



Makale / Research Paper

Elektrokardiyografi Yardımıyla Hipertansiyonun Otomatik Belirlenmesinde Ampirik Kip Ayırışımının Gürültülü ve Gürültüsüz Sinyaller Üzerindeki Performansının Karşılaştırılması

Yunus Emre Erdoğan^{1,2a}, Ali Narin^{2b}

¹Elektronik Otomasyon Müdürlüğü, Ereğli Demir Çelik AŞ, Zonguldak, Türkiye

²Elektrik Elektronik Mühendisliği Bölümü, Zonguldak Bülent Ecevit Üniversitesi, Zonguldak, Türkiye

yeerdogan@erdemir.com.tr

Received/Geliş: 14.10.2021

Accepted/Kabul: 14.12.2021

Öz: Hipertansiyon (HPT), kalpten vücuda taşınan kanın atardamar duvarlarına uyguladığı kuvvetin, bazı hastalıklara sebebiyet verecek kadar yüksek olduğu duruma verilen isimdir. HPT'ye bağlı hastalıklar sonucunda her yıl dünyada birçok insan hayatını kaybetmektedir. Bu sebepten dolayı HPT'nin erken teşhis edilmesi oldukça kritik bir öneme sahiptir. Bu çalışma, elektrokardiyogram (EKG) sinyalleri kullanılarak HPT hastalarının otomatik ve en az hata ile tespit edilmesi amacıyla yapılmıştır. EKG sinyalleri 4 farklı grupta incelenmiştir. Bu incelemeler sırasıyla, normalize edilmiş gürültülü EKG sinyalleri, normalize edilmiş gürültüsüz EKG sinyalleri, normalize edilmemiş gürültülü EKG sinyalleri ve son olarak normalize edilmemiş gürültüsüz EKG sinyalleridir. Ampirik Kip Ayırışımı metodu kullanılarak 5 katmanlı iç mod fonksiyon (İMF) sinyalleri elde edilmiştir. Her bir İMF sinyali üzerinden elde edilen entropi ölçümleri yardımıyla HPT hastalarının Normal kişilerden ayırt edilme performansları analiz edilmiştir. 10-kat çapraz doğrulama metodundan yararlanılarak destek vektör makineleri (DVM), k-en yakın komşu (k-NN) ve karar ağaçları algoritmaları ile sonuçlar elde edilmiştir. En yüksek doğruluk değeri gürültüsüz ve normalize edilmiş sinyaller kullanılarak, DVM ve k-NN algoritmaları ile %99,99 olarak elde edilmiştir. Gürültüsüz ve normalize edilmiş veriler ile elde edilen böylesine yüksek performans değerlerinin HPT hastalarının tespit edilmesinde, uzman kişilere karar destek sistemi olarak fayda sağlayacağını düşünmekteyiz.

Anahtar Kelimeler: Hipertansiyon; EKG; Ampirik Kip Ayırışımı; İç Mod Fonksiyonları; Sınıflandırma.

Comparison of Performance of Empirical Mode Decomposition on Noisy and Noiseless Signals in Automatic Detection of Hypertension with the Assistance of Electrocardiography

Abstract: Hypertension (HPT) is the name given to the condition in which the force exerted by the blood carried from the heart to the body on the arterial walls is high enough to cause some diseases. As a result of HPT-related diseases, many people die every year around the world. For this reason, early diagnosis of HPT is of critical importance. This study was conducted to detect HPT patients automatically and with minimum error by using electrocardiogram (ECG) signals. ECG signals were analyzed in 4 different groups. These examinations are respectively normalized noisy ECG signals, normalized noiseless ECG signals, unnormalized noisy ECG signals and finally unnormalized noiseless ECG signals. Using the Empirical Mode Decomposition method, 5-layer intrinsic mode function (IMF) signals are obtained. With the help of entropy measurements obtained from each IMF signal, HPT patients' discrimination performances from Normal people were analyzed. By using the 10-fold cross validation method, the results were obtained with support vector machines (SVM), k-nearest neighbor (k-NN) and decision tree algorithms. The highest accuracy value was obtained as 99.99% with SVM and k-NN algorithms using noise-free and normalized signals. We think that such high performance values obtained with noiseless and normalized data will be beneficial to experts as a decision support system in the detection of HPT patients.

Keywords: Hypertension; ECG; Empirical Mode Decomposition; Intrinsic Mode Functions; Classification

Bu makaleye atf yapmak için

Erdogan, Y.E., Narin, A., "Elektrokardiyografi Kullanarak Hipertansiyonun Otomatik Tespitinde Deneysel Mod Ayırışımının Gürültülü ve Gürültüsüz Sinyaller Üzerindeki Performansının Karşılaştırılması" El-Cezerî Fen ve Mühendislik Dergisi 2022, 9 (2); 788-800.

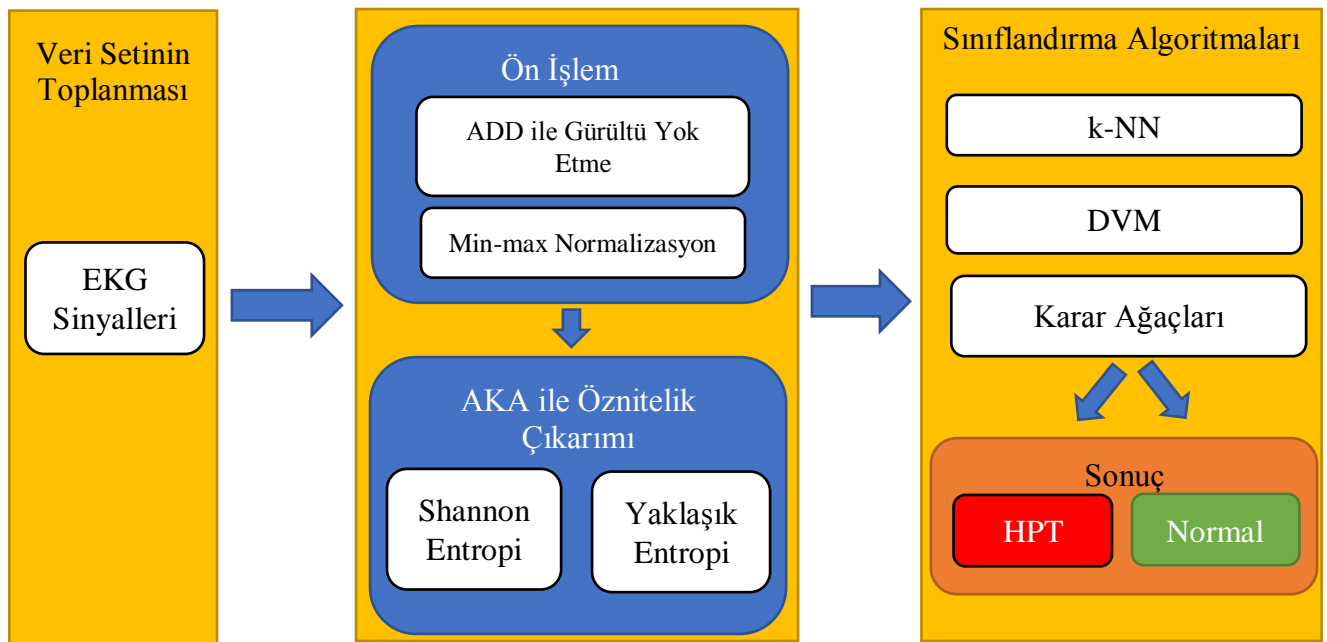
How to cite this article

Erdogan, Y.E., Narin, A., "Comparison of the Performances of Empirical Mode Decomposition on Noisy and Noiseless Signals for Automated Detection of Hypertension using Electrocardiography" El-Cezerî Journal of Science and Engineering, 2022, 9 (2); 788-800.

ORCID ID: ^a0000-0003-3677-5564; ^b0000-0003-0356-2888

1. Giriş

Hipertansiyon (yüksek tansiyon) (HPT), atardamar içerisindeki kan basıncının aşırı yükselmesine verilen bir sağlık problemidir [1]. Ülkemizde her üç yetiştikenden birinde HPT görülmektedir [2]. Erken tanı ve tedavinin yapılmaması durumunda HPT bazı sağlık sorunlarına ve organ tahribatlarına yol açabilir. HPT'nin ortaya çıkış sebepleri arasında kalıtsal eğilim ve yüksek miktarda tuz tüketimi araştırmalarda başı çekmektedir. Fakat, hastaların çoğunda HPT'nin hangi sebeple ortaya çıktığı bilinmemektedir [3]. Mutlaka kontrol altına alınması gereken bir rahatsızlık olan HPT, ani tansiyon yükselme durumlarında beyin kanamasına ve hatta felce neden olabilmektedir [4]. En yaygın HPT belirtileri baş ağrısı, sersemleme, soluk darlığı, çarpıntı, göğüs ağrısı ve görme sorunları olarak sıralanabilir [5]. HPT'nin teşhisinde öne çıkan tanı yöntemlerinden birisi yirmi dört saat ayaktan kan basıncının takip edilmesidir [6]. Fakat bu yöntem ile bazı hipertansiyon çeşitlerini tespit etmek zordur. Bu sebepten dolayı HPT tespitinde elektrokardiyogram (EKG) ve ekokardiyografi gibi ilave ölçümler yapılması gereklidir. Kişilerin gündelik hayatını etkilemesi ve yaşam kalitesini düşürmesi sebebiyle HPT'nin doğru bir şekilde önceden teşhisi oldukça kritik bir öneme sahiptir. Literatürde biyomedikal işaret işleme ve makine öğrenmesi yöntemleriyle HPT'nin otomatik tespiti üzerine yapılmış araştırmalar yer almaktadır. Khan ve arkadaşları, HPT hastası kişileri tespit etmek amacıyla nabız pletismograf sinyallerini kullanarak bir çalışma gerçekleştirdiler. Çalışmada, Ampirik Kip Ayırımı (AKA) yöntemi kullanarak öznelik çıkarma işlemi gerçekleştirdiler. K-en yakın komşuluk (k-NN) sınıflandırıcı algoritması ile tespit performanslarını ortaya koydular. En yüksek başarımların %99,4 olarak elde ettiler [7]. Rajput ve arkadaşları, HPT hastası kişileri belirlemek için EKG sinyallerini kullandılar. 5 seviyeli dalgacık dönüşüm yöntemini kullanarak dalgacık entropisi ve örnek entropi gibi doğrusal olmayan yöntemleri kullanarak öznelik çıkardılar. k-NN, destek vektör makinaları (DVM), topluluk öğrenme algoritması (ensemble bagged tree (EBT)) ve ağaç algoritmaları ile performans sonuçlarını elde ettiler. En yüksek başarımları EBT algoritması ile %99,95 olarak buldular [8]. Bir başka çalışmada, Soh ve arkadaşları EKG sinyalleri yardımıyla HPT hastalarının otomatik tespiti üzerine yaptıkları çalışmada, evrişimli sinir ağı (ESA) modellerini kullanarak en yüksek %99,99 doğruluk değerine ulaştıklarını belirttiler [9]. Poddar ve arkadaşları, HPT sorunu bulunan kişileri belirlemek için kalp hızı değişkenliği (KHD) sinyallerinden yararlandılar. Zaman bölgesi, frekans bölgesi ve doğrusal olmayan entropi ölçümleriyle çalıştılar. DVM algoritmasını kullanarak %100 oranında bir başarımlar elde ettiler [10].



Şekil 1. Gerçekleştirilen çalışmanın akış diyagramı.

Bu çalışmada, literatürdeki çalışmalarla sağlıklı bir karşılaştırma yapabilmek için Physionet.org web sitesinden ücretsiz elde edilen halka açık verilerden yararlanılmıştır. Veri seti içerisinde, sağlıklı ve HPT hastası olmak üzere iki sınıfa ait EKG verileri bulunmaktadır. Diğer çalışmalardan farklı olarak EKG sinyalleri dört farklı yaklaşım ile incelenmiştir. Bu yaklaşımlar sırasıyla; normalize edilmiş gürültülü EKG sinyalleri, normalize edilmiş gürültüsüz EKG sinyalleri, normalize edilmemiş gürültülü EKG sinyalleri ve son olarak normalize edilmemiş gürültüsüz EKG sinyalleridir. Her bir sinyal türü üzerinden AKA metodu kullanılarak 5 katmanlı İç Mod Fonksiyonları (İMF) elde edilmiştir. Bu fonksiyonlar üzerinden Shannon ve yaklaşık entropi ölçümleri kullanılarak öznitelikler çıkarılmıştır. Burada, Shannon ve yaklaşık entropi ölçümlerinin gürültülü ve gürültüsüz sinyaller üzerindeki performans analizleri gerçekleştirilmiştir. Şekil 1’de çalışmanın akış diyagramı detaylı bir şekilde verilmiştir.

Çalışmanın ilerleyen bölümlerinde sırasıyla, çalışmada kullanılan veri setinden, veriler üzerine uygulanan ön işlemlerden, öznitelik çıkarımı için kullanılan yöntemlerden, sınıflandırıcı algoritmalarından, performans ölçütlerinden bahsedilip son kısımda ise elde edilen sonuçlar ve tartışmaya yer verilmiştir.

2. Materyal ve Metot

2.1. Veri Seti

Tablo 1’de verilen HPT hastalarına ve Normal kişilere ait veriler herkesin erişimine açık olan Phsionet.org sitesinden temin edilmiştir. HPT hastası kişilere ait veriler “Smart Health for Assessing the Risk of Events via ECG (SHAREE) database” veri tabanından alınan EKG sinyalleridir [11]. HPT sinyalleri 49’u kadın ve 90’ı erkek olmak üzere 55 yaş ve üzeri 139 adet EKG verisinden meydana gelmektedir. Normal bireylerin bulunduğu sınıfa ait EKG sinyalleri “Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) normal sinus rhythm database” veri tabanından temin edilmiştir. Toplamda 18 adet EKG sinyalinden meydana gelmektedir [12]. Her bir kayıt yaklaşık 24 saatlik periyotlardan meydana gelmektedir. Bu kayıtlar 8 bit duyarlılıkla saniyede 128 örnek ile örneklenmiştir.

Tablo 1. Çalışmada Kullanılan Veriler ve Özellikleri

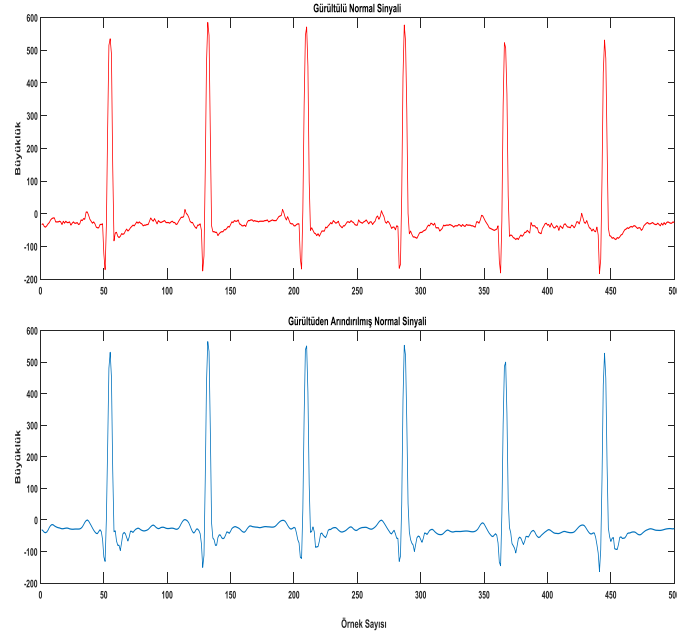
Sınıf	Veri Sayısı	Kadın	Erkek	Yaş Aralığı	Veri Uzunluğu
HPT	139	49	90	46-92	1.000.000
Normal	18	13	5	20-50	1.000.000

Bu çalışmada, bir milyon uzunluğa sahip verilerden oluşan her bir sinyal on bin uzunluklu kısımlara ayrılarak kullanıldı. Bu işlem sonunda 1800 adet sağlıklı ve 13622 adet HPT hastasına ait veri oluşturuldu. Her bir HPT sinyalinin ilk 0 ile 20000 arasında kalan sinyal parçaları bozuk EKG sinyali içermesi nedeniyle çalışma dışında tutulmuştur.

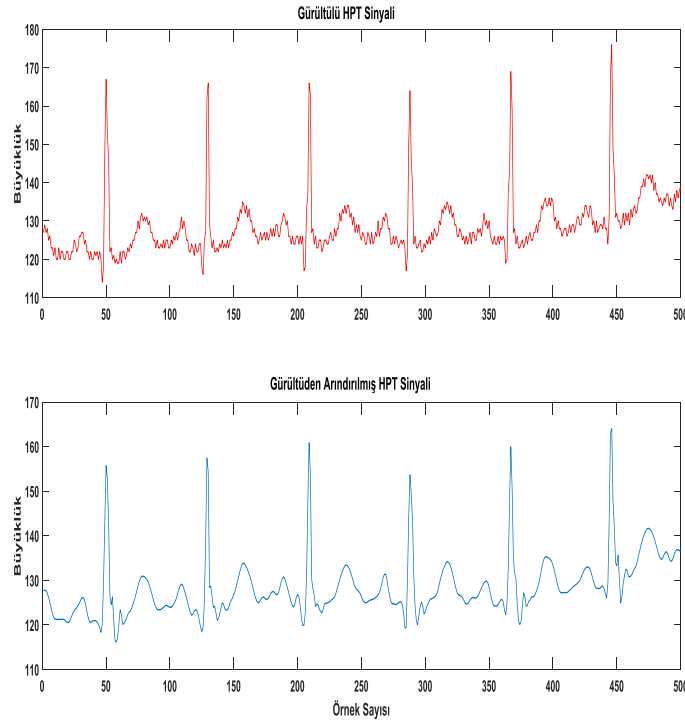
2.2. Ayrık Dalgacık Dönüşümü (ADD) ile Gürültü Eliminasyonu

EKG sinyalini gürültüden arındırmak için, dalgacık dönüşümü uygulayarak gürültülü sinyali ayrıştırma işlemine başlanır. Dalgacık dönüşümü sayesinde, sinyali farklı frekans seviyelerinde katsayı gruplarına ayırmak mümkündür [13]. Sinyalin farklı frekans segmentlerinde nasıl davrandığını anlamak, bir sonraki aşamada en uygun eşik seçilmesine imkân tanıyacaktır. Bir sonraki adım, en iyi eşik değerlerini belirlemek ve bu katsayılara eşik değerleri uygulayarak istenmeyen verileri ortadan kaldırmaktır. Algoritmanın son adımı ise, filtrelenmiş katsayı kümelerini kullanarak EKG sinyalini yeniden oluşturmaktır. Şekil 2’de sağlıklı bireylere ait

gürültülü sinyal ADD kullanılarak gürültüden arındırılarak gösterilmiştir. Şekil 3'te ise aynı işlem gürültülü HPT sinyali için tekrarlanarak verilmiştir.



Şekil 2. Sağlıklı bireylere ait orijinal EKG sinyali ve gürültüden arındırılmış EKG sinyali.



Şekil 3. HPT hastalarına ait orijinal EKG sinyali ve gürültüden arındırılmış EKG sinyali.

2.3. Min-Max Normalizasyon

Gürültüden ayrılan veriler min-max normalizasyon yöntemiyle aşağıdaki denkleme göre 0 ve 1 arasında normalize edilmiştir.

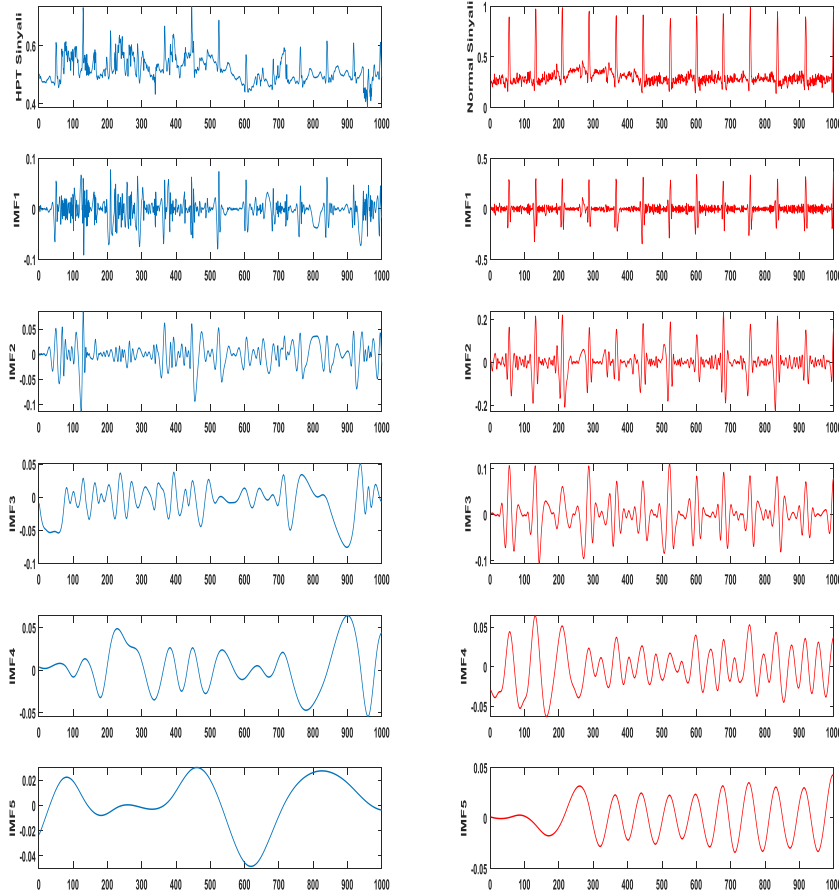
$$v' = \frac{v - \min_x}{\max_x - \min_x} \quad (1)$$

Burada, $\min_x X$ sinyalinin minimum değeri ve \max_x ise X sinyalini maksimum değeridir. X sinyalinin orijinal ve normalize edilmiş değerleri sırasıyla v ve v' ile gösterilir. Yukarıdaki denklemden minimum ve maksimum sinyal değerlerinin sırasıyla 0 ve 1'e eşitlendiği görülebilir.

2.4. Ampirik Kip Ayrışımı

Ampirik kip ayrışımı (AKA) , lineer olmayan ve durağan olmayan işaretleri çözümlenmek amacıyla yararlanılan bir metottur. Bahsedilen bu metot 1998 tarihinde Huang'ın tavsiyesi ile göz önüne alınmıştır [14]. AKA'da ilk sinyal döngü şeklinde ayrıştırılarak yüksek frekanstan düşük frekansa doğru İMF'lere bölünür. İMF'ler iki önemli prensibe dayanır [15]. Bu prensipler:

- 1) Tüm veri kümesinde lokal ekstremum noktaların ve sıfır geçişlerin sayısı birbirine eşit olmalıdır ya da en fazla 1 noktada farklılaşmalıdır.
- 2) Herhangi bir noktada, lokal maksimumlar tarafından tanımlanan zarfın ortalama değeri ve lokal minimumlar tarafından tanımlanan zarfın ortalama değeri sıfırdır.



Şekil 4. HPT ve Normal kişilere ait sinyallerin 5 katmanlı İMF'lere ayrılmış gösterimi.

Bir $x(t)$ sinyaline AKA işlemi aşağıda verilen maddelere göre uygulanır:

- 1) Orijinal $x(t)$ sinyalinin tüm lokal ekstremum noktaları belirlenir.
- 2) $x(t)$ sinyalinin kübik spline interpolasyon yöntemi ile üst ve alt zarfı meydana getirilir.
- 3) Üst ve alt zarflardan her noktada ortalama $m(t)$ değeri hesaplanır.
- 4) Orijinal sinyalden ortalama değer çıkarılarak fark sinyali hesaplanır. $[d(t)=x(t)-m(t)]$

- 5) $d(t)$ bir İMF ise, $c(t) = d(t)$ olarak atanır ve $d(t)$ bir İMF değilse, $d(t)$ orijinal sinyal olarak kabul edilir ve ilk 4 adım $d(t)$ değeri bir İMF olana kadar tekrarlanır.
- 6) İlk bileşeni bulduktan sonra, ilk bileşen orijinal sinyalden çıkarılır ve kalıntı $r(t) = x(t) - c(t)$ şeklinde bulunur.
- 7) Son adımda $r(t)$ orijinal veri olarak kabul edilir ve ilk 6 adım tekrarlanır. AKA, kalıntı $r(t)$ monotonik bir fonksiyon veya daha fazla İMF bileşeninin çıkarılamayacağı bir sabit olduğunda süreç durdurulur. HPT ve Normal kişilere ait 5 katmanlı İMF işaretleri Şekil 4'de verilmiştir. Bu çalışmada, elde edilen her bir İMF işareti için entropi ölçümlerinden yaklaşık (approximate) ve Shannon entropi kullanılmıştır.

2.5. Yaklaşık Entropi

Yaklaşık entropi (Approximate entropy-ApEn), bir zaman serisinin karmaşıklığını veya düzensizliğini ortaya koyan bir ölçümdür. ApEn değeri büyüdükçe, sinyal daha karmaşık ve düzensiz bir hal alır. Kısa, gürültülü veri kümeleri kullanılarak doğru bir ApEn tahmini yapılabilir. ApEn'i ortaya koymak için kullanılan genel formüller aşağıda açıklanmıştır [16].

1. N veri noktası içeren orijinal sinyalimiz $Y(n)$ olmak üzere, $Y(n)=[x(1),x(2),\dots,x(N)]$.
2. m -vektörleri $Y(1),\dots,Y(N-m+1)$, $Y(i)=[y(i),y(i+1),\dots,y(i+m-1)]$ $i=1,2,\dots,N - m + 1$ göre tanımlanır. Bu vektörler, i 'inci noktadan başlanarak m tane ardışık y değerini belirtir.
3. $Y(i)$ ve $Y(j)$ arasındaki mesafe $d[Y(i), Y(j)]$ ile belirtilir, ilgili skaler bileşenleri arasındaki maksimum mutlak fark, diğer bir deyişle maksimum norm:

$$D[X(i),X(j)] = \max_{k=1,2,\dots,m} |y(i+k-1)-y(j+k-1)| \quad (2)$$

formülü ile hesaplanır.

4. Verilen bir $y(i)$ için, ($j = 1, \dots, N-m + 1, j \neq i$) sayısı bulunur, böylece $d[y(i), y(j)] \leq r$, $N^m(i)$ olarak gösterilir. O zaman, $i = 1$ için, $\dots, N-m + 1$:

$$C_r^m(i) = \frac{N^m(i)}{N - m + 1} \quad (3)$$

5. $C_r^m(i)$ bir r toleransı içinde m uzunluk penceresi tarafından verilene benzer modellerin frekansını ölçer.
6. Her bir $C_r^m(i)$ 'nin doğal logaritmasını hesaplanır ve i üzerinden ortalaması alınır:

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_r^m(i) \quad (4)$$

7. Boyutu $m+1$ 'e yükseltilir. $C_r^m(i)$ and $\phi^{m+1}(r)$ elde etmek için 2'den 6'ya kadar olan adımlar tekrarlanır. Son olarak, ApEn aşağıdaki formüle göre hesaplanır:

$$\text{ApEn}(m,r,N) = \phi^m(r) - \phi^{m+1}(r) \quad (5)$$

Uzunluğu N olan sinyalin ApEn değerini hesaplamadan önce iki parametre belirtilmelidir: m , gömme boyutu ve r , bir tolerans penceresi.

2.6. Shannon Entropi

Shannon Entropi 1940'ların sonunda Shannon tarafından geliştirilmiş bir yöntemdir. Bu yöntemde asıl amaç sinyalin kararsızlığını ve düzensizliğini belirlemektir [17]. H_{sh} ile gösterilen Shannon

entropi, sinyalin olasılık yoğunluk fonksiyonunu karakterize eder ve sürekli sinyal için aşağıdaki şekilde tanımlanır:

$$H_{sh} = - \sum_i P_i \log P_i \quad (6)$$

Burada i sinyalin tüm genlikleri boyunca değişiklik gösterir. P_i ise sinyalin olasılığını gösterir.

2.7. Destek Vektör Makineleri

Destek Vektör Makineleri (DVM) makine öğreniminde, özellikle sınıflandırmada kullanılan oldukça etkili ve kolay bir denetimli öğrenme yöntemidir. DVM yöntemi Vapnik tarafından geliştirilmiştir [18]. DVM sınıfların birbirinden ayrıldığı sınırı, bir optimum hiper düzlem çizebilecek fonksiyonu bulmayı amaçlamaktadır. Her bir sınıf için bu sınıra en yakın olan elemanların üzerinden geçen bir hiper düzlem çizilir. Çizilen hiper-düzlemler desteklik sağladığı için destek vektör makineleri ismi buradan gelmektedir. Bu destek hiper düzlemlerine eşit uzaklıkta olacak şekilde çok sayıda optimum hiper düzlem çizilebilir. Ancak bu optimum hiper düzlemlerden sınıflar arası ayrımın en uzak olduğu optimum hiper düzlem çözümü verir. Bu optimum hiper düzlem üzerinde hiçbir sınıfa ait eleman bulunmaz. Bu metot lineer olarak ayrışabilen veriler için sıkça kullanılmaktadır. Fakat lineer bir şekilde ayrıştırılamayan veriler çekirdek fonksiyonları aracılığıyla lineer bir şekilde ayrıştırılabilir hale getirilerek çözüm bulunur.

2.8. K En Yakın Komşu

K-en yakın komşu (k-NN) yöntemi, basit uygulaması ve önemli sınıflandırma performansı nedeniyle veri madenciliği ve istatistikte kullanılan popüler bir sınıflandırma algoritmasıdır. Bazı öznelik vektörleri tarafından temsil edilen bilinmeyen bir örneği öznelik uzayında bir nokta olarak sınıflandırmak için, k-NN sınıflandırıcı eğitim veri setindeki nokta ve noktalar arasındaki mesafeleri hesaplar. Uzaklık ölçüsü olarak genellikle Öklid uzaklığı kullanılır. Ardından, sınıfı bilinmeyen noktayı en yakın komşusuna atar (burada k bir tamsayıdır). Örneğin, k = 1 ise, nokta en yakın sınıfa aittir, k = 5 ise, nokta en yakın beş noktayı içine alan sınıfa aittir.

2.9. Karar Ağaçları

Bir karar ağacı, düğümler ve kenarlardan oluşur. Ağaç, gelen kenarı olmayan kök düğüm adı verilen bir düğümden başlar. Giden kenarları olan düğümlere iç düğümler veya test düğümleri denir ve diğer tüm düğümlere yaprak veya uç düğümler denir. Dahili düğümler, girdi nitelikleri değerlerinin belirli bir ayrık fonksiyonuna göre girdi uzaylarının iki veya daha fazla alt uzaya bölünmesidir. Ayrıca, her yaprak düğüm, hedef özelliğin belirli bir değere sahip olma olasılığını gösteren bir olasılık vektörü tutabilir. Dahili düğümler tarafından verilen kararlara göre, örnekler ağacın kökünden bir yaprağa doğru yönlendirilir ve ardından sınıflandırılır.

2.10. Başarım Ölçütleri

Bu çalışmada sınıflandırıcı performansını değerlendirmek için genel doğruluk (ACC-Accuracy), hassaslık (REC-Recall), özgünlük (SPE-Specificity), kesinlik (PRE-Precision), ve F1-Skoru (F1-score) kullanılmıştır [19,20]. Kullanılan performans ölçütlerinin matematiksel karşılıkları denklem (7), (8), (9), (10) ve (11)'de verilmiştir [21-23].

$$\text{Doğruluk (ACC)} = \frac{GP + GN}{GP + YN + YP + GN} \quad (7)$$

$$\text{Hassaslık (REC)} = \frac{GP}{GP + YN} \quad (8)$$

$$\text{Özgünlük (SPE)} = \frac{GN}{GN + YP} \quad (9)$$

$$\text{Kesinlik (PRE)} = \frac{GP}{GP + YP} \quad (10)$$

$$\text{F1 Ölçütü (F1)} = \frac{2 * PRE * REC}{PRE + REC} \quad (11)$$

Burada GP gerçek pozitif, GN gerçek negatif, YP yanlış pozitif ve YN yanlış negatif ifadelerini göstermektedir.

Sınıflandırıcının normalde HPT olup HPT olarak belirlediklerinin sayısını GP, yanlışlıkla sağlıklı olarak tespit ettiklerinin sayısı YN, normalde hasta olmayıp sınıflandırıcının da sağlıklı olarak bulduklarının sayısı GN, yanlışlıkla HPT olarak belirlediklerinin sayısını da YP olarak ifade edilmektedir [24].

3. Bulgular ve Tartışma

HPT'nin otomatik tespit modeli geliştirmek için MATLAB (2020a) programlama ortamından yararlanılmıştır. Tablo 2'de verilen sayılarda 10000 (on bin) uzunluklu, normalize edilmiş gürültülü EKG sinyalleri, normalize edilmiş gürültüsüz EKG sinyalleri, normalize edilmemiş gürültülü EKG sinyalleri ve son olarak normalize edilmemiş gürültüsüz EKG sinyalleri kullanılmıştır.

Tablo 2. Çalışmada kullanılan veri sayıları.

Sınıf	Veri Sayısı
HPT	13622
Normal	1800

5 farklı İMF sinyali üzerinden elde edilen 2 adet doğrusal olmayan ölçüm ile sonuçlar karşılaştırılmıştır. Üç farklı sınıflandırma algoritması kullanarak sonuçlara ulaşılmıştır. Daha güvenilir ve kararlı sonuçlara ulaşmak amacıyla 10-kat çapraz doğrulama metoduyla bütün verilerin eğitimi ve testi sağlanmıştır. Gürültüden arındırılmış sinyaller için tüm İMF ölçümlerinden alınan özneliliklerin kullanılmasıyla bulunan başarımlar Tablo 3'de gösterilmiştir.

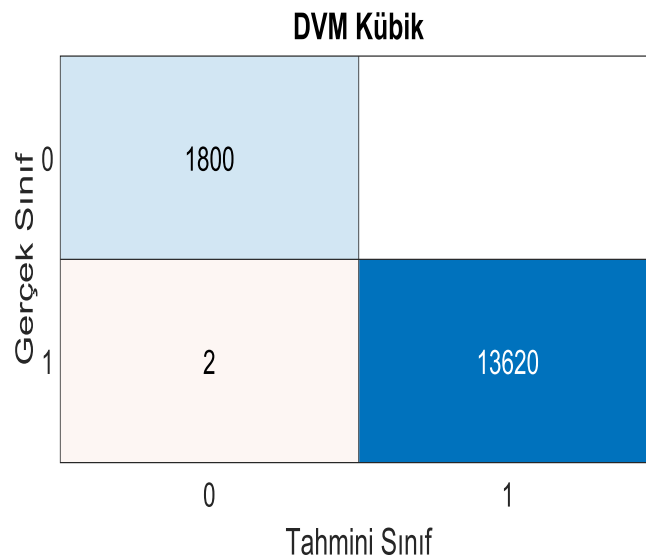
Tablo 3'den elde edilen sonuçlar incelendiğinde, en yüksek başarımlar DVM kübik algoritması kullanılarak %99,99 olarak gürültüsüz normalize edilmiş sinyallere ait olduğu görülmektedir. Aynı zamanda ikinci en düşük eğitim süresi de 14,55 saniye ile bu sinyale aittir. Yine aynı şekilde normalize edilmiş ve DVM Lineer algoritması kullanılarak sınıflandırılmış olan sinyalinde başarımlar %99,99 olarak elde edilmiştir. Eğitim süresi de benzer şekilde 15,32 saniye olarak karşımıza çıkmaktadır. Normalize edilmemiş sinyaller arasında ise k-NN algoritması ile sınıflandırılmış olan sinyal %99,99'luk yüksek bir başarımlara sahiptir. Eğitim süresi ise 60,17 saniyedir. Eğitim süresi en fazla olan sinyal ise normalize edilmemiş ve DVM kübik algoritması kullanılarak sınıflandırılmış olan sinyale aittir. Her ne kadar yüksek bir başarımlara sahip olmuş olsa da eğitim süresinin önemli olduğu sınıflandırma problemlerinde bu sinyalin kullanılması araştırmacıya

zaman kaybettireceği açıktır. Buradan hareketle gürültüden arındırılmış sinyaller üzerinde sınıflandırma yaparken normalizasyonun avantajları, başarımları ve eğitim süresi açısından Tablo 3’de açıkça görülmektedir.

Tablo 3. Gürültüsüz sinyallerdeki performans değerleri

Normalizasyon Var/Yok	Algoritma	Eğitim Süresi (sn)	Doğruluk (%)	Hassaslık (%)	Özgünlük (%)	Kesinlik (%)	F1 Ölçütü (%)
Yok	DVM Lineer	36,49	99,81	99,87	99,33	99,91	99,89
Yok	DVM Kuadratik	35,68	99,98	99,99	99,94	99,99	99,99
Yok	DVM Kübik	394,03	99,94	99,96	99,78	99,97	99,97
Yok	k-NN	60,17	99,99	100,00	99,94	99,99	99,99
Yok	Karar Ağaçları	10,76	99,94	99,96	99,72	99,96	99,96
Var	DVM Lineer	15,32	99,99	99,99	100	100	99,99
Var	DVM Kuadratik	20,10	99,98	99,98	100	100	99,99
Var	DVM Kübik	14,55	99,99	99,99	100	100	100
Var	k-NN	61,45	99,98	100,00	99,83	99,98	99,99
Var	Karar Ağaçları	3,44	99,97	99,98	99,89	99,99	99,98

Her bir algoritmaya ait tespit edilen en yüksek performanslara ait hata matrisleri Şekil 4, Şekil 5 ve Şekil 6’da gösterilmiştir.



Şekil 4. DVM için ulaşılan en yüksek başarıma ait hata matrisi (normalize edilmiş gürültüsüz veriler için)

KNN

Gerçek Sınıf	0	1
0	1799	1
1		13622
	0	1
	Tahmini Sınıf	

Şekil 5. k-NN için ulaşılan en yüksek başarıma ait hata matrisi (normalize edilmemiş gürültüsüz veriler için)

Karar Ağaçları

Gerçek Sınıf	0	1
0	1798	2
1	3	13619
	0	1
	Tahmini Sınıf	

Şekil 6. Karar Ağaçları için ulaşılan en yüksek başarıma ait hata matrisi (normalize edilmiş gürültüsüz veriler için)

Şekil 4’te DVM kübik algoritması için verilen hata matrisi incelendiğinde, sağlıklı (0) olan kişilerin %100 doğrulukla belirlendiği görülürken, HPT (1) hastası kişilerin ise yalnızca 2 adedinin yanlış kalan 13620 adedinin doğru bir şekilde belirlendiği gözlemlenmiştir. Şekil 5’te k-NN (k=1) algoritması için verilen hata matrisi incelendiğinde, HPT (1) hastası kişilerin %100 doğrulukla belirlendiği görülürken, sağlıklı (0) olan kişilerin ise yalnızca 1 adedinin yanlış kalan 1799 adedinin doğru bir şekilde belirlendiği gözlemlenmiştir. Şekil 6’da Karar Ağaçları algoritması için verilen hata matrisine bakıldığında, HPT (1) hastası kişilerin yalnızca 3 adedinin yanlış kalan 13619 adedinin doğru olarak belirlendiği görülürken, sağlıklı (0) olan kişilerin ise yalnızca 2 adedinin yanlış kalan 1798 adedinin doğru olarak belirlendiği tespit edilmiştir.

Gürültülü sinyaller için tüm İMF ölçümlerinden alınan özneliklerin kullanılması ile tespit edilen başarımlar Tablo 4’te verilmiştir. Tablo 4’de verilen gürültülü sinyaller üzerinden elde edilen sonuçlar değerlendirildiğinde, en yüksek başarımlar k-NN (k=1) algoritması için %99,98 ile normalize edilmemiş sinyale aittir. Eğitim süresi olarak 64,1 saniye olarak kayıt edilmiştir. Normalize edilmemiş sinyaller için ise karar ağaçları algoritması %99,94 olarak elde edilmiştir. Eğitim süresi olarak 3,87 saniye ile en düşük eğitim süresi olarak karşımıza çıkmaktadır. Eğitim süresi en fazla olan sinyal ise normalize edilmemiş ve DVM kübik algoritması kullanılarak

sınıflandırılmış olan sinyale aittir. Buradan hareketle gürültülü sinyaller üzerinde sınıflandırma yaparken normalizasyonun çokta fazla bir avantajı olmadığı açıkça görülmektedir.

Tablo 4. Gürültülü sinyallerdeki performans değerleri.

Normalizasyon Var/Yok	Algoritma	Eğitim Süresi (sn)	Doğruluk (%)	Hassaslık (%)	Özgünlük (%)	Kesinlik (%)	F1 Ölçütü (%)
Yok	DVM Lineer	195,07	97,64	99,85	80,94	97,54	98,68
Yok	DVM Küadratik	37,09	99,91	99,92	99,83	99,98	99,95
Yok	DVM Kübik	5216,10	58,29	60,10	44,61	89,14	71,80
Yok	k-NN	64,1	99,98	99,98	100,00	100,00	100,00
Yok	Karar Ağaçları	3,87	99,94	99,98	99,67	99,96	99,97
Var	DVM Lineer	214,91	96,03	99,63	68,83	96,03	97,79
Var	DVM Küadratik	2755,80	99,02	99,57	94,83	99,32	99,45
Var	DVM Kübik	5087,40	34,81	30,20	69,67	88,28	45,01
Var	k-NN	59,85	99,88	99,93	99,44	99,93	99,93
Var	Karar Ağaçları	9,09	99,74	99,86	98,83	99,85	99,85

Hipertansiyonun otomatik tespit edilmesi konularında literatürde yapılmış çalışmaların karşılaştırmaları Tablo 5'te verilmiştir.

Tablo 5. Literatürdeki diğer çalışmalar ile başarımların karşılaştırılması.

Yazarlar	Metot	Sınıflandırma Algoritması	Doğruluk(%)
Khan ve arkadaşları [7]	AKA	k-NN	99,40
Rajput ve arkadaşları [8]	Dalgacık Dönüşümü	k-NN, DVM, EBT	99,95
Soh ve arkadaşları [9]	ESA	TBK	99,99
Poddar ve arkadaşları [10]	Zaman Alanı ve Hızlı Fourier Dönüşümü Ölçümleri	DVM	100
Bu Çalışma	AKA	DVM	99,99
Bu Çalışma	AKA	k-NN	99,99
Bu Çalışma	AKA	Karar Ağaçları	99,97

Bu çalışmalarda daha çok geleneksel makine öğrenmesi yaklaşımları kullanılmaktadır. Tablo 5'ten elde edilen sonuçlara bakıldığında Khan ve arkadaşları AKA metodunu ve k-NN sınıflandırıcısını kullanarak %99.4'lük doğruluk elde etmişlerdir [7]. Rajput ve arkadaşları dalgacık dönüşümü yöntemini ve k-NN, DVM, EBT sınıflandırıcılarını kullanarak %99.95'lik doğruluk elde etmişlerdir [8]. Soh ve arkadaşları ESA ve tam bağımlı katman (TBK) kullanarak %99,99'lük doğruluk elde etmişlerdir [9]. Poddar ve arkadaşları KHD sinyalleri üzerinden zaman alanı ölçümleri ve hızlı Fourier dönüşümü ölçümleri ile DVM sınıflandırıcısını kullanarak %100'lük doğruluk değeri elde ettiler [10]. Bu çalışmada ise, her bir veri on bin uzunluklu segmentlere ayrılarak artırılmıştır. Gürültülü ve gürültüsüz, normalize edilmiş ve normalize edilmemiş olarak elde edilen veriler kullanılmıştır. AKA yöntemiyle elde edilen 5 seviyeli İMF'ler üzerinden elde edilen öznelikler ile, DVM, k-NN ve karar ağaçları algoritmaları kullanılarak %99,99'lük doğruluk değeri elde edilmiştir.

HPT hastalarının tespit edilmesinde benzer çalışmaların birçoğundan daha başarılı sonuçlar verdiği ortaya konulmuştur. Gürültüsüz ve normalize edilmiş veriler ile çalışmak hem başarıyı arttırdığı hem eğitim süresini azalttığı tespit edilmiştir. Bu çalışmayı sınırlandıran en kritik parametre verilerin dengeli dağılmasıdır. Gelecekteki çalışmalarda, manuel öznelik çıkarımı olmadan uçtan uca yaklaşım gösteren derin öğrenme yöntemlerinin HPT hastalarının tespitindeki performanslarının araştırılması planlanmaktadır.

Çıkar Çatışması

Yazarlar, çıkar çatışması olmadığını beyan eder.

Yazar(lar)ın Katkıları

YEE ve AN birlikte makaleyi yazdıç Makalenin son halini okudu ve onayladı.

Kaynaklar

- [1].Tabassum N., Ahmad F., "Role of natural herbs in the treatment of hypertension, *Pharmacogn Rev.*, 2011 Jan;5(9):30-40.
- [2].Sarıışık A, Oğuz A, Uzunlulu M., "Control of hypertension in Turkey--is it improving? The Kocaeli 2 study., *Turk Kardiyol Dern Ars.*, 2009,37(6):13-16
- [3].Chalmers J.,"Implementation of guidelines for management of hypertension, *Clinical and Experimental Hypertension*, 1992, 21:647–657.
- [4].Drozd D.,Kawecka-Jaszcz K., "Cardiovascular changes during chronic hypertensive states, *Pediatr. Nephrol.*, 29 (9) (2014) 1507–1516.
- [5].Ni H.,Wang Y.,Xu G.,Shao Z.,Zhang W., and Zhou X., "Multiscale Fine-Grained Heart Rate Variability Analysis for Recognizing the Severity of Hypertension," *Comput. Math. Methods Med.*, vol. 2019, 2019, doi: 10.1155/2019/4936179.
- [6].Hermida R.C., Smolensky M.H., Ayala D.E., Portaluppi F., "Ambulatory Blood Pressure Monitoring (ABPM) as the reference standard for diagnosis of hypertension and assessment of vascular risk in adults, *Chronobiol. Int.* 32 (10) (2015) 1329–1342.
- [7].Khan M.U.,Aziz S.,Akram T.,Amjad F.,Iqtidar K.,Nam Y.,Khan M.A., "Expert Hypertension Detection System Featuring Pulse Plethysmograph Signals and Hybrid Feature Selection and ReductionScheme,*Sensors*,21,2021,1,247,doi:10.3390/s21010247,htGPs://www.mdpi.com/1424-8220/21/1/247
- [8].Rajput J.S.,Sharma M.,Tan R.S., Acharya UR., "Automated detection of severity of hypertension ECG signals using an optimal bi-orthogonal wavelet filter bank, *Comput Biol Med.*,2020 Aug;123:103924. doi: 10.1016/j.combiomed.2020.103924. Epub 2020 Jul 23. PMID: 32768053.

- [9]. Soh D.C.K., Ng E.Y.K., Jahmunah V., Oh S.L., Tan R.S., Acharya U.R., "Automated diagnostic tool for hypertension using convolutional neural network, Computers in Biology and Medicine, Volume 126, 2020, 103999, ISSN 0010-4825, htGPs://doi.org/10.1016/j.compbiomed.2020.103999.
- [10]. Poddar M.G., Kumar V., Sharma Y.P., "Linear-nonlinear heart rate variability analysis and SVM based classification of normal and hypertensive subjects, Journal of Electrocardiology, Volume 46, Issue 4, 2013, Page e25, ISSN 0022-0736.
- [11]. Moody G.B., Mark R.G., Goldberger A.L., "PhysioNet: physiologic signals, time series and related open source software for basic, clinical, and applied research, Conf. Proc. IEEE Eng. Med. Biol. Soc., 2011 (2011) 8327–8330. Cambridge, MA 20139, USA.
- [12]. Sannino G, Melillo P, Stranges S, De Pietro G, Pecchia L. Short term Heart Rate Variability to predict blood pressure drops due to standing: a pilot study. BMC Med Inform Decis Mak. 2015;15 Suppl 3(Suppl 3):S2. doi:10.1186/1472-6947-15-S3-S2
- [13]. Patil S. S., Pawar M. K., "Quality advancement of EEG by wavelet denoising for biomedical analysis, 2012 International Conference on Communication, Information & Computing Technology (ICCICT), 2012, pp. 1-6, doi: 10.1109/ICCICT.2012.6398151.
- [14]. Huang N. E., Shen Z., Long S. R., Wu M. C., Shih H. H., Zheng Q., Yen N.-C., Tung C. C., and Liu H. H., "The empirical mode decomposition and the Hilbert spectrum for non-linear and non stationary time series analysis, Proc. Royal Soc. London A, vol. 454, pp. 903–995, Mar. 1998.
- [15]. Rilling G., Flandrin P., and Gonçalves P., "On empirical mode decomposition and its algorithms, in Proc. IEEE-EURASIP Workshop Nonlinear Signal Image Process., Jun. 2003, pp. 1–5.
- [16]. Sabeti M., Katebi S., Boostani R., "Entropy and complexity measures for EEG signal classification of schizophrenic and control participants, Artificial Intelligence in Medicine, Volume 47, Issue 3, 2009, Pages 263-274, ISSN 0933-3657, htGPs://doi.org/10.1016/j.artmed.2009.03.003.
- [17]. Lin J., "Divergence measures based on the Shannon entropy, in IEEE Transactions on Information Theory, vol. 37, no. 1, pp. 145-151, Jan. 1991, doi: 10.1109/18.61115.
- [18]. Vapnik V. N., "An overview of statistical learning theory, in IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 988-999, Sept. 1999, doi: 10.1109/72.788640.
- [19]. Kaya D., Turk M., Kaya T., "Examining the Effect of Dimension Reduction on EEG Signals by K-Nearest Neighbors Algorithm, El-Cezerî Journal of Science and Engineering, 2018, 5(2); 591-595.
- [20]. Narin A., Özer M., İşler Y., "Effect of linear and non-linear measurements of heart rate variability in prediction of PAF attack, 2017 25th Signal Processing and Communications Applications Conference (SIU), 2017, pp. 1-4, doi: 10.1109/SIU.2017.7960358.
- [21]. Erdoğan Y. E., Narin A., "Performance of Empirical Mode Decomposition in Automated Detection of Hypertension Using Electrocardiography, 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9477887.
- [22]. Erdoğan Y. E., Narin, "COVID-19 detection with traditional and deep features on cough acoustic signals, Computers in Biology and Medicine, Volume 136, 2021, 104765, ISSN 0010-4825, htGPs://doi.org/10.1016/j.compbiomed.2021.104765.
- [23]. İşler Y., Narin A. and Ozer M., "Comparison of The Effects of Cross-Validation Methods on Determining Performances of Classifiers Used in Diagnosing Congestive Heart Failure, Measurement Science Review, 2015 15 (4): 196-201.
- [24]. Yargı V., Postalcıoğlu S., "EEG İşareti Kullanılarak Bağımlılığa Yatkınlığın Makine Öğrenmesi Teknikleri ile Analizi, El-Cezerî Fen ve Mühendislik Dergisi, 2021, 8(1): 142-154.