



**Makale / Research Paper**

**EEG İşareti Kullanılarak Bağımlılığa Yatkınlığın  
Makine Öğrenmesi Teknikleri ile Analizi**

Veysel YARĞI<sup>1a</sup>, Seda POSTALCIOĞLU<sup>2b</sup>

<sup>1</sup>Bolu Abant İzzet Baysal Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü,  
Bolu/TÜRKİYE

<sup>2</sup>Bolu Abant İzzet Baysal Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Bolu/TÜRKİYE  
postalcioglu\_s@ibu.edu.tr

**Received/Geliş:** 29.08.2020

**Accepted/Kabul:** 12.10.2020

**Öz:** Alkol bağımlılığının Elektroensefalografi (EEG) sinyalleri ile teşhisi, hem kişisel açıdan hem de toplum açısından önemli bir konudur. Günümüzde birçok insan bu bağımlılıktan etkilenmektedir. Başta beyin, kalp ve bağışıklık sistemi olmak üzere fizyolojik etkileri olduğu gibi, psikolojik etkileri de söz konusudur. Bu etkileri gözlemleyebilmek için EEG sinyalleri etkin bir şekilde kullanılmaktadır. Bu çalışmada, alkolizme yatkınlığın EEG sinyalleri kullanılarak teşhisi yapılmıştır. Veri tabanı aracılığı ile elde edilen EEG sinyal üzerinde öncelikle veri analizi yapılmıştır. Özyinelemeli öznetelik seçimi gerçekleştirilmiştir. Sınıflandırma için Çok Katmanlı Yapay Sinir Ağları (ÇKYSA), Evrişimsel Sinir Ağları (ESA), XGBoost Algoritması (XGBA), Rassel Orman Algoritması (ROA), K-En Yakın Komşu Algoritması (K-EKA) kullanılmıştır. Python ortamında çalışılmıştır. Sınıflandırma başarımları için doğruluk, kesinlik, duyarlılık ve F1 Skor kullanılmıştır. Algoritmalar çalışma süreleri açısından karşılaştırılmıştır. Sınıflandırma başarımları açısından ÇKYSA ve ESA en iyi sonuçları vermiştir. Algoritmaların çalışma süreleri açısından bakıldığında XGBA en hızlı çalışan algoritma olduğu görülmüştür.

**Anahtar Kelimeler:** Elektroensefalografi (EEG); K-En Yakın Komşu; Rassel Orman Algoritması; XGBoost; Yapay Sinir Ağları

**Analysis of Predisposition to Addiction with Machine Learning  
Techniques Using EEG Signals**

**Abstract:** Diagnosis of alcohol dependence with Electroencephalography (EEG) signals is an important issue both personally and society. Today, many people are affected by this addiction. It has physiological effects, especially the brain, heart and immune system, as well as psychological effects. EEG signals are used effectively to observe these effects. In this study, genetic predisposition to alcoholism is diagnosed using EEG signals. Firstly, data analysis was performed on the EEG signal data obtained through the database. Recursive feature selection is used. For the classification, Multilayer Artificial Neural Networks (MLPNN), 1D-Convolutional Neural Networks (CNN), XGBoost Algorithm (XGBA), Random Forest Algorithm (RFA), K-Nearest Neighbor Algorithm (K-NN) are used. It has been studied in Python environment. Accuracy, precision, sensitivity and F1 Score are used for classification performance criteria. Algorithms are compared according to the running time. In terms of classification success, MLPNN and CNN gives the best results. In terms of running time of algorithms, XGBA is the fastest running algorithm.

**Keywords:** Elektroensefalografi (EEG); K-Nearest Neighbour; Neural Network; Random Forest Algorithm; XGBoost

*Bu makaleye atıf yapmak için*

Yarğı V., Postalcioglu S., "EEG İşareti Kullanılarak Bağımlılığa Yatkınlığın Makine Öğrenmesi Teknikleri ile Analizi" El-Cezerî Fen ve Mühendislik Dergisi, 2021, 8(1); 142-154.

*How to cite this article*

Yarğı V., Postalcioglu S., "Analysis of Predisposition to Addiction with Machine Learning Techniques Using EEG Signals" El-Cezerî Journal of Science and Engineering, 2021, 8(1); 142-154.

ORCID ID: <sup>a</sup>0000-0002-5722-5740, <sup>b</sup>0000-0002-3188-8116

## 1. Giriş

Alkol bađımlılıđından her yıl milyonlarca insan maddi ve manevi etkilenmektedir. Hem bađımlı olanın kendi fiziksel ve psikolojik zarar görmesinin yanı sıra bařta ailesi ve yakın çevresi olmak üzere tüm toplumu etkileyen büyük bir sorun olarak göze çarpmaktadır. Bu yüzden hastalığın teřhisi bu noktada çok önemlidir. Alkol bađımlılıđından en çok beyin fonksiyonları etkilenmektedir. Beyin görüntüleme ve sinyalleri sayesinde bu bađımlılıđın beyne ne denli zarar verdiđi günümüzde daha çok anlařılmaya başlanmıřtır. Alkol bađımlılıđının Elektroensefalografi (EEG) sinyalleri ile teřhisi, hem kiřisel açıdan hem de toplum açısından önemli bir konudur. Günümüzde birçok insan bu bađımlılıktan etkilenmektedir. Bařta beyin, kalp ve bađıřıklık sistemi olmak üzere fizyolojik etkileri olduđu gibi, psikolojik etkileri de söz konusudur. Bu etkileri gözlemleyebilmek için EEG sinyalleri etkin bir şekilde kullanılmaktadır. Bu çalıřmada EEG verisi ile genetik olarak alkole bađımlılıđın tespiti üzerinde durulmuřtur. Beř farklı makine öğrenmesi teknikleri ile sınıflandırma başarımlarının analizi ve bu başarımlar dođrultusunda algoritmaların çalıřma süreleri incelenmiřtir. Kullanılan teknikler; Çok Katmanlı Yapay Sinir Ađları (ÇKYSA), Evriřimsel Sinir Ađı (ESA), K-En Yakın Komřu Algoritması (K-EKA), XGBoost Algoritması (XGBA), Rassal Orman Algoritması (ROA)'dır.

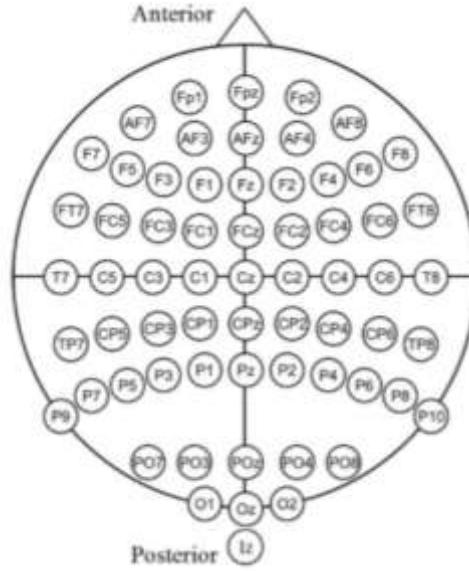
## 2. Literatür Özeti

Bu çalıřma kapsamında yapılan literatür taramasında EEG iřaretleri ve makine öğrenmesi teknikleri kullanarak gerçekleřtirilen çalıřmalar sunulmuřtur. Güler ve arkadaşları, çalıřmalarında ağır beyin yaralanmalarında travmanın řiddetinin tespit edilmesi için EEG sinyallerine ayrık zamanlı fourier dönüşümü uygulayıp, yapay sinir ađlarıyla sınıflandırma yapmıřtır. %87 oranında bir sınıflandırma performansı elde edilmiřtir [1]. Ersöz ve Özřen, yaptıkları çalıřmada EEG sinyallerini yapay sinir ađları yöntemiyle analiz edip uyku evrelerini tespit etmiřlerdir. Neticede dalgacık dönüşümü ve hızlı fourier dönüşümü için yakın sonuçlar elde edilmiřtir [2]. Sharanreddy ve Kulkarni, beyin tümörü olan hastalara ait EEG sinyallerini ve epilepsi nöbeti esnasında oluřan EEG sinyallerini analiz etmiřlerdir. Yapay sinir ađları ve çoklu dalgacık tekniđi kullanılmıřtır [3]. Kalaivani ve arkadaşları, yaptıkları çalıřmada beyindeki anormal aktiviteleri tespit etmek için, EEG sinyallerinden ayrık dalgacık dönüşümü tekniđi ile özellik çıkartımı yapıp kümeleme algoritmasıyla sınıflandırma iřlemi yapmıřtır [4]. Gajic ve arkadaşları, epileptik nöbetlerin teřhis edilmesi için istatistiksel örüntü tanıma ve dalgacık dönüşümü teknikleriyle EEG sinyalleri sınıflandırmıřtır [5]. O'Shea ve arkadaşları, yaptıkları çalıřmada konvölüsyonel sinir ađları ve destek vektör makineleri algoritmalarını yenidođan nöbet teřhisinde kullanmıřlardır. İki algoritmada da benzer başarımlar elde etmiřlerdir [6]. Vieira ve arkadaşları, beyin ile ilgili hastalıkların sınıflandırılmasında derin sinir ađlarıyla ilgili yapılan çalıřmaları incelemiřlerdir. Neticede derin sinir ađlarının psikolojik ve nörolojik hastalıkların teřhis edilmesinde kullanılabileceđini belirlemiřlerdir [7]. Chambon ve arkadaşları, uyku karakteristiklerini sınıflandırmak için, uyku anında ölçülen EEG sinyallerini derin sinir ađlarıyla analiz etmiřlerdir [8]. İpek, yaptıđı çalıřmasında, EEG sinyallerini konvölüsyonel sinir ađları ve destek vektörleri makinesi algoritmalarını sentezleyerek incelemiřtir [9]. Gül, çalıřmasında lösemili ve sađlıklı bireylere ait EEG sinyallerini dalgacık dönüşümü tekniklerini ve ileri beslemeli yapay sinir ađlarını kullanarak sınıflandırmıřtır. Bu çalıřmada dalgacık ailelerinin sınıflandırmaya etkisi ortaya koyulmuřtur [10]. Edenberg ve arkadaşları, genetik varyasyonların alkol bađımlılıđını ve aşırı tüketimini nasıl etkilediđini incelemiřlerdir [11]. Rieg ve arkadaşları alkol bađımlılıđını tespit eden bir araştırma yapmıřlardır. Rassal Orman algoritması ile sınıflandırılmıřtır [12]. Bavkar ve arkadaşları, yaptıkları çalıřmada alkolizm tespiti için EEG sinyallerini dar bant geçiren ve butterworth filtreleriyle daha kullanıřlı hale getirip özellikle alfa ve gamma bandlarındaki sinyallerin özelliklerinden faydalanarak hibrit bir özellik çıkarım metodu geliřtirmiřtir [13]. Wang ve arkadaşları, yaptıkları çalıřmada alkol bađımlılıđını tespit etmek için konvölüsyonel sinir ađlarını kullanarak başarımlarını artırmaya yönelik çalıřmalar yapmıřlardır [14]. Postalcıođlu ve Tepecik çalıřmalarında EEG ile kiřiselleřtirilmiř müzik listesi

Tasarımını gerçekleştirmişlerdir [15]. Kaya ve arkadaşları En Yakın Komşu Algoritması kullanarak EEG sinyallerine boyut azaltmanın etkilerini incelemişlerdir [16].

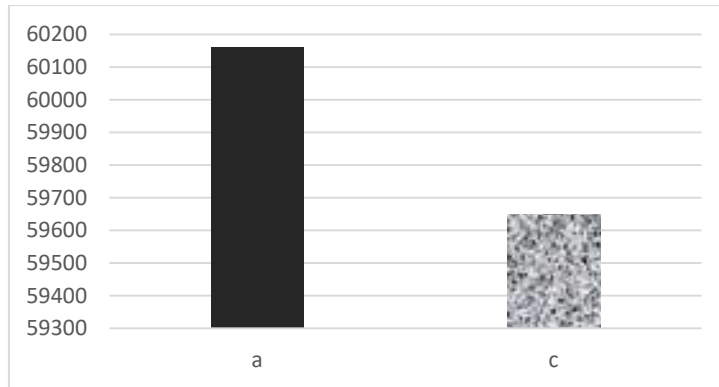
### 3. Materyal ve Metot

Çalışmada kullanılan Veri seti UCI veri tabanı üzerinden sağlanmıştır [17]. Bu veriler University of New York Health Center’da bulunan Neurodynamics Laboratory bölümünden elde edilmiştir. Deneklerin kafatası yüzeyine yerleştirilen elektrotların pozisyonu (uluslararası 10-20 sistemine göre) Şekil 1’deki gibi konumlandırılmıştır [18]. Ölçüm değerleri 256 Hz’de örneklenmiştir.



Şekil 1. Sensör konumları [18]

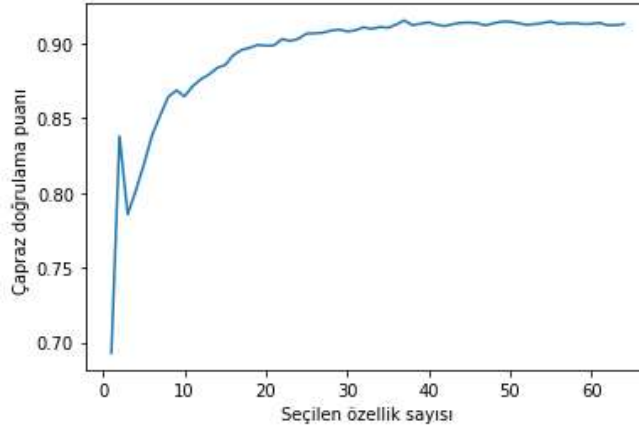
Alkolik grup (a) ve kontrol grubu (c) olmak üzere iki grup bulunmaktadır. EEG verileri için deneklere farklı resimler göstererek beyin dalgaları incelenmiştir. Şekil 1’deki sensörlerden gelen veriler kullanılmıştır. Eğitim verisindeki alkol grubu (a) ve kontrol grubu (c) sayıları Şekil 2’de gösterilmiştir. Alkolik özellik gösteren veri sayısı 60160 ve kontrol grubu olan veri sayısı ise 59648’dir.



Şekil 2. Eğitim veri setindeki alkol ve kontrol grubu dağılımı

Özyinelemeli Öznitelik Seçimi (ÖÖS) yöntemi oldukça popülerdir. ÖÖS sarmal (wrapper) tipi bir özellik seçim algoritmasıdır. Kullanımı ve konfigürasyonu basittir. Eğitim verilerinde bulunan sütunlardaki özelliklerden en önemlileri seçmede ve hedef değişkeni tahmin etmede oldukça etkili olan bir tekniktir. ÖÖS, eğitim verilerindeki tüm özellikleri analiz ederek, gereksiz tüm özellikleri kaldırır ve sadece en önemli özellikler kalana kadar devam eder. Böylece kullanılan makine

öđrenme algoritmalarının daha başarılı bir sınıflandırma yapmasını sağlar [19]. Çalışmada, çapraz doğrulama ile öz-yinelemeli öznitelik seçim yöntemi kullanılmıştır. Puanlama olarak doğruluk oranı temel alınmıştır. Şekil 3’de seçilen özellik sayısına göre çapraz doğrulama puanı görölmektedir.



Şekil 3. Özellik sayısına göre çapraz doğrulama puanı

Ham veri, öznitelik seçim öncesi 61,34 MB iken, öznitelik seçim sonrası 35,46 MB’a düşmüştür. Öz-yinelemeli öznitelik seçimine göre ‘AF1’, ‘AF2’, ‘AF7’, ‘C5’, ‘CP3’, ‘CP4’, ‘CP5’, ‘CP6’, ‘F1’, ‘F3’, ‘F4’, ‘F5’, ‘F6’, ‘F7’, ‘FC4’, ‘FC6’, ‘FPZ’, ‘FT8’, ‘OZ’, ‘P2’, ‘P6’, ‘P7’, ‘PO1’, ‘PO7’, ‘T7’, ‘T8’, ‘TP7’ en kötü özelliđe sahip kolonlar kaldırılmıştır.

Sınıflandırma işleminde Çok Katmanlı Yapay Sinir Ağları (ÇKYSA), Evrişimsel Sinir Ağları (ESA), K-Enyakın Komşu algoritması (K-EKA), karar ağaçlarından olan Rassal Orman Algoritması (ROA) ve eXtreme Gradient Boosting Algoritması (XGBA) kullanılmıştır.

ESA, klasik yapay sinir ağları modelinde kullanılan vektörler yerine matrisleri kullanır. Bunun en önemli avantajı ise, komşu tabakalar arasındaki bağlantıyı ve iletilen bilgilerin kaybolmasını engellemektir. Bir boyutlu ESA, ÇKYSA ve birbiriyle tamamen bađlı olan veya havuzlama işlemi gören bir veya daha fazla konvolüsyonel tabakaları kullanır. Konvolüsyon işlemi konvolüsyon tabakaları arasındaki her basamakta uygulanır. Bundan dolayı daha derin tabakalar her bir giriş bilgisinin özelliđine bađlı olarak çıkışı etkiler. Konvolüsyon tabakası matrisleri kullanarak özellik çıkarımı yapar. Bu tabakanın çıkışı bir özellik haritası çıkarır. Havuzlama katmanı önemli bilgileri elde etmek için özelliklerin sayısını düşürür [20].

K-En Yakın Komşu (K-EK) algoritması, veriler arasındaki mesafenin ölçülmesi aracılıđıyla sınıflandırma yapar. Algoritmanın performansı k en yakın komşu sayısı, mesafe ölçütü ve öznitelik sayısı gibi deđişkenlere bađlı olarak deđişiklik gösterir. K en yakın komşu sayısının nasıl seçileceđine dair belli bir standart bulunmamaktadır. Deneme yanılma yöntemiyle veya sezgisel olarak belirlenebilir [21]

Karar ağaçları sınıflandırma işleminde kullanıldıđı zaman, yařanan en büyük problemlerden biri dođru ağaç yapısını her zaman için oluřturamamaktır. Bu yüzden sınıflandırma sonucu istenilen başarı elde edilemeyebilir. Bu olumsuz etkiyi ortadan kaldırmak veya minimize etmek için Rassal Orman Algoritması geliştirilmiştir. Rassal Orman Algoritması ilk 1996 yılında Breiman tarafından ortaya atılmıştır. Rassal Ormanı oluřtururken, rasgele vektörler kullanılarak ağaçlar geliştirilmiş ve veri setinden rastgele seçim için ise bagging (torbalama) tekniđi kullanılmıştır [22].

XGBoost (eXtreme Gradient Boosting) algoritması Chen ve He tarafından 2015’de geliştirilen kollektif bir ağaç algoritmasıdır. XGBoost algoritması farklı derinlikte karar ağaçları oluřturmadaki

ve aynı anda farklı algoritmaları optimize edebilmesindeki performansı sayesinde regresyon ve sınıflandırma problemlerinde etkili ve verimli çözümler sunar [23].

### 3.1 Sınıflandırma Başarım Ölçütleri

Karışıklık matrisi verilerin sınıflandırılması sonucu çıktılarının en az iki veya daha fazla olduğu durumlarda performans ölçütü olarak kullanılabilir. Tablo 1 karışıklık matrisini ifade etmektedir.

**Tablo 1.** Karışıklık matrisi

Tahmin Edilen Değerler	Gerçek Değerler	
	Pozitif	Negatif
P	Gerçek Pozitif (GP)	Yanlış Pozitif (YP)
N	Yanlış Negatif (YN)	Gerçek Negatif (GN)

Gerçek Pozitif (GP), Verideki alkolik birine ait bilgiyi algoritma doğru tahmin ederse. Yani “bu veri alkolizme yatkın birine aittir” diye tahmin yaparsa, buna gerçek pozitif denir. Gerçek Negatif (GN), Verilerdeki kontrol grubuna ait birini, algoritma doğru tahmin ederse, yani “bu veri alkolik olmayan birine aittir” diye tahmin yaparsa, buna gerçek negatif denir. Yanlış Pozitif (YP), veri setindeki alkolik olmayan birini algoritma sınıflandırırken, “bu kişi alkoliktir” diye tahminde bulunursa, buna yanlış pozitif denir. Yanlış Negatif (YN), verilerdeki alkolik olan birisini algoritma sınıflandırırken, “bu kişi alkolik değildir” diye tahmin yaparsa, bu durum da yanlış negatif olarak değerlendirilir. Doğruluk değeri, modelde doğru tahmin ettiğimiz alanların toplam veri kümesine oranı ile hesaplanmaktadır. Eşitlik (1) ‘de doğruluk denklemi gösterilmiştir [24].

$$\text{Doğruluk} = \frac{GP+GN}{GP+YP+GN+YN} \quad (1)$$

Kesinlik, pozitif olarak tahmin edilen değerlerin gerçekten kaç adedinin pozitif olduğunu ifade eder. Eşitlik (2) ‘de kesinlik denklemi gösterilmiştir [24].

$$\text{Kesinlik} = \frac{GP}{GP+YP} \quad (2)$$

Duyarlılık ise pozitif olarak tahmin edilenlerin, işlemlerin ne kadarını pozitif olarak tahmin ettiğimizi gösteren bir metriktir [24]. Eşitlik (3) ‘de duyarlılık denklemi verilmiştir.

$$\text{Duyarlılık} = \frac{GP}{GP+YN} \quad (3)$$

Kesinlik ve Duyarlılık değerlerinin harmonik ortalamasını göstermektedir. Harmonik ortalama kullanılmasının sebebi çok uç durumları göz önünde bulundurma zorunluluğudur. Basit bir aritmetik ortalama bizi yanıltabilir [25, 24]. Eşitlik (4) ‘de ifade verilmiştir.

$$F1 - \text{Skor} = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4)$$

## 4. Bulgular

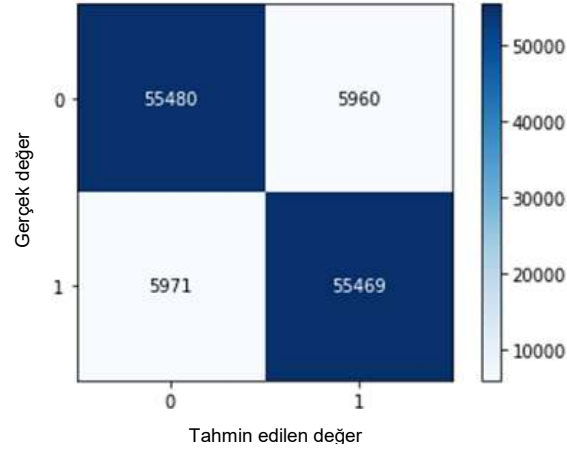
Çalışmalar, işlemcisi Intel core-i7, 6500U CPU 2.50 GHz, 8GB Ram’i olan dizüstü bilgisayarda gerçekleştirilmiştir. Çok Katmanlı Yapay Sinir Ağı (ÇKYSA) 4 katmandan oluşmaktadır. İlk 3 katmanda ReLu, çıkış katmanında ise sigmoid fonksiyonu kullanılmıştır. Ağda kayıp fonksiyonu olarak ortalama karesel hata değeri kullanılmıştır. Eğitim süresi 17dk 16sn, test süresi 1,32sn

sürmüştür. Ortalama karesel hata değeri ise 0.0870 olarak hesaplanmıştır. Tablo 2’de başarıml sonuçları verilmiştir.

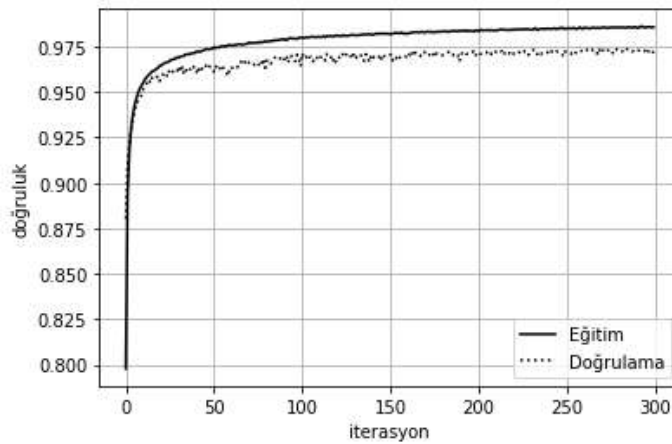
**Tablo 2.** ÇKYSA için başarıml sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skor	Dođruluk
a-0	%90	%90	%90	
c-1	%90	%90	%90	%90

Şekil 4’de ÇKYSA için Karışıklık matrisi verilmiştir. Şekil 5 ve Şekil 6’da ise modelin dođruluk ve kayıp eğrileri görülmektedir.



**Şekil 4.** ÇKYSA için karışıklık matrisi

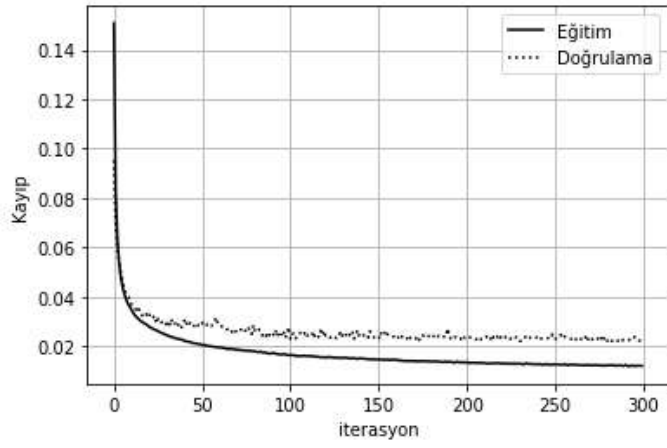


**Şekil 5.** ÇKYSA için modelin dođruluk eğrisi

1D-Evrişimsel Sinir Ağ (ESA) yapısı kullanılmıştır. Giriş ve ara katmanlarda Relu, çıkışta sigmoid fonksiyonu tercih edilmiştir. Ağda kayıp fonksiyonu olarak ortalama karesel hata değeriinden faydalanılmıştır. Ortalama karesel hata değeri ise 0.0864 olarak hesaplanmıştır. Eğitim süresi 21 dk 53sn, test süresi 1.68sn sürmüştür. Tablo 3’de başarıml sonuçları verilmiştir.

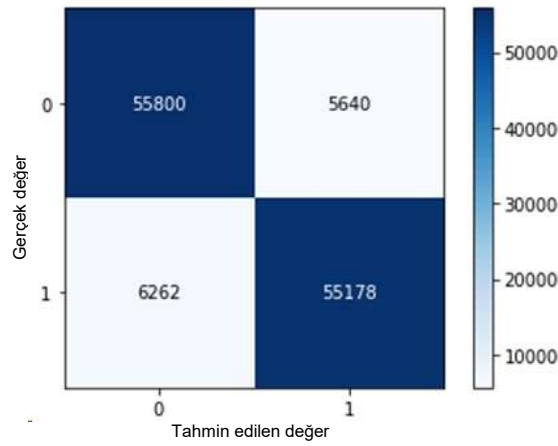
**Tablo 3.** ESA için başarıml sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skor	Dođruluk
a -0	%90	%91	%90	
c -1	%91	%90	%90	%90

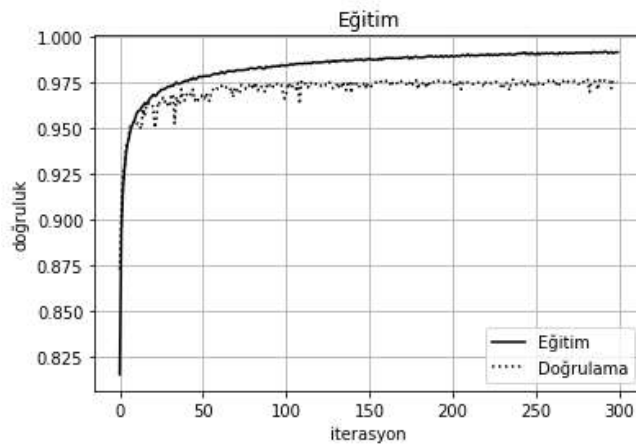


Şekil 6. ÇKYSA için modelin kayıp eğrisi

Şekil 7' de karışıklık matrisi verilmiştir. Şekil 8 ve Şekil 9'da ESA için modelin doğruluk ve kayıp eğrisi gösterilmiştir.



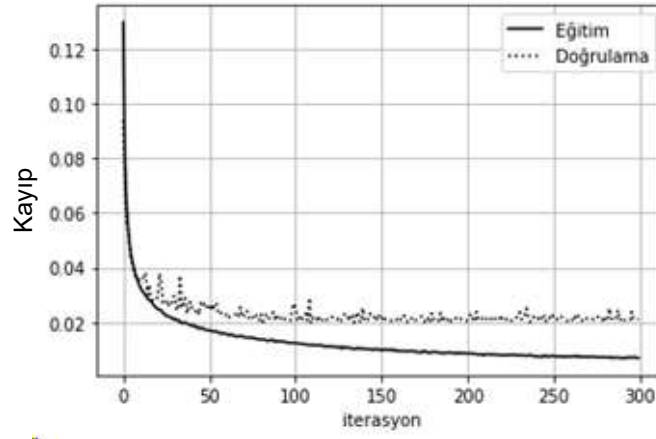
Şekil 7. ESA için karışıklık matrisi



Şekil 8. ESA için modelin doğruluk eğrisi

Tablo 4. XGBoost için başarımlar sonuçları

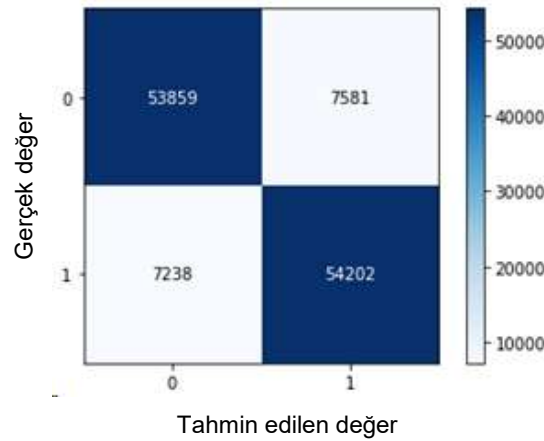
Sınıf	Kesinlik	Duyarlılık	F1-Skor	Doğruluk
a - 0	%88	%88	%88	
c - 1	%88	%88	%88	%88



Şekil 9. ESA için modelin kayıp eğrisi

Karar Ağacı temelli XGBoost algoritma (XGBA) yapısı kullanılmıştır. Derinlik 50 olarak alınmıştır. Ortalama karesel hata değeri ise 0.120 olarak hesaplanmıştır. Eğitim süresi 3 dk 65sn, test süresi 1.44 sn sürmüştür. Tablo 4'te XGBoost için başarımları sonuçları verilmiştir.

Şekil 10'da, XGBoost için karışıklık matrisi verilmiştir.



Şekil 10. XGBoost için karışıklık matrisi

Rassal Orman Algoritması (ROA) için ağaç sayısı 400 olarak atanmıştır. Yapılan benzetimler sonucu ortalama karesel hata değeri ise 0.141 olarak hesaplanmıştır. Eğitim süresi 9 dk 34sn, test süresi 20.59 sn sürmüştür. Tablo 5'te Rassal Orman Algoritması için başarımları sonuçları verilmiştir.

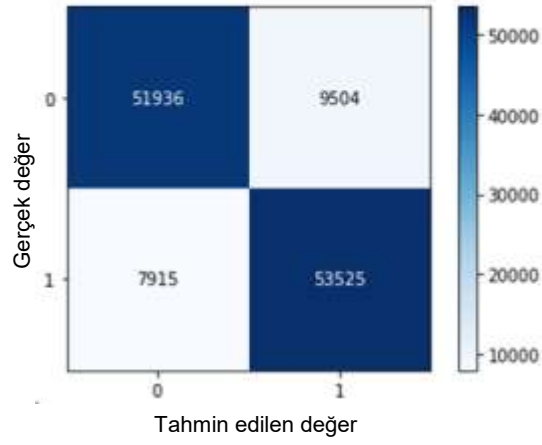
Tablo 5. Rassal Orman algoritması için başarımları sonuçları

Sınıf	Kesinlik	Duyarlılık	F1-Skor	Doğruluk
a -0	%87	%85	%86	%86
c -1	%85	%87	%86	

Şekil 11'de karışıklık matrisi gösterilmiştir.

K En Yakın Komşuluk (K-EKA) algoritması için komşu sayısı 5 olarak atanmıştır. Ortalama karesel hata değeri 0.2132 olarak hesaplanmıştır. Eğitim süresi 0.554 sn, test süresi 13dk 69 sn sürmüştür. Tablo 6'da başarımları sonuçları sunulmuştur.



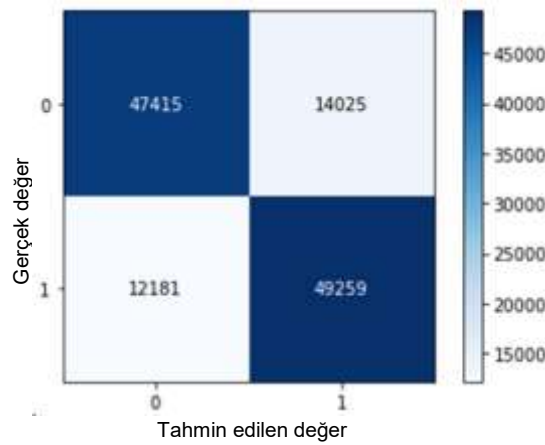


Şekil 11. ROA için karışıklık matrisi

Tablo 6. K-EKA başarımları sonuçları

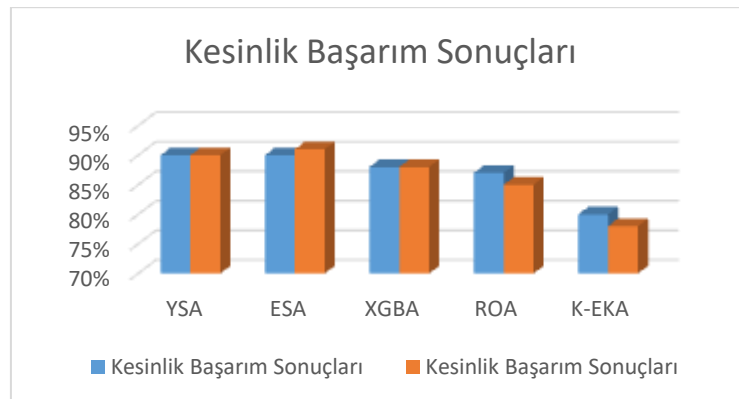
	Kesinlik	Duyarlılık	F1-Skor	Doğruluk
<b>a-0</b>	%80	%77	%78	%79
<b>c-1</b>	%78	%80	%79	

Şekil 12’de karışıklık matrisi verilmiştir.



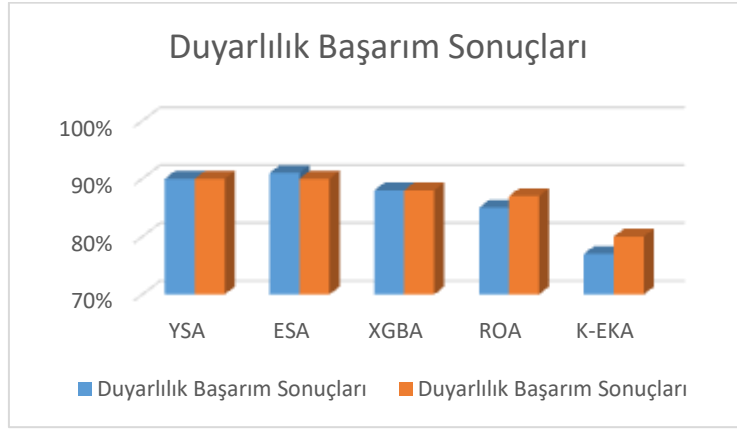
Şekil 12. K-EKA için karışıklık matrisi

Şekil 13’de kullanılan algoritmaların kesinlik başarımları sonuçları verilmiştir. Kesinlik başarımları en yüksek ESA ve YSA olduğu görülmektedir.



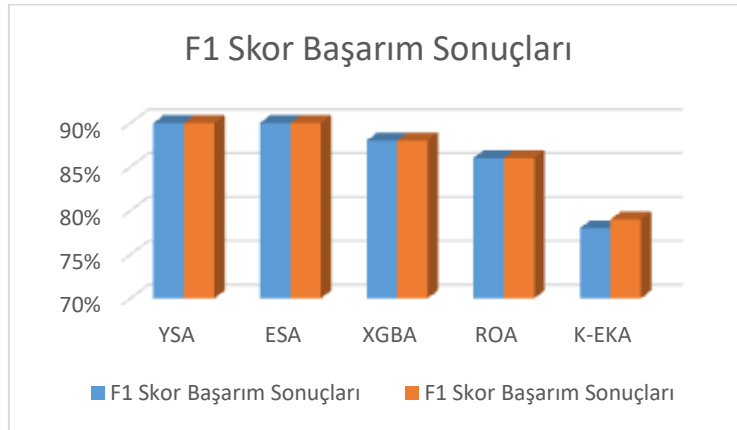
Şekil 13. Kesinlik başarımları sonuçları

Őekil 14’de kullanılan algoritmaların duyarlılık başarımları gösterilmiştir. YSA ve ESA duyarlılık başarımları en yüksektir.



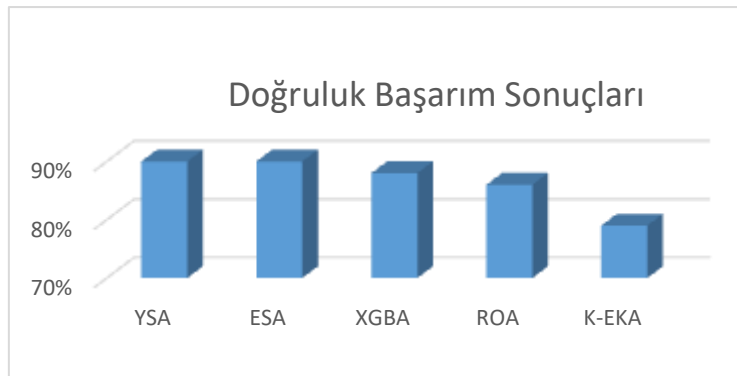
**Őekil 14.** Duyarlılık başarımları

Őekil 15’de kullanılan algoritmaların F1 Skor sonuçları gösterilmiştir. YSA ve ESA başarımları en yüksektir.



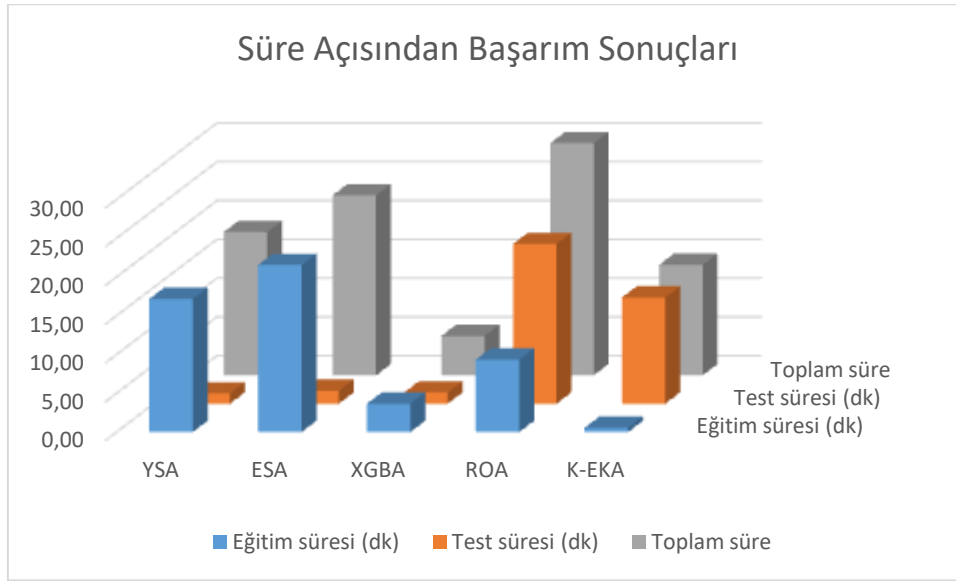
**Őekil 15.** F1 Skor başarımları

Őekil 16’da doğruluk başarımları verilmiştir. Őekil 16’ya bakıldığında YSA ve ESA doğruluk açısından diğerlerine göre %90 oranında başarımları göstermiştir. En düşük başarımları ise K-EKA göstermiştir.



**Őekil 16.** Dođruluk başarımları

Şekil 17’de algoritmaların eğitim ve test süreleri açısından başarımları verilmiştir. Eğitim süreleri açısından analiz edildiğinde, ESA eğitim süresi en yüksek, K-EKA eğitim süresi en düşük çıkmıştır. Test süreleri açısından ise ROA en fazla süreyi içermektedir. Toplam süre açısından bakıldığında ise en hızlı XGBA çıkmıştır.



Şekil 17. Süre açısından başarımları

## 5. Sonuç ve Öneriler

Bu çalışmada EEG işaretleri kullanılarak bağımlılığa yatkınlığın farklı makine öğrenmesi teknikleri ile sınıflandırma başarımları ve çalışma süreleri değerlendirilmiştir. EEG verileri alkolizme olan yatkınlığın EEG ile ilişkisini içermektedir. Veriseti, alkolik ve kontrol olmak üzere iki sınıftan oluşmaktadır. Çalışmada ilk olarak veri analiz edilmiştir. Dağınık ve çok sayıda olan csv dosyaları tek bir csv dosyasına dönüştürülerek daha düzenli bir veriseti elde edilmiştir. Veriler arasındaki korelasyon incelenmiş, öznelik seçimi yapılmıştır. L2 normalizasyon kullanılmıştır. Öznelik seçiminde öz yinelenmeli öznelik seçimi kullanılmıştır. Kullanıma hazır olan veri için beş farklı makine öğrenme tekniği kullanılarak sınıflandırma yapılmıştır. Kullanılan teknikler; Çok Katmanlı Yapay Sinir Ağları (ÇKYSA), Evrimsel Sinir Ağı (ESA), K-En Yakın Komşu Algoritması (K-EKA), XGBoost Algoritması (XGBA), Rastgele Orman Algoritması (ROA)’dır. ÇKYSA, eğitim süresi 17dk 16sn, test süresi 1,32sn sürmüştür. Ortalama karesel hata değeri ise 0.0870 olarak hesaplanmıştır. Doğruluk oranı %90’dır. ESA için ortalama karesel hata değeri ise 0.0864 olarak hesaplanmıştır. Eğitim süresi 21 dk 53sn, test süresi 1.68sn sürmüştür. Doğruluk oranı %90 elde edilmiştir. XGBA için ortalama karesel hata değeri ise 0.120 olarak hesaplanmıştır. Eğitim süresi 3 dk 65sn, test süresi 1.44 sn sürmüştür. Doğruluk oranı %88 olarak elde edilmiştir. ROA, ortalama karesel hata değeri ise 0.141 olarak hesaplanmıştır. Eğitim süresi 9 dk 34sn, test süresi 20.59 sn sürmüştür. Doğruluk oranı %86 çıkmıştır. K-EKA, ortalama karesel hata değeri 0.2132 olarak hesaplanmıştır. Eğitim süresi 0.554 sn, test süresi 13dk 69 sn sürmüştür. Doğruluk oranı %79 elde edilmiştir. Sonuçlara bakıldığında ortalama karesel hata en düşük ESA’da elde edilmiştir. Eğitim süresi en uzun ESA’da, en kısa ise K-EKA’da elde edilmiştir. Test süresi en fazla süren ROA olurken, en az süren ise ÇKYSA’dır. Toplam süre açısından bakıldığında ise en hızlı XGBA çıkmıştır. Doğruluk oranı açısından en iyi sonuçları ÇKYSA ve ESA vermiştir. Sonraki çalışmalar olarak farklı öznelik çıkarım teknikleri uygulanabilir, tasarlanan modellerdeki hiper parametreler değiştirilerek başarımları daha da artırmaya yönelik çalışmalar yapılabilir.

EEG sinyalleri üzerinde yapılan alıřmalar incelendiđinde sıklıkla hastalıđı daha dođru tespit etmek iin alıřmalar yrtlmřtr. Bu alıřmayı diđer alıřmalardan ayıran fark ise bađımlılık zerine bir arařtırma gerekleřtirmemizdir. İnsan beyni grdđ ve duyduđu her řeyden etkilenmektedir. Bađımlılık tedavisinde genetik yatkınlıkların EEG ile iliřkisi belirlenebilirse hangi sinyallerden beynin olumlu etkilendiđi tespit edilebilir ve bađımlı olan ya da bađımlı olmaya aday bireyleri bađımlılık yaratacak kt alıřkanlıklardan koruyabilir veya iyileřtirebiliriz.

## Teřekkr

Bu alıřma Bolu Abant İzzet Baysal niversitesi, Lisansst Eđitim Enstits'nde yapılan "EEG İřaretleri Kullanılarak Bađımlılıđa Yatkınlıđın Makine đrenmesi Teknikleri ile Deđerlendirilmesi" adlı yksek lisans tezinden retilmiřtir.

## Kaynaklar

- [1]. Gler, İ., Gkil, Z., Glbandılar, E., Evaluating of Traumatic Brain Injuries using Artificial Neural Networks, Expert Systems with Applications, 2009, 36(7):10424-10427.
- [2]. Ersz, A., zřen, S., Uyku EEG Sinyalinin Yapay Sinir Ađ Modeli İle Sınıflandırılması, Elektrik-Elektronik ve Bilgisayar Sempozyumu, 2011, Elazıđ, Trkiye, ss. 298-301.
- [3]. Sharanreddy, M., Kulkarni, P., Automated EEG Signal Analysis for Identification of Epilepsy Seizures and Brain Tumour, Journal of Medical Engineering & Technology, 2013, 37(8): 511-519.
- [4]. Kalaivani M., Kalaivani V., Devi V.A., Analysis of EEG Signal for the Detection of Brain Abnormalities, International Journal of Computer applications (IJCA), 2014, 2: 1-6.
- [5]. Gajic D., Djurovic Z., Gennaro S.D., Gustafsson F., Classification of EEG Signals For Detection of Epileptic Seizured Based on Wavelets and Statistical Pattern Recognition, Biomedical Engineering Applications, Basis and Communications, 2014, 26,(2):145021.
- [6]. O'shea A., Neonatal Seizure Detection Using Convolutional Neural Networks, <https://arxiv.org/abs/1709.05849>, eriřim tarihi: 30 Ađustos 2020.
- [7]. Vieira S., Pinaya W.H.L., Mechelli A., Using Deep Learning to Investigate the Neuroimaging Correlates of Psychiatric and Neurological Disorders Methods and Applications, Neuroscience& Biobehavioral Reviews, 2017, 74: 58-75.
- [8]. Chambon S., Galtier MN., Arnal P.J., Wainrib G., Gramfort A., A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018, 26(4):758-769.
- [9]. İpek B., EEG Sinyallerinin Epileptik Rahatsızlıkların Teřhisi iin Konvolsyonel Sinir Ađları ve Destek Vektr Makineleri ile Tasnif Edilmesi, Yksek Lisans Tezi, Karatay niversitesi, Fen Bilimleri Enstits, 2018, Konya.
- [10]. Gl E., EEG Sinyallerinin Wavelet Yntemiyle Dnřtrlerek Yapay Sinir Ađları ile Sınıflandırılması, Yksek Lisans Tezi, Eskiřehir Osmangazi niversitesi, Sađlık Bilimleri Enstits, 2018, Eskiřehir.
- [11]. Edenberg H.J., Gelernter J., Agrawal A., Genetics of Alcoholism, Current Psychiatry Reports, 2019, 21(26):1-7.
- [12]. Rieg T., Frick J., Hitzler M., Buettner R., High-performance detection of alcoholism by unfolding the amalgamated EEG spectra using the Random Forests method, Proceedings of the 52<sup>nd</sup> Hawaii International Conference on Systems Sciences, 2019, Honolulu United States, pp. 3769-3777.
- [13]. Bavkar S., Iyer B., Deosarkar S., Detection of Alcoholism: An EEG Hybrid Features and Ensemble Subspace K-NN Based Approach, International Conference on Distributed Computing and Internet Technology, Lecture Notes in Computer Science, 2019, 11319: 161-168.

- [14]. Wang S.H., Muhammad K., Hong J., Sangalah A.K., Zhang Y.D., Alcoholism identification via convolutional neural network based on parametric ReLU dropout and batch normalization, *Neural Computing and Applications*, 2020, 32: 665-680.
- [15]. Postalcioglu S., Tepecik H.H., EEG ile Kişiselleştirilmiş Müzik Listesi Tasarımı, 3rd International Conference on Data Science and Applications, June 25-28 2020, Istanbul, Turkey, ss.197-201.
- [16]. Kaya, D., Türk, M., Kaya, T., En Yakın Komşu Algoritması Kullanılarak EEG Sinyallerine Boyut Azaltmanın Etkilerinin İncelenmesi, *El-Cezeri Journal of Science and Engineering* 5, 2018: 591-595.
- [17]. UCI, <https://archive.ics.uci.edu/ml/datasets/eeg+database>, erişim tarihi: 10 Haziran 2020.
- [18]. Ge S., Yang Q., Wang R., Lin P., Gao J., Leng Y., Yang Y., Wang H., A Brain-Computer Interface Based on a Few-Channel EEG-fNIRS Bimodal System, *IEEE Access*, 5, 2017: 208-218.
- [19]. Brownlee J., Recursive Feature Elimination (RFE) for Feature Selection in Python, <https://machinelearningmastery.com/rfe-feature-selection-in-python/>, erişim tarihi: 28 Haziran 2020.
- [20]. Choubey R.N., Amar L., Khare S., Internet traffic classifier using artificial neural network and 1D-CNN, *International Conference on Information Technology (ICIT)*, 2019, Bhubaneswar, India, pp. 291-296.
- [21]. Bhatia N., Vandana A., Survey of Nearest Neighbor Techniques , *International Journal of Computer Science and Information Security*, 2010, 8(2): 302-305.
- [22]. Breiman L., Bagging Predictors, *Machine Learning*, 1996, 24: 123-140.
- [23]. Zhou J., Li E., Wang M., Chen X., Shi X., Jiang L., Feasibility of Stochastic Gradient Boosting Approach for Evaluating Seismic Liquefaction Potential Based on SPT and CPT Case Histories, *Journal of Performance of Constructed Facilities*, 2019, 33: 04019024.
- [24]. Powers W., Ailab A., Evaluation: from precision, recall and F-measure to ROC informedness, markedness and correlation, *J. Mach. Learn. Technolgy*, 2008, 2: 2229-3981.
- [25]. Gülcan O., Doğruluk (Accuracy) Kesinlik (Precision) Duyarlılık (Recall) F1 Score, <https://medium.com/@gulcanogundur/do%C4%9Fruluk-accuracy-kesinlik-precisionduyarl-%C4%B11%C4%B1k-recall-ya-da-f1-score-300c925feb38>, erişim tarihi: 24 Haziran 2020.