

KNN, NN, BAYES, DT ve SVM Kullanılarak EKG Vurularının Sınıflandırılması

*¹Yasin Kaya ve ²Hüseyin Pehlivan

¹Karadeniz Teknik Üniversitesi, Enformatik Bölümü, Trabzon, Türkiye

²Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği, Trabzon, Türkiye

Abstract

Today, one of the most important problems that cause death is heart disorders. Early detection and treatment of heart disease may prevent sudden death. Detected through the human body and seen as a result of activity of the heart's electrical signals is called electrocardiogram (ECG). The large percentage of human death is composed by heart diseases. ECG signal, which can be easily obtained through patient's body without damaging, is a good indicator of the disorder during operation of the hearth. In this study, Normal beats (N), premature ventricular contraction (PVC) beat classified and the classification performance has been analyzed. Time series of the signal is used as an input vector for classification algorithms instead of extracting feature from the signal. Neural Networks (NN), K-Nearest Neighbor (KNN), BAYES, Decision Trees (DT), and Support Vector Machines (SVM) classification algorithms results were discussed.

Key words: ECG Classification, NN, KNN, BAYES, SVM, DT

Özet

Günümüzde ölüme neden olan en önemli problemlerden biri kalp rahatsızlıklarıdır. Kalp rahatsızlıklarının erken teşhisi ve tedavisi ani ölümleri önleyebilir [1]. İnsan vücudu üzerinden algılanan ve kalbin aktivitesi gösteren elektriksel işaretlere elektrokardiyogram (EKG) adı verilir. İnsan ölümlerinin büyük bir yüzdesi kalp rahatsızlıklarından oluşmaktadır. Bu yüzden kalbin çalışması sırasındaki bozuklukların iyi bir göstergesi olan ve hastaya zarar vermeden, vücut üzerinden kolaylıkla elde edilebilen EKG işaretleri, işleme ve yorumlama açısından büyük önem taşımaktadır[1]. Bu çalışmada EKG vurularından Normal vuru ve Erken karıncık kasılması vurusu sınıflandırılmış ve sınıflandırma başarımları irdelenmiştir. Sınıflandırma için sinyalden özellik çıkarmak yerine sinyalin zaman serisi sınıflandırma algoritmalarına giriş veriri olarak verilmiştir. Yapay sinir ağları(NN), K-ene yakın komşuluk (KNN), Bayes, Karar ağaçları (DT) ve Destek vektör makinesi (SVM) sınıflandırma algoritmalarının sonuçları tartışılmıştır.

Anahtar Kelimeler: EKG Sınıflandırma, NN, KNN, SVM, DT, Neive Bayes

1. Giriş

EKG'deki farklı türden kalp vurularının algılanması ve sınıflandırılması kardiyolojik fonksiyon bozukluklarının tanısında hayati önem taşımaktadır. Bazı aritmiler sık olmayan aralıklarla görüldüğü için klinisyenler onları yakalamak için Holter kayıt cihazlarını kullanmaktadırlar. EKG kayıtlarının analizinde kullanılacak özel algoritmalar çok uzun Holter EKG kayıtlarının analizini kolaylaştıracaktır. Bu çalışmada normal ve erken karıncık kasılması vurularının sınıflandırılması için sınıflandırma algoritmaların başarımları incelenmiştir. Literatürde EKG vurularının

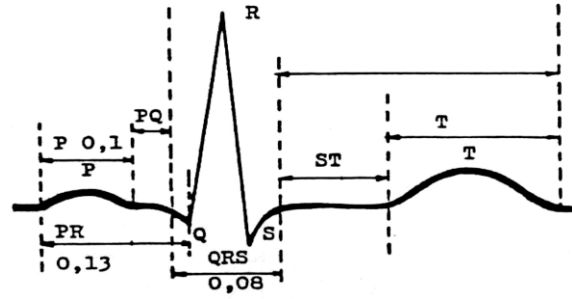
*İlgili yazar: Adres: Karadeniz Teknik Üniversitesi, Enformatik Bölümü, 61080, Trabzon, TÜRKİYE. E-mail: yasin@ktu.edu.tr, Phone: +904623773102

sınıflandırılması için çalışmalar mevcuttur[2]. Engin çalışmasında EKG vurularının fuzzy-hibrid yapay sinir ağlarını kullanarak sınıflandırmaya çalışmıştır. Sınıflandırma işlemlerinde EKG vurusu yerine otoregresiv model katsayıları (autoregressive model coefficients), yüksek mertebeden birikinti (higher-order cumulant) ve dalgacık dönüşüm varyanslarını özellik olarak kullanmıştır. Sınıflandırma için ise fuzzy c-means algoritması ve MLP yapay sinir ağlarını kullanarak yüksek başarımlar elde etmiştir[3]. Erdoğan çalışmasında MIT-BIH veri tabanından alınan EKG sinyallerine Ayırık Dalgacık Dönüşümü uygulanmıştır. Dalgacık fonksiyonlarından Daubechies-2 (db2) dalgacıkları kullanılmıştır. Özellik çıkarımında 4. seviyede CD4 katsayıları kullanılmıştır. Sınıflandırma için yapay sinir ağlarını (YSA) kullanmıştır ve dört grup özellik katsayısı test işlemine tabi tutulmuştur. Test sonuçlarında normalizasyon ve QRS algılama yapılmamasına rağmen sınıflandırma başarısının, Hızlı Fourier Dönüşümüne göre daha yüksek çıktığı belirtilmiştir[4]. Yu çalışmasında bağımsız bileşen analizi (ICA) ve yapay sinir ağlarını EKG vurularını sınıflandırmak için kullanmıştır. ICA, ECG sinyallerinin istatistiksel olarak karşılıklı olarak bağımsız temel bileşenlerinin hesaplanmasında kullanılmıştır. Bu bileşenlerin izdüşümleri, RR aralığı ile beraber takip eden sınıflandırıcının özellik vektörünü oluşturmaktadır. Olasılıklı yapay sinir ağı (PNN) ve geri-beslemeli yapay sinir ağını (BPNN) içeren iki yapay sinir ağı sınıflandırıcı olarak kullanılmıştır[5]. Güler çalışmasında EKG vurularını sınıflandırmak için birleşik yapay sinir ağı modeli tasarlanmıştır. ECG sinyalleri zaman-frekans domenine kesikli dalgacık dönüşümü (DWT) kullanılarak ayrıştırılmış ve dağılımını anlamak için istatistiksel özellikler hesaplanmıştır. Birinci seviye ağlar istatistiksel özellikleri girişler olarak kullanarak ECG vurularının sınıflandırılması için uygulanmıştır. Tanısal doğruluğu geliştirmek için ikinci seviye ağlar birinci seviye ağların çıktısını giriş verisi olarak kullanarak eğitilmiştir. Physiobank veritabanından alınan dört tür ECG vurusu (Normal beat, congestive heart failure beat, ventricular tachyarrytmia beat, atrial fibrillation beat) %96.94 doğrulukla birleştirilmiş ağ modeli tarafından sınıflandırılmıştır[6].

2. Materyal ve Metod

2.1. EKG Verileri

Normal EKG işareti, kalbin dinlenme durumundaki taban seviyesi üzerine sıralanan belli başlı P, Q, R, S ve T adları verilen dalgalardan oluşur. Bazen T dalgasını takiben küçük genlikli bir U dalgası da olabilir. P dalgası olarak isimlendirilen kısım atriumların kasılması sonucu oluşur. PQ aralığı his demeti iletim zamanını gösterir. QRST dalgası, ventriküler kompleks olarak isimlendirilir. QRS, ventriküler depolarize olması anlamına gelir. His demeti ve kollarındaki iletim bozuklukları QRS dalgasında değişikliklere yol açar[7].



Şekil 1. Normal bir EKG vurusundaki önemli dalgalar

Normal bir kalp atımı, dalganın şekli, süresi, RR aralığı gibi temel parametreler içerir. Bu parametrelerdeki değişimler kalpte bir rahatsızlık olduğunu gösterir. EKG sinyalindeki bu düzensizlik **aritmî** olarak isimlendirilir ve aritmîler hasta için tehlikelidir[1], [7].

2.2. EKG Veritabanı

Massachusetts Teknoloji Enstitüsü – Beth İsrail Hastanesi (Massachusetts Institute of Technology – Beth Israel Hospital MIT-BIH) Aritmî Veri tabanında iki kanallı yaklaşık 30dk.’lık süreyi içeren 48 kayıt vardır. Bu kayıtların tamamı BIH Aritmî laboratuvarı tarafından 1975 ile 1979 yılları arasında çalışılan 47 denek üzerinden oluşturulmuştur. Bunların 25’i erkek ve 22’si kadındır. Erkekler 32-89 yaş aralığında, kadınlar ise 23-89 yaş aralığındadır. 201 ve 202 numaralı kayıtlar tek bir erkek denekten elde edilmiştir. Veri tabanı ayakta tedavi edilen hastalardan 24 saatlik 4000 kayıttan rasgele seçilmiş 23 tanesini içermektedir. Bu kayıtlar 100-124 aralığında numaralandırılmıştır. Kayıtların yaklaşık %60’ı hastanede yatan hastalardan, %40’ı ise ayakta tedavi edilen hastalardan elde edilmiştir. Kayıtların geri kalanı (200-234 arasında etiketlenmiş 25 kayıt.) kayıtlardan rasgele seçilmiş herhangi bir küçük örnek ile temsil edilemeyecek önemli klinik olayların kaydını korumak için aynı EKG setlerinden seçilmiştir[8], [9].

Çalışmada MIT-BIH aritmî veritabanından alınan 48 EKG kaydı kullanılmıştır. Bu kayıtlardan elde edilmiş Normal vuru (N) ve Erken karıncık kasılması vurusu (V) sınıflandırma algoritmalarına giriş verisi olarak hazırlanmıştır.

2.3. EKG verilerinin sınıflandırılması

EKG verilerinin sınıflandırılmasında karar verme işlemi genellikle üç adımda gerçekleştirilmektedir. Ön işleme ve normalizasyon işlemi, öznelik çıkarma işlemi ve sınıflandırma işlemi. Literatürde yapılan araştırmada ön işleme ve normalizasyon işlemi için genelde tüm araştırmacılar benzer yöntemlerle işlemlerini gerçekleştirirken öznelik çıkarma ve sınıflandırma işlemleri için farklı modeller üzerinde çalışmışlardır.

EKG işaretlerinin genliklerindeki değişim, hesaplanacak öznelik vektörlerini olumsuz yönde etkilemektedir. Farklı hastalardan alınan aynı tür EKG işaretlerinde dahi dikkate değer bir değişim gözlenebilmektedir. Normalizasyon ve ön işleme yapılarak EKG sinyalindeki değişiklikler (gürültü ve taban hattından sapma) en aza indirilir. İşlem yükünü azaltmak için

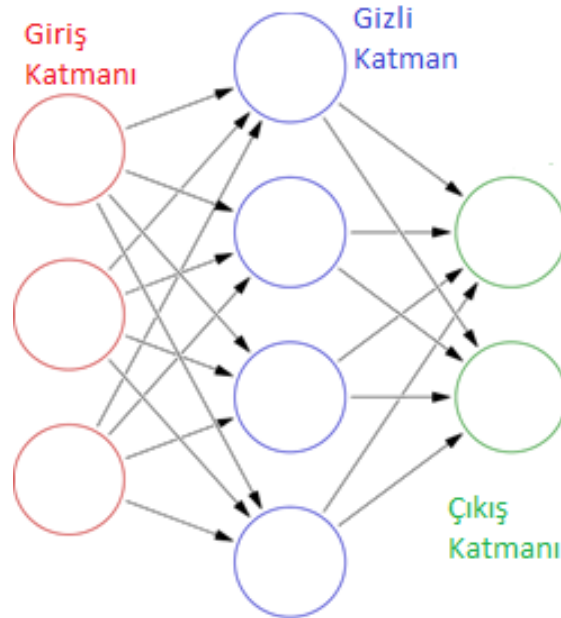
sinyalin ortalaması sıfırlanır. Böylece EKG öznitelik vektörünün hastanın yaşına, cinsiyetine ve ölçüm sisteminin parametrelerine olan bağımlılık minimum düzeye indirilmiş olur.

Yapılan bu çalışmada öznitelik çıkarma işlemi yapılmamış ve normalizasyon ve ön işleme adımından çıkan bir vuruluk EKG sinyali sınıflandırma algoritmalarına giriş verisi olarak sunulmuştur.

2.4. Yapay sinir ağları

Yapay sinir ağları (YSA), insan beyninin bilgi işleme tekniğinden modellenerek geliştirilmiş bir bilişim teknolojisidir. Biyolojik sistemlerde öğrenme nöronlar arasındaki sinaptik bağlantıların ayarlanması ile gerçekleştirilir. Bu işlem zaman içerisinde insanın tecrübeleri ile gerçekleştirilir. Öğrenme işlemi YSA için eğitime yoluyla örnekler kullanılarak yapılır. Eğitim işleminde girdi ve çıktı işlenerek nöronların ağırlıkları değiştirilir. Bu işlem hatanın belirlenmiş bir eşik değeri altına düşünceye kadar tekrarlanarak devam edilir.

YSA'lar, ağırlıklandırılmış şekilde birbirlerine bağlanmış birçok işlem biriminden (nöronlar) oluşan matematiksel sistemlerdir. Bir işlem birimi, aslında sık sık transfer fonksiyonu olarak adlandırılan bir denklemdir. Bu işlem birimi, diğer nöronlardan sinyalleri alır; bunları birleştirir, dönüştürür ve sayısal bir sonuç ortaya çıkartır. Genelde, işlem birimleri kabaca gerçek nöronlara karşılık gelirler ve bir ağ içinde birbirlerine bağlanırlar; bu yapı yapay sinir ağlarını oluşturmaktadır.



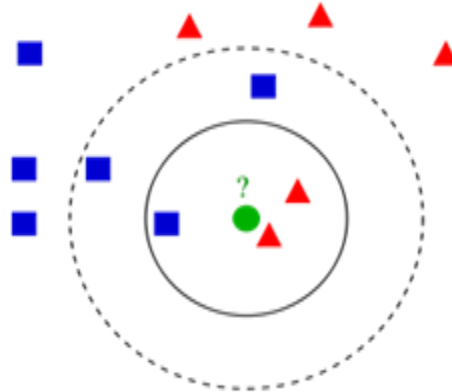
Şekil 2. Normal bir EKG vurusundaki önemli dalgalar

İlk defa 1950'li yıllarda ortaya atılan yapay zekâ terimi zaman içinde oldukça yoğun ilgi görmüş ve 40-50 yıllık bir zaman diliminde hayatın vazgeçilmez parçası olan sistemlerin doğmasına neden olmuştur. Yapay sinir ağları yapay zekâ çalışmalarının da ivmesini artırmıştır[10].

Model olarak birbirinden farklı birçok yapay sinir ağı vardır. Çalışma içerisinde sınıflandırma işlemleri için geri yayılım algoritması (çok-katmanlı algılayıcı – multi-layer perceptron) tarafından eğitilmiş bir ileri beslemeli yapay sinir ağı (feed forward neural network) kullanılmıştır.

2.5. K-En yakın komşu algoritması (KNN - K-Nearest Neighbour)

K en yakın komşuluk algoritması, tüm makine öğrenme algoritmaları arasında en basit, denetimli öğrenme algoritmasıdır. Algoritmanın eğitim aşamasında sınıf özelliği daha önceden belirlenmiş verilerden faydalanılır. K en yakın komşuluk algoritmasını uygulayabilmek için, öznitelik vektörlerinin oluşturulması gerekir. Aşağıdaki şekle göre yeşil daire sınıflandırılmak istenmektedir. Burada komşuluk için k parametresine ihtiyaç duyulur. K parametresi komşuluk sayısını belirtir. $k=3$ seçildiğinde birinci daire içerisindeki komşulara bakılmış olur. Çünkü birinci dairede 3 tane belirli sınıf var (iki tane kırmızı üçgen ve bir mavi kare). Bu durum için yani $k=3$ komşuluğunda kırmızı üçgen fazla olduğundan vektör, üçgen sınıfına atanır. Kesik çizgili daire alınırsa yani $k=5$ için, 3 kare 2 üçgen olduğundan vektör kare sınıfına atanır. Seçilen k komşuluğunda en fazla sınıflandırılmış veri hangi sınıftaysa, sorgu vektörü de o sınıfa atanır[11].

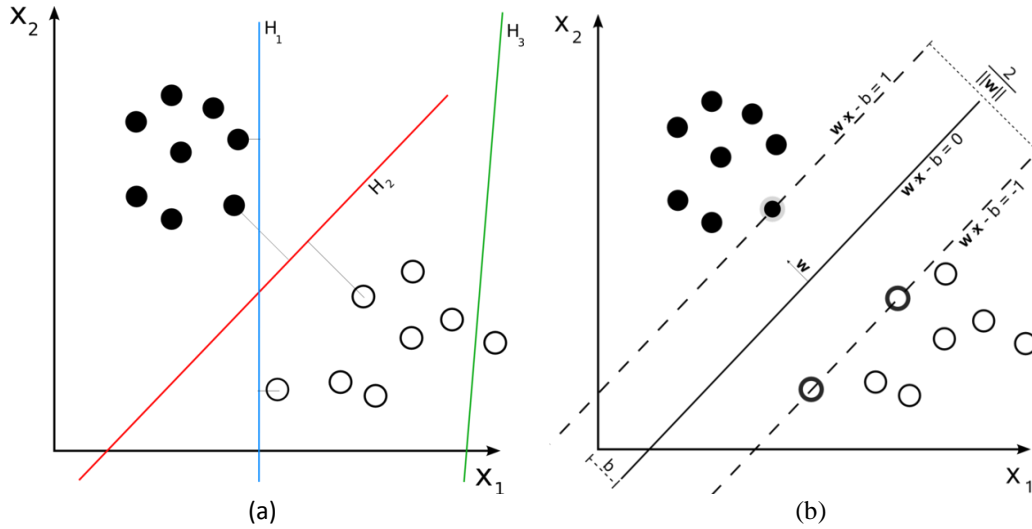


Şekil 3. K-En Yakın komşuluğun tespiti

2.6. Destek Vektör Makinesi (SVM – Support Vector Machine)

Destek Vektör Makinesi Algoritması, sınıflandırmada kullanılan, bir denetimli öğrenme algoritmasıdır. Bu algoritma son 15 yıl içinde geliştirilmiştir. Sınıflandırma konusunda kullanılan oldukça etkili ve basit yöntemlerden birisidir. Bir haritadaki iki bölgeyi en iyi şekilde birbirinden ayıracak bir sınır çizgisi çizilmek istenmektedir. En doğru çizgi çekilmelidir ki bölgeler doğru şekilde ayrılsın. SVM algoritmasının temelinde bu problemin çözümü vardır. Sınıflandırma için bir düzlemde bulunan iki grup arasında bir sınır çizilerek iki grubu ayırmak mümkündür. Bu sınırın çizileceği yer ise iki grubun da üyelerine en uzak olan yer olmalıdır. İşte SVM bu sınırın nasıl çizileceğini belirler. Farz edelim ki pozitif ve negatif örnekleri birbirinden ayıran bir aşırı düzlem var, bu düzlem üzerindeki noktalar $w \cdot x + b = 0$ eşitliğini sağlayacaktır, burada w aşırı düzleme olan normal ve $|b|/||w||$ aşırı düzlemden orijine olan dik uzaklıktır. Aşırı düzleme en yakın pozitif ve negatif örnekler arasındaki mesafeye ayırıcı aşırı düzleminin “tolerans”ı dersek,

destek vektör yöntemi bu “tolerans”ın en yüksek olduğu bir aşırı düzlemi bulmaya çalışır.



Şekil 4. (a) Toleransın belirlenmesi, (b) Grupların düzlemde gösterilmesi

Bu işlemin yapılması için iki gruba da yakın ve birbirine paralel iki sınır çizgisi çizilir ve bu sınır çizgileri birbirine yaklaştırılarak ortak sınır çizgisi üretilir[11].

2.6. Naive Bayes Metodu

Naive Bayes sınıflandırma algoritması, bayes olasılık kuralına dayanır. Eldeki verilere göre hipotezlerin doğru olma olasılığına göre hareket eder. Gelen verilere göre maksimum olasılığa sahip hipotez seçilir. Niteliklerin hepsi aynı derecede önemli ve birbirinden bağımsızdır.

Naive Bayes sınıflandırıcılar, bilinen bir sınıf için terim olasılıklarının hesaplanma yöntemine göre çok terimli (multinomial) ve çok değişkenli (multivariate) olmak üzere ikiye ayrılırlar. Çok terimli yöntemde terimlerin geçiş sayıları da dikkate alınırken, çok değişkenli yöntemde terimlerin sadece var olup olmadıklarına bakılır[11].

Naive Bayes sınıflandırıcısı Bayes teoreminin bağımsızlık önermesiyle basitleştirilmiş halidir. Bayes teoremi aşağıdaki denklemle (1) ifade edilir;

$$P(A|B) = (P(B|A)P(A))/P(B) \quad (1)$$

$P(A|B)$: B olayı gerçekleştiği durumda A olayının meydana gelme olasılığıdır.

$P(B|A)$: A olayı gerçekleştiği durumda B olayının meydana gelme olasılığıdır

$P(A)$, $P(B)$: A ve B olaylarının önsel olasılıklarıdır.

Burada önsel olasılık Bayes teoremine öznellik katar. Diğer bir ifadeyle örneğin $P(A)$ henüz elde veri toplanmadan A olayı hakkında sahip olunan bilgidir. Diğer taraftan $P(B|A)$ ardıl olasılıktır.

Çünkü veri toplandıktan sonra, A olayının gerçekleşmiş olduğu durumlarda B olayının gerçekleşme ihtimali hakkında bilgi verir[12].

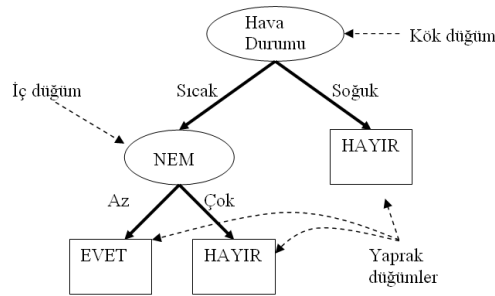
2.7. Karar Ağaçları (DT – Decision Trees)

Veri madenciliğinde bir karar ağacı, veriyi değil kararları temsil eder. Karar ağaçları, bir ağaç grafiği veya bir kararlar modeli olarak, olası sonuçları yorumlamayı sağlayan bir karar destek aracıdır. Hedef değer ile ilgili sonuçlar gözlemlerle eşleştirilir. Karar ağaçları için iki model vardır. Sınıflandırma ağacı (kesikli sonuç) veya Regresyon ağacı (sürekli sonuç). Karar ağaçları düğümler ve dallardan oluşan, anlaşılması oldukça kolay olan bir tekniktir. En basit şekilde her dal kendinden sonra evet veya hayır şeklinde en az iki dala ayrılmaktadır. Karar ağacında bulunan her bir dalın belirli bir olasılığı mevcuttur. Bu sayede son dallardan köke veya istediğimiz yere ulaşana dek olasılıklar birbiriyle çarpılarak hedefe giden en yüksek olasılıklı yol, hesaplanmış olur. Hesaplamanın verimliliği ağacın belirli dalları kesilerek ya da belirleyici özellikler değiştirilerek, yani daha az yararlı kurallardan kurtularak geliştirilebilir[11][13].

Ağacın 3 tür düğümü vardır:

- Kendisinden önce bir dal olmayan ve kendisinden sıfır veya daha fazla dal çıkabilen Kök düğümü
- Kendisinden önce ona doğru gelen sadece bir dal olan ve kendisinden iki veya daha fazla dal çıkan İç düğümler
- Kendisinden önce ona doğru gelen sadece bir dal olan ve kendisinden hiç dal çıkmayan Yaprak veya kutup (terminal) düğümler

Aşağıdaki şekilde “Hava durumu tenis oynamaya uygun mu?” sorusunun karar ağacı ile çözümü yer almaktadır.



Şekil 5. Hava Durumu problemi için örnek karar ağacı

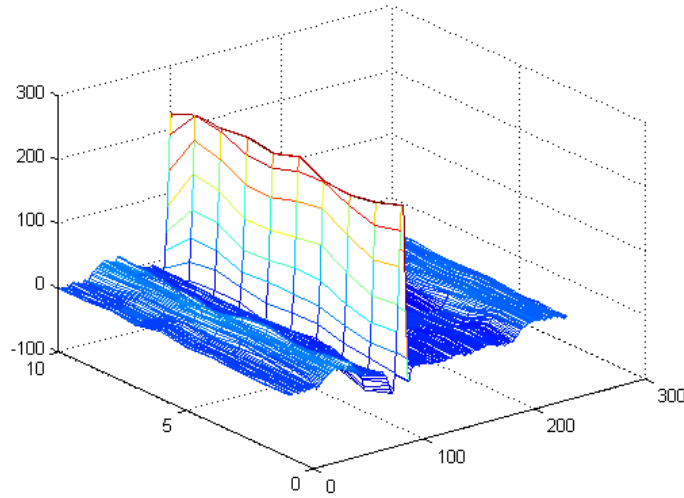
Yaygın olarak kullanılan birçok özel karar ağacı algoritması vardır.

Bunlardan en önemlileri:

- ID3 (Iterative Dichotomiser 3)
- C4.5 (ID3'ün geliştirilmiş hali)
- CART (Classification And Regression Tree)
- CHAID (CHi-squared Automatic Interaction Detector)
- MARS [13]

3. Sonuçlar

Çalışmada Matlab paket programı kullanılmıştır. Bu çalışmada kullanılan EKG verileri MIT-BIH aritmi veritabanından elde edilmiştir[8]. Matlab’da ön işlemeden geçmiş, gürültü ve taban hattından sapmalar giderilmiş sinyaller bir vuruluk zaman serisine dönüştürülmüştür. Sinyali ilk olarak pencere genişliği 3 olan bir ortalama filtresi kullanılarak sinyal üzerindeki gürültüler azaltılmıştır. Sinyalde bulunan daha büyük sapmaları gidermek için (özellikle ana çizgiden sapmalar - baseline drift) sinyal filtrelenmiştir. Burada ana çizgiden sapmanın giderilmesi için 2Hz’in altındaki frekans bileşenleri sinyalden alçak geçiren (low pass) filtre kullanılarak çıkarılmıştır. Bir vuru sinyalini elde etmek için sinyalde R tepeleri tespit edilmiş ve 100 örnek önce ve 150 örnek sonra olmak üzere 251 örneklilik bir vektör oluşturulmuştur. QRS yapılarındaki R tepesinin tespiti için veritabanında bulunan konum bilgisinden faydalanılmıştır. Bu vektöre ek olarak 1 tane sınıf etiketi eklenmiş ve sınıflandırma vektörümüzün uzunluğu 252 olarak belirlenmiştir. Şekil 6’da 10 vurdan oluşan örnek giriş verisi görülmektedir. Her vuru tipinde eğitim ve test için kullanılan vuru sayısı Tablo 1’de gösterilmiştir.



Şekil 6. On vurdan oluşan giriş verisi

Tablo 1. Vuru türleri, eğitim ve test için kullanılan vuru sayıları

Vuru Tipi	Açıklama	Veri Sayısı	
		Eğitim	Test
NORMAL (N)	Normal Vuru	1843	71320
PVC (V)	Erken Karıncık Kasılması	157	6055
Toplam		2000	77375

Oluşturulmuş bu öznetelik vektörü yapay sinir ağları, k-en yakın komşu algoritması, bayes sınıflandırıcı algoritması, karar ağaçları algoritması ve destek vektör makinesi algoritması kullanılarak sınıflandırılmış ve sonuçları irdelenmiştir.

4. Tartışma

Sınıflandırma modelinin başarımlarının karşılaştırılması için başarımlar ölçütü olarak duyarlılık (Sensitivity), özgüllük (Specificity) ve doğruluk (Accuracy) kullanılmıştır. Modelin başarısı, doğru sınıflandırılan örnek sayısı ve yanlış sınıflandırılan örnek sayıları ile ilgilidir. Test sonucunda ulaşılan başarımlar bilgilerine göre bu ölçütler aşağıdaki (2) (3) (4) formüllere göre hesaplanırlar.

$$\text{Doğruluk} = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

$$\text{Özgüllük} = \frac{TN}{TN+FP} \quad (3)$$

$$\text{Duyarlılık} = \frac{TP}{TP+FN} \quad (4)$$

Burada TP:(True Pozitif) Doğru sınıflandırılmış normal vuru, TN: (True Negatif) Doğru sınıflandırılmış erken karıncık kasılması vurusu, FP:(False Pozitif) Yanlış sınıflandırılmış normal vuru, FN:(False Negatif) yanlış sınıflandırılmış erken karıncık kasılması vurusu anlamına gelmektedir. Sınıflandırma sonuçları doğruluk, duyarlılık ve özgüllük kriterlerine göre değerlendirilmiş ve sonuçlar Tablo 2’de sunulmuştur. Sonuçlarda yapay sinir ağı ve k-en yakın komşu sınıflandırma modellerinin diğerlerine göre daha iyi sonuç verdiği gözlenmektedir. Burada YSA-1 modelinde 1 adet gizli katman mevcuttur ve gizli katmandaki düğüm sayısı 2 olarak belirlenmiştir ve başarımları 98,67% olarak yüksek bir doğruluk değerine ulaşılmıştır. YSA-2 modelinde 3 adet gizli katman mevcuttur ve gizli katmanlardaki düğüm sayısı sırasıyla 4, 4 ve 2’dir. İkinci YSA’nın başarımları; doğruluk değeri 98,61%, duyarlılık değeri 99,21% ve özgüllük değeri 91,56% olarak hesaplanmıştır. YSA-3 modelinde yine 3 adet gizli katman mevcut olup katmanlardaki düğüm sayısı sırasıyla 4, 20 ve 2’dir. Algoritma başarımlarında doğruluk değeri 98,85% olarak hesaplanmıştır. KNN-1 modelinde K=1 olarak belirlenmiş ve algoritma başarımları olarak doğruluk 99,31%, duyarlılık 99,76% ve özgüllük 94,05% olarak hesaplanmıştır. Çalışma göstermiştir ki KNN algoritması bu çalışma için en iyi sınıflandırma sonucunu vermiştir. Ancak K değerinin 1’den büyük olduğu değerlerde sınıflandırma başarımları düşmektedir. KNN-2 modelinde K=3 alınmış ve başarımların az da olsa KNN-1 sistemine göre düşük hesaplandığı görülmüştür.

Table 2. Sınıflandırma algoritmaları ve sınıflandırma başarımları (%)

Algoritma - Başarım	Doğruluk	Duyarlılık	Özgüllük
YSA-1	98,67	99,33	90,87
YSA-2	98,61	99,21	91,56
YSA-3	98,85	99,49	91,31
KNN-1	99,31	99,76	94,09
KNN-2	99,11	99,79	91,05
NAIVE BAYES	92,28	92,59	88,28
DT	98,76	99,43	90,93
DT - ID3	98,65	99,1	93,31
SVM	95,91	99,52	53,34

Yapılan çalışmada karar ağaçları (DT, DT – ID3) modellerinin de yüksek başarımlar elde ettiği görülmüştür. Naive Bayes ve destek vektör makinesi modellerinin diğer modellere göre daha düşük sınıflandırma başarımları sergiledikleri görülmüştür.

Sınıflandırıcı modellerinin farklı olması, sınıflandırılan vuruların türlerinin farklı olması, sınıflandırmada kullandıkları öznitelik vektörlerinin farklı olması gibi nedenlerden dolayı literatürde karşılaşılan yöntemler ile gerçekleştirilen çalışmaların başarımlarını karşılaştırmak oldukça zordur.

Kaynaklar

- [1] Z. Dokur, “Yapay Sinir Ağları ve Genetik Algoritmalar Kullanılarak EKG Vurularının Sınıflandırılması,” İstanbul Teknik Üniversitesi, 1999.
- [2] G. Bortolan, I. Jekova, and I. Christov, “Comparison of four methods for premature ventricular contraction and normal beat clustering,” in *Computers in Cardiology, 2005*, 2005, pp. 921–924.
- [3] M. Engin, “ECG beat classification using neuro-fuzzy network,” *Pattern Recognit. Lett.*, vol. 25, no. 15, pp. 1715–1722, Nov. 2004.
- [4] P. Erdoğmuş and A. Pekçakar, “Dalgacık dönüşümü ile ekg sinyallerinin özellik çıkarımı ve yapay sinir ağları ile sınıflandırılması,” pp. 13–15, 2009.
- [5] S. Yu and K. Chou, “Integration of independent component analysis and neural networks for ECG beat classification,” *Expert Syst. Appl.*, vol. 34, pp. 2841–2846, 2008.
- [6] İ. Güler and E. D. Übeyli, “ECG beat classifier designed by combined neural network model,” *Pattern Recognit.*, vol. 38, no. 2, pp. 199–208, Feb. 2005.
- [7] E. Yazgan and M. Korurek, *Tıp Elektronikliği*, 1. Baskı. İstanbul: İstanbul Teknik Üniversitesi Ofset Baskı Atolyesi, 1996.
- [8] G. Moody and R. Mark, “The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it,” in *[1990] Proceedings Computers in Cardiology*, 1990, pp. 185–188.
- [9] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001.
- [10] E. Öztemel, *Yapay Sinir Ağları*. İstanbul: Papatya Yayıncılık, 2003.
- [11] F. Gürcan, “Web içerik madenciliği ve konu sınıflandırması,” Karadeniz Teknik Üniversitesi, 2009.

- [12] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4. Edition. Burlington: Academic Press, 2009.
- [13] L. Rokach and O. Maimon, *Data mining with decision trees: theory and applications*. Singapore: World Scientific Publishing Co., 2008.