



AR Model Katsayıları ve Makine Öğrenme Teknikleri Kullanarak Epileptik Nöbet Tahmini

Epileptic Seizure Forecasting Using AR Model Coefficients and Machine Learning Techniques

Muharrem Çelebi¹, M. Kemal Güllü²

^{1,2} Elektronik ve Haberleşme Mühendisliği Bölümü, Kocaeli Üniversitesi, Kocaeli, Türkiye
muharrem.celebi@kocaeli.edu.tr; kemalg@kocaeli.edu.tr

Özetçe — Epilepsi, insan yaşamını olumsuz etkileyen nörolojik bir rahatsızlıktır. Bu çalışmada, epileptik nöbet tahminini gerçekleştirmek için AR model katsayıları ve makine öğrenmesi tekniklerinden yararlanılarak, EEG işaretlerinde epileptik nöbet tahmini çalışması gerçekleştirilmiştir. Bu amaç için, AR model katsayıları EEG işaretlerinden çıkarılır ve öznelik olarak kullanılır. Ardından elde edilen bu öznelikler, üç farklı sınıflandırıcıya uygulanmış ve başarımları sonuçları elde edilmiştir. Amerikan Epilepsi Derneği'nin veri seti üzerinde testler gerçekleştirilmiştir. Elde edilen bulgular, AR model katsayılarının epileptik nöbet tahmini çalışmaları için uygun olduğunu göstermektedir.

Anahtar Kelimeler — Epilepsi; EEG; AR model katsayıları; makine öğrenmesi.

Abstract — Epilepsy is a neurological disorder that affects human life negatively. In this study, epileptic seizure prediction study on EEG signals is performed by using AR model coefficients and machine learning techniques in order to estimate the epileptic seizure. For this purpose, AR model coefficients are extracted from the EEG signals and used as features. Then, these features are applied to three different classifiers and the performance results are obtained. Experiments are conducted on the American Epilepsy Society dataset. The results show that AR model coefficients are suitable features for epileptic seizure prediction studies.

Keywords — Epilepsy; EEG; AR model coefficients; machine learning.

I. GİRİŞ

Epilepsi, dünya nüfusunun yaklaşık %1'ini etkileyen nörolojik bir hastalıktır [1]. Epileptik kriz, beyin hücrelerinin ani olarak elektriksel boşalmaları sonucunda ortaya çıkan ve kısa süreli olarak hastanın bilincini kaybetmesine neden olmaktadır [2]. Epilepsi krizi esnasında, hastanın yere

düşmesi, bilincini kaybetmesi, çevresine zarar vermesi, dişlerini veya dilini ısırması, ani kas kasmaları şeklinde davranışlar gözlemlenir. Bazı nöbetler, hastanın ayakta olduğu durumda gerçekleşmekte, bazı nöbetlerde ise hasta yere düşebilmektedir.

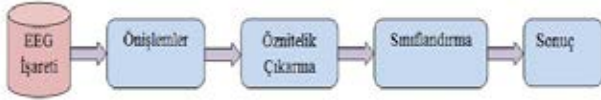
Hans Berger [3], 1929 yılında, ilk defa insan beyni ile ilgili elektriksel kayıtları elde etmiştir. Kafatası üzerine yerleştirilen, elektrotlar yardımıyla elde Elektroensefalografi (EEG) işaretleri insan beynindeki elektriksel değişimler hakkında bilgi sağlamaktadır. Epilepsi ile ilgili çalışmalar, EEG işaretinin dışında, Elektrokardiyografi (EKG) ve Elektromiyografi (EMG) gibi biyolojik işaretler de tercih edilmektedir. Fakat, epilepsi nöbetleri, insan beyninde gerçekleştiği için, EEG işaretleri temel biyolojik işaretleri oluşturmaktadır. EEG işareti, gürültü seviyesi yüksek, bozucu etkileri ise fazla olan durağan olmayan, genliği ve frekansı sürekli olarak değişen bir işarettir. Bu nedenden dolayı, EEG ile ilgili çalışmalarda, anlamlı bir sonuç elde edebilmek için uzun süreli veri kayıtlarına ihtiyaç duyulmaktadır.

Uluslararası standart 10-20 sistem genellikle anlık EEG kaydı için kullanılır. 21 elektrot kafatasının yüzeyine yerleştirilmiştir [4]. EEG elektrotların kafatası üzerine yerleştirilirken, yüzeysel elektrotlar (scalp EEG veya surface EEG, sEEG) ve kafaiçi elektrotlar (intracranial EEG veya inside EEG, iEEG) olmak üzere iki çeşit elektrot mevcuttur.

Uzmanlar, EEG kayıtlarına bakarak epileptik nöbet kayıtlarını dört farklı sınıfa ayırmıştır. Bu sınıflar sırasıyla, inter-iktal; iki epileptik nöbet arasında, pre-iktal; epileptik nöbet öncesi dönemi, iktal; epileptik nöbet anını, post-iktal; epileptik nöbet sonrasında ifade etmektedir. Nöbet tahmin çalışmaları, inter-iktal ve pre-iktal bölgenin birbirinden ayrıştırılmasını amaçlar.

II. MATERYAL VE YÖNTEM

Gerçekleştirilen sistemin blok diyagramı şekil 1'de görülmektedir. Gerçekleştirilen yöntem, ön işlemler, öz nitelik çıkarma ve sınıflandırma olmak üzere üç aşamadan oluşmaktadır. Öz nitelik çıkarma yöntemi olarak otoregresif (autoregressive - AR) model katsayıları tercih edilmiştir. Sonrasında ise elde edilen öz nitelikler, 3 farklı sınıflandırıcı girişine uygulanmıştır. Doğrusal Ayırtaç Analizi (DAA), Destek Vektör Makineleri (DVM) ve Karar Ağaçları (KA) sınıflandırıcı olarak tercih edilmiştir. Tüm bu işlemler MATLAB ortamında gerçekleştirilmiştir. İlerleyen paragraflarda, sırasıyla kullanılan materyaller ve yöntemler açıklanmıştır.



Şekil 1. Çalışmanın Blok Diyagramı

A. Veriseti

Amerikan Epilepsi Derneği'nin (AED) [5] veri setinde bulunan, kayıtlar üzerinde işlem gerçekleştirilmiştir. Her bir kayıt dosyası, 10 dakikalık süreler halinde inter-iktal ve pre-iktal sınıfları olarak ayrıştırılmıştır. Her bir dosya *mat-uzantılı* olarak sunulmuştur. Bu durum araştırmacılara verileri bilgisayar ortamına aktarmada kolaylık sağlamaktadır. Tablo 1'de veri setinin sayısal değerleri gösterilmiştir. EEG kayıtları, 400 Hz ve 5000 Hz olmak üzere iki farklı örnekleme frekansında örneklenmiştir. Bazı denek kayıtları 15 kanallı EEG işareti olarak kaydedilmiş, bazıları ise 16 ve 24 kanallı EEG kayıtlarından oluşmaktadır.

Tablo 1. Kullanılan Veri Seti Özellikleri

Denek Adı	Kanal Sayısı	Örnekleme Frekansı	İnter-iktal Segment Sayısı	Pre-iktal Segment Sayısı
Köpek-1	16	400	480	24
Köpek-2	16	400	500	42
Köpek-3	16	400	1440	72
Köpek-4	16	400	804	97
Köpek-5	15	400	450	30
Hasta-1	15	5000	50	18
Hasta-2	24	5000	42	18

B. Ön işlemler

Her bir segment için DC bileşen değeri ve 60 Hz şehir şebeke gürültüsü bastırılmıştır. EEG ile ilgili çalışmalarda kayan pencere analizi en çok tercih edilen yöntemdir. Pencere

süresi olarak 30 saniyelik süreler ve %50 örtüşme tercih edilmiştir. Veri setinde bulunan, iki sınıfın (inter-iktal ve pre-iktal) sayıları eşit değildir. Bu durum dengesiz veri seti problemi oluşturmaktadır. Bu sorunu çözmek pre-iktal sınıfta aşırı örnekleme yaparak üstesinden gelinmiştir. Aşırı örnekleme yapmak için %50'den daha küçük örtüşme oranları tercih edilmiştir. Bu sayede her iki sınıfın nokta sayıları eşitlenmiş olur.

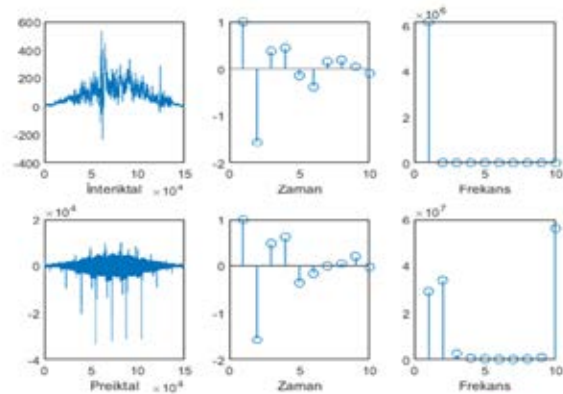
C. Öz nitelik Çıkarma

Öz nitelik çıkarma, makine öğrenmesinde en önemli aşamalardan birisidir. Tercih edilen öz niteliğin problemdeki sınıfları başarılı bir şekilde ayırt etme yeteneğinin olması gerekmektedir. Bu nedenle literatürde çokça kullanılan otoregresif model katsayıları tercih edilmiştir [6, 7].

AR model katsayıları, zaman serileri analizde çokça tercih edilir. Özellikler istatistik, ekonomi gibi alanlarda sıklıkla kullanılmaktadır [8]. Bir zaman serisinin geçmiş değerleri kullanılarak, şu andaki değerlerinin tahmin edilmesi esasına dayanmaktadır. Lineer denklem sistemi için eşitlik (1)' de görülmektedir.

$$X(t) = \sum_{i=1}^p a(i)X(t-i) + e(t) \quad (1)$$

$e(t)$, eşitlikteki, beyaz gürültüyü ifade etmektedir. $a(i)$ değerleri, AR model parametrelerini oluşturmaktadır. AR model katsayıları her bir EEG işaretinin kanalları için çıkarılarak öz nitelik vektörleri çıkarılır.



Şekil 2. AR Model Katsayıları

Şekil 2' de Hasta-1 kaydına ait 30 saniyelik EEG inter-iktal ve pre-iktal sınıflarının ham hali görülmektedir. Ayrıca bu EEG kayıtlarından elde edilen hem zaman hem de frekans eksenindeki AR model katsayıları görülmektedir. Şekil 2 incelendiği zaman, her iki sınıf arasındaki farklılık zaman

ekseninde az iken, frekans ekseninde daha belirgin bir fark bulunmaktadır. Bundan dolayı, AR model katsayıları iyi bir öznelik olarak sistemin başarısına olumlu katkı sağlayacağı anlaşılmaktadır.

AR model katsayılarının çıkartılması esnasında, çoğunlukla Burg ve Yule-Walker parametre kestirim yöntemleri kullanılmaktadır. Bu çalışmada biz, Yule-Walker parametre kestirim yöntemini tercih ettik. EEG işaretinin hem zaman ekseninde hem de frekans eksenindeki, AR model katsayıları hesaplanmıştır.

D. Sınıflandırma

Bu bölümde, bu çalışmada kullanılan üç popüler sınıflandırma yöntemi kısaca özetlenmiştir. Doğrusal Ayırtaç Analizi (DAA), sınıflar arasındaki sınırı tahmin etmek için aritmetik ortalama ve kovaryans hesaplamaları yapar [9]. DAA, karar bölgesi oluşturmak için aşağıdaki adımları takip eder. İlk olarak, bireysel sınıfların aritmetik ortalamaları ve tüm veri kümesinin aritmetik ortalaması hesaplanır. İkincisi, sınıf içi dağılım hakkında bilgi edinmek için her bir sınıfın kovaryans matrisleri hesaplanır. DAA karar sınır çizgisini belirlemek için iki sınırlama getirmiştir. Bunlardan ilki, sınıfların birbirine en uzak olması sınırlamasıdır (sınıflar arası değişkenlik). İkincisi ise, her sınıftaki varyansın (sınıf içi varyans) en küçük hale getirilmesidir.

Destek Vektör Makinesi (DVM), ikili sınıflandırmada kullanılan denetimli bir yöntemdir. Sınıfları birbirinden ayırmak için, DVM karar sınırında hiper düzlem oluşturur [10]. Sınır bölgesindeki en yakın noktalara destek noktaları denir. Öznelik uzayının non-lineer yapıda olduğu durumlar için DVM doğrusal olmayan kernel fonksiyonlarından faydalanılır. Bu çalışmada polinomial çekirdek fonksiyonu tercih edilmiştir.

Karar Ağaçları (KA) üç temel bileşenden oluşur: kök düğümü, iç düğüm ve yaprak şeklinde bir ağaç ve aşağı doğru uzanan bir yapıya sahiptir. Verileri en iyi ayırt eden ilk boyut belirlenir. İlk özellik boyutu, ilk kök düğümü ile oluşturulur. İkinci en iyi ayırt edici özelliği kullanarak, ikinci iç düğüm yaratılır. Bu düğümler, alt alta hizalanarak KA piramit şeklinde aşağı doğru genişler. Düğümlerin bitiş noktasına yaprak denir. Sonuç olarak, her bir yaprak karar sınıfını temsil eder [10].

III. DENEY BULGULARI

Sınıflandırma aşamasında, her bir hasta kaydı için elde edilen AR model katsayıları kullanılarak elde edilen öznelik uzayı, k=10 çapraz geçişleme ile DAA, DVM ve KA

sınıflandırıcılarının girişlerine uygulanmıştır. Her bir sınıflandırıcı çıkışından ise 4 farklı performans ölçüm kriteri hesaplanmıştır. Bunlar sırasıyla, duyarlılık, belirlilik, doğruluk ve eğri altında kalan alan (Area Under Curve - AUC)'dir. Bu performans ölçüm kriterlerinin eşitlikleri (2), (3), (4)' de görülmektedir.

$$\text{Duyarlılık (Sensitivity)} = \frac{DP}{(DP+YN)} \quad (2)$$

$$\text{Belirlilik (Specificity)} = \frac{DN}{(DN+YP)} \quad (3)$$

$$\text{Doğruluk (Accuracy)} = \frac{DN+DP}{(DN+YN+DP+YP)} \quad (4)$$

Doğru Pozitif (DP); gerçekte pre-iktal dönemde bulunurken sınıflandırma sonunda pre-iktal dönemde belirlenen nokta sayısı, Yanlış Pozitif (YP); gerçekte inter-iktal dönemde olmasına rağmen sınıflandırma sonunda pre-iktal dönemde belirlenen nokta sayısı, Doğru Negatif (DN); gerçekte inter-iktal dönemde bulunurken sınıflandırma sonunda inter-iktal dönemde belirlenen nokta sayısı, Yanlış Negatif (YN); gerçekte pre-iktal dönemde olmasına rağmen sınıflandırma sonunda inter-iktal dönemde belirlenen nokta sayısını ifade eder.

Duyarlılık; pre-iktal dönemdeki doğru sınıflandırma oranını, belirlilik; inter-iktal dönemdeki doğru sınıflandırma oranını, doğruluk; her iki dönemlerdeki, doğru sınıflandırma oranını ifade eder.

Tablo 2. DAA Sınıflandırıcısı Performans Sonuçları

Denek Adı	Duyarlılık	Belirlilik	Doğruluk	AUC
Köpek-1	0.97	0.87	0.92	0.92
Köpek-2	1.00	0.98	0.99	0.99
Köpek-3	0.95	0.93	0.94	0.94
Köpek-4	0.95	0.85	0.90	0.90
Köpek-5	1.00	0.96	0.98	0.98
Hasta-1	0.96	0.99	0.98	0.98
Hasta-2	1.00	0.99	1.00	1.00
Ortalama Başarım	0.98	0.94	0.96	0.96

DAA sınıflandırıcı sonuçları Tablo 2' de gösterilmiştir. Tablo 2 incelendiği zaman, en düşük başarı oranı %0.92 ile köpek-1 kaydı için olmakta, en yüksek başarı oranı ise %1.00 ile hasta-2 kaydı için elde edilmiştir. Sonuç olarak tüm kayıtları için ortalama başarı oranı ise %0.96 değeri elde edilmiştir.

Tablo 3. DVM Sınıflandırıcısı Performans Sonuçları

Denek Adı	Duyarlılık	Belirlilik	Doğruluk	AUC
Köpek-1	1.00	0.99	1.00	1.00
Köpek-2	1.00	1.00	1.00	1.00
Köpek-3	1.00	1.00	1.00	1.00
Köpek-4	1.00	0.98	0.99	0.99

Köpek-5	1.00	1.00	1.00	1.00
Hasta-1	0.87	0.98	0.92	0.92
Hasta-2	0.90	1.00	0.95	0.95
Ortalama Başarım	0.97	0.99	0.98	0.98

DVM sınıflandırıcı sonuçları Tablo 3' de gösterilmiştir. Tablo 3 incelendiği zaman, en düşük başarımları oranı %0.92 ile hasta-1 kaydı için olmakta, en yüksek başarımları oranı ise %1.00 ile köpek-1, köpek-2, köpek-3 ve köpek-5 kayıtları için elde edilmiştir. Sonuç olarak tüm kayıtları için ortalama başarımları oranı ise %0.98 değeri elde edilmiştir.

Tablo 4. KA Sınıflandırıcısı Performans Sonuçları

Denek Adı	Duyarlılık	Belirlielik	Doğruluk	AUC
Köpek-1	0.96	0.92	0.94	0.94
Köpek-2	0.98	0.97	0.98	0.98
Köpek-3	0.98	0.96	0.97	0.97
Köpek-4	0.93	0.90	0.92	0.92
Köpek-5	0.98	0.96	0.97	0.97
Hasta-1	0.98	0.98	0.98	0.98
Hasta-2	0.99	0.99	0.99	0.99
Ortalama Başarım	0.97	0.95	0.96	0.96

KA sınıflandırıcı sonuçları Tablo 4' de gösterilmiştir. Tablo 4 incelendiği zaman, en düşük başarımları oranı %0.92 ile köpek-4 kaydı için olmakta, en yüksek başarımları oranı ise %0.99 ile hasta-2 kaydı için elde edilmiştir. Sonuç olarak tüm kayıtları için ortalama başarımları oranı ise %0.96 değeri elde edilmiştir.

Her üç sınıflandırıcı sonuçları karşılaştırıldığı zaman en yüksek başarımları oranı, %0.98 için DVM sınıflandırıcı için ulaşılmıştır. Bunun yanında, her üç sınıflandırıcı sonucu AED veri setinde yüksek başarımları oranı elde ettiği görülmektedir.

Cui ve ark. [11] çalışmalarında, AED veri seti üzerinde, Bag-of-Wave yöntemini kullanarak öznelilik çıkarmışlar ve ardından, Aşırı Öğrenme Makinesi (AÖM) girişine uygulamışlardır. Ortalama başarımları oranı olarak AUC=%1.00 elde edilmiştir.

Zhang ve Keshab [12] çalışmalarında, AED veri seti üzerinde, spektral güç değerleri, EEG kanalları arasındaki korelasyon gibi öznelilikler çıkarılmıştır. Ardından, karar ağaçları ile öznelilik seçme işlemi gerçekleştirmişlerdir. Sonrasında seçilen öznelilikler ikinci dereceden Kalman filtreye uygulanmıştır. AdaBoost, DVM ve Yapay Sinir Ağları (YSA) olmak üzere üç farklı sınıflandırıcıya uygulamışlardır. En yüksek ortalama başarımları oranı olarak YSA için AUC = %0.89 değerini elde etmişlerdir.

Tablo 5. Diğer Çalışmalar ile Karşılaştırma

Denek Adı	Öznelilik	Sınıflandırıcı	AUC
Cui [11]	Bag-of-Wave	AÖM	1.00
Önerilen Yöntem	AR Model Katsayıları	DVM	0.98
Zhang [12]	Spektral Güç Hesaplama Kanallar Arası Korelasyon	YSA	0.89

Tablo 5' de diğer araştırmacıların aynı veri seti üzerinde elde ettikleri sonuçlar sunulmuştur. AR model katsayılarını kullanarak çıkarılan öznelilikler, DVM sınıflandırıcı ile sınıflandırılmış ve ortalama başarımları oranı için %0.98 sağlanmıştır. Elde edilen sonuç, aynı veri setini kullanan diğer çalışmalar ile kıyaslandığı zaman, Cui' nin çalışmasına yaklaşılmaktadır. Bunun yanında ise Zhang' ın çalışmasından daha iyi bir performans sunmaktadır.

IV. SONUÇ

Bu çalışmada, EEG sinyalleri üzerinde, AR model katsayılarını ve örutü tanıma tekniklerini kullanarak, epileptik nöbet tahmin çalışması gerçekleştirilmiştir. En yüksek ortalama başarımları oranı AUC=%0.98 değeri elde edilmiştir. Elde edilen bulgular neticesinde, AR model katsayılarının bu problem için faydalı olduğu görülmektedir. AR model katsayıları, EEG işaretini hem zaman hem de frekans ekseninde iyi bir şekilde temsil ettiği söylenebilir. Ayrıca kolay hesaplanan ve hızlı bir yöntemdir. Gelecek çalışmalarda, bu yapının yanında farklı veri setleri ile öznelilikler çıkarılacak ve derin öğrenme temelli yaklaşımlara uygulanacaktır.

KAYNAKLAR

- [1] <http://www.turkepilepsi.org.tr/menu/29/epilepsi-nedir>, (Ziyaret Tarihi: 15 Haziran 2019).
- [2] Schelter B., Timmer J., Bonhage A., *Seizure prediction in Epilepsy*, Wiley, Almanya, 2008.
- [3] Yazgan E., Korürek M., *Tıp Elektronikliği*, İTÜ, İstanbul, 1996.
- [4] Sanei S., Chambers J.A., *EEG SIGNAL PROCESSING*, John Wiley, UK, 2007.
- [5] <https://www.kaggle.com/c/seizure-prediction>, (Ziyaret Tarihi: 15 Haziran 2019).
- [6] Mousavi, S. R., M. Niknazar, and B. Vosoughi Vahdat. "Epileptic seizure detection using AR model on EEG signals.", *Cairo International Biomedical Engineering Conference*. IEEE, 2008.
- [7] Chisci, Luigi, et al. "Real-time epileptic seizure prediction using AR models and support vector machines." *IEEE Transactions on Biomedical Engineering*, (2010): 1124-1132.
- [8] Işıklar, Zeynep Ergen. "İMKB Ulusal 100 endeksi getiri volatilitésinin analizi üzerine bir araştırma." *Selçuk Üniversitesi Sosyal ve Teknik Araştırmalar Dergisi* 12 (2016): 245-260.
- [9] Alpaydın E., *Yapay Öğrenme*, Boğaziçi Üniversitesi Yayınevi, İstanbul, 2013.
- [10] Z. Zhou, *Ensemble Methods Foundations and Algorithms*, CRC Press, 2012.
- [11] Cui, Song, et al. "Seizure Prediction for iEEG Signal with Bag-of-Wave Model and Extreme Learning Machine." *International Conference on Extreme Learning Machine*. Springer, Cham, 2017.
- [12] Zhang, Zisheng, and Keshab K. Parhi. "Seizure prediction using long-term fragmented intracranial canine and human EEG recordings." *2016 50th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2016.