değişkenlerini içeren modeller arasında, HR2 değişkenini içeren modelin (model no: 9) en düşük STH değerini ürettiği gözlemlenmiştir.

5. MIN ve HR değişkenlerinin ikili kombinasyonlarından oluşan modellerde, 20. modelin (yaş, cinsiyet, boy, vücut ağırlığı, MIN2, HR2, MIN3 ve HR3) en düşük STH değerini ürettiği gözlemlenmiştir.

6. En düşük STH değerini üreten modelin 17.model (yaş, cinsiyet, boy, vücut ağırlığı, MIN3, HR3) olduğu gözlemlenmiştir.

7. En yüksek STH değerini üreten modelin ise 14. model (yaş, cinsiyet, boy, vücut ağırlığı, HR1, HR2, HR3) olduğu gözlemlenmiştir.

Çizelge 2. MIN değişkenleri içeren modellere ait sonuçlar

No	Model Değişkenleri	ÇDR		ÇKA		DVM	
		KK	STH	R	STH	KK	STH
1	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1	0.77	3.89	0.79	3.65	0.81	3.46
2	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN2	0.78	3.81	0.80	3.55	0.83	3.31
3	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN3	0.80	3.62	0.82	3.38	0.83	3.28
4	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, MIN2	0.77	3.97	0.79	3.70	0.81	3.46
5	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, MIN3	0.78	3.83	0.80	3.55	0.83	3.32
6	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN2, MIN3	0.77	3.88	0.79	3.61	0.83	3.33
7	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, MIN2, MIN3	0.75	4.11	0.78	3.78	0.81	3.49

Çizelge 3. HR değişkenlerini içeren modellere ait sonuçlar

No	Model Değişkenleri	ÇDR		ÇKA		DVM	
		KK	STH	KK	STH	KK	STH
8	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR1	0.66	4.55	0.71	3.73	0.73	4.03
9	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR2	0.68	4.47	0.71	3.40	0.73	4.03
10	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR3	0.66	4.55	0.70	3.13	0.73	4.07
11	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR1, HR2	0.66	4.57	0.71	4.35	0.72	4.08
12	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR1, HR3	0.63	4.79	0.69	4.52	0.71	4.13
13	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR2, HR3	0.64	4.78	0.70	4.38	0.72	4.13
14	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, HR1, HR2, HR3	0.60	5.09	0.67	4.71	0.72	4.15

Çizelge 4. MIN ve HR değişkenlerini içeren modellere ait sonuçlar

No	Model Değişkenleri	ÇDR		ÇKA		DVM	
		KK	STH	KK	STH	KK	STH
15	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, HR1	0.75	4.08	0.79	3.73	0.81	3.46
16	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN2, HR2	0.80	3.62	0.80	3.40	0.83	3.29
17	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN3, HR3	0.84	3.25	0.85	3.13	0.87	2.90
18	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, HR1, MIN2, HR2	0.86	4.12	0.78	3.84	0.80	3.49
19	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, HR1, MIN3, HR3	0.80	3.84	0.81	3.58	0.83	3.32
20	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN2, HR2, MIN3, HR3	0.82	3.52	0.84	3.26	0.87	2.96
21	Yaş, Cinsiyet, Boy, Vücut Ağırlığı, MIN1, HR1, MIN2, HR2, MIN3, HR3	0.75	4.30	0.78	3.94	0.80	3.57

## 5. SONUÇLAR

Bu çalışmada submaksimal değişkenler aracılığı ile  $VO_2max$  tahmini yapabilmek için DVM, ÇDR ve ÇKA yöntemleri kullanılmıştır. Veri kümesinde bulunan değişkenlerin (cinsiyet, yaş, boy, vücut ağırlığı, MIN1, HR1, MIN2, HR2, MIN3 ve HR3) kombinasyonlarından oluşan 21 farklı tahmin modelinin performansı KK ve STH değerleri hesaplanarak karşılaştırılmıştır. En iyi  $VO_2max$  tahmin modeli yaş, cinsiyet, boy, vücut ağırlığı, MIN3 ve HR3 değişkenlerinin bulunduğu modeldir. DVM tabanlı tahmin modellerinin ÇDR ve ÇKA tabanlı tahmin modellerine göre daha az hatalı sonuçlar verdiği gösterilmiştir. Bu nedenle DVM tabanlı modeller  $VO_2max$ 'ın tahminin de geçerli bir yöntem olabilmektedir.

## TEŞEKKÜR

Bu çalışma Çukurova Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından desteklenmiştir (Proje No: MMF2013D17 ve Proje No: FYL-2014-1961).

## KAYNAKÇA

- Dalleck L. C., Kravitz L., Robergs R. A. (2006): "Development of a Submaximal Test to Predict Elliptical Cross-Trainer  $VO_2max$ ", National Strength and Conditioning Association, Cilt 20, No. 2, s.278-283.
- Friedrichs F., Igel C. (2005): "Evolutionary Tuning of Multiple SVM Parameters", Neurocomputing, Cilt 64, s.107-117.
- George J. D., Paul S. L., Hyde A., Bradshaw D. I., Vehrs P. R., Hager R. L. (2009): "Prediction of Maximum Oxygen Uptake Using Both Exercise and Non-Exercise Data", Measurement in Physical Education and Exercise Science, Cilt 13, No. 1, s.1-12.

- Hsu C. W., Chang C. C., Lin C. J. (2003): “A Practical Guide to Support Vector Classification”, <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- Kale G., Akay, M. F., Aktürk E., Tunçdemir A. E. (2013): “Yapay Sinir Ağları Kullanılarak Egzersiz ve Egzersize Dayalı Olmayan Verilerden VO<sub>2</sub>max tahmini”, 21. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SİU2013, 24-26 Nisan 2013, KKTC.
- Larsen G. E., George J. D., Alexander J. L., Fellingham G. W., Aldana S. G., Parcell A. C., (2002): Prediction of Maximum Oxygen Consumption from Walking, Jogging, or Running”, *Research Quarterly for Exercise and Sport*, Cilt 71, No. 1, s.66–72.
- McComb J. J. R., Roh D., Williams J. S. (2006): “Explanatory Variance in Maximal Oxygen Uptake”, *Journal of Sports Science and Medicine*, Cilt 5, No. 2, s.296–303.
- Nielson E. D. (2009): “Predicting VO<sub>2</sub>max in College-Aged Participants Using Cycle Ergometry and Nonexercise Measures”, Department of Exercise Sciences Brigham Young University, Master of Science, December.
- Vehrs P. R., George J. D., Fellingham G. W., Plowman S. A., Dustman-Allen K. (2007): “Submaximal Treadmill Exercise Test to Predict VO<sub>2</sub>max in Fit Adults”, *Measurement in Physical Education and Exercise Science*, Cilt 11, No. 2, s.61–72.
- William L., Wilkins (2000): “ACSM’s Guideline for Exercise Testing and Prescription (6<sup>th</sup> ed.)”, American College of Sports Medicine, Philadelphia.