



# Nöromusküler Hastalıkların Teşhisinde SVM, K-EYK ve DA Algoritma Tabanlı Sınıflandırma Bölüm-II

## Neuromuscular Disease Diagnosis of SVM, K-NN and DA Algorithm Based Classification Part-II

Hanife Küçük<sup>1</sup>, İlyas Eminoğlu<sup>2</sup>

<sup>1,2</sup>Elektrik ve Elektronik Mühendisliği Bölümü, Biyomedikal Araştırma Lisansüstü Laboratuvarı: (BAL-LAB)  
Ondokuz Mayıs Üniversitesi, Samsun, Türkiye  
hanifesms@gmail.com, ilyaseminoglu@hotmail.com

### I. GİRİŞ

**Özetçe**—Bu çalışma nöromusküler hastalıkların en yaygın çeşidi olan ALS ve myopati hastalıklarının teşhisi için beş işlem aşamasından oluşan bir sınıflandırma yapısının ikinci bölümünü içermektedir. Bu çalışmada zaman, frekans uzayındaki toplam 14 öznelik vektörü ve bunların farklı birleşimleri ile elde edilen çoklu öznelik vektörleri denenmiştir. Sınıflandırma aşamasında ise Destek Vektör Makinesi (DVM), K-En Yakın Komşu (K-EYK) algoritması ve Diskriminant Analiz (DA) algoritmaları denenmiştir. Deneysel sonuçlar önerilen öznelik vektörleri arasında çoklu öznelik vektörlerinin tekli öznelik vektörlerine göre daha başarılı olduğunu göstermiştir. Bu çalışma ile sınıflandırma başarımının, öznelik vektörünün sınıflar arası ayrılabilirliğinin yüksek derecede olmasına bağlı olduğu anlaşılmıştır.

**Anahtar Kelimeler** — EMG; ALS; DVM; K-EYK; DA

**Abstract**— This study includes a classification structure consisting of second part for the automatic diagnosis of the neuromuscular disease of ALS (Amyotrophic Lateral Sclerosis) and myopathy being a muscular disease. In this study feature vectors containing time domain parameters, frequency domain parameters (a total of 25 feature vectors) as well as feature vectors composed of combination of these parameters were used. In the classification stage, Support Vector Machines (SVM), K-Nearest Neighbors (K-NN) and Discriminant Analysis (DA) algorithms were employed. Experimental results showed that the multiple feature vectors proved to be more successful compared to the individual feature vectors. It is understood with this study; the classification performance depends highly on separability of feature vectors between different classes.

**Keywords** — EMG; ALS; DVM; K-NN; DA

Kaynaklarda biyomedikal uygulamalarda yorumlama ve sınıflandırma problemlerinin çözümleri birçok sinyal işleme, yapay zeka (yapay sinir ağları, vb.) metotları ile elde edilmeye çalışılmaktadır. 1970'li yıllarda yapay zeka tekniklerinin biyomedikal alanındaki kullanımı artmış ve o günden sonra gelişimi hızlanmıştır [1].

Rutin klinik teşhislerde EMG işaretlerinin analizine duyulan ihtiyaç ise otomasyon ve bilgisayar tekniklerinin kullanımını zorunlu kılmıştır. Bu nedenle objektif bir değerlendirme yapabilmek için EMG işaretlerinin farklı yöntemlerle analiz edilmesi yoluna gidilmiştir. Son yıllarda geliştirilen bu yöntemlerin bazıları işaretlerin istatistiksel analizinin yapılmasında, diğerleri ise özel dalga şekillerinin belirlenip, ayrıştırılıp incelenmesinde kullanılmaktadır. Bu yöntemlere örnek olarak; zaman analizi, frekans analizi, zaman-frekans analizi yöntemi ve dalga şekillerinin ayrıştırılması ve sınıflandırılması sayılabilir [2].

1979 yılında De Luca tarafından yazılan ve bu konuda bir ilk olan EMG içerik bilgisi ile ilgili makalesi EMG işaretlerinin fizyolojisi ve matematiksel metotlarını birleştiren bir klasik sayılmaktadır. Sonrasında birçok araştırmacı bilgisayar tabanlı EMG analiz algoritması geliştirmiştir [3].

Son zamanlarda yapılan çalışmalara baktığımızda işaret işleme teknikleri ve yapay zeka tekniklerinin gelişmesine paralel olarak bilgisayar tabanlı EMG analiz algoritmaları gelişmiştir.

Tablo 1' de nöromusküler bozuklukların tanısı için EMG sinyalleri sınıflandırma alanında literatürde bulunan bir kısım çalışmaların yayın yılı, kullanılan yöntemler, MUAP sayısı ve başarımları verilerek bir özet sunulmuştur.



Kullanılan Yöntemler		Denek Sayısı (MUAP Sayısı)	Doğruluk (%)
1999 [4]	Öz Denetleyici Harita (ÖDH), Vektör Nicelendirmeli Öğrenme (VNÖ) ve Öklid uzaklığına dayalı istatistiksel yöntemler	(1213)	90
2006 [5]	Dalgacık Tabanlı Sinir Ağı (DTSA)	(1200)	90.7
2012 [6]	ÖDH ve VNÖ	11	97.6
2006 [7]	Uyarlamalı Bulanık K-EYK Sınıflandırıcı (UBKS)	11	96.6
2005 [8]	DVM	59	92.3
2010 [9]	İkili DVM	12	100
2012 [10]	Dalgacık Tekniği ile Öznitelik Çıkarımı ve DVM	300	99.4
2012 [11]	Ayrık dalgacık dönüşümü ile öznitelik çıkarma ve Bulanık Destek Vektör Makinesi (BDVM)	27	97.67

**Tablo 1.** Nöromusküler Bozuklukların Tanısı için EMG Sinyallerinin Sınıflandırılması Alanında Yapılan Bazı Çalışmalar [12]

Bu çalışmada Amiyotrofik Lateral Skleroz (ALS) ve myopati hastalığının teşhisinde kullanılan öznitelik vektörleri zaman uzayı ve frekans uzayında hesaplanmıştır. Daha sonra öznitelik vektörleri sırasıyla Destek Vektör Makinesi, K-En Yakın Komşu algoritması ve Diskriminant Analiz algoritmalarına giriş olarak verilerek verilerin sınıflandırılması sağlanmıştır.

## II. MATERYAL VE YÖNTEM

### A. EMG Verisinde Öznitelik Çıkarma

EMG verisinden gizlenmiş faydalı bilgileri çıkarmak ve verinin içinden istenmeyen bölümleri, girişimleri kaldırmak adına EMG verilerinde özellik çıkarımı önemli bir aşamadır [13], [14]. Bu çalışmada Tablo 2' de verilen zaman uzayı ve frekans uzayındaki toplam 14 tane öznitelik vektörleri her bir EMG verisi için hesaplanmıştır. Yüksek sınıflandırma performansı elde etmek için etkili öznitelik vektörü kullanmak gerekir [15]. Bu amaç doğrultusunda sınıflandırma doğruluğunu arttırmak için öznitelik vektörlerinin çeşitli şekilde birleştirilmesi ile elde edilen çoklu öznitelik vektörleri kullanılmıştır.

### B. Destek Vektör Makinesi (DVM) İle Sınıflandırma

Destek vektör makinesi olarak adlandırılan yaklaşım istatistiksel bir öğrenme algoritmasıdır. DVM' nin temel amacı iki sınıf arasında en uygun ayrımı yapabilen hiperdüzlemi tanımlamaktır [16].

Çoğu gerçek problemlerde verilerin doğrusal bir ayırım fonksiyonu ile hatasız ayrılabilmesi söz konusu değildir. DVM de doğrusal olarak ayrılamayan her bir giriş vektörü  $\Phi$ : dönüşüm fonksiyonu ya da koordinat dönüşüm fonksiyonu ile yüksek boyutlu bir uzaya haritalanarak doğrusal ayrılabilir duruma gelmektedir. DVM optimizasyon tabanlı bir eğitim algoritması kullanarak öğrenebilen sistemlerdir.

### C. K-En Yakın Komşu (K-EYK) Algoritması

K-En Yakın Komşu (K-EYK) algoritması öznitelik uzayındaki en yakın eğitim örneklerine dayanarak nesnelere sınıflandıran, en basit örüntü tanıma yöntemlerinden birisidir [17]. Bu algoritma verilen K

değeri kadar en yakın komşunun sınıfına göre sınıflandırma işlemi yapmaktadır.

Zaman Uzayındaki Öznitelik Vektörleri		Frekans Uzayındaki Öznitelik Vektörleri
<i>Standart Sapma</i> $STD = \sqrt{VAR}$	<i>Ort Mutlak Değer</i> $MAV = \frac{1}{N} \sum_{i=1}^N  x_i $	<i>Tepe Frekans</i> $TFRE = \max(P_j), j=1, \dots, M$
<i>EMG Varyans</i> $VAR = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$	<i>Tümleşik EMG</i> $IEMG = \sum_{i=1}^N  x_i $	<i>Ortalama Frekans</i> $MF = \frac{\sum_{i=1}^M f_i P_i}{\sum_{i=1}^M P_i}$
<i>Sıfır Geçiş Sayısı (SGS)</i> $\{x_i > 0 \text{ ve } x_{i+1} < 0\}$ veya $\{x_i < 0 \text{ ve } x_{i+1} > 0\}$ ve $ x_i - x_{i+1}  \geq \text{esik}$		<i>Ortalama Güç</i> $MNP = \frac{OCF}{M} \sum_{i=1}^M P_i$
<i>Etkin Değer</i> $RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$		<i>Ortanca Frekans (MDF)</i> $\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^M P_j = \frac{1}{2} \sum_{j=1}^M P_j$
<i>Willison Genliği</i> $WAMP = \sum_{i=1}^{N-1} [f( x_n - x_{n+1} )]$ $f(x) = \begin{cases} 1, & x \geq \text{esik} \\ 0, & \text{aksi durumda} \end{cases}$		<i>1., 2. ve 3. Dereceden Spektral Hareket</i> $SM1 = \sum_{j=1}^M P_j f_j$ $SM2 = \sum_{j=1}^M P_j f_j^2$ $SM3 = \sum_{j=1}^M P_j f_j^3$

**Tablo 2.** Zaman ve Frekans Uzayında Öznitelik Vektörleri

### D. Diskriminant Analiz

Bir sınıflandırma yöntemi olan diskriminant analiz (DA) farklı sınıfların farklı Gaussian dağılımına sahip oldukları ilkesine dayanmaktadır [18]. Doğrusal Diskriminant Analiz (LDA) ve Quadratik Diskriminant Analiz (QDA) olarak iki temel gruba ayrılan diskriminant



## Sinyal İşleme 3

2. Gün / 28 Ekim 2016, Cuma

analizinin, ayırma ve sınıflama şeklinde iki işlevi bulunmaktadır.

### III. TARTIŞMA

Bu çalışmada kullanılan EMG işaretleri 10 normal, 7 Myopati ve 8 ALS hastası olan her bir veri grubunda 8 sn süren 230' ar tane veri bulunmaktadır. Sınıflandırıcı başarımını belirlemek için çapraz geçerlilik yöntemi ile toplam 3x230 tane veri seti rastgele olarak "k" gruba ayrılmıştır.

DVM ile sınıflandırmada Matlab'ın BioinformaticsToolbox™ (MathWorks, 2007) adlı yazılımı kullanılmıştır. DVM ile yapılan sınıflandırmada kernel fonksiyonu olarak seçilen RTF (Radyal Tabanlı Fonksiyon)'nin Gauss genişliği:  $r=1.5$  ve ceza faktörü:  $C=400$  seçilmiştir.

Sınıflandırma başarımının üstünlüğünün arkasında yatan ana neden öznitelik vektörlerinin sınıflar arası ayrılabilirliğinin yüksek derecede olmasıdır [19].

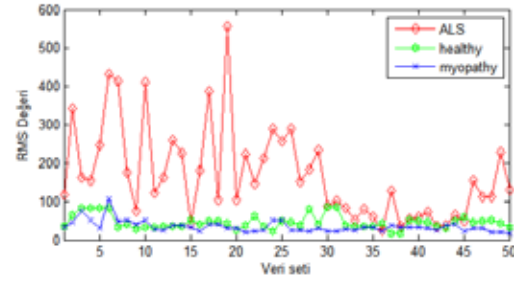
Deneysel sonuçlar bu çalışmanın sınıflandırma yeteneğinin sadece iki sınıflı problemleri ayırabildiğini göstermiştir. Bu duruma en iyi örnek Şekil 1' de verilmiştir. Grafikte her üç sınıftaki 50 veriye ait MUAP'ların RMS öznitelik vektörlerinin ortalama değerleri çizdirilmiştir. Burada ALS hastalığına karşılık gelen RMS özniteliklerinin ortalama değerleri diğer iki sınıfa (Myopati, sağlıklı) göre çok daha yüksek olmasına rağmen, RMS öznitelik vektörünün Myopati ile sağlıklı verileri ayırma derecesi çok tatmin edici değildir. Bu çalışmada kullanılan öznitelik vektörlerinin sadece sağlıklı veri ile ALS verilerini ya da ALS ile myopati verileri arasında daha iyi bir sınıflandırma doğruluğunu mümkün kılmaktadır.

Bu durum diğer tekli ve çoklu öznitelik vektörleri içinde geçerli olması nedeniyle sınıflandırma başarımları Tablo 4' de ikili sınıflandırma olarak sunulmuştur. Sonuçlar göstermiştir ki önerilen özellikler sadece iki sınıfın (ALS-Sağlıklı ve ALS-Myopati) yüksek derecede ayrılabilirliğini sağlamıştır.

Bu çalışmada elde edilen sınıflandırma sonuçlarını daha önce yapılan farklı tip veri setini kullanan çalışmalarla kıyaslamak yerine; ortak bir veri tabanını kullanan [20] ve [21] çalışmaları ile karşılaştırmak daha doğru olacaktır. 2015 yılında yayınlanan bu iki çalışmada veriler (<http://www.EMGLab.net>, 2011) internet adresindeki veri tabanından alınmıştır.

Ancak bu iki makalede; bu çalışmadaki veri setine göre daha az veri seti ile işlem yapılmıştır. Bu durumlar göz önünde bulundurulduğunda bu çalışmada elde edilen başarımların aynı veri setini kullanan çalışmalara göre daha iyi olduğu görülmüştür (Tablo 3). Bu tez çalışmasında ALS-Sağlıklı, ALS-Myopati ve Myopati-Sağlıklı sınıflarının başarımlarını değerlendirmek amacıyla bazı istatistiksel ölçütler Tablo 4' de hesaplanmıştır.

Tablo 4' de öznitelik vektörü IEMG+SGS+WAMP+MF seçilmesi durumunda her bir sınıflandırma algoritması için yukarıda anlatılan ölçütler hesaplanmıştır. ALS-Sağlıklı ve ALS-Myopati ikili sınıflarda DVM-RTF sırasıyla %97.39, %94.57 en iyi başarımları sağlarken; Myopati-Sağlıklı ikili sınıfta en iyi başarımları %86.74 ile DVM-Doğrusal ve DDA sağlamıştır.



Şekil 1. ALS, Myopati ve Sağlıklı kişilere ait veriler RMS öznitelik vektörlerinin ortalama değer değişimi

### IV. SONUÇLAR

Bu çalışmada önerilen öznitelik vektörleri arasında çoklu öznitelik vektörlerinin tekli öznitelik vektörlerine göre daha başarılı olduğunu göstermiştir. Bu çalışma ile sınıflandırma başarımının öznitelik vektörünün sınıflar arası ayrılabilirliğinin yüksek derecede olmasına bağlı olduğu anlaşılmıştır. Bu amaç doğrultusunda IEMG+SGS+WAMP+MF çoklu öznitelik vektörünün kullanılması durumunda ALS-Sağlıklı ve ALS-Myopati ikili sınıflarda DVM-RTF sırasıyla %97.39, %94.57 en iyi başarımları sağlarken; Myopati-Sağlıklı ikili sınıfta en iyi başarımları %86.74 ile DVM-Doğrusal ve DDA sağlamıştır.

Çalışmalar	Kullanılan Yöntemler		Doğruluk (%)
	Öznitelik Çıkarma	Sınıflandırma	
[20]	ADD	K-EYK	ALS-Sağlıklı; %96.5 Myopati-Sağlıklı; %83.5
[21]	ADD	DVM ve YSA	ALS için DVM; %98.0 Myopati için YSA; %86.6
Yazarın önerisi	Çoklu Öznitelik Vektörü	DVM	ALS-Sağlıklı; %97.39 Myopati-Sağlıklı; %86.74

Tablo 3. Aynı Veri Setini Kullanan Çalışmaların Karşılaştırılması



ALS ve SAĞLIKLI (Özellik Vektörü: IEMG+SGS+WAMP+MF)							
Sınıflandırıcı	DP	YP	YN	DN	Duyarlılık (%)	Özgüllük (%)	Toplam Doğruluk (%)
DVM(Doğrusal)	203	23	27	207	88.26	90.00	89.13
DVM(RTF)	224	6	6	224	97.39	97.39	97.39
K-EYK(K=1)	198	21	32	209	86.09	90.87	88.48
K-EYK(K=11)	198	14	32	216	86.09	93.91	90.00
DDA	203	23	27	207	88.26	90.00	89.13
KDA	198	5	32	225	86.09	97.83	91.96
MİYOPATİ ve SAĞLIKLI (Özellik Vektörü: IEMG+SGS+WAMP+MF)							
Sınıflandırıcı	DP	YPz	YN	DN	Duyarlılık (%)	Özgüllük (%)	Toplam Doğruluk (%)
DVM(Doğrusal)	186	17	44	213	80.87	92.61	86.74
DVM(RTF)	189	31	41	199	82.17	86.52	84.35
K-EYK(K=1)	185	36	45	194	80.43	84.35	82.39
K-EYK(K=11)	193	26	37	204	83.91	88.70	87.39
DDA	186	17	44	213	80.87	92.61	86.74
KDA	186	37	44	193	83.91	80.87	82.39
ALS ve MİYOPATİ (Özellik Vektörü: IEMG+SGS+WAMP+MF)							
Sınıflandırıcı	DP	YP	YN	DN	Duyarlılık (%)	Özgüllük (%)	Toplam Doğruluk (%)
DVM(Doğrusal)	210	29	20	201	91.30	87.39	89.35
DVM(RTF)	214	9	16	221	93.04	96.09	94.57
K-EYK(K=1)	200	31	30	199	86.96	86.52	86.74
K-EYK(K=11)	198	31	32	199	86.09	86.52	86.30
DDA	210	29	20	201	91.30	87.39	89.35
KDA	200	9	30	221	86.96	96.09	91.52

DP: Doğru Pozitif      YP: Yanlış Pozitif      YN: Yanlış Negatif      DN: Doğru Negatif

**Tablo 4.** Performans Değerlendirme

#### KAYNAKÇA

- [1] Xiang, Y., 1993. Multiply sectioned Bayesian networks for neuromuscular diagnosis, *Intell. Med.*, 5, 293.
- [2] Bozkurt, M. R., 2007. EMG İşaretlerinin Modern Yöntemlerle Önışlemesi ve Sınıflandırılması, Doktora Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 216254.
- [3] Stalberg, E., Falck, B., Lang, H., Rosenfalck, A., Trojaborg, W., 1986. Quantitative Analysis Of Individual Motor Unit Potentials: A Proposition For Standardized Terminology And Criteria For Measurement, *J. Clinic Neurophysiology*, 3: 313-348 .
- [4] Christodoulou, C., Pattichis, C., 1999. Unsupervised pattern recognition for the classification of EMG signals, *IEEE Trans. Biomed. Eng.*, 46, 169.
- [5] Subasi, A., Yilmaz, M., Ozcalik, H., 2006. Classification of EMG signals using wavelet neural network, *J. Neurosci. Methods*, 156 (1-2), 360-367.
- [6] Bhardwaj, A., Manish, A., Arora, K., 2012. Classification of MUAPs by using ANN pattern recognition technique, *Int. J. Sci. Res. Eng. Technol. (IJSRET)*, 1:10-14.
- [7] Rasheed, S., Stashuk, D., Kamel, M., 2006. Adaptive fuzzy k-NN classifier for EMG signal decomposition, *Med. Eng., Phys.* 28, 694.
- [8] Guler, N. F., Koçer, S., 2005. Classification of EMG signals using PCA and FFT, *J. Neurosci. Methods*, 29 (3), 241-250.
- [9] Kaur, G., Arora, A., Jain, V., 2010. EMG diagnosis via ar modelling and binary support vector machine classification, *Int. J. Eng. Sci. Technol.*, 2, 1767-1772.
- [10] Dobrowolski, A., Wierzbowski, M., Tomczykiewicz, K., 2012. Multiresolution MUAPs decomposition and DVM-based analysis in the classification of neuromuscular disorders, *Comput. Methods Prog. Biomed.*, 107 (3), 393-403.
- [11] Subasi, A., 2012. Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines, *Comput. Biol. Med.*, 42, 806-815
- [12] Yousefi, J., Hamilton, A., 2014. Characterizing EMG data using machine-learning tools, *Computers in biology and medicine* 51, 1-13.
- [13] MathWorks, Matlab Toolbox of Bioinformatics: User's Guide. Mathworks Inc, 2007..
- [14] Fisher, 1936. R. A. The Use of Multiple Measurements in Taxonomic Problems..
- [15] Weinstein S, Obuchowski NA, Lieber ML. 2005. Clinical evaluation of diagnostic tests. *AJR Am J Roentgenol.* 184,14-19.
- [16] Vapnik, V.N., "The Nature of Statistical Learning Theory", Second Edition, Springer-Verlag, New York, 2000.
- [17] N Güler, Antenatal Fetüs Verilerinin Örüntü Tanıma Yöntemleriyle Tanımlanması, Doktora Tezi. İstanbul: Yıldız teknik üniversitesi fen bilimleri enstitüsü, Matematik müh. Anabilim dalı, 2002.
- [18] Macit, E G. Çok Kriterli Sıralı Sınıflandırma Problemi Ve Sağlık Sektöründe Bir Uygulama. Y. Lisans Tezi, Erciyes Ün. , 2014.
- [19] Fattah, Shaikh Anowarul, et al. "Evaluation of Different Time and Frequency Domain Features of Motor Neuron and Musculoskeletal Diseases." *International Journal of Computer Applications* 43.23 (2012): 34-40.
- [20] Krishna V, A., Thomas P., 2005. Classification of EMG Signals Using Spectral Features Extracted from Dominant Motor Unit Action Potential, *International Journal of Engineering and Advanced Technology (IJEAT)*, ISSN: 2249 - 8958, Volume-4 Issue-5.
- [21] Maboud N., F., A., Elbagoury M., B., Roushdy M., Salem M., A., 2015. Hybrid Classifier for Neuromuscular Disorders Diagnoses, *Egyptian Computer Science Journal*, Vol. 39 No. 1.