Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi
Afyon Kocatepe University – Journal of Science and Engineering
https://dergipark.org.tr/tr/pub/akufemubid

Araştırma Makalesi / Research Article
e-ISSN: 2149-3367
DOI: https://doi.org/10.35414/akufemubid.1574539
AKÜ FEMÜBİD **25** (2025) 047201 (956-968)
AKU J. Sci. Eng. **25** (2025) 047201 (956-968)

Epileptik Nöbetlerin Tespiti için Öznitelik Çıkartma ve Sınıflandırma Yöntemleri

Makale Bilgisi / Article Info
Alındı/Received: 27.10.2024
Kabul/Accepted: 01.02.2025
Yayımlandı/Published: 04.08.2025

Feature Extraction and Classification Methods for Epileptic Seizure Detection

Nuri İKİZLER*, Güneş EKİM

Karadeniz Teknik Üniversitesi, Trabzon Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Trabzon, Türkiye

© 2025 The Authors | Creative Commons Attribution-Noncommercial 4.0 (CC BY-NC) International License

Öz

Bu çalışmada, Bonn Üniversitesi Epileptoloji Bölümü'nden alınan EEG veri seti kullanılarak, epileptik nöbet tespitine yönelik bir model sunulmuştur. Frekans analizinde daha hassas sonuçlar elde etmek için ikiye ayrılan sinyaller ile sınıflandırma algoritmalarının performansının artırılmasına katkı sağlanan çalışmada, her sınıf için oluşturulan referans sinyaller, sınıf içi tutarlılık ve anormal sinyallerin tespiti açısından önemli bir rol oynamıştır. Sinyallerin spektral özelliklerinin karşılaştırılmasının, sınıflar arasındaki farkların belirginleşmesine ve sınıflandırma performansının artmasına yardımcı olduğu modelde, öznitelik vektöründe spektral farklar, enerji, entropi ve frekans sapmaları gibi özellikler kullanılarak, özellikle Kullback-Leibler sapması ve Euclidean mesafesi gibi metrikler sayesinde sınıflar arası spektral farklılıkların tespit edilmesi sağlanmıştır. Sınıflandırma aşamasında kullanılan beş farklı sınıflandırma algoritması içinde k-EYK (k=1) ve LMA en yüksek performansı göstererek, on üç sınıflandırma görevi içerisinde yedi görevde epileptik nöbetlerin %100 doğrulukla tespit edilmesi sağlamıştır. Tüm sınıflandırma görevlerinde k-EYK ile LMA sınıflandırıcıları için bulunan ortalama %98,03 ve %98,23 doğruluk değerleri, modelin epileptik nöbet tespiti için çok başarılı ve güvenilir bir yöntem olduğunu göstermektedir.

Anahtar Kelimeler: *Epileptik nöbet tespiti; EEG; Spektral analiz; Welch yöntemi; k-En yakın komşu; Lojistik model ağacı.*

Abstract

In this study, a model for detecting epileptic seizures is presented using the EEG dataset from the Epileptology Department of the University of Bonn. To obtain more precise results in frequency analysis, the signals are split into two parts, which contributed to improving the performance of the classification algorithms. The reference signals is obtained for each class played a significant role in assessing intra-class consistency and detecting abnormal signals. The comparison of spectral features among signals is helped to highlight differences between classes and enhance classification performance. In the feature vector, spectral differences, energy, entropy, and frequency deviations are employed, with metrics such as Kullback-Leibler divergence and Euclidean distance are used to detect spectral differences between classes. Among the five different classification algorithms used, k-NN (k=1) and LMT showed the highest performance, successfully detecting epileptic seizures with 100% accuracy in seven out of thirteen classification tasks. The average accuracy values of 98.03% for k-NN and 98.23% for LMT across all classification tasks demonstrate that the model is a highly effective and reliable method for detecting epileptic seizures.

Keywords: *Epileptic seizure detection; EEG; Spectral analysis; Welch method; k-Nearest neighbor; Logistic model tree*

1. Giriş

Günümüzde nörolojik hastalıkların en önemlilerinden biri olarak kabul edilen epilepsi beyinde bulunan nöronların anormal aktivitesinden kaynaklanmalıdır. Beyindeki bu elektriksel aktivite hastalarda bilinç kaybı, istemsiz hareketler, duyusal veya bilişsel bozukluklarla nöbetler şeklinde kendini gösterebilir. Hastadan hastaya farklılık gösteren nöbetlerle yaşamak zorunda olan dünya çapında 70 milyondan fazla insan vardır (İnt.Kay.-1). Her yıl yükselen bu rakam, özellikle az gelişmiş ülkelerde daha yüksek bir oranda artış göstermektedir (İnt.Kay.-2, Alotaiby 2016).

Epileptik nöbetler, fokal ve jeneralize olmak üzere ikiye ayrılır. Fokal nöbetler beynin belirli bölgesinde, tek taraflı bir etkiye sahipken, jeneralize nöbetler beynin bütününü etkiler ve fokal nöbetlerde görülen basit duyusal ve bilişsel semptomlara ek olarak ani kasılmalar ve düşme gibi daha geniş çaplı semptomlarla kendini gösterir. Beyinde oluşan bu anormal elektriksel aktivitenin ölçümünde ve epilepsi teşhisinin konulmasında, non-invaziv bir teknik olan Elektroensefalografi (EEG)'den faydalanılır (Milligan 2021). Beyin aktivitesinin zamana göre değişimi gösteren EEG epileptik hastalarda, pre-iktal, post-iktal, iktal ve inter-iktal adı verilen nöbet öncesi, nöbet sonrası, nöbet ve nöbet arası aralıklarının tespit edilmesinde önemli rol oynar.

EEG kullanılarak epileptik nöbetlerin tespit edilmesi uzun bir süreçtir. Hastadan alınan uzun EEG kayıtları tek tek incelenerek EEG işaretleri üzerinde pre-iktal, post-iktal, iktal ve inter-iktal bölümler tespit edilir. Uzman doktorlar

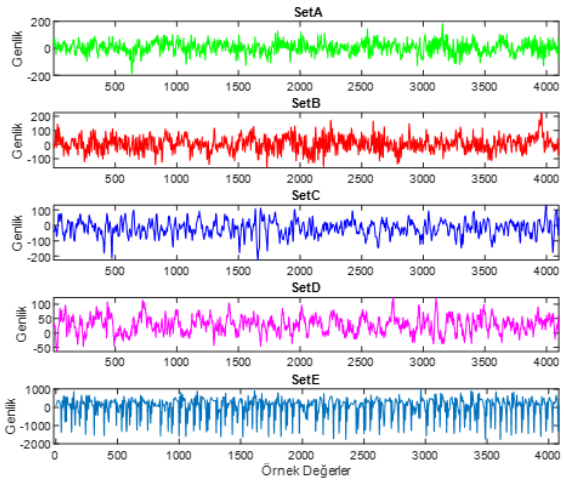
*Sorumlu Yazar/Corresponding Author: Nuri İKİZLER e-posta/e-mail: nikizler@ktu.edu.tr

tarafından yapılan bu incelemeler oldukça zahmetli ve zaman alıcıdır. Özellikle hastanın uykusunda alınan EEG kayıtları saatler sürmektedir. Bunların incelenmesinde büyük bir dikkat harcanarak epileptik nöbet bölümlerinin gözden kaçırılmaması gerekir. Bu bilgiler hastanın tedavisinde oldukça önemlidir. Ancak bir işlemin defalarca ve çok uzun sürelerde yapılması, insan doğası gereği sıkılmayı buna bağlı olarak dikkatin azalmasını ve bunun sonucunda da hatayı beraberinde getirir. Bilgisayarların ise kendilerine verilen görev ve emirleri bıkmadan, sıkılmadan defalarca hatasız yerine getirip, hayatımızın birçok alanında bize yardımcı olmaktadırlar. Önemli olan nokta doğru programlanmalarıdır. Bu nedenlerle epileptik nöbet tanıma üzerinde oldukça fazla araştırma yapılan önemli bir konudur. Araştırmacılar, farklı yöntemleri derleyerek, EEG kayıtları üzerinde en yüksek doğruluk değeri ile epileptik nöbetlerin tespiti konusunda çalışmaktadırlar. Bu alanda yapılan araştırmalar için en büyük zorluk, EEG nöbetlerini tespit edebilmek için EEG sinyallerine duyulan ihtiyaçtır. Bunun yanında tüm çalışmalarda bulunan doğruluk değerlerinin referans olarak alınacak EEG sinyallerine göre yapılması da araştırmaların başarasının değerlendirilmesinde önemli rol oynamaktadır. Günümüzde epileptik nöbetlere ait daha fazla sayıda EEG sinyallerinin yer aldığı ve araştırmacılar tarafından ilgi odağı olmuş iki önemli EEG veri tabanı mevcuttur. CHB-MIT ve Bonn Üniversitesi EEG veri tabanları kullanılarak yapılan çalışmalar literatürde önemli yer tutmaktadır. Literatürde yapılan 190 deneysel çalışmanın 89 tanesi Bonn üniversitesi, 20 tanesi CHB-MIT veri tabanı kullanılarak gerçekleştirilmiştir (Miltiadous vd. 2022). Çizelge 1'de bu iki veri tabanı kullanılarak yapılan çalışmaların bir kısmı yer almaktadır. Bu çalışmalar farklı zaman veya frekans bölgesi öznitelik çıkarım yöntemleri ve sınıflandırma yöntemleri içermektedir. Çeşitli sınıflandırma görevlerinde elde edilen doğruluk yüzdeleri bu çizelgede görülmektedir. Burada araştırmaların amacı hiç kuşkusuz en yüksek doğruluk yüzdesini elde etmektir. Ancak gerçekleştirilen işlemin bir sınıflandırma işlemi olduğu unutulmamalıdır. EEG veri tabanında yer alan farklı gruplardaki EEG verilerinin birbirinden ayrılmasından ziyade, araştırmacılar çalışmalarında pre-iktal, post-iktal, iktal ve inter-iktal bölümlerine ait sinyallerin normal EEG işaretlerinden ayrılmasına odaklı sınıflandırma görevleri üzerine Çizelge 1'de görüldüğü gibi yoğunlaşmışlardır. Çalışmamızda literatürde yer alan çalışmalardan öznitelik çıkarımı aşamasında farklı bir yaklaşım göstererek hazırladığımız modelin açık erişimli Bonn Üniversitesi EEG verilerini kullanarak otomatik nöbet tespitindeki doğruluk yüzdesinin bulunması hedeflenmiştir. Burada, Welch yöntemi ile elde edilen sinyallere ait güç spektral yoğunluğu (GSY) kullanılarak, veri setinde yer alan sinyallerin, referans güç spektral yoğunluklarına göre Kullback-Leibler sapması, Euclidean mesafesi, Jensen-Shannon sapması, Bhattacharya mesafesi, spektral enerji farkı, spektral entropi farkı, spektral düzlük farkı, spektral bant genişliği farkı, maksimum spektral sapma ve sıfır geçiş değerleri öznitelik vektörü çıkarımında kullanılmıştır. Elde edilen öznitelik vektörü kullanılarak Lojistik Model Ağacı sınıflandırma yöntemi ile farklı sınıflandırma görevleri için sınıflandırma doğruluk yüzdesi bulunmuş ve önerilen modelin otomatik nöbet tespitindeki başarısı araştırılmıştır.

2. Materyal ve Metot

2.1 Veri tabanı

Bu çalışmada Bonn Üniversitesi Epileptoloji Bölümü'nden alınan EEG veri seti kullanılmıştır. Kamuya açık olan bu veri seti, 173,61 Hz örnekleme frekansında 128 kanallı, 12 bitlik bir EEG kayıt sistemi ile örneklenerek kaydedilmiş ve bu 128 kanal, sinyallerin ortalama değeri alınarak tek bir kanala indirgenmiştir. Her set, 23,6 saniye boyunca 100 adet tek kanallı EEG segmenti içerir ve her segment 4097 ayrı örnek (A, B, C, D ve E) içerir. EEG sinyalleri 0,53- 40 Hz bandında filtrelenmiş ve tüm deneyler göz ve kas hareketlerinden kaynaklanan eserlerden temizlenmiştir. Kayıt, uluslararası 10-20 kanal sistemine göre yapılmıştır. Set A, gözleri açık ve kafatasının dışından kaydedilen 5 sağlıklı gönüllüden alınan EEG kayıtlarından oluşmaktadır. Set B, uyanık ve sakin bir durumda gözleri kapalı 5 gönüllü sağlıklı bireyden alınan yüzey EEG kaydıdır. C ve D setleri kafatasının içinden alınan kayıtlardır, beş epileptik hastada epileptik nöbet geçirmeyen bireylerde ölçülmüştür. E seti ise, aynı 5 hastadan nöbet döneminde kafatası boşluğunun içinden yapılan kayıtlardır (Miltiadous vd. 2022, Paul 2018). Veri setlerine ait genlik-zaman grafikleri Şekil 1'de ve istatistiksel özellikler ise Çizelge 2'de gösterilmiştir.



Şekil 1. Bonn Üniversitesi Epileptoloji Bölümü EEG veri tabanı içerisinde yer alan setlerin birleştirilmiş genlik-zaman gösterimleri

Çizelge 1. Farklı veri tabanları kullanarak, frekans/zaman bölgesi öznitelik çıkarımları ile otomatik epileptik nöbet tespitinde son yıllarda yapılan çalışmalar ve doğruluk oranları (Paul 2018).

Yazar ve Veri tabanı	Öznitelik Çıkarımı	Sınıflandırma	Sınıflandırma Görevi	Doğruluk (%)
Li vd. 2019 Bonn Üni. Veri tabanı	KZFD-Spektogram-Skalogram	DVM	AB-CD-E	99,6
Zhao vd. 2019 Bonn Üni. Veri tabanı	Entropi özellikleri	YSA	AB-CD-E	93,3
Choubey ve Pandey 2019 Bonn Üni. Veri tabanı	FFT	k-EYK	AD-E	97
Saini ve Dutta 2018 Bonn Üni. Veri tabanı	İstatistiksel ve Entropi özellikleri	YSA	A-D-S	99,3
Sharmila ve Mahalakshmi 2017 Bonn Üni. Veri tabanı	ADD	k-EYK	A-E	100
Eltrass vd. 2021 Bonn Üni. Veri tabanı	Sinyal enerjisi	Nicemlenmiş Çekirdek En Küçük Ortalama Kare	A-B-C-D-E	97,88
Jana vd. 2023 Bonn Üni. Veri tabanı	Frekans ve zaman bölgesi özellikleri	DVM	ABCD-E	97,63
Mandhouj vd. 2021 Bonn Üni. Veri tabanı	KZFD-Spektogram	Geleneksel Sinir Ağları (GSA)	AB-E	98,88
Wang vd. 2020 Bonn Üni. Veri tabanı	FFT	Ağırlıklandırılmış k-EYK	A-D-E	99
Al-Hadeethi vd. 2020 Bonn Üni. Veri tabanı	İstatistiksel özellikler	Adaboost LS-DVM	Tüm kombinasyonlar	99
Harpale ve Bairagi 2021 CHB-MIT Veri tabanı	Zaman bölgesi özellikleri-FFT	Bulanık Mantık Sınıflandırıcısı	İktal-interiktal	96,02
Zhang vd. 2020 CHB-MIT Veri tabanı	Ayrık FD- Bant enerjileri	AttVGGNet	İktal-interiktal	95,6
Gabr vd. 2020 CHB-MIT Veri tabanı	KZFD-Spektogram-Skalogram	GSA	İktal-preiktal-interiktal	97

Çizelge 2. Bonn Üniversitesi Epileptoloji Bölümü EEG veri tabanı içerisinde yer alan setlerine ait istatistiksel özellikler.

	Set A	Set B	Set C	Set D	Set E
Ortalama	-6,260	-12,513	-8,879	-6,200	-4,739
Medyan	-6	-12	-7	-6	-5
Std.Sapma	48,338	70,682	59,386	90,345	341,123
Min	-288	-424	-412	-1147	-1885
Max	294	360	623	2047	2047
Çarpıklık	0,035	-0,035	-0,045	3,970	-0,385
Basıklık	3,285	3,580	4,821	79,674	5,889

2.2 Welch yöntemi

Welch yöntemi, bir sinyalin güç spektral yoğunluğunu (GSY) hesaplamak için kullanılan bir tekniktir. Bu yöntem, temel periodogram yöntemine göre daha güvenilir sonuçlar sağlar. Çünkü sinyali küçük parçalara ayırıp bu parçalardaki sonuçları ortalayarak spektral tahmindeki rastgele dalgalanmaları azaltır. Aşağıdaki adımlar izlenir (Alkan ve Kiymik 2006).

1- Sinyalin Parçalara Bölünmesi: Uzun sinyal x(n) belirli uzunluklarda, birbirleriyle örtüşen segmentlere ayrılır. Örneğin, %50 örtüşme ile N-uzunlukta segmentler elde edilir. Her segment, $x_i(n)$ ile gösterilir.

2- Pencereleme: Her segmente bir pencereleme fonksiyonu w(n) uygulanır. Genellikle Hamming ya da Hanning penceresi tercih edilir. Pencereli segment:

$$x_\omega(n) = x_i(n) \cdot \omega(n) \quad (1)$$

olarak hesaplanır.

3- Fourier Dönüşümü: Pencere uygulanmış her bir segment için Hızlı Fourier Dönüşümü (FFT) uygulanarak frekans domenine geçilir. Her segmentin Fourier dönüşümü şu şekilde hesaplanır:

$$X_\omega(k) = \sum_{n=0}^{N-1} x_\omega(n) e^{-j2\pi kn/N} \quad (2)$$

Burada k, frekans bileşenini temsil eder.

4- Periodogram Hesaplama: Her segment için periodogram, güç spektral yoğunluğu $P_x(k)$, Fourier dönüşümünün karesinin büyüklüğü olarak hesaplanır:

$$P_x(k) = \left(1/N|\omega|^2\right) |X_\omega(k)|^2 \quad (3)$$

Burada $|\omega|^2$, pencereleme fonksiyonunun enerji düzeltme faktörüdür.

5- Ortalama Alma: Tüm segmentlerin periodogramları ortalanarak nihai güç spektral yoğunluğu elde edilir:

$$P_{Welch}(k) = \frac{1}{M} \sum_{i=1}^{M} P_{x_i}(k) \quad (4)$$

Burada M, segment sayısını temsil eder.

Welch yöntemi, bu adımları izleyerek her segmentin periodogramını hesaplar ve bu periodogramları ortalayarak daha az gürültülü bir GSY tahmini üretir. Yöntemin avantajı, segmentleme ve ortalama alma yoluyla varyansı düşürmesi ve daha güvenilir bir spektral tahmin sağlamasıdır.

2.3 Öznitelik vektörü çıkarımı

2.3.1 Kullback-Leibler sapması (KLS)

Özellikle bilgi teorisi ve istatistikte kullanılan iki olasılık dağılımı arasındaki farklılığı ölçen bir yöntemdir. Asimetrik bir ölçümdür. KLS, iki olasılık dağılımı arasındaki bilgi kaybını veya bir dağılımın, diğer dağılıma göre ne kadar farklı olduğunu gösterir. Bir nevi, iki dağılımın ne kadar benzer ya da farklı olduğunu anlamaya çalışır (Ji vd. 2020).

Matematiksel olarak, Kullback-Leibler sapması şu şekilde ifade edilir:

$$D_{\{KL\}}(P||Q) = \sum_i P(i) \log\left(\frac{P(i)}{Q(i)}\right) \quad (5)$$

Bu formülde, P(i) ve Q(i), iki olasılık dağılımı olan P ve Q'nun ilgili olasılıklarıdır. $D_{\{KL\}}(P||Q)$, Q'yu temel alarak P'nin bilgi içeriğini ölçer; eğer P ve Q çok benzerse, KL sapma değeri sıfıra yakın olur.

2.3.2 Euclidean mesafesi

İki nokta arasındaki doğrusal mesafeyi ölçen geometrik uzaklık ölçülerinden biridir. İki vektör veya nokta arasındaki farkı ölçerken kullanılan bu yöntem, genellikle veri analizi, makine öğrenmesi ve istatistikte iki veri noktası arasındaki farklılıkları veya benzerlikleri değerlendirmek için kullanılır. Fiziksel uzaydaki en kısa mesafeyi temsil ettiği için, genellikle "doğru mesafe" olarak da adlandırılır (Krislock ve Wolkowicz 2012). Matematiksel olarak, n boyutlu bir uzayda iki vektör $x = (x_1, x_2, \cdots x_n)$ ve $y = (y_1, y_2, \cdots y_n)$ arasındaki Euclidean mesafesi aşağıdaki formül ile hesaplanır:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (6)$$

Bu formülde, x_i ve y_i, x ve y vektörlerinin her bir bileşenidir. Euclidean mesafesi, genellikle verilerin birbirine ne kadar yakın veya uzak olduğunu belirlemede, kümelerin içindeki veya arasındaki mesafeleri hesaplamakta kullanılır.

2.3.3 Jensen-Shannon Sapması (JSS)

Kullback-Leibler sapması (KLS) yöntemine dayalı, simetrik ve daha kararlı bir bilgi teorisi ölçüsüdür. JSS, iki olasılık dağılımı arasındaki benzerliği ölçmek için kullanılır ve KLS'nin aksine her iki dağılım arasındaki mesafeyi eşit şekilde dikkate alır, bu nedenle simetriktir ve her zaman pozitif bir değer verir. KLS'nin asimetrik yapısındaki dezavantajları gidermek için geliştirilmiş olan JSS hem KLS hem de bu iki dağılımın ortalamasını kullanarak mesafe hesaplar (Menéndez vd. 1997).

Matematiksel olarak iki olasılık dağılımı P ve Q arasındaki JSS şu şekilde tanımlanır:

$$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M) \quad (7)$$

Burada, M iki dağılımın ortalamasıdır.

$$M = \frac{1}{2}(P + Q) \quad (8)$$

JSS, olasılık dağılımlarının birbirine ne kadar benzediğini ölçmek için daha dengeli bir yöntemdir ve genellikle farklı uygulamalarda KLS'ye tercih edilir. Çünkü daha iyi bir simetri ve sınırlı değer aralığı sunar (0 ile 1 arasında).

2.3.4 Bhattacharyya mesafesi

İki olasılık dağılımı arasındaki benzerliği ölçmek için kullanılan bir yöntemdir. Bhattacharyya mesafesi, olasılık dağılımları arasındaki örtüşme miktarını belirler. Dağılımlar arasındaki örtüşme ne kadar fazlaysa, Bhattacharyya mesafesi o kadar küçük olur, bu da dağılımların birbirine daha yakın olduğunu gösterir. Bhattacharyya mesafesi, genellikle model doğruluğunu ölçmek ve sınıflandırma görevlerinde sınıflar arasındaki ayrımı değerlendirmek için kullanılır (Nwe vd. 2013). Bhattacharyya mesafesi, iki sürekli olasılık dağılımı P(x) ve Q(x) arasında şu şekilde tanımlanır:

$$D_B(P,Q) = -\ln\left(\int \sqrt{P(x)Q(x)}dx\right) \quad (9)$$

Burada $\sqrt{P(x)Q(x)}dx$ Bhattacharyya katsayısıdır. Bhattacharyya mesafesi, bu katsayının logaritmasıyla negatif bir işaretle alınarak hesaplanır. Mesafe arttıkça dağılımlar arasındaki fark da artar. Bu yöntem, görüntü tanıma, sinyal işleme ve makine öğrenimi gibi alanlarda yaygın olarak kullanılır.

2.3.5 Spektral enerji

İki güç spektral yoğunluğu (GSY) arasındaki enerji farkı, sinyallerin toplam enerji seviyelerindeki farkı belirlemek için hesaplanır. Enerji, bir sinyalin genliğinin karesi alınarak hesaplanır ve belirli bir frekans bandındaki toplam enerjiye odaklanır (Vidyaratne ve Iftekharuddin 2017). Enerji farkını bulmak için, her iki GSY'nin toplam enerjisi şu formülle hesaplanır:

$$E = \int_{f_1}^{f_2} GSY(f)\, df \quad (10)$$

Bu formül, f_1 ve f_2 arasındaki frekans bandında GSY'nin integralini alarak sinyalin enerjisini hesaplar. İki GSY arasındaki enerji farkı ise aşağıdaki gibi bulunur:

$$\Delta E = |E_1 - E_2| \quad (11)$$

Burada E_1 ve E_2 iki farklı sinyalin toplam enerjilerini temsil eder.

2.3.6 Spektral entropi

Spektral entropi, iki güç spektral yoğunluğu (GSY) arasındaki bilgi düzensizliğini veya belirsizliği ifade eder. Yüksek spektral entropi, sinyalin daha düzensiz ve karmaşık olduğu anlamına gelirken, düşük entropi, sinyalin daha düzenli ve belirli bir yapıya sahip olduğunu gösterir. Aşağıdaki adımlar izlenerek hesaplanır (Mirzaei vd. 2010).

1- GSY'nun Normalizasyonu: Öncelikle GSY'nun toplam alanı 1 olacak şekilde normalize edilmesi gerekir. Bu, her bir frekans bileşeninin olasılık dağılımı olarak yorumlanmasına olanak tanır.

$$P(f) = \frac{GSY(f)}{\sum_i GSY(f_i)} \quad (12)$$

2- Spektral Entropi Formülü: Normalleştirilmiş GSY kullanılarak spektral entropi hesaplanır. Formül şu şekildedir:

$$H = -\sum_i P(f_i) \log\big(P(f_i)\big) \quad (13)$$

Burada H, spektral entropiyi; $P(f_i)$, her bir frekans bileşeninin olasılığını temsil eder.

2.3.7 Spektral düzlük

İki güç spektral yoğunluğu (GSY) arasındaki düzensizlik ve düzlük derecesini ifade eder. Yüksek spektral düzlük, bir sinyalin daha düz (ve dolayısıyla daha az karmaşık) bir spektral yapıya sahip olduğunu gösterirken, düşük spektral düzlük, daha düzensiz ve zengin bir spektral yapıyı işaret eder (Boubchir vd. 2017). Spektral düzlük genellikle aşağıdaki formül ile hesaplanır:

$$Düzlük = \frac{10^{\frac{1}{N}\sum_{i=1}^{N} \log_{10}(GSY(f_i))}}{\frac{1}{N}\sum_{i=1}^{N} GSY(f_i)} \quad (14)$$

Burada; N, toplam frekans bileşenlerinin sayısını ve $GSY(f_i)$, i-inci frekans bileşeninin gücünü temsil eder.

2.3.8 Spektral bant genişliği

İki güç spektral yoğunluğu arasında, spektrumda ne kadar geniş bir frekans aralığına yayıldığını hesaplamak için kullanılır. Daha geniş bir spektral bant, genellikle daha fazla bilgi veya karmaşıklık içerirken, dar bir bant daha basit bir yapı veya az sayıda frekans bileşeni barındırır. Spektral bant genişliği genellikle aşağıdaki formül ile hesaplanır (Liu vd. 2023).

$$Bant\ genişliği = \sqrt{\frac{\sum_{i=1}^{N}(f_i-\mu)^2 \cdot GSY(f_i)}{\sum_{i=1}^{N} GSY(f_i)}} \quad (15)$$

$$\mu = \sqrt{\frac{\sum_{i=1}^{N} f_i \cdot GSY(f_i)}{\sum_{i=1}^{N} GSY(f_i)}} \quad (16)$$

Burada; N, toplam frekans bileşenlerinin sayısını; f_i, i-inci frekans bileşeni; μ, GSY'nin ağırlıklı ortalamasını temsil eder.

2.3.9 Maksimum spektral sapma

İki güç spektral yoğunluğu arasındaki maksimum farkı ifade eden bu ölçüm, sinyallerin frekans bileşenlerinin nasıl dağıldığını ve bir sinyalin diğerine göre ne kadar farklı olduğunu değerlendirmek için kullanılır. İki GSY arasındaki maksimum spektral sapma şu şekilde hesaplanır (Niknazar vd. 2020).

$$Max\ Spektral\ Sapma = \max_i |GSY_1(f_i) - GSY_2(f_i)| \quad (17)$$

Burada; $GSY1(f_i)$ ve $GSY2(f_i)$, sırasıyla iki farklı sinyalin f_i frekansındaki güç spektral yoğunluğunu temsil eder ve i, toplam frekans bileşenlerinin indeksidir.

2.3.10 Sıfır geçiş oranı (SGO)

Bir sinyalin zaman içinde sıfır değerini geçiş sayısını ifade eder. İki sinyal arasındaki sıfır geçiş hızı farkı, bu sinyallerin değişim hızını ve dalga biçimini karşılaştırarak, hangi sinyalin daha fazla dalgalanma veya değişim gösterdiğini anlamaya yardımcı olur. Aşağıdaki şekilde hesaplanır (Kumar ve Kolekar 2014).

$$SGO = \frac{1}{N}\sum_{i=1}^{N-1} sgn(x[i]) \neq sgn(x[i+1]) \quad (18)$$

Burada; x[i], sinyalin i-inci örneği ve sgn(x), sinyalin işaretini (pozitif veya negatif) belirler.

2.4 Sınıflandırıcılar

2.4.1 Çok katmanlı algılayıcı (ÇKA)

ÇKA, yapay sinir ağlarının bir türüdür ve veri setlerindeki karmaşık ilişkileri öğrenmek için kullanılan güçlü bir modeldir. Çok katmanlı yapısı sayesinde doğrusal olmayan problemlerde başarılıdır. Girdi verisini gizli katmanlar aracılığıyla işler ve her bir nöron, doğrusal bir dönüşümün ardından bir aktivasyon fonksiyonu ile veriyi işler. Geri yayılım algoritması ile ağırlıkları günceller ve öğrenir (Sriraam vd. 2018).

2.4.2 k-En yakın komşu (k-EYK)

k-EYK, basit ve sezgisel bir sınıflandırma algoritmasıdır. Yeni bir örneğin sınıfını tahmin etmek için, en yakın k komşusunun sınıflarına bakar ve çoğunlukla hangi sınıf varsa o sınıfı tahmin eder. Mesafe ölçümüne dayalı olduğu için yüksek boyutlu veri setlerinde dikkatle kullanılması gerekir. Parametre olarak seçilen k değeri modelin performansını önemli ölçüde etkiler (Hasan vd. 2017).

2.4.3 Rastgele orman (RO)

RO, birden çok karar ağacının oluşturduğu bir topluluk öğrenme modelidir. Her ağaç veri setinin rastgele bir alt

kümesinde eğitilir ve sonuçlar birleştirilir. Bu yöntem, aşırı öğrenmeyi önler ve genellikle yüksek doğruluk sağlar. Aynı zamanda gürültüye dayanıklıdır ve önemli özellikleri belirlemek için de kullanılabilir (Basri ve Arif 2021).

2.4.4 Lojistik model ağacı (LMA)

LMA, karar ağaçları ile lojistik regresyonu birleştiren bir sınıflandırma yöntemidir. Her yaprak düğümde lojistik regresyon modelleri kullanarak sınıflandırma yapar. Karar ağacının bölme gücünü lojistik regresyonun doğrusal sınıflandırma yeteneğiyle birleştirir, bu da daha karmaşık ilişkilerin öğrenilmesine olanak tanır (Kabir ve Zhang 2016).

2.4.5 Destek vektör makineleri (DVM)

DVM, iki sınıf arasındaki veriyi en iyi ayıran hiper düzlemi bulmaya çalışır. Amacı, sınıflar arasındaki marjini maksimize etmektir. Doğrusal olmayan verilerde de başarılıdır çünkü çekirdek fonksiyonları kullanarak veriyi daha yüksek boyutlara taşıyabilir ve burada ayrım yapabilir. Gürültüye karşı dayanıklı ve genellikle güçlü bir modeldir (Shiao vd. 2017).

2.5 Sınıflandırma sonuçlarının değerlendirilmesi

Sınıflandırıcıların doğruluk ve kesinlik değerleri, bir modelin performansını değerlendirmek için kullanılan temel ölçülerdir. Sınıflandırıcılar, verilen bir veri kümesinde doğru ve yanlış sınıflandırmalar yapabilir. Bu metrikler, modelin başarısını farklı yönlerden ölçer (Juba ve Le 2019). Doğruluk, modelin tüm sınıflandırma tahminlerinin ne kadarının doğru olduğunu gösterir. Yani, doğru pozitif (DP) ve doğru negatif (DN) tahminlerinin toplam tahminlere oranını hesaplar. Genel olarak modelin doğru sınıflandırma oranını gösterir. Modelin başarısını genel bir perspektiften değerlendirir. Ancak, dengesiz veri setlerinde (bir sınıfın çok daha fazla olduğu durumlarda) yanıltıcı olabilir. Bu tür durumlarda model, çoğunluk sınıfını tahmin ederek yüksek doğruluk elde edebilir ama diğer metriklerde düşük performans gösterebilir.

$$Doğruluk = \frac{DP+DN}{DP+DN+YP+YN} \quad (19)$$

Burada;

DP (Doğru Pozitif): Modelin pozitif olarak doğru tahmin ettiği örnekler.

DN (Doğru Negatif): Modelin negatif olarak doğru tahmin ettiği örnekler.

YP (Yanlış Pozitif): Modelin pozitif olarak yanlış tahmin ettiği örnekler.

YN (Yanlış Negatif): Modelin negatif olarak yanlış tahmin ettiği örnekler.

Kesinlik, modelin pozitif olarak tahmin ettiği verilerden ne kadarının gerçekten pozitif olduğunu ölçer. Yani, yanlış pozitiflerin etkisini değerlendirir. Yanlış pozitiflerin maliyetli olduğu durumlarda, kesinlik önemli bir metriktir.

$$Kesinlik = \frac{DP}{DP+YP} \quad (20)$$

3. Sonuçlar

Şekil 2'deki önerilen çalışmanın blok diyagramında, Bonn Üniversitesi'nden alınan halka açık EEG veri setleri kullanılmıştır. Bu veri setleri A, B, C, D, E sınıflarına ait toplam 500 adet veri dosyasından oluşmaktadır. Sınıflandırma doğruluğunu ve kesinliğini arttırmak amacıyla, her bir sınıf 2048 örneklik 200 dosyaya ayrılmıştır.

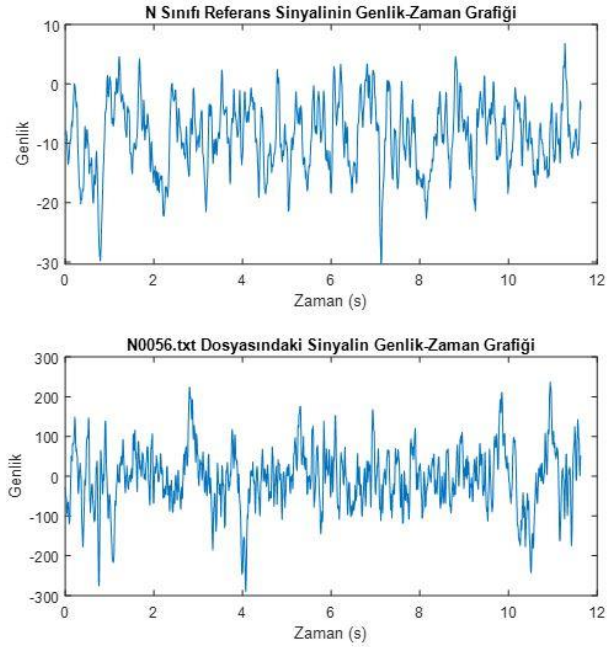


Şekil 2. Önerilen çalışmanın blok diyagramı

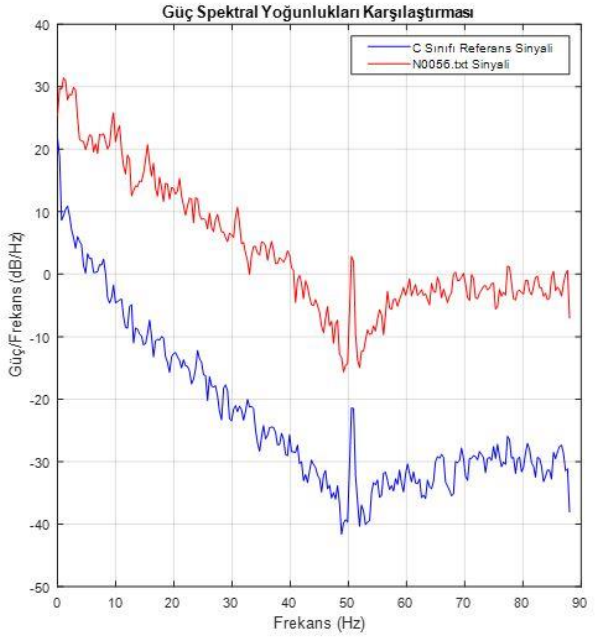
Dosyalar şu şekilde adlandırılmıştır: A sınıfı Z0001.txt-Z0200.txt, B sınıfı O0001.txt- O0200.txt, C sınıfı N0001.txt-N0200.txt, D sınıfı D0001.txt- D0200.txt, E sınıfı S0001.txt-S0200.txt.

A, B, C, D, E sınıfları için sinyallerin ortalaması alınarak 5 adet referans sinyal (Ref_A, Ref_B, Ref_C, Ref_D ve Ref_E) oluşturulmuştur. Şekil 3'te C sınıfına ait Ref_C ile rastgele

seçilen N0056.txt dosyasının genlik-zaman grafikleri verilmiştir. Çalışmanın 3. adımında, Welch yöntemiyle referans sinyalleri ve tüm sinyallerin güç spektral yoğunlukları hesaplanmıştır. Şekil 4'te Ref_C ile N0056.txt dosyasının güç spektral yoğunluklarının karşılaştırılması gösterilmektedir.



Şekil 3. C sınıfına ait referans sinyalinin ve rastgele seçilmiş N0056.txt dosyasındaki sinyalin genlik-zaman grafikleri



Şekil 4. C sınıfı referans sinyaline ait güç spektral yoğunluğunun C sınıfında yer alan rastgele seçilmiş N0056.txt sinyaline ait güç spektral yoğunluğuyla karşılaştırılması

Çalışmanın 4. adımında, veri setindeki sinyaller ve referans sinyallerine ait güç spektral yoğunlukları üzerinden; D_1: Kullback-Leibler sapması, D_2: Euclidean mesafesi, D_3: Jensen-Shannon sapması, D_4: Bhattacharya mesafesi, D_5: Spektral enerji farkı, D_6: Spektral entropi farkı, D_7: Spektral düzlük farkı, D_8: Spektral bant genişliği farkı, D_9: Maksimum spektral sapma, D_{10}: Sıfır geçiş oranı farkı hesaplanmıştır. Öznitelik vektörleri, bu analizlerden elde edilen D_1-D_{10} özelliklerini içermekte olup, Çizelge 3'te örnek değerler sunulmuştur.

Sınıflandırma aşamasında, 5 farklı yöntem kullanılarak öznitelik vektörleri ile eğitilen sınıflandırıcılar 10-kat çapraz doğrulama ile test edilmiştir. Çizelge 4, 13 farklı sınıflandırma görevinin doğruluk ve kesinlik sonuçlarını göstermektedir. Bu görevlerden 7'sinde %100 doğruluk sağlayan k-EYK yöntemi için farklı k değerlerinin sonuçları Çizelge 5'te, k=1 durumunda A-B-C-D-E sınıflandırmasına ait karışıklık matrisi ise Çizelge 6'da sunulmuştur.

Çizelge 4'teki sonuçlara göre;

- ÇKA ortalama %97,29 doğruluk ile iyi bir performans göstermiş, A-E, B-E ve C-E görevlerinde %100 doğruluk sağlamıştır.
- k-EYK (k=1) %98,04 ortalama doğruluk ile en iyi performansı sergileyen algoritma olup, 7 görevde %100 başarı elde etmiştir.
- RO %97,93 ortalama doğruluk ile k-EYK'nun hemen ardından gelir. Ancak özellikle D-E görevinde diğer yöntemlere kıyasla daha düşük sonuçlar almıştır.
- LMA %98,23 ortalama doğruluk ile en yüksek performansı sunmuştur. A-E, B-E ve ABCD-E görevlerinde %100 doğruluk sağlarken, karmaşık görevlerde performansı bir miktar düşmüştür.
- DVM %90,68 ortalama doğrulukla diğer algoritmaların gerisinde kalmıştır. D-E ve A-B-C-D-E gibi görevlerde düşük doğruluk (%74,7'ye kadar) elde etmiştir.

Çoklu sınıflı görevlerde, k-EYK ve LMA tutarlı sonuçlar verirken, DVM ve RO performans açısından geride kalmıştır. Özellikle DVM, A-B-C-D-E görevinde en düşük doğruluk değerine sahiptir.

Çizelge 5'te, k-EYK sınıflandırıcısının farklı k değerleri için doğruluk ve kesinlik sonuçları karşılaştırılmaktadır. k parametresi, komşu sayısını belirler ve bu değer arttıkça sınıflandırma performansında değişiklikler gözlenir.

- k=1: En yüksek doğruluk ve kesinlik değerleri elde edilmiştir. A-E, B-E ve C-E gibi basit sınıflandırma görevlerinde %100 doğruluk ve kesinlik sağlanmıştır. Küçük k değerleri, sınıflar arasındaki komşuluk ilişkilerini daha iyi yansıtarak üstün performans göstermektedir.

Çizelge 3. A, B, C, D, E veri setlerinde rastgele seçilmiş ikişer sinyal için hesaplanan öznitelik vektörleri.

Veri seti	Öznitelik Vektörü									
	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
Set A	11072	486	724	-5,728	2312	2135	0,0191	5,8362	191	0,1099
Set A	13372	928	956	-6,056	3075	3268	0,0237	3,4767	660	0,0889
Set B	41317	2733	2813	-7,180	9124	11790	0,0294	4,9428	1907	0,1060
Set B	36561	1390	2197	-6,782	6951	7774	0,1226	7,6117	501	0,1523
Set C	32507	2907	2345	-6,929	7506	10544	0,0026	0,8843	1656	0,0215
Set C	39855	3789	2892	-7,068	9167	13413	0,0004	0,6114	2269	0,0137
Set D	61903	3182	4144	-7,189	12813	16691	0,0345	2,9879	1517	0,0308
Set D	50137	4059	3682	-7,233	11545	16146	0,0078	0,6435	2940	0,0264
Set E	53224	3586	4606	-8,331	16796	23116	0,0083	1,3982	1402	0,0273
Set E	57632	3470	5006	-8,386	18171	24368	0,0014	0,4685	1428	0,0444

Çizelge 4. Farklı sınıflandırma görevleri için ÇKA, k-EYK, RO, LMA ve DVM (lineer kernel) sınıflandırıcılarının doğruluk ve kesinlik değerleri.

Sınıflandırma görevi (Setler) (2048 örnek)	Sınıflandırıcı									
	ÇKA		k-EKY (k=1)		RO		LMA		DVM	
	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik
A-E	100	1,000	100	1,000	100	1,000	100	1,000	100	1,000
B-E	100	1,000	100	1,000	100	1,000	100	1,000	100	1,000
C-E	100	1,000	100	1,000	99,75	0,998	100	1,000	100	1,000
D-E	99,75	0,998	100	1,000	98,50	0,985	99,50	0,995	99,50	0,995
AB-CD-E	95,60	0,957	99,10	0,991	98,70	0,987	99,50	0,995	84,50	0,846
A-B-E	100	1,000	100	1,000	99,50	0,998	99,83	0,998	88,33	0,885
ABCD-E	99,90	0,999	100	1,000	99,60	0,996	100	1,000	99,90	0,999
A-B-C-D-E	89,20	0,892	93,00	0,931	93,70	0,936	94,70	0,947	74,70	0,747
A-C	100	1,000	100	1,000	98,50	0,985	99,25	0,993	99,75	0,998
C-D	84,75	0,848	83,50	0,836	88,00	0,880	85,75	0,858	57,75	0,578
CD-E	99,83	0,998	100	1,000	99,16	0,992	100	1,000	99,66	0,997
A-D	99,25	0,993	99,50	0,995	99,00	0,990	99,50	0,995	98,25	0,983
AB-CDE	96,60	0,967	99,50	0,995	98,70	0,987	99,00	0,990	76,60	0,784
Ortalama doğruluk (%) (Tüm görevler)	97,29		98,04		97,93		98,23		90,68	

Çizelge 5. Farklı k değerleri için k-EYK sınıflandırma yönteminin sunulan çalışmada doğruluk ve kesinlik değerleri.

Sınıflandırma görevi (Setler) (2048 örnek)	Sınıflandırıcı									
	k-EKY (k=1)		k-EKY (k=3)		k-EKY (k=5)		k-EKY (k=7)		k-EKY (k=9)	
	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik	Doğruluk (%)	Kesinlik
A-E	100	1,000	100	1,000	100	1,000	100	1,000	100	1,000
B-E	100	1,000	100	1,000	100	1,000	100	1,000	100	1,000
C-E	100	1,000	100	1,000	99,50	0,995	99,25	0,993	99,25	0,993
D-E	100	1,000	99,00	0,990	98,75	0,988	98,50	0,985	98,50	0,985
AB-CD-E	99,10	0,991	97,60	0,977	97,20	0,972	96,20	0,962	95,90	0,959
A-B-E	100	1,000	99,83	0,998	99,50	0,995	98,83	0,989	98,33	0,983
A-B-C-D-E	93,00	0,931	91,40	0,916	89,90	0,902	89,40	0,899	87,30	0,880

Çizelge 6. A-B-C-D-E beşli sınıflandırma görevi için k-EYK (k=1) sınıflandırıcıda elde edilen karışıklık matrisi.

		Tahmin Edilen Sınıf				
	k-EYK	**A**	**B**	**C**	**D**	**E**
Gerçek Sınıf	**A**	197	0	3	0	0
	B	0	198	1	1	0
	C	1	0	169	30	0
	D	0	0	34	166	0
	E	0	0	0	0	200

- k=3: Daha karmaşık görevlerde (ör. AB-CD-E), doğruluk %99,10'dan %97,60'a gerilemiştir.
- k=5 ve üzeri: Karmaşık görevlerde performans daha belirgin şekilde düşmektedir. Örneğin, A-B-C-D-E görevinde doğruluk k=1'de %93,00 iken k=9'da %87,30'a kadar gerilemiştir.

Basit görevlerde (A-E, B-E, C-E, D-E), tüm k değerlerinde genellikle %100 doğruluk ve kesinlik sağlanmıştır. Ancak, k arttıkça özellikle C-E ve D-E görevlerinde küçük doğruluk düşüşleri görülmektedir, örneğin D-E görevi için k=9'da doğruluk %98,50'dir. Küçük k değerleri, özellikle çok sınıflı görevlerde daha yüksek doğruluk ve kesinlik sağlamaktadır. Büyük k değerlerinde ise sınıflar arası sınırlar bulanıklaşmakta ve performans düşmektedir. Çizelge 6'daki karışıklık matrisi incelendiğinde sınıflar arasındaki performans ve karışıklıklar şu şekilde özetlenebilir:

- A Sınıfı: 200 örnekten 197'si doğru sınıflandırılmış, 3 örnek ise yanlışlıkla C sınıfına atanmıştır. Bu durum, A sınıfının genelde doğru sınıflandırıldığını, ancak bazı durumlarda C sınıfıyla karışabileceğini göstermektedir.
- B Sınıfı: 200 örnekten 198'i doğru sınıflandırılmıştır. Yanlış sınıflandırılan 2 örnekten biri C sınıfına, diğeri ise D sınıfına atanmıştır. Bu sınıf için sınıflandırma oldukça başarılıdır.
- C Sınıfı: 200 örnekten 169'u doğru sınıflandırılmıştır. 30 örnek D sınıfına, 1 örnek ise A sınıfına atanmıştır. Bu, C ve D sınıfları arasında belirgin bir karışıklık olduğunu ve sınırların net olmadığını ortaya koymaktadır.
- D Sınıfı: 200 örnekten 166'sı doğru sınıflandırılmış, 34 örnek ise yanlışlıkla C sınıfına atanmıştır. D sınıfının da C sınıfıyla karıştığı ve modelin bu iki sınıfı ayırt etmekte zorlandığı gözlemlenmiştir.
- E Sınıfı: 200 örneğin tamamı doğru sınıflandırılmıştır (%100 doğruluk). Bu, modelin en iyi performansını E sınıfında gösterdiğini göstermektedir.

Model, A, B ve E sınıflarında yüksek doğruluk elde ederken, C ve D sınıfları arasında belirgin bir karışıklık bulunmaktadır. Bu, özellikle bu iki sınıfın ayrımını güçlendirmek için daha hassas özellikler veya farklı sınıflandırma yöntemlerinin gerekli olabileceğini göstermektedir.

Sunulan çalışmada seçilen sınıflandırma görevlerinde örnek dengesizliği olmayan görevler A-E, B-E, C-E, D-E, A-B-E, A-C, C-D, A-D, A-B-C-D-E şeklindedir. Bu görevlerde her sınıf eşit sayıda örnek içerdiği için, sınıf içi performansları eşit ağırlıkta değerlendiren makro ortalama tercih edilmiştir. Örnek dengesizliği bulunan görevler ise AB-CD-E, ABCD-E, CD-E, AB-CDE şeklindedir. Bu görevlerde ise sınıf örnek sayılarındaki farklılıkları dikkate almak için, sınıfların örnek sayılarına göre ağırlıklandırma yapan ağırlıklı ortalama yöntemi kullanılmıştır.

WEKA programı tarafından gerçekleştirilen özellik önemi analizi ise Çizelge 7 sunulmuştur. Bu çizelgeden spektral düzlük farkı dışında seçilen tüm öznitelik vektörlerinin sınıflandırmaya dengeli bir katkı sağladıklarını, Bhattacharya mesafesinin ise en önemli öznitelik olarak öne çıktığını söylemek mümkündür.

Çizelge 7. WEKA programının sıralanmış özellikleri ile öznitelik vektörlerinin A-B-C-D-E görevi için önem puanları.

Öznitelik (D) Vektörü	Özellik Önem Sıralaması	Önem Puanı
Kullback-Leibler sapması (D_1)	6	0,725
Euclidean mesafesi (D_2)	7	0,677
Jensen-Shannon sapması (D_3)	3	0,757
Bhattacharya mesafesi(D_4)	1	1,064
Spektral enerji farkı (D_5)	4	0,753
Spektral entropi farkı (D_6)	5	0,746
Spektral düzlük farkı (D_7)	10	0,189
Spektral bant genişliği farkı (D_8)	8	0,652
Maksimum spektral sapma (D_9)	9	0,632
Sıfır geçiş oranı farkı (D_{10})	2	0,791

4. Tartışma

Bonn Üniversitesi Epileptoloji Bölümü'nden alınan EEG veri seti kullanılarak, otomatik olarak epileptik nöbetlerin tespiti için sunulan bu çalışmada, ilk olarak gerçekleştirilen veri setinde yer alan sinyallerin ikiye bölünmesi işlemi önemlidir. Yapılan bu bölme işlemi sinyalin frekans ve zaman analizinde önemli bir rol oynar. Sinyali 200 parçaya bölmek, 100 parçaya bölmeye kıyasla daha fazla veri noktası sağladığı için, sinyalin frekans bileşenlerini daha ince detaylarla yakalamaya olanak tanır. Bu, özellikle düşük frekanslardaki bileşenleri daha iyi anlamaya ve analiz etmeye yardımcı olabilir. Sınıflandırma algoritmalarında, daha fazla örnek kullanmak modelin genelleme yeteneğini artırabilir. 200 bölüm, 100 bölüme göre daha fazla veri sağlayarak makine öğrenme algoritmalarının daha iyi performans göstermesine olanak tanıyacaktır. Ayrıca sinyali daha küçük parçalara bölmek, zaman içerisinde sinyalin nasıl değiştiğini daha iyi inceleme şansı verir, bu da sinyaldeki önemli örüntüleri tespit etmek açısından faydalı olacaktır.

Sunulan çalışmada diğer önemli bir nokta, her sınıf için bir referans sinyal oluşturulması ve bu referans sinyalin spektral özelliklerini kendi sınıfı içerisinde yer alan sinyal parçaları ile karşılaştırılması, sınıf içi tutarlılığın ölçülmesi, anormal sinyallerin tespiti, sınıflar arası ayrımı kolaylaştırma, özellik azaltımı ve model basitleştirme ile sinyal dinamiklerinin iyi kavranması açılarından önemlidir. Referans sinyal, her sınıfın genel spektral özelliklerini temsil eder. Sınıftaki bireysel parçaların bu referans sinyale ne kadar benzediğini karşılaştırmak sınıf içindeki sinyallerin tutarlılığı hakkında bilgi verecektir. Bu da sınıflandırma süreçlerinde sınıf içi değişmenin ne kadar düşük olduğunun anlaşılmasını kolaylaştıracaktır.

Referans sinyalden önemli ölçüde sapma gösteren sinyal parçaları, potansiyel olarak anormal ya da farklı özellikler taşıyan sinyaller olarak işaretlenebilir. Özellikle epilepsi gibi sinyallerin anormal örüntüler içerdiği durumlarda bu karşılaştırma, nöbet sinyallerini tespit etmede kullanılabilir. Her sınıf için oluşturulan referans sinyallerin spektral özellikleri, sınıflar arasında belirgin farklar olup olmadığını gösterebilir. Bu farklar, sınıfların birbiriyle karışmamasını sağlamak için kullanmak ve sınıflandırma algoritmaları için daha güçlü ayrıştırıcı özellikler elde etmeyi sağlayacaktır. Tüm parçalar yerine sadece referans sinyallerin spektral özellikleri üzerinden bir model oluşturmak ise, daha az veri kullanarak benzer performans elde edebilmeyi mümkün kılar. Referans sinyal, her sınıfın genel dinamiklerini yansıttığı için, bu dinamiklerin anlaşılması ve bu dinamiklere göre sinyal parçalarının karşılaştırılması ile, sinyalin sınıfa özgü örüntülerinin daha iyi kavranması mümkündür. Örneğin, nöbet öncesi ve nöbet anı sinyalleri referans sinyale göre farklılık gösterebilir. Sonuç olarak bu tür bir yaklaşım, spektral analiz ve sınıflandırma aşamalarında daha güvenilir ve güçlü bir model elde etmeye katkı sağlayacaktır.

Gerçekleştirilen çalışmada yöntemini tercih etmemizin sebepleri arasında, gürültü ve varyans azaltımı, pencerelenmiş veri kullanması, frekans çözünürlüğü ile kısa ve uzun sinyal boyutlarında esneklik sağlaması gösterilebilir.

Sunulan çalışmada D öznitelik vektörünü oluşturan özellikler sırasıyla; D_1: Kullback-Leibler sapması, D_2: Euclidean mesafesi, D_3: Jensen-Shannon sapması, D_4: Bhattacharya mesafesi, D_5: Spektral enerji farkı, D_6: Spektral entropi farkı, D_7: Spektral düzlük farkı, D_8: Spektral bant genişliği farkı, D_9: Maksimum spektral sapma, D_{10}: Sıfır geçiş oranı farkı olarak alınmıştır. Bu öznitelik vektörünü oluşturan 10 farklı özellik içinde Kullback-Leibler sapması özellikle sinyallerin frekans bileşimindeki küçük değişiklikleri iyi yakalar, bu da farklı sınıflar arasındaki spektral değişiklikleri belirlemekte çalışmada fayda sağlamıştır. Euclidean mesafesi ise spektral yoğunluklar arasında büyük farklar varsa, bu farkları iyi yansıtacaktır. Jensen-Shannon sapması sinyallerin genel benzerliğini iyi tespit etmesi nedeniyle tercih edilmiştir. Bhattacharya mesafesi özellikle spektral yoğunlukların büyük ölçüde örtüştüğü durumlarda farkları anlamamıza yardımcı olmuştur. Spektral enerji farkı, sinyalin genel güç yapısındaki farklılıkları incelememizi sağlamıştır. Spektral entropi, sinyallerin kaotik ya da düzenli olup olmadığını analiz eden bu teknikten EEG sinyallerinde, yüksek karmaşıklık içeren epileptik nöbet anlarının tespitinde yararlanılmıştır. Spektral düzlük, sinyalin spektrumunun ne kadar düz ya da tepecikli olduğunu ölçen bir yöntemdir. Daha düz bir spektrum, beyaz gürültü gibi sinyalleri işaret ederken, daha belirgin tepecikler daha düzenli frekans bileşenlerine işaret eder. Bu farkları analiz ederek, sınıflar arasındaki spektral yapıların belirlenmesinde spektral düzlükten faydalanılmıştır. Spektral bant genişliğinden, farklı sınıfların farklı frekans bantlarında yoğunlaşan sinyallerinin ayrılmasında yararlanılmıştır. Maksimum spektral sapma, referans sinyale göre sinyalin hangi frekans bileşenlerinde en büyük sapmayı gösterdiğini tespit eder. Bu özellik, sınıflar arasındaki en belirgin farkları ortaya koymaktadır. Zaman bölgesinde gerçekleştirdiğimiz tek analiz olan sıfır geçiş oranı ise sinyalin ne kadar sıklıkla sıfırdan geçtiğini ölçer ve bu da sinyaldeki yüksek frekanslı bileşenlerin varlığını işaret eder. Özellikle nöbet anlarındaki hızlı değişim içeren sinyalleri tespit etmekte kullanılmıştır.

Epileptik nöbetler ve sağlıklı durum arasındaki farklılıklar, beynin elektriksel aktivitesindeki anormalliklerden kaynaklanır. Bu durumlar genellikle inhibisyon-eksitasyon dengesinin bozulması, nöronal senkronizasyonun artması ve beyin hücrelerindeki iyon dengesizlikleri ile ilişkilidir. Özellikle nöbetlerin türüne, beynin etkilenen bölgesine ve bireyin fizyolojik durumuna bağlı olarak bu farklılıklar değişkenlik gösterebilir. Epileptik nöbetler sırasında sinyallerde ani değişimler ve frekans yoğunluğunda farklılıklar gözlemlenir. Bu durum, doğrusal sınıflandırıcılar yerine doğrusal olmayan ilişkileri öğrenebilen algoritmaların daha iyi performans göstermesini sağlamıştır. Veri setindeki E sınıfının (epileptik nöbet) diğer sınıflardan daha belirgin farklılıklar göstermesi, bazı sınıflandırıcıların bu farkları daha kolay algılamasına olanak tanımıştır. Örneğin, RO ve k-EYK (k=1) bu ayırımları daha hassas bir şekilde yakalamıştır. RO, karar ağaçlarından oluşan bir topluluk yöntemidir ve veri setindeki varyans ile hatayı azaltır. Özellikle karmaşık sınıflandırma görevlerinde (AB-CD-E, A-B-C-D-E gibi), sınıflar arasındaki sınırları etkili bir şekilde öğrenebilir. Çok sayıda özellikten bilgi çıkarabilmesi ve aşırı öğrenme riskini minimize etmesi nedeniyle, spektral özelliklerin yoğun kullanıldığı EEG sinyallerinde yüksek doğruluk sağlamıştır. k-En Yakın Komşu (k-EYK), parametrik olmayan yapısı sayesinde, veri setindeki ayrımları net bir şekilde yansıtmaktadır. Özellikle küçük k değerlerinde, sınıf sınırlarına yakın örnekleri daha hassas şekilde sınıflandırmıştır. Ancak k değerinin artırılmasıyla sınıflar arasındaki farklar bulanıklaşmıştır, bu da karmaşık görevlerde doğruluğun düşmesine neden olmuştur. Lojistik Model Ağacı (LMA), lojistik regresyonun esnekliği ile karar ağacının açıklanabilirliğini birleştirerek özellikle

düşük boyutlu veri setlerinde ve daha basit sınıflandırma görevlerinde güçlü performans göstermiştir. Çok Katmanlı Algılayıcı (ÇKA)'nın, çok katmanlı yapısı, özellikler arasındaki karmaşık ilişkileri öğrenme yeteneğini arttırmıştır. Ancak, eğitim sırasında parametre ayarlarının hassasiyet gerektirmesi nedeniyle bazı görevlerde performans dalgalanmaları görülmüştür. Destek Vektör Makineleri (DVM), doğrusal karar sınırlarına dayalı bir algoritmadır ve yüksek boyutlu veri setlerinde performans düşüşü yaşamıştır. Bu nedenle özellikle çok sınıflı görevlerde, sınıf sınırlarının bulanıklaşması nedeniyle doğruluk oranları düşmüştür. Epileptik nöbet sırasında sinyallerde yüksek frekanslı bileşenler, spektral enerji artışı ve düşük entropi gibi özellikler ön plana çıkmaktadır. Bu farklılıklar, özellikle RF ve k-EYK gibi algoritmalar tarafından daha iyi yakalanmıştır. Spektral düzeydeki bu belirgin farklar, nöbetli (E) ve nöbetsiz (A, B, C, D) sinyallerin kolay ayrılmasını sağlamaktadır. Sunulan çalışmada RO ve k-EYK'nin en yüksek performansı göstermesi, yukarıda açıklanan nedenlere dayanmaktadır. Özellikle nöbetli sinyallerin spektral enerji, entropi ve diğer istatistiksel özellikler açısından belirgin farklar içermesi, bu algoritmaların doğruluğunu artırmıştır.

Bazı sınıflandırma görevlerinde elde edilen %100 doğruluk oranı, modelin veri setindeki sınıflar arasındaki belirgin farkları başarılı bir şekilde yakaladığını göstermektedir. Özellikle nöbetli (E) ve nöbetsiz (A, B, C, D) sinyallerin spektral özelliklerindeki bariz farklılıklar, bu sonuçların temel nedenlerinden biridir. Bu durum, nöbetli sinyallerin güçlü frekans değişimleri, yüksek enerji yoğunluğu veya spektral entropi gibi özelliklerle nöbetsiz sinyallerden kolaylıkla ayrılabilmesine dayanmaktadır. Ayrıca, kullanılan veri setinin homojen ve gürültüsüz yapısı, doğru öznitelik çıkarma yöntemleri ve dengeli örnek dağılımı, modelin hata yapma olasılığını azaltarak bu yüksek başarıya katkı sağlamıştır.

Ancak, mevcut veri setinde her sınıf için 200 örnek bulunması ve bu örneklerin sınıf içi özelliklerinin oldukça tutarlı olması, modelin öğrenme sürecini kolaylaştırmış olabilir. Örnek sayısının artırılması, özellikle gerçek dünya uygulamalarında gözlemlendiği gibi, veri çeşitliliğinin artmasına ve potansiyel olarak daha karmaşık sinyal örüntülerinin sisteme dahil edilmesine yol açacaktır. Bu durumda, modelin %100 doğruluk oranını sürdürmesi daha zor hale gelebilir. Çünkü daha fazla veri, sınıf içi çeşitliliğin artmasına ve sınıflar arası ayrımların daha az belirgin hale gelmesine neden olabilir.

Bununla birlikte, veri sayısının artması, modelin genelleme yeteneğini ve gerçek dünya verilerine uygulanabilirliğini de geliştirebilir. Bu nedenle, %100 doğruluk oranlarının düşme ihtimali, modelin daha geniş ve çeşitli veri kümelerinde daha güvenilir bir performans sergileyebilmesi için bir gereklilik olarak da görülebilir.

Elde edilen sonuçların ışığı altında, çalışmada sunulan modelin otomatik nöbet tespitinde epileptik nöbetleri %100 doğruluk ile ayırarak çok başarılı olduğunu söylemek mümkündür. Genel olarak LMA ve k-EYK (k=1) algoritmaları, sınıflandırmada en yüksek performansı sunmuştur. Özellikle içinde epileptik nöbet anlarının yani E sınıfının bulunduğu 7 sınıflandırma görevinde %100 doğruluk ve 1,000 kesinlik değerlerinin elde edildiği k-EYK sınıflandırıcısının seçimi, modelin amacına daha uygun görülmektedir.

Etik Standartlar Bildirgesi

Yazarlar tüm etik standartlara uyduklarını beyan ederler.

Yazarlık Katkı Beyanı

Yazar 1: Kavramsallaştırma, Kaynaklar, Araştırma, Metodoloji, Doğrulama, Yazılım, Yazma – orijinal taslak, Görselleştirme, Biçimsel analiz, İnceleme ve düzenleme

Yazar 2: Kavramsallaştırma, Kaynaklar, Araştırma, Metodoloji, Deneyleme, Biçimsel analiz, Doğrulama, Görselleştirme, İnceleme ve düzenleme

Çıkar Çatışması Beyanı

Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

Verilerin Kullanılabilirliği

Bu çalışmada kullanılan veriler açık kaynaklıdır.

5. Kaynaklar

Al-Hadeethi, H., Abdulla, S., Diykh, M., Deo, R.C. and Green, J.H., 2020. Adaptive boosts LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications. *Expert Systems with Applications*, **161**, 113676. https://doi.org/10.1016/j.eswa.2020.113676

Alkan, A. and Kiymik, M.K., 2006. Comparison of AR and Welch methods in epileptic seizure detection. *Journal of Medical Systems*, **30**, 413-419. https://doi.org/10.1007/s10916-005-9001-0

Alotaiby, T.N., Alshebeili, S.A., Abd El-Samie, F.E., Alabdulrazak, A. and Alkhnaian, E., 2016. Channel selection and seizure detection using a statistical approach. *In 2016 5th international conference on electronic devices, systems and applications (ICEDSA)*. Ras Al Khaimah, United Arab Emirates, 1-4. https://doi.org/10.1109/ICEDSA.2016.7818505

Basri, A. and Arif, M., 2021. Classification of seizure types using random forest classifier. *Advances in Science and Technology Research Journal*, **15**(3), 167–178. https://doi.org/10.12913/22998624/140542

Birjandtalab, J., Pouyan, M.B., Cogan, D., Nourani, M. and Harvey, J., 2017. Automated seizure detection using

limited-channel EEG and non-linear dimension reduction. *Computers in biology and medicine*, **82**, 49-58. https://doi.org/10.1016/j.compbiomed.2017.01.011

Boubchir, L., Daachi, B. and Pangracious, V., 2017. A review of feature extraction for EEG epileptic seizure detection and classification. *In 2017 40th International Conference on Telecommunications and Signal Processing (TSP)*. Barcelona, Spain, 456-460. https://doi.org/10.1109/TSP.2017.8076027

Choubey, H. and Pandey, A., 2019. A new feature extraction and classification mechanisms For EEG signal processing. *Multidimensional Systems and Signal Processing*, **30**, 1793-1809. https://doi.org/10.1007/s11045-018-0628-7

Eltrass, A.S., Tayel, M.B. and EL-qady, A.F., 2021. Automatic epileptic seizure detection approach based on multi-stage Quantized Kernel Least Mean Square filters. *Biomedical Signal Processing and Control*, **70**, 103031.

Gabr, R.H., Shahin, A.I., Sharawi, A.A. and Aouf, M., 2020. A deep learning identification system for different epileptic seizure disease stages. *Journal of Engineering and Applied Science*, **67**(4), 925-944.

Gao, X., Yan, X., Gao, P., Gao, X. And Zhang, S., 2020. Automatic detection of epileptic seizure based on approximate entropy, recurrence quantification analysis and convolutional neural networks. *Artificial intelligence in medicine*, **102**, 101711. https://doi.org/10.1016/j.artmed.2019.101711

Harpale, V., and Bairagi, V., 2021. An adaptive method for feature selection and extraction for classification of epileptic EEG signal in significant states. *Journal of King Saud University-Computer and Information Sciences*, **33**(6), 668-676. https://doi.org/10.1016/j.jksuci.2018.04.014

Hasan, M.K., Ahamed, M.A., Ahmad, M. and Rashid, M.A., 2017. Prediction of epileptic seizure by Analysing time series EEG signal using k-NN classifier. *Applied bionics and biomechanics*, **2017**(1), 6848014. https://doi.org/10.1155/2017/6848014

Jana, G.C., Praneeth, M.S. and Agrawal, A., 2023. A multi-view SVM approach for seizure detection from single channel EEG signals. *IETE Journal of Research*, **69**(6), 3120-3131. https://doi.org/10.1080/03772063.2021.1913074

Ji, S., Zhang, Z., Ying, S., Wang, L., Zhao, X. and Gao, Y., 2020. Kullback–Leibler divergence metric learning. *IEEE transactions on cybernetics*, **52**(4), 2047-2058. https://doi.org/10.1109/TCYB.2020.3008248

Jiang, Y., Chen, W. and You, Y., 2020. Scattering transform-based features for the automatic seizure detection. *Biocybernetics and Biomedical Engineering*, **40**(1), 77-89. https://doi.org/10.1016/j.bbe.2019.11.002

Juba, B. and Le, H.S., 2019. Precision-recall versus accuracy and the role of large data sets. *In Proceedings of the AAAI conference on artificial intelligence*. Honolulu, Hawaii, USA, 4039-4048. https://doi.org/10.1609/aaai.v33i01.33014039

Kabir, E.S. and Zhang, Y., 2016. Epileptic seizure detection from EEG signals using logistic model trees. *Brain informatics*, **3**, 93-100. https://doi.org/10.1007/s40708-015-0030-2

Krislock, N. and Wolkowicz, H., 2012. Euclidean Distance Matrices and Applications. In Handbook on Semidefinite, Conic and Polynomial Optimization. International Series in Operations Research & Management Science, vol 166. Springer, New York, NY, 879-914. https://doi.org/10.1007/978-1-4614-0769-0_30

Kumar, A. and Kolekar, M.H., 2014. Machine learning approach for epileptic seizure detection using wavelet analysis of EEG signals. *In 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, Greater Noida, India, 412-416. https://doi.org/10.1109/MedCom.2014.7006043

Li, M., Sun, X., Chen, W., Jiang, Y. and Zhang, T., 2019. Classification epileptic seizures in EEG using time-frequency image and block texture features. *IEEE Access*, **8**, 9770-9781. https://doi.org/10.1109/ACCESS.2019.2960848

Liu, S., Wang, J., Li, S. and Cai, L., 2023. Epileptic seizure detection and prediction in EEGS using power spectra density parameterization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **31**, 3884-3894. https://doi.org/10.1109/TNSRE.2023.3317093

Mandhouj, B., Cherni, M.A. and Sayadi, M., 2021. An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis. *Analog integrated circuits and signal processing*, **108**(1), 101-110. https://doi.org/10.1007/s10470-021-01805-2

Menéndez, M.L., Pardo, J.A., Pardo, L. and Pardo, M.C., 1997. The jensen-shannon divergence. *Journal of the Franklin Institute*, **334**(2), 307-318. https://doi.org/10.1016/S0016-0032(96)00063-4

Milligan, T.A., 2021. Epilepsy: a clinical overview. *The American Journal of Medicine*, **134**(7), 840-847. https://doi.org/10.1016/j.amjmed.2021.01.038

Miltiadous, A., Tzimourta, K.D., Giannakeas, N., Tsipouras, M.G., Glavas, E., Kalafatakis, K. and Tzallas, A.T., 2022. Machine learning algorithms for epilepsy detection based on published EEG databases: A systematic review. *IEEE Access*, **11**, 564-594. https://doi.org/10.1109/ACCESS.2022.3232563

Mirzaei, A., Ayatollahi, A., Gifani, P. and Salehi, L., 2010. Spectral entropy for epileptic seizures detection. *In 2010 2nd International Conference on Computational Intelligence, Communication Systems and Networks*. Liverpool, UK, 301-307. https://doi.org/10.1109/CICSyN.2010.84

Niknazar, H., Mousavi, S.R., Niknazar, M., Mardanlou, V. and Coelho, B.N., 2020. Performance analysis of EEG seizure detection features. *Epilepsy Research*, **167**, 106483. https://doi.org/10.1016/j.eplepsyres.2020.106483

Nogay, H.S. and Adeli, H., 2021. Detection of epileptic seizure using pretrained deep convolutional neural network and transfer learning. *European neurology*, **83**(6), 602-614. https://doi.org/10.1159/000512985

Nwe, T.L., Hieu, N.T. and Limbu, D.K., 2013. Bhattacharyya distance based emotional dissimilarity measure for emotion classification. *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver*, BC, Canada, 7512-7516. https://doi.org/10.1109/ICASSP.2013.6639123

Pal, H. and Kumar, A., 2023. Stability analysis of multiscale bubble entropy and power metric-based seizure detection technique with MLA. *IETE Journal of Research*, **69**(6), 3455-3466. https://doi.org/10.1080/03772063.2021.1912650

Polat, K. and Nour, M., 2020. Epileptic seizure detection based on new hybrid models with electroencephalogram signals. *IRBM*, **41**(6), 331-353. https://doi.org/10.1016/j.irbm.2020.06.008

Paul, Y., 2018. Various epileptic seizure detection techniques using biomedical signals: a review. *Brain informatics*, **5**, 1-19. https://doi.org/10.1186/s40708-018-0084-z

Quintero-Rincón, A., d'Giano, C. and Batatia, H., 2020. A quadratic linear-parabolic model-based EEG classification to detect epileptic seizures. *Journal of Biomedical Research*, **34**(3), 205. https://doi.org/10.7555/JBR.33.20190012

Saini, J. and Dutta, M., 2018. Epilepsy classification using optimized artificial neural network. *Neurological Research*, **40**(11), 982-994. https://doi.org/10.1080/01616412.2018.1508544

Sharmila, A. and Mahalakshmi, P., 2017. Wavelet-based feature extraction for classification of epileptic seizure EEG signal. *Journal of medical engineering & technology*, **41**(8), 670-680. https://doi.org/10.1080/03091902.2017.1394388

Shiao, H.T., Cherkassky, V., Lee, J., Veber, B., Patterson, E. E., Brinkmann, B.H. and Worrell, G.A., 2016. SVM-based system for prediction of epileptic seizures from iEEG signal. *IEEE Transactions on Biomedical Engineering*, **64**(5), 1011-1022. https://doi.org/10.1109/TBME.2016.2586475

Siddiqui, M.K., Huang, X., Morales-Menendez, R., Hussain, N. and Khatoon, K., 2020. Machine learning based novel cost-sensitive seizure detection classifier for imbalanced EEG data sets. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 1491-1509. https://doi.org/10.1007/s12008-020-00715-3

Sriraam, N., Raghu, S., Tamanna, K., Narayan, L., Khanum, M., Hegde, A. S. and Kumar, A. B., 2018. Automated epileptic seizures detection using multi-features and multilayer perceptron neural network. *Brain Informatics*, **5**(2), 10. https://doi.org/10.1186/s40708-018-0088-8

Vidyaratne, L.S. and Iftekharuddin, K.M., 2017. Real-time epileptic seizure detection using EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **25**(11), 2146-2156. https://doi.org/10.1109/TNSRE.2017.2697920

Wang, Z., Na, J. and Zheng, B., 2020. An improved k-NN classifier for epilepsy diagnosis. *IEEE Access*, **8**, 100022-100030. https://doi.org/10.1109/ACCESS.2020.2996946

Zhang, J., Wei, Z., Zou, J. and Fu, H., 2020. Automatic epileptic EEG classification based on differential entropy and attention model. *Engineering Applications of Artificial Intelligence*, **96**, 103975. https://doi.org/10.1016/j.engappai.2020.103975

Zhao, X., Zhang, R., Mei, Z., Chen, C. and Chen, W., 2019. Identification of epileptic seizures by characterizing instantaneous energy behavior of EEG. *IEEE Access*, 70059-70076. https://doi.org/10.1109/ACCESS.2019.2919158

İnternet kaynakları

1- National Institute of Neurological Disorders and Stroke. http://www.ninds.nih.gov/(23.08.2024)

2- World Health Organization, https://www.who.int/, (23.08.2024)