



KALP RİTİM BOZUKLUKLARININ ÇOK SINIFLI SINIFLANDIRILMASINDA RELİEFF YÖNTEMİ VE MAKİNE ÖĞRENİMİ TABANLI YAKLAŞIM

Çiğdem Gülüzar ALTINTOP^{1*}

¹Erciyes University, Faculty of Engineering, Department of Biomedical Engineering, 38039, Kayseri, Türkiye

Özet: Kardiyovasküler hastalıklar (KVH), dünya genelinde ciddi sağlık tehditleri arasında yer almakta ve tüm ölümlerin yaklaşık %32'sine neden olmaktadır. Bu nedenle, KVH'ların erken tanısı ve uygun tedaviye başlanması hayati önem taşımaktadır. Elektrokardiyografi (EKG), kalbin elektriksel aktivitesinin kaydedilmesiyle elde edilen önemli bir tanı yöntemidir. Ancak, aritmi gibi kalp rahatsızlıklarının tanısı, uzman klinisyenlerin gözle incelemesine dayanmakta ve bu süreç zaman alıcı ve zahmetli olabilmektedir. Bu çalışmada, EKG sinyallerinden otomatik aritmi tespiti için bir bilgisayar destekli sistem geliştirilmesi amaçlanmıştır. Normal sinüs ritmi, pacemaker ritmi ve 15 farklı aritmi olmak üzere toplam 17 farklı kardiyak aktivite; EKG sinyallerinden çok çeşitli öznelik çıkarımı, ReliefF kullanılarak öznelik seçimi ve farklı makine öğrenimi algoritmalarını kullanılması ile sınıflandırılmıştır. Elde edilen sonuçlar, K-En Yakın Komşu ve Rastgele Orman algoritmalarının %93,4 ve %99,23 doğruluk oranları ile en yüksek başarıyı gösterdiğini ortaya koymuştur. Bu çalışma, aritmilerin çok sınıflı sınıflandırmasında EKG sinyallerinden morfolojik, zaman, frekans, entropi ve karmaşıklık özneliklerini bir arada çıkararak, farklı makine öğrenimi algoritmaları kullanarak 17 farklı kardiyak aktivite sınıfını yüksek doğruluk oranlarıyla sınıflandırmıştır. Böylece, literatürdeki çalışmalardan farklılaşarak aritminin otomatik sınıflandırılmasına önemli bir katkı sağlamıştır.

Anahtar kelimeler: Elektrokardiyogram, Aritmi, Öznelik seçimi, ReliefF, Makine öğrenimi, Sınıflandırma

The ReliefF Method and Machine Learning-Based Approach in the Multi-Class Classification of Cardiac Arrhythmias

Abstract: Cardiovascular diseases (CVDs) represent a significant health threat worldwide and account for approximately 32% of all deaths. Therefore, early diagnosis and timely treatment of CVDs are crucial. Electrocardiography (ECG) is an important diagnostic method that involves recording the electrical activity of the heart. However, the diagnosis of heart disorders, such as arrhythmias, relies on the visual examination of expert clinicians, making this process time-consuming and labor-intensive. This study aims to develop a computer-assisted system for automatic arrhythmia detection from ECG signals. A total of 17 different cardiac activities, including normal sinus rhythm, pacemaker rhythm, and 15 different arrhythmias, were classified using various feature extraction methods from ECG signals, feature selection through ReliefF, and different machine learning algorithms. The results obtained show that the K-Nearest Neighbors and Random Forest algorithms achieved the highest accuracies of 93.4% and 99.23%, respectively. This study distinguishes itself from existing literature by extracting morphological, temporal, frequency, entropy, and complexity features from ECG signals for multi-class classification of arrhythmias and successfully classifying the 17 different cardiac activity classes with high accuracy. Thus, it provides a significant contribution to the automatic classification of arrhythmias.

Keywords: Electrocardiogram, Arrhythmia, Feature selection, ReliefF, Machine learning, Classification

*Sorumlu yazar (Corresponding author): Erciyes University, Faculty of Engineering, Department of Biomedical Engineering, 38039, Kayseri, Türkiye

E mail: cigdemacer@erciyes.edu.tr (Ç. G. ALTINTOP)

Çiğdem Gülüzar ALTINTOP



<https://orcid.org/0000-0001-8632-3385>

Gönderi: 13 Ekim 2024

Kabul: 22 Şubat 2025

Yayınlanma: 15 Mayıs 2025

Received: October 13, 2024

Accepted: February 22, 2025

Published: May 15, 2025

Cite as: Altintop ÇG. 2025. The ReliefF method and machine learning-based approach in the multi-class classification of cardiac arrhythmia. BSJ Eng Sci, 8(3): 627-641.

1. Giriş

Kardiyovasküler hastalıklar (KVH) günümüzde insan sağlığını tehdit eden ciddi hastalıklardan biridir. Dünya Sağlık Örgütü'nün (DSÖ) raporlarına göre, gerçekleşen tüm ölümlerin yaklaşık %32'sine tekabül eden 17,9 milyon insan KVH'lardan ölmüştür. Bu nedenle günümüzde tüm ölüm nedenleri arasında ilk sırada yer almaktadır. KVH'lardan meydana gelen ölümlerin %75'inden fazlası gelişmekte olan ülkelerde meydana gelmektedir. Çünkü düşük ve orta gelirli ülkelerde

yaşayan insanlar genellikle KVH için risk faktörleri olan kişilerin erken teşhisi ve tedavisi için gerekli sağlık programlarından yararlanamamaktadır (Organization, 2021). Ayrıca DSÖ'nün raporuna göre KVH yaygınlığı ve mortalitesi hala artmaktadır. Bu nedenle hastalığın erken tanısı ve uygun tedaviye başlanması hayati önem taşımaktadır. KVH tanısı, çok yönlü bir değerlendirme gerektirmekte ve genellikle hastanın öyküsü, kan testleri, stetoskop ile kalbin dinlenmesi, Elektrokardiyografi (EKG), ekokardiyografi (EKO) gibi yöntemler bir arada kullanılarak konulmaktadır.



Bu yöntemlerden EKG kalbin elektriksel aktivitesinin kayıt altına alınmasıdır. EKG, milyonlarca kardiyak hücrenin depolarizasyon potansiyelinin toplamı sonucu üretilir. Normal bir EKG sinyali, kalp atışının belirli kısımlarını temsil eden P, Q, R, S, T ve U dalgalarından oluşmaktadır. P dalgası atriyumlar (kulakçık) kasıldığında ortaya çıkmaktadır. Normal bir EKG'de P dalgasının genliği 0,25 mV'nin altındadır. Bu nedenle, daha yüksek genliğe sahip bir P dalgası atriyumlarda işlev bozukluğunun işareti olabilmektedir (Kossmann, 1953). Aritmi (kalp ritim bozukluğu), kalbin anormal elektriksel aktivitesinden kaynaklanan bir kalp rahatsızlığıdır. Her aritmi yaşamı tehdit edici değildir. Ventriküler flutter ve ventriküler fibrilasyon gibi birkaç kardiyak aritmi ani ölüme, hemodinamik çöküşe ve kalp durmasına neden olabilmektedir (Huikuri vd., 2001). Farklı aritmi türleri, kalpte genellikle farklı mekanizmalarla oluşur ve her biri için farklı bir tedavi yaklaşımı gerektirir (Gopinathannair vd., 2015). Aritmiler, girişimsel olmayan, ucuz ve uygulaması kolay olan EKG kullanılarak tanı almaktadır. Tanı uzman klinisyen tarafından EKG sinyalindeki dalgaların genliği, süresi ve düzenliliği ölçülerek ve normal sınırlarla karşılaştırılarak konmaktadır. Ancak, genliğin ve frekans bileşenlerinin, manuel yöntem kullanan bir klinisyen için doğru ve tutarlı bir şekilde tespit edilmesi çok zahmetlidir. Ayrıca doğru tanının konması için uzun süreli EKG kayıtlarının alınması, hekimin incelemesi açısından zaman alıcı ve uğraştırıcıdır. Daha da önemlisi, aritmi tespitinin uzun çalışma saatlerinin yorgunluğu ve gözle hekim tarafından incelenmesi nedeniyle gözden kaçırılabilir (Shilaskar ve Ghatol, 2013). Sonuç olarak görsel inceleme yanlış tanıya ve yanlış sınıflandırmaya yol açabilmektedir. Hekimler, karmaşık yapıları, bilinmeyen mekanizmaları ve klinik ilişkileri nedeniyle de aritmilerin teşhisinde zorlukla karşılaşmaktadır (Goldberger vd., 2017). Ani kardiyak ölümüne neden olabileceği için ölümcül aritmileri mümkün olan en erken zamanda doğru bir şekilde tespit etmek ve teşhis etmek çok önemlidir.

Klinik ortamlarda çekilen EKG'ler genellikle doktorların kalbin aktivitesini kapsamlı bir şekilde incelemeleri için yeterli değildir. Bazı aritmiler, EKG veya Efor gibi yöntemlerle teşhis edilememektedir. Bu nedenle, şüpheli aritmilerin teşhisi genellikle hastaların günlük aktiviteler sırasında kalbin işleyişini sürekli olarak izlemek için holter cihazı ile birkaç gün (bir veya iki gün) kalp atımı kayıt altına alınmaktadır (Zimetbaum ve Goldman, 2010). Tabi holter cihazıyla alınan sonuçlar ve EKG kayıtlarının yorumlanması anormal kalp atışlarının tanımlanmasına dayanır ve bu da manuel olarak yoğun bir işlemdir. Bu nedenle, kısa süreli EKG sinyalleri kullanılarak aritmilerin bilgisayar destekli tanısı, klinisyenlerin kardiyovasküler hastalıkları önemli ölçüde yönetmesine yardımcı olacaktır. Geliştirilecek bilgisayar destekli bu sistemlerin maliyet açısından avantaj sağlayacağı, teşhis doğruluğunu ve hasta iyileşme sonuçlarını önemli ölçüde iyileştirdiği bildirilmektedir (Dinakarrao vd., 2019). Bu

amaçla araştırmacılar EKG sinyallerinden otomatik aritmi tespiti için pek çok çalışma gerçekleştirmiştir (Wang vd., 2020; Ojha vd., 2022; Wan vd., 2024). Çalışmalar genel olarak iki aşamadan oluşmaktadır: öznitelik çıkarımı ve sınıflandırma. Öznitelik çıkarımı ve analiz uygulamalarında zaman analizi (Yeh vd., 2012), frekans analizi yöntemleri (Gupta ve Mittal, 2020; Mazaheri ve Khodadadi, 2020); morfolojik özellikler (Mazaheri ve Khodadadi, 2020; Yang ve Wei, 2020), istatistiksel özellikler (Mian Qaisar ve Hussain, 2023; Nascimento vd., 2020) kullanılmıştır. EKG sinyalleri durağan olmayan sinyaller olduğu için zaman-frekans analizi yöntemlerinden Kısa Zamanlı Fourier Dönüşümü (KZFD) (Daydulo vd., 2023), S dönüşümü (Das ve Ari, 2014; Mishra vd., 2021) veya Dalgacık Dönüşümü (Sharma vd., 2023; Allabun, 2024) yöntemleri kullanılarak da öznitelik çıkarımı gerçekleştirilmiştir.

Aritmi sınıflandırılmasında genel olarak Sinir Ağları (Mitra ve Samanta, 2013), Destek Vektör Makinesi (DVM) (Kohli ve Verma, 2011), Karar Ağacı (Sahoo vd., 2020), K-EnYakın Komşu (K-NN) ve Naive Bayes (Devadas, 2021) makine öğrenimi yöntemleri kullanılmıştır. Çalışmaların birçoğu sınıflandırma aşamasında makine öğrenimi yaklaşımı kullansa da derin öğrenme ile aritminin tespiti ve otomatik sınıflandırması için de pek çok çalışma bulunmaktadır. Vasquez vd. (2024) gerçekleştirdikleri derleme çalışma da aritmi sınıflandırmasında kullanılan derin öğrenme çalışmalarına kapsamlı bir şekilde yer verilmiştir. Bilinen bir gerçek var ki derin öğrenme çalışmalarında her ne kadar öznitelik çıkarımı, seçimi yapılmadan sınıflandırma ham sinyallerden gerçekleştirilebilse de, derin öğrenme mimarilerinin oluşturulmasının zorluğu, hesaplama yükünün fazla olması, zaman alıcı olması ve donanımı iyi bilgisayarlar gerektirmesi nedeniyle dezavantaja sahiptir (Afaq ve Manocha, 2024). Aritmi sınıflandırmasında doğruluk oranı yüksek bazı çalışmalardan bahsetmek gerekirse: Sangaiah vd. (2020) kalp aritmilerinin hızlı tespitine yönelik Dalgacık (Wavelet) Dönüşümü (DD) tabanlı öznitelik çıkarımı ve gizli Markov modeli kullanarak sınıflandırma gerçekleştirmiştir. Model, MIT-BIH aritmi veri tabanından elde edilen Normal Sinüs Ritmi (NSR), Sağ Dal Bloğu (right bundle branch block/ RBBB), Sol Dal Bloğu (left bundle branch block/ LBBB), Prematüre Ventriküler Kasılma (PVC) ve Atriyal Prematüre Kasılma (APC) aritmileri (5 sınıf) %99,7 doğruluk ile ayırt edebilmiştir. Yang ve Yan (2020), aynı veri setinde EKG aritmilerinin sınıflandırılması için çok boyutlu öznitelik çıkarma ve seçme yöntemi ortaya koymuştur. İlk olarak, görüntü şekil özniteliklerini çıkarmak için yeni çok aralıklı simetrik nokta deseni yöntemi önerilmiş ve bu öznitelikler zaman-frekans, morfolojik ve RR aralık öznitelikleriyle birleştirilerek yeni bir öznitelik seti oluşturulmuştur (toplam 316 öznitelik). Ardından, genetik algoritma (GA) ve DVM ile entegre edilen, ağırlık tabanlı ReliefF (WReliefF) yöntemi geliştirilmiş ve bu yöntem öznitelik seçimi için kullanılmıştır. Beş sınıflı sınıflandırma sonucu yöntem %99,42 duyarlılık, %99,83

özgüllük ve %99,74 doğruluk oranlarına ulaşmıştır. Dhyani vd. (2023), EKG sinyallerini analiz etmek için 3 Boyutlu (B) Ayırık Dalgacık Dönüşümü (ADD) ve DVM yöntemlerini kullanmıştır. 3B dalgacık dönüşümü, sinyalden gürültünün giderilmesi ve öznel çıkarmı için kullanılmış, SVM ise EKG sinyallerini dokuz farklı kalp atım türüne sınıflandırmıştır. China Physiological Signal Challenge (CPSC) 2018 aritmi veri seti kullanılarak yapılan çalışmada, Symlet-8 (Sym8) filtresiyle seviyelendirilmiş 4. düzey yaklaşık katsayılarıyla en yüksek doğruluk oranı %99.02 olarak elde edilmiştir.

Literatür incelendiğinde aritminin otomatik tespiti için gerçekleştirilen çalışmaların çok olması, bu konunun kardiyovasküler hastalıkların erken teşhisi ve tedavisi için büyük önem taşıdığını göstermektedir. Çalışmaların çoğunda az aritmi türü ile çalışılmıştır. Ayrıca EKG'nin sadece genişliğinde ve süresindeki değişiklikleri belirleyerek kardiyak anormalliği teşhis etmek ve analiz etmek doğrusal olmayan ve durağan olmayan sinyalde bulunan gizli bilgiyi çözmek açısından zordur ve gürbüz analiz sağlamamaktadır. Bu çalışmada, EKG sinyallerindeki 17 farklı kardiyak aktiviteyi otomatik olarak tanımlayan bir sistem geliştirilmiştir. Bu sistem, zaman, frekans ve doğrusal olmayan analiz yöntemlerinden elde edilen özelliklerin birleştirilmesiyle aritmi tanısında bilgisayar tabanlı bir yaklaşım sunmayı ve önemli klinik bilgileri ortaya çıkarmayı hedeflemektedir. Bu amaçla, EKG sinyallerinden öznel çıkarmı ve Relief ile öznel seçimi işlemi gerçekleştirilerek farklı makine öğrenimi algoritmaları ile 17 farklı sınıfın sınıflandırma başarıları karşılaştırılmıştır.

2. Materyal ve Yöntem

2.1. Materyal

Bu çalışmada kullanılan EKG sinyalleri, PhysioNet (Goldberger vd., 2000) ve MIT-BIH Aritmi veri tabanından (Moody ve Mark, 2001) elde edilmiştir. Verilere açık erişimle "https://data.mendeley.com/datasets/7dybx7wyfn/3" bağlantısından ulaşılabilir (Plawiak, 2017). Veritabanı, 45 hastadan alınan EKG sinyallerini içermektedir. EKG sinyalleri, normal sinüs ritmi, pacemaker ritmi ve 15 farklı aritmi olmak üzere toplam 17 sınıftan oluşmaktadır ve her bir sınıf için en az 10 sinyal bulunmaktadır. EKG sinyalleri tek kanallı olarak Derivasyon-II'den 360 Hz. örnekleme frekansı ile örneklenecek elde edilmiştir. Analiz için, rastgele seçilen ve üst üste binmeyen, 10 saniyelik (3600 örnek) 1000 adet EKG sinyali kullanılmıştır.

Bu çalışmada kullanılan EKG ritimleri ve kalp rahatsızlıkları şu şekildedir: Normal Sinüs Ritmi (NSR- Normal Sinus Rhythm), kalbin normal bir sıklıkta döngüsel kasılmasıdır. Atrial Prematüre Atım (APB- Atrial Premature Beat), atriyumda zamanından önce oluşan erken atımı belirtirken, Atrial Flutter (AFL- Atrial Flutter) ise, atriyumda hızlı elektrik impulslarının etkisi ile ilişkili düzenli oluşan ritim bozukluğudur. Atrial

Fibrilasyon (AFIB- Atrial Fibrillation) ise düzensiz ve hızlı atriyum kasılmaları olarak tanımlanmaktadır. Supraventriküler Taşikardi (SVTA- Supraventricular Tachycardia) neredeyse her zaman kalbin üst bölgelerinde ortaya çıkan anormal hızlı aktivasyonu belirtirken, Wolff-Parkinson-White Sendromu (WPW - Wolff-Parkinson-White syndrome) ise kalp elektrik sinyallerinde, ek bir iletim yolunun bulunmasından kaynaklanmaktadır. Ventriküler Prematüre Atım (PVC- Premature Ventricular Contraction) ise ventriküllerde zamanından önce oluşan erken atımı belirtmektedir (Goldberger vd., 2017). Ventriküler Bigemini (Bigeminy) ve Ventriküler Trigemini (Trigeminy) ise her iki veya üç atımda bir erken ventriküler atımların olduğu ritim bozukluğu durumudur. Ventriküler Taşikardi (VT- Ventricular Tachycardia) ise ventriküldeki kalp atışlarının sayıca çok olmasını ifade etmektedir. İdioventriküler Ritim (IVR - Idioventricular Rhythm) ventriküllerde yavaş ancak düzenli atım olarak tanımlanırken, Ventriküler Flatter (VFL -Ventricular Flutter) ise hızlı ve düzenli atımı belirtmektedir. Füzyon Atımları (Fusion Beats), normal ve anormal atımların bir arada olduğu ritim bozukluğudur. Sol Dal Bloğu (LBBB- Left Bundle Branch Block) ve Sağ Dal Bloğu (RBBB- Right Bundle Branch Block), kalbin elektrik sinyallerinin sol veya sağ dalda iletiminde blokaj olduğu durumlar iken Ani Dal Bloğu (SDHB- Sudden Bundle Branch Block) ani gelişen blokaj durumudur. Prematüre Ritim (PR- Premature Rhythm) ise kalbin erken atımlarını ifade etmektedir (Goldberger vd., 2017). Tablo 1'de çalışmada kullanılan 17 sınıf için EKG sinyali sayısı (veri sayısı) verilmiştir.

2.2. Önileme

EKG sinyallerinin öncelikle kas hareketi gibi yüksek frekanslı bileşenlerden arındırılması için 4. Dereceden bant geçiren [1-100 Hz] Butterworth filtre ile gürültü eliminasyonu gerçekleştirilmiştir. Şekil 1'de NSR ve AF için ham ve filtrelenmiş sinyaller gösterilmektedir. Şekil 2'de ise çalışmada yer alan 17 farklı kalp ritmine ait 2 saniyelik filtrelenmiş EKG sinyalleri verilmiştir.

2.3. Öznel çıkarmı

Öznel çıkarmı, ham verilerin çok boyutlu, karmaşık ve gürültülü olabilmesi nedeniyle veriden anlamlı bilgiler elde ederek, veriyi yorumlama ve makine öğrenimi modelleri ile doğruluğu yüksek sınıflandırma amacıyla kullanılmaktadır. Özellikle öznel çıkarmı EEG veya EKG gibi fizyolojik sinyallerinden frekans eksenini, zaman eksenini veya istatistiksel özellikler çıkararak, modellerin karmaşık hastalıkları tanıması açısından önemlidir.

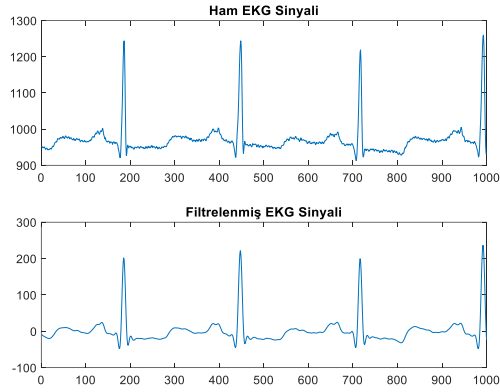
EKG sinyallerinden öznel çıkarmında yaygın olarak RR aralığı, QRS bileşik dalgası ve T dalgası sürelerine ek olarak, morfolojik özellikler zaman ekseninde elde edilen özellikler olarak kullanılmaktadır (Martis, Acharya ve Adeli, 2014; Mondéjar-Guerra, Novo, Rouco, Penedo ve Ortega, 2019). Zaman ekseninde öznel çıkarmı yapılması basit ve hızlıdır, ancak bu özellikler EKG sinyallerinin ince değişikliklerini ve gizli bilgilerini ortaya çıkarmakta yeterli değildir. Bazı istatistiksel

hesaplamalar ise EKG sinyallerinin daha yüksek dereceli kümülatif özelliklerini çıkararak gizli bilgi elde edebilmektedir (Martis vd., 2013). Öznitelik çıkarımında

frekans eksenli analizi yöntemleri de kullanılmakla birlikte öznitelik çıkarımında yeni araştırmalarda gerçekleştirilmektedir (Marinho vd., 2019).

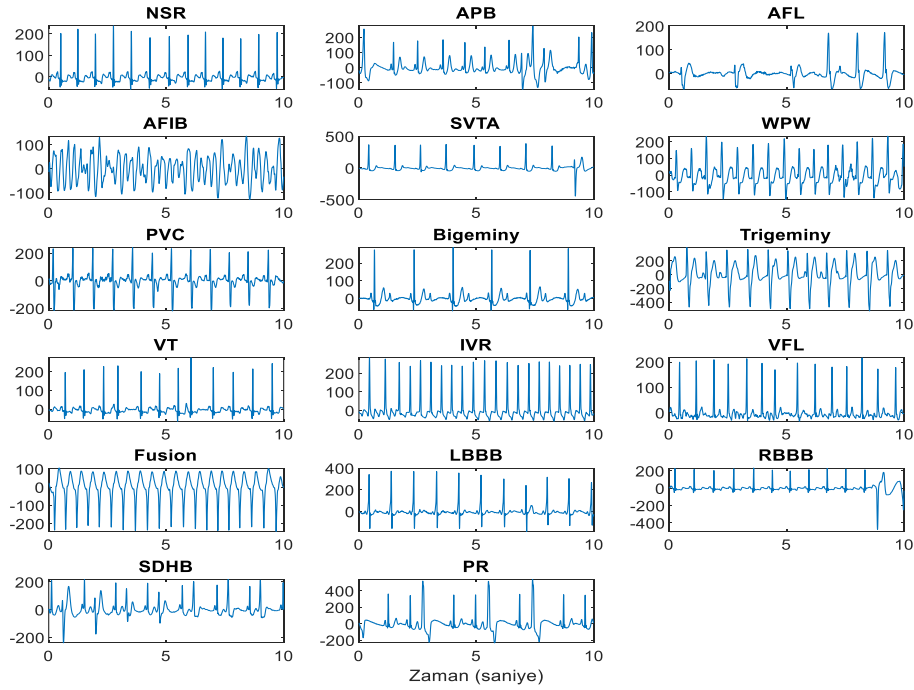
Tablo 1. 17 sınıf için veri sayısı

NSR	APB	AFL	AFIB	SVTA	WPW	PVC	Bigeminy	Trigeminy	VT	IVR	VFL	Fusion	LBBB	RBBB	SDHB	PR	Toplam
283	66	20	135	13	21	133	55	13	10	10	10	11	103	62	10	45	1000



Şekil 1. Önışlemden geçirilmiş EKG sinyalleri, normal sinüs ritmi.

Şekil 2. Önışlemden geçirilmiş EKG sinyalleri, atriyal flutter.



Şekil 3. 17 farklı kalp ritmine ait EKG sinyalleri.

Bu çalışmada, 17 farklı kalp ritmine ait EKG sinyallerinden anlamlı bilgi elde etmek için çeşitli öznitelikler çıkarılmıştır. R-R aralıklarından elde edilen öznitelikler kalp ritmi ve otonom sinir sistemi aktivitesini yansıtırken, QRS kompleksinden hesaplanan öznitelikler kalp atış hızını ve QRS genliklerinin istatistiksel özelliklerini içermektedir. Güç spektrumundan çıkarılan öznitelikler sinyalin frekans bileşenlerini ve enerjisini, filtrenmiş EKG sinyalinden elde edilen öznitelikler ise

sinyalin genlik, entropi ve istatistiksel özelliklerini temsil etmektedir. Dalgacık dönüşümü ve DFA ile hesaplanan öznitelikler ise sinyalin frekans bantları ve kısa vadeli korelasyonlarını analiz etmektedir. Tablo 2’de çıkarılan öznitelikler kısaltmaları ve kategorilerine göre özetlenmiştir. Her EKG sinyali için 78 adet öznitelik elde edilmiştir

I. R-R aralıklarından hesaplanan öznitelikler: R-R aralıklarının varyansı ile kalp ritmindeki değişkenlik

(RR_variance), HRV (Heart Rate Variability)/Kalp Hızı Değişkenliği (KHD) otonom sinir sistemi aktivitesi hakkında bilgi vermektedir. RR_mean, R-R aralıklarının ortalamasıyla ortalama kalp atış hızını temsil eder. RMSSD (Root Mean Square of Successive Differences), ardışık R-R aralıklarındaki değişikliklerin kök ortalama karesi ile kısa vadeli HRV'yi ölçmektedir. NN50, ardışık RR aralıkları arasındaki sürenin 50 ms'den büyük olduğu çiftlerin sayısını gösterirken, pNN50, bu sayının toplam R-R aralıklarına oranını verir. HRV_Triangular_Index, KHD üçgen endeksi olarak, R-R histogramının taban genişliğinden türetilmektedir ve R-R histogramının taban genişliği ile toplam R-R aralıklarının oranıdır.

II. QRS kompleksinden hesaplanan öznitelikler: bpm (beats per minute), dakika başına düşen kalp atış sayısını (nabız), qrs_mean, qrs_var ve qrs_min ise sırasıyla QRS kompleksinin genliklerinin ortalamasını, varyansını ve minimum değerini göstermektedir. qrs_skewness ve qrs_kurtosis, ise QRS kompleksinin çarpıklığını ve basıklığını analiz etmektedir. Bu öznitelikler Pan-

Tompkins algoritması ile tespit edilmiş QRS komplekslerine ait genlik değerlerinden hesaplanmıştır. Ayrıca bu yöntemdeki filtreleme işlemi olmadan ham QRS kompleksine ait değerlerde hesaplanmıştır (raw olarak belirtilen öznitelikler).

III. Güç Spektral Yoğunluğu (GSY) grafiğinden hesaplanan öznitelikler: TotalP, sinyaldeki toplam gücü temsil ederken (GSY grafiğinin toplam alanı), relative_energy, güç spektrumunun göreceli enerjisini göstermektedir. spectral_entropy, frekans bileşenlerinin düzensizliğini ve sinyalin karmaşıklığına ait bilgi veren hesaplamadır. Şekil 3'te verilen GSY grafiğinde görüldüğü gibi Area1 (A) ve Area2 (B), güç spektrumundaki farklı alan değerleridir. peak_freq, en yüksek güç yoğunluğuna sahip frekansı, max_power ve min_power ise spektral güç yoğunluğunun maksimum ve minimum değerlerini göstermektedir. mean_power ve std_power, sırasıyla güç spektrumunun ortalamasını ve standart sapmasını verirken, kurtosis_psd ve skewness_psd, güç spektrumunun basıklığını ve çarpıklığının ölçüsüdür.

Tablo 2. Çıkarılan öznitelikler

Kategori	Öznitelik kısaltması	Adet
R-R Aralıklarından		
Hesaplanan	RR_variance, HRV, RR_mean, RMSSD, NN50, pNN50, HRV_Triangular_Index	7
Öznitelikler		
QRS		
Kompleksinden	bpm, qrs_mean, qrs_var, qrs_min, qrs_skewness, qrs_kurtosis, qrs_raw_mean,	
Hesaplanan	qrs_raw_var, qrs_raw_max, qrs_raw_min, qrs_raw_skewness, qrs_raw_kurtosis	12
Öznitelikler		
GSY Grafiğinden	TotalP, relative_energy, spectral_entropy, Area1, Area2, peak_freq, max_power,	
Hesaplanan	min_power, diff_maxpower_minpower, mean_power, std_power, kurtosis_psd,	13
Öznitelikler	skewness_psd	
EKG Sinyalinden	ecg_mean, ecg_max, ecg_min, ecg_ShanEn, ecg_LogEnergyEn, ecg_ApEn, ecg_SampEn,	
Çıkarılan	ecg_skewness, ecg_kurtosis, ecg_onesec_ApEn, ecg_onesec_SampEn, ecg_onesec_ShanEn,	19
Öznitelikler	ecg_onesec_LogEnergyEn, energy, power, activity, mobility, complexity, Alpha1	
Dalgacık (Wavelet)	LF_power, HF_power, LF_HF_ratio, D3_max, D3_min, D3_diff, D3_dist, D3_mean, D3_std,	
Dönüşümüyle	D3_skew, D3_rel_energy, D4_max, D4_min, D4_diff, D4_dist, D4_mean, D4_std, D4_skew,	27
Hesaplanan	D4_rel_energy, D5_max, D5_min, D5_diff, D5_dist, D5_mean, D5_std, D5_skew,	
Öznitelikler	D5_rel_energy	27

I. Filtrelenmiş EKG sinyalinin çıkarılan öznitelikler: ecg_mean, EKG sinyalinin genliğinin ortalama değerini, ecg_max ve ecg_min ise sırasıyla maksimum ve minimum değerleridir. ecg_ShanEn, Shannon Entropisi ile sinyalin rastgeleliliğini ölçerken, ecg_LogEnergyEn, sinyalin enerji temelli logaritmik entropisini analiz etmektedir. ecg_ApEn (Approximate Entropy), Yaklaşık Entropi algoritması ile sinyaldeki düzenlilik ve öngörülebilirliği, ecg_SampEn (Sample Entropy) Örnek Entropi algoritması ile sinyaldeki düzensizliği ölçmektedir. Entropi hesaplamaları her bir saniyelik veri için gerçekleştirilerek elde edilen sonucun ortalaması alınmıştır. Entropi algoritmaları için gereken gömülü boyut m=2, ve tolerans değeri r=2(*zaman serisinin standart sapması) olarak literatür incelemesi sonucu belirlenmiştir (Liou vd., 2022). Ayrıca, ecg_skewness ve

ecg_kurtosis, sinyalin çarpıklığını ve basıklığını analiz ederken, energy ve power, EKG sinyalinin toplam enerjisi ve gücünü göstermektedir. Hjorth parametreleri aktivite, mobilite ve kompleksite (activity, mobility, ve complexity) ise sırasıyla bir sinyalin varyansını; mobilite, sinyalin zaman içindeki frekans bileşenlerinin yayılımını; kompleksite ise sinyalin şekil ve yapısındaki karmaşıklık düzeyini vermektedir.

II. Dalgacık dönüşümünden hesaplanan öznitelikler: Öncelikle EKG sinyali "Daubechies 4" ana dalgacık fonksiyonu kullanılarak 5 seviyeye ayrıştırılmıştır. Üçüncü, dördüncü ve 5. Seviyedeki detay katsayıları yeniden yapılandırıldıktan sonra bazı öznitelikler elde edilmiştir. LF_power ve HF_power, düşük ve yüksek frekans bandındaki gücü gösterirken, LF_HF_ratio, düşük frekans bandı gücünün yüksek frekans bandı gücüne

oranıdır ve otonom dengeyi ifade etmektedir. D3_max, D3_min, ve D3_diff, üçüncü seviyedeki dalgacık katsayılarının maksimum, minimum ve fark (maksimum-minimum) değerlerini temsil ederken, D4_max ve D5_max gibi öznitelikler de dördüncü ve beşinci seviyedeki dalgacık katsayılarına ait öznitelikleri vermektedir.

III. Eğilimsiz Dalgalanma Analizi/Detrended Fluctuation Analysis (DFA) yöntemiyle hesaplanan Alpha1 değeri, sinyalin kısa vadeli korelasyon özelliklerini ortaya koymaktadır. Bu yöntemi kısaca açıklamak gerekirse:

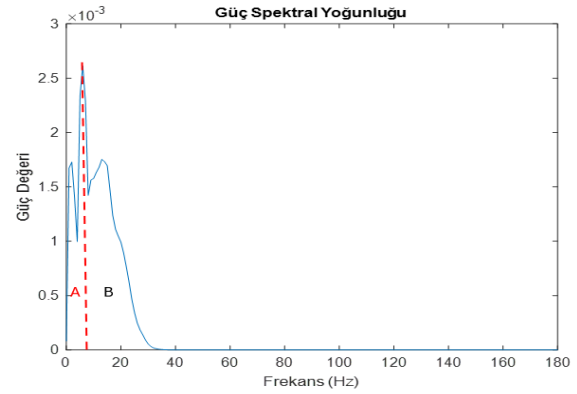
Eğilimsiz Dalgalanma Analizi 1995 yılında Peng vd. (Peng vd., 1995) tarafından önerilen durağan olmayan verilerin analizinde zaman dizilerindeki kısa ve uzun vadeli korelasyonları ve fraktal özellikleri incelemek için yaygın olarak kullanılan bir yöntemdir. Biyomedikal sinyallerde düzenlilik ve karmaşıklık analizinde tercih edilmektedir. Korelasyonlar, zaman serisinin fraktal öz-yakınlık boyutunun bir tahminini sağlayan bir ölçekleme üssü açısından ölçülmektedir. Bu yöntem için rasgele yürüyüşlü analitik yaklaşımda denilmektedir çünkü zaman dizisini kümülatif bir yürüyüşe dönüştürmektedir (Mizobuchi vd., 2021). Öncelikle, belirli bir zaman serisi $y(i)$, $i = 1, 2, 3, \dots, N$ adet örneğe sahip ve örnekleme periyodu f_s için DFA yöntemi, zaman dizisini eşit büyüklükteki n adet pencereye böler; bu $\tau = n \cdot f_s$ zaman ölçeğine karşılık gelmektedir (eşitlik 1).

$$F(n) = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y(i) - Y_n(i))^2} \quad (1)$$

Ardından belirlenen pencere boyutuna göre verileri uydurmak için doğrusal bir fonksiyon kullanır ve dalgalanmalara ait trend çıkarılır. Pencerelere ayrılmış verinin dalgalanma büyüklüğü hesaplanır ve pencere boyutuna karşı log-log ölçeğinde grafiği çizilmektedir. Elde edilen grafiğin eğimi Alfa değeridir ve sinyalin fraktal özellikleri hakkında bilgi verir. Alfa1 değeri kısa süreli korelasyonlar, alfa2 değeri ise uzun süreli korelasyonlar hakkında bilgi vermektedir (Mizobuchi vd., 2021). Bu çalışmadaki sinyaller 10 saniyelik olduğu için alfa1 değeri incelenmiştir. Alfa değeri $\alpha = 0,5$ ise bu sinyalin beyaz gürültü olduğu anlamına gelir. Öte yandan, $\alpha < 0,5$ ise sinyaldeki korelasyonlar antikalcıdır (yani bir artışın ardından bir azalma gelmesi çok olasıdır ve bunun tersi de geçerlidir) ve $\alpha > 0,5$ ise sinyaldeki korelasyonlar kalıcıdır (yani bir artışın ardından bir artış gelmesi çok olasıdır ve bunun tersi de geçerlidir) (Rodriguez, Lerma, Echeverria ve Alvarez-Ramirez, 2008). Bu nedenle $\alpha < 0,5$ ise sinyallerdeki düzensizliğin daha fazla olduğu, $\alpha > 0,5$ ise sinyallerin daha düzenli bir örüntüye sahip olduğu bilgisine ulaşılmaktadır.

Eşitlik 1'de verilen $F(n)$, belirli pencere boyutundaki dalgalanma fonksiyonunu, $Y(i)$ zaman dizisini, $Y_n(i)$ trend çıkarılmış diziyi ifade etmektedir. Şekil 4'te NSR'e ait EKG sinyali için pencere boyutu 100'der adım değiştirilerek

elde edilmiş log-log trend grafiği ve hesaplanan alfa değeri verilmiştir.



Şekil 4. NSR için GSY.

Aritmi tespiti çalışmalarının büyük bir kısmı, genellikle yalnızca zaman analizi üzerine odaklanmıştır; bunu takiben frekans analizi yöntemleri de kullanılmaktadır. Bu çalışmada ise hem zaman hem de frekans eksenindeki analizler bir araya getirilerek, entropi ve karmaşıklık tabanlı yaklaşımlarla desteklenen hibrit bir yöntem geliştirilmiştir. Böylece, otomatik aritmi tespitinde daha kapsamlı ve etkili bir model sunulmuştur.

2.4. Relief ile Öznitelik Seçimi

Öznitelik seçimi yöntemleri Filtreleme, Sarmalayıcı (wrapper) ve hibrit yöntemler olmak üzere üç ana kısma ayrılmaktadır. Relief yöntemi, Kira ve Rendell (Kira ve Rendell, 1992) tarafından iki sınıflı veriler için; Kononenko (Kononenko, 1994) tarafından ise çok sınıflı veriler ve eksik değerler içeren veri kümelerinde öznitelik seçimi yapmak için geliştirilmiş bir filtre tabanlı öznitelik seçimi yöntemidir. Relief yöntemi, her bir özniteliğin sınıflar arasındaki ayırt etme yeteneğini ölçerek özniteliklerin ağırlığını hesaplamaktadır. Bu işlemi verilerin sınıf içi en yakın komşuları ve sınıf dışı en yakın komşuları arasındaki uzaklığın analizine dayalı gerçekleştirmektedir. Özniteliklerin ağırlığı Denklem (2)'deki gibi hesaplanmaktadır ve hesaplama sırasında gerçekleştirilen işlemler şu şekildedir: İlk olarak, eğitim setinden rastgele bir R örneği seçilir ve R'nin ait olduğu sınıftan en yakın komşu örneği Y seçilir. Ardından, diğer sınıflardan en yakın komşu örneği U seçilir. Eğer R ve Y arasındaki mesafe, R ve U arasındaki mesafeden küçükse, bu durum bu özniteliğin sınıflandırma için uygun olduğunu gösterir ve bu özniteliğin ağırlığı artırılmalıdır; tersine, eğer mesafe büyükse, bu özniteliğin sınıflandırma için uygun olmadığını gösterir ve ağırlığı azaltılmalıdır. Bu süreç, bütün özniteliklerin ağırlığı elde edilene kadar tekrarlanmaktadır (Yang ve Yan, 2020). Eşitlik 2'de $W(F_i)$ F_i özniteliğinin ağırlığını, $dist(F_i, R, Y)$ F_i özniteliği için R ile komşu Y arasındaki uzaklık, $P(a)$ a sınıfının olasılığı, m ise tekrarlamaya sayısı (örnek sayısını) temsil etmektedir.

$$W(F_i) = W(F_i) - \frac{dist(F_i, R, Y)}{m} + \sum_{a \neq \text{sınıf}(R)} \frac{P(a)}{1 - P(\text{sınıf}(R))} \cdot \frac{dist(F_i, R, U)}{m} \quad (2)$$

Öznitelik seçimi süreci, veri boyutunu azaltarak hesaplama yükünü, süresini kısaltmakta ve modelin performansını iyileştirmektedir. Bu çalışmada basit, etkili ve daha az hesaplama karmaşıklığı olması nedeniyle ReliefF yöntemi tercih edilmiştir (Huang vd., 2018). Deneysel olarak öznitelik seçiminde LASSO (Least Absolute Shrinkage and Selection Operator) ve Kruskal-Wallis yaklaşımları da denenmiş fakat ReliefF yönteminden daha kötü sonuç verdikleri için çalışmaya eklenmemiştir.

2.5. Makine Öğrenimi

Bu çalışmada seçilen özniteliklerin sınıflandırılmasında RO, DVM, K-NN ve DA algoritmaları kullanılmıştır. RO, Leo Breiman (Breiman, 2001) tarafından geliştirilen hızlı ve gürültüye dayanıklı bir sınıflandırma yöntemidir. Bu yöntem, makine öğrenmesinde özniteliklerin rastgele seçildiği topluluk tabanlı bir yaklaşımdır. Her bir karar ağacı, rastgele seçilen bir öznitelik alt kümesi kullanarak bağımsız bir sınıflandırma yapar ve bu sınıflar için oy verir. Birden fazla ağaç kullanılması, tek bir ağacın çıktısına göre daha az gürültü ve aykırılık sağlar, bu da tahminlerin sağlamlığını artırır ve değişkenliği azaltır. Bu nedenle, özellikle tıbbi veri setlerinde tercih edilmektedir (Altıntop vd., 2022). DVM, doğrusal olmayan verileri etkili bir şekilde sınıflandırmak için kullanılan güçlü bir yöntemdir. DVM, sınıflar arasındaki ayrım mesafesini maksimize eden karar hiperdüzlemleri oluşturmayı hedefler. Eğer veriler doğrusal olarak ayrılmıyorsa, DVM, verileri daha yüksek boyutlu bir uzaya haritalar ve burada optimal bir sınıflandırma hiperdüzlemi bulur (Cortes ve Vapnik, 1995). Kullanılan çekirdek fonksiyonu, verileri daha yüksek boyutlara taşımak ve sınıflandırma doğruluğunu artırmak için kritik öneme sahiptir. Polinom, Gauss radyal tabanlı ve Sigmoid çekirdek fonksiyonları en yaygın olanlarıdır. K-NN, sınıflandırma ve regresyon için kullanılan parametrik olmayan bir öğrenme yöntemidir. Bu algoritma, yeni bir örneği sınıflandırmak için mevcut örnekler arasındaki en yakın K komşuyu dikkate alır. K değeri seçildiğinde, hedef noktaya en yakın komşuların sınıfları arasında çoğunluk oyu alınarak sınıflandırma yapılır. Örneğin, K=1 olduğunda, örnek en yakın komşusunun sınıfına atanır. Bu yöntem, tembel öğrenme olarak adlandırılan bir yaklaşımı benimser ve tüm hesaplamaları sınıflandırma aşamasında gerçekleştirir (Gou vd., 2019). Diskriminant Analizi, iki veya daha fazla sınıfı ayırt etmek için geliştirilmiş istatistiksel bir yöntemdir. Amaç, farklı sınıflar arasında maksimum ayrım sağlayacak bir karar sınırı oluşturmaktır. Bu yöntem, verileri ayıran doğrusal bir hiperdüzlem tanımlayarak çalışır. Diskriminant fonksiyonu, sınıflar arasındaki değişkenliği artırırken, aynı zamanda her bir sınıf içindeki değişkenliği azaltmaya çalışır. Genellikle doğrusal olarak ayrılabilen verilerde yüksek bir başarı sağlar ve sınıflandırma,

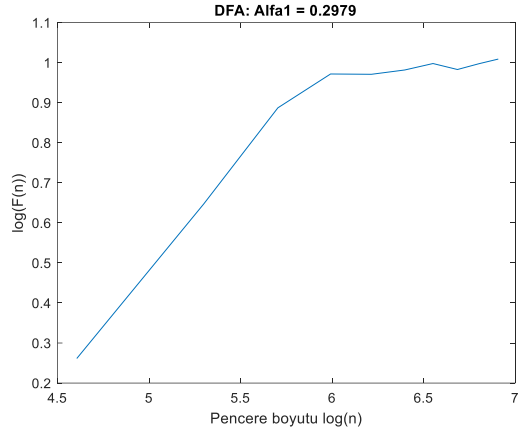
modelleme ve özellik seçimi gibi uygulamalarda yaygın olarak tercih edilmektedir (Bishop ve Nasrabadi, 2006). Sınıflandırma işlemi, 5-fold çapraz doğrulama ve 10-fold çapraz doğrulama ile gerçekleştirilmiştir. 10-fold çapraz doğrulama için veri seti 10 eşit parçaya bölünür ve her seferinde bir parça test verisi olarak kullanılırken, kalan 9 parça modelin eğitimi için kullanılmaktadır. Bu işlem, tüm veri seti için tekrarlandığından, her bir veri örneği bir kez test verisi olarak kullanılmış olmaktadır. Sonuç olarak, model performansı 10 defa gerçekleştirilen test sonuçlarının ortalaması alınarak hesaplanmaktadır. Modellerin hiper-parametreleri Bayes optimizasyonu kullanılarak otomatik olarak belirlenmiştir.

3. Bulgular

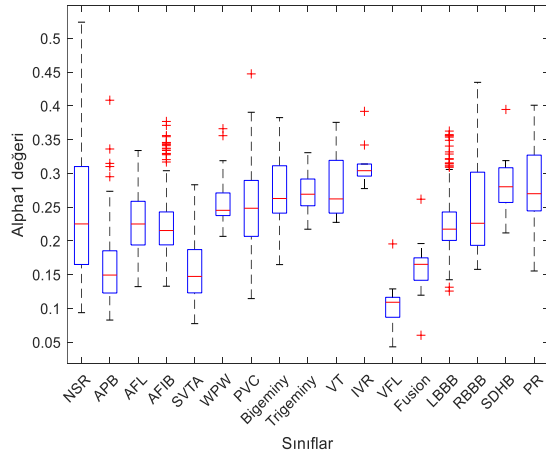
Bu bölümün alt başlıklarında öznitelik seçimi sonuçları ve makine öğrenimi algoritmaları ile sınıflandırma sonucu verilmiştir.

3.1. ReliefF ile Öznitelik Seçimi Sonuçları

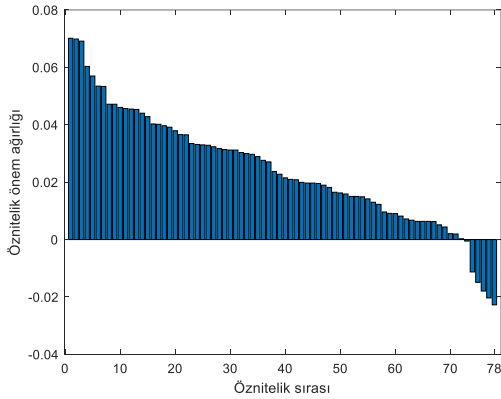
ReliefF ile öznitelik seçiminde literatürde çok kullanılması nedeni ile K=10 en yakın komşuya bakılarak işlem gerçekleştirilmiştir. Şekil 5'te EKG sinyallerinden çıkarılan toplam 78 adet öznenin ağırlıkları gösterilmiştir. Sınıflandırma işlemi için özniteliklerin seçiminde ağırlık değerlerine göre öznitelikler seçilerek iteratif olarak makine öğrenimi algoritmalarının eğitiminde kullanılmıştır. Şekil 6'da ise RO algoritması ile en yüksek doğruluk (accuracy) değerinin elde edildiği 41 adet öznenin ismi ve ağırlık bilgileri gösterilmiştir. Buna göre ReliefF ile öznitelik önem sırasına ilk olarak GSY grafiğinden elde edilen maksimum gücü veren frekans değeri (peak_freq), ikinci olarak DD analizi sonucu 3.seviye detay katsayısına ait çarpıklık değerleri (D3_skew), üçüncü olarak DFA yöntemi sonucu elde edilen alfa1 değeri (Alpha1) olduğu görülmektedir. Bu üç öznenin önem ağırlığı birbirine yakındır. Öznitelik seçiminde ReliefF yönteminin kullanılması ile normal sinüs ritmi, pacemaker ritmi ve 15 farklı aritmi olmak üzere toplam 17 sınıfı birbirinden ayırmada özniteliklerin ne derece başarılı olabileceği görülmüştür. DFA ile elde edilen Alfa değerlerinin gruplar arasındaki ilişkisini incelemek için kutu-grafiğini (Şekil 7) olarak gözlemlemek gerekirse; genel olarak, kutu-grafiğindeki alfa değerleri sınıflar arasında belirgin farklılıklar göstermektedir. SVTA, APB, Fusion ve VT gibi ritim bozuklukları daha düşük alfa değerlerine sahipken, NSR, RBBB, LBBB, PVC ve WPW gibi sınıflar daha geniş dağılımlar ve daha yüksek alfa değerleri ile dikkat çekmektedir. Bu durum, bazı patolojik durumlarda sinyallerin daha düzenli ve tahmin edilebilir olduğunu, diğerlerinde ise sinyallerin karmaşıklığının arttığını işaret etmektedir. Özellikle AFIB ve LBBB gibi bozukluklar geniş bir alfa dağılımı göstererek, sinyalin hem düzenli hem de düzensiz dönemler sergileyebileceğini ortaya koymaktadır. Sonuç olarak, bu alfa değerleri sinyallerin farklı ritim ve karmaşıklık özelliklerini yansıtarak DFA analizi kullanılarak sınıfların birbirinden ayırt edilebileceğini göstermektedir.



Şekil 5. NSR için DFA eğrisi



Şekil 6. Kalp rahatsızlıklarının DFA analizi sonucu alfa değerlerinin çizimi



Şekil 7. Relief sonucu tüm özniteliklerin önem ağırlığı

3.2. Sınıflandırma Sonucu

Bu çalışmada sınıflandırma aşamasında farklı metodolojilere sahip algoritmaların çıkarılan ve seçilen özniteliklerden 17 sınıfı ne kadar öğrenebildiği ve sınıflandırabildiğini karşılaştırabilmek için RO, K-NN, DVM ve DA algoritmaları kullanılmıştır. Verilerin eğitim ve test sürecinde 5-fold ve 10-fold çapraz doğrulama kullanılmıştır. Bazı sınıflarda veri sayısının az olması durumundan dolayı bu örneklerin hiç eğitime dahil olamayacağı endişesinden veri 10 eşit parçaya ayrılırken (10-fold için) her sınıftan verinin olması sağlanmıştır.

Ayrıca en yüksek örnek sayısına sahip 283 örnek ile NSR iken, VT, IVR, VFL, SDHB ise 10 örnek ile en az sayıda EKG sinyali içeren gruplardır. Veri setinde sınıflar arasındaki dengesizlik, sınıflandırma çalışmalarında istenmeyen bir durum olup, doğru sonuçlar elde edilmesini zorlaştırmaktadır. Özellikle azınlık (minority) ve çoğunluk (majority) sınıfları arasındaki bu dengesizlik, sınıflandırma performansını olumsuz etkileyerek yanıltıcı sonuçlara yol açabilmektedir. Bu durumun önüne geçebilmek için veri sayısını dengeli yapabilmek amacıyla bu çalışmada Sentetik Azınlık Aşırı Örnekleme Yöntemi (Synthetic Minority Oversampling Technique-SMOTE) algoritması kullanılmıştır. SMOTE algoritması, azınlık sınıfındaki örneklerin özellik uzayındaki en yakın komşularını (bu komşular rastgele seçilir) temel alarak, benzer özelliklere sahip sentetik veriler oluşturmaktadır (Chawla vd., 2002). Bu çalışmada K=5 en yakın örneğe bakılarak sentetik veriler oluşturulmuş ve 17 sınıfın hepsinde 283 adet örnek (toplam 4811 örnek) elde edilmiştir. Tablo 3'te öznitelik seçimi sırasında en yüksek doğruluk değerinin elde edildiği öznitelik sayısı algoritmalarla göre verilmiştir.

Relief algoritması ile öznitelik seçiminde sınıflandırma algoritmalarının farklı performans sergilediği ve SMOTE yöntemi ile dengelenmiş verinin sınıflandırılmasında öznitelik sayısının LDA algoritması hariç artırdığı görülmektedir. LDA algoritması dengesiz veri için en fazla öznitelikli seçerken, K-NN ve DVM algoritmaları ile sınıflandırmada daha az sayıda öznitelik seçilmiştir. SMOTE kullanımı, genel olarak tüm algoritmalarla seçilen öznitelik sayısında artışa yol açmıştır. Sınıflandırma sonuçları her fold'un doğruluğunun ortalama±standart sapma değerleri ile dengeli ve dengesiz veri için ayrı verilmiştir. Bu çalışmada sınıflandırma metriği olarak sadece doğruluk değeri verilmiştir. Çünkü 17 sınıf olduğu için hassasiyet (sensitivity), özgüllük (specificity) gibi ölçümlerin her sınıf için ayrı değerlendirilmesi gerekmektedir. Bu da çalışmada karmaşıklık oluşturabilecek bir durumdur ve anlamlı bilgi içermeyecektir.

Tablo 4, dört farklı sınıflandırma algoritmalarının 5-fold ve 10-fold çapraz doğrulama ile modelin eğitimi ve test süreci sonrası doğruluk sonuçlarını karşılaştırmaktadır. DVM algoritması, 5-fold çapraz doğrulamada %90,70 doğruluk ile düşük bir standart sapma (%1,2) göstermektedir. 10-fold çapraz doğrulamada doğruluk oranı %91,90'a çıkarak daha yüksek bir performans sunmakta, ancak standart sapma %3,07 ile biraz artmaktadır. K-NN algoritması, her iki çapraz doğrulama yönteminde de en yüksek doğruluk oranlarına sahip olup, 5-fold'ta %92,90, 10-fold'ta ise %93,40 doğruluk sağlamaktadır. Aynı zamanda düşük standart sapma değerleri (%1,78 ve %1,96) ile istikrarlı bir performans sergilemektedir. LDA algoritması ise en düşük doğruluk oranlarını göstermekte; hem 5-fold hem de 10-fold çapraz doğrulama sonuçlarında %87,80 doğruluğa sahiptir ve standart sapma oranları nispeten düşük kalmaktadır (%1,75 ve %2,49). Son olarak, RO

algoritması da yüksek doğruluk oranlarına sahiptir; 5-fold çapraz doğrulamada %92,90, 10-fold çapraz doğrulamada ise %93,10 doğruluk göstermekte olup, oldukça düşük standart sapma oranları (%0,74 ve %1,68) ile en tutarlı sonuçları sunmaktadır. Genel olarak, K-NN ve RO algoritmaları en yüksek doğruluk oranlarına ve tutarlılığa sahipken, LDA en düşük doğruluk oranını göstermiştir. Genel olarak sınıflandırma başarısı 10-fold çapraz doğrulama için başarılı olmasından ve 5-fold ile yakın sonuçlar vermesinden dolayı Tablo 5'te SMOTE ile dengelenmiş veri için sınıflandırma sonuçları verilmektedir.

Tablo 5 ve Şekil 8 incelenirse; RO algoritması dengelenmiş veri için %99,23 ile en yüksek doğruluk değerine sahip olup, aynı zamanda düşük bir standart sapma (%0,38) göstererek en tutarlı ve güvenilir performansı sağlamaktadır. K-NN algoritması ise %98,48 doğruluk ile ikinci sırada yer almakta ve yine düşük bir standart sapma (%0,42) ile nispeten istikrarlı sonuçlar

sunmaktadır. LDA algoritması, %97,40 doğruluk ile en düşük performansı göstermektedir ve bu algoritmanın standart sapması diğerlerine göre daha yüksektir (%0,95), bu da daha değişken bir performansa işaret etmektedir. DVM algoritması ise %97,86 doğruluk ile LDA'dan daha iyi, ancak RO ve K-NN'den daha düşük performans göstermekte olup, düşük standart sapma (%0,40) değeri ile istikrarlı bir sonuç sunmaktadır. Genel olarak, RO algoritması en yüksek doğruluğu ve tutarlılığı sergilerken, LDA algoritması en düşük doğruluk ve en yüksek değişkenliği göstermiştir.

Şekil 9 ve Şekil 10'da RO, K-NN algoritmalarının 10-fold çapraz doğrulama kullanılarak öznelilik sayısının ReliefF yöntemi ile belirlenen ağırlık değerlerine göre sırasıyla seçilmesiyle elde edilen doğruluk değerleri ve en yüksek doğruluk değerine sahip modelin karmaşıklık matrisi gösterilmiştir. Algoritmaların Bayes optimizasyonu sonucu elde edilen hiper-parametreleri Tablo 6'da verilmiştir.

Tablo 3. ReliefF ile seçilen öznelilik sayısı

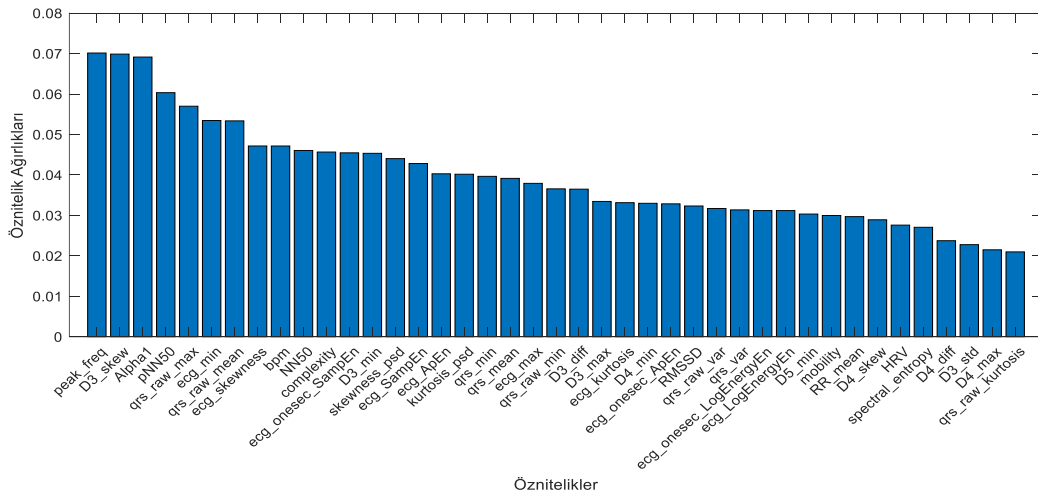
Algoritma	Dengesiz veri 5-fold	Dengesiz veri 10-fold	SMOTE ile Dengelenmiş veri (10-fold)
RO	48	41	63
K-NN	27	24	52
DVM	37	40	52
LDA	69	72	30

Tablo 4. Dengesiz veri ile sınıflandırma sonuçları

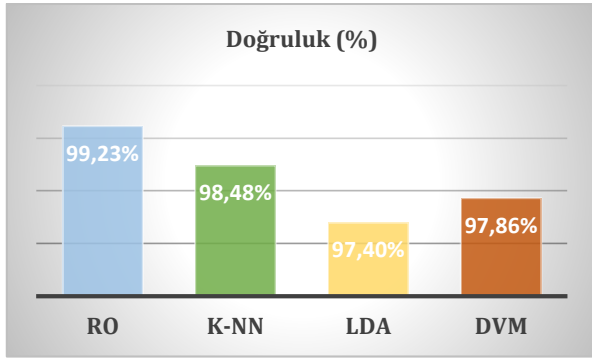
Algoritma	5-fold çapraz doğrulama (Doğruluk Değeri) $\mu \pm \sigma$	10-fold çapraz doğrulama (Doğruluk Değeri) $\mu \pm \sigma$
DVM	0,9070±0,0120	0,9190±0,0307
K-NN	0,9290±0,0178	0,9340±0,0196
LDA	0,8780±0,0175	0,8780±0,0249
RO	0,9290±0,0074	0,9310±0,0168

Tablo 5. Dengeli verinin sınıflandırılması sonucu

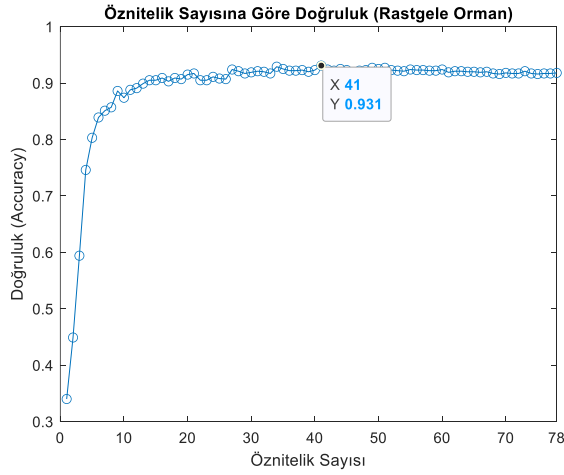
Doğruluk değeri 10-fold için	RO	K-NN	LDA	DVM
Ortalama	0,9923	0,9848	0,9740	0,9786
Standart sapma	0,0038	0,0042	0,0095	0,0040



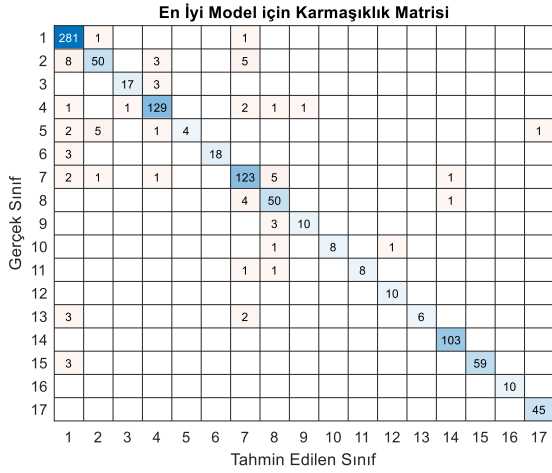
Şekil 8. En yüksek doğruluğun elde edildiği öznelilikler (41 öznelilik).



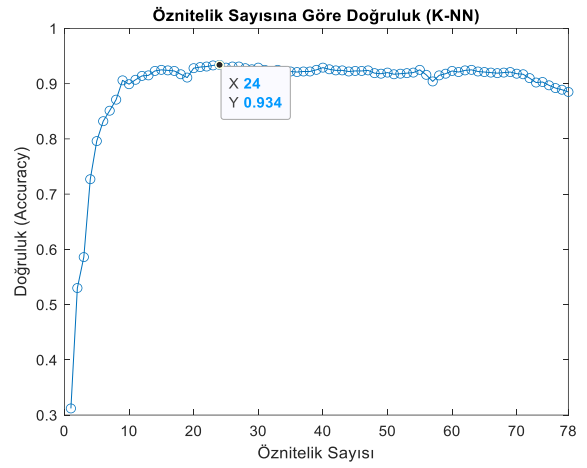
Şekil 9. SMOTE ile dengelenmiş verinin sınıflandırma sonucu



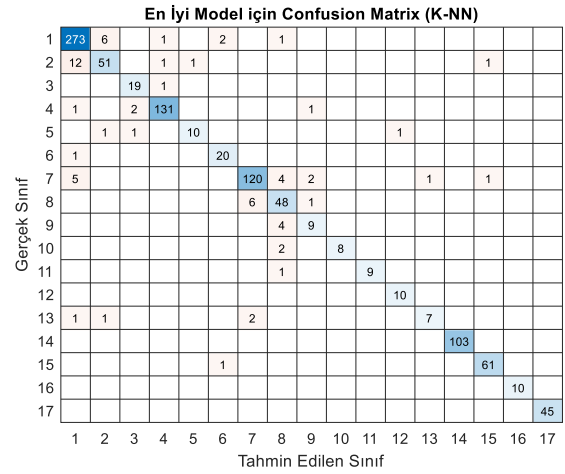
Şekil 10. Dengesiz veri için RO algoritması sonuçları; Öznitelik sayısına göre doğruluk grafiği.



Şekil 11. Dengesiz veri için RO algoritması sonuçları; En yüksek doğruluk değeri için karmaşıklık matrisi.



Şekil 12. Dengesiz veri için K-NN algoritması sonuçları; Öznitelik sayısına göre doğruluk grafiği.



Şekil 13. Dengesiz veri için K-NN algoritması sonuçları; En yüksek doğruluk değeri için karmaşıklık matrisi.

Tablo 6. En yüksek doğruluğu veren modele ait optimizasyon ile belirlenen parametreler

Model	5-fold	10-fold	SMOTE ile 10-fold
DVM	BoxConstraint= 0.1067 Kernelfunction= 'polynomial' NumNeighbors=1 Distance=	BoxConstraint=0.0956 Kernelfunction= 'polynomial' NumNeighbors=1 Distance=	BoxConstraint= 0.6589 Kernelfunction= 'polynomial' NumNeighbors=1 Distance=
K-NN	'euclidean' NSMethod= 'exhaustive'	'euclidean' NSMethod= 'exhaustive'	'euclidean' NSMethod= 'exhaustive'
LDA	DiscrimType= "linear" Gamma=4.1295e-04 Delta=0.003	DiscrimType= "linear" Gamma= 0.0025 Delta=2.1447e-05	DiscrimType= "pseudoQuadratic" Gamma=0 Delta=0
RO	Ağaç sayısı:500	Ağaç sayısı:500	Ağaç sayısı:100

4. Tartışma

Kalp anomalisi EKG dalgası kullanılarak teşhis edilebilmektedir. Ancak, EKG dalgalarının genlik ve frekans bileşenlerinin, manuel yöntem kullanan bir klinisyen için doğru ve tutarlı bir şekilde tespit edilmesi çok incelikli ve hata yapılması ihtimali yüksek bir durumdur. Bu nedenle, EKG sinyalinin incelenmesinde bilgisayar destekli yaklaşım kullanılması, klinisyenlerin kardiyovasküler hastalıkları önemli ölçüde yönetmesine yardımcı olacaktır. Bu amaçla birçok çalışma gerçekleştirilmiştir:

Osowski ve Linh (2001), hibrit bulanık sinir ağı kullanarak yedi EKG sınıfını (normal, LBBB, RBBB, APC, PVC, VFL ve ventriküler kaçış atımı) ayırt etmek için yüksek dereceli istatistik özelliklerini kullandılar ve %96,06'lık doğruluk ile 7 sınıfı sınıflandırabildiler. Sınıflandırıcının eğitimi ve testi için sırasıyla 4035 ve 3150 atım kullandılar (toplamda 7185 atım). İnan vd. (2006) PVC, normal ve diğer atımları %95,16 doğrulukla sınıflandırmak için sinir ağı tabanlı bir sınıflandırıcı ve zaman özellikleri, DD özellikleri kullanmıştır. Normal, PVC ve diğer sınıfları içeren MIT-BIH aritmi veritabanının 40 dosyasından 93.281 EKG atımı kullanmışlardır. Normal atımlar sırasıyla %98,33 ve %97,58 duyarlılık ve pozitif tahmin değeri ile sınıflandırılmıştır. Benzer şekilde, PVC atımları sırasıyla %82,57 ve %93,42 duyarlılık ve pozitif tahmin değeri ile sınıflandırılmıştır. Oh vd. (2019) ise 5 farklı aritminin (NSR, APB, PVC, LBBB ve RBBB) sınıflandırılmasında toplam 94667 EKG atımını değiştirilmiş U-Net mimarisi kullanmıştır. Model, 10 katlı çapraz doğrulama ile %97,32 sınıflandırma doğruluğu ve %99,3 R tepe noktası tespit doğruluğu sağlamıştır. Shi vd. (2019)'nin çalışmasında, EKG kalp atışlarının sınıflandırılması için ağırlıklı extreme gradient boosting (XGBoost) tabanlı hiyerarşik yöntem önerilmiştir. Çalışmada 168 özellik çıkarılmış ve recursive feature elimination (RFE) yöntemiyle özellik seçimi yapılmıştır. İki katmanlı hiyerarşik sınıflandırıcıda, ilk katmanda XGBoost ile kalp atışları normal (N), supraventriküler (S), ventriküler (V), füzyon (F) ve bilinmeyen (Q) sınıflarına ayrılmış, ikinci katmanda ise ek bir eşik sınıflandırıcı ve XGBoost ile daha ayrıntılı sınıflandırma yapılmıştır. N,S,V,F sınıflarında duyarlılıklar sırasıyla %92,1, %91,7, %95,1 ve %61,6 olarak elde edilmiştir. Çalışmada toplam 10630 atım kullanılmıştır ve tekli kalp atışlarını

sınıflandırmada XGBoost'un ilk kez kullanıldığını belirtmişlerdir. Daha fazla çalışma için aritminin otomatik tespitine dair yakın zamana ait (Ebrahimi vd., 2020; Singhal ve Kumar, 2023) derleme çalışmaları incelenebilir.

Bu çalışmalarda her bir nabız atımı ayrı değerlendirilmiştir, veri sayısı oldukça yüksektir ve sınıflandırılan aritmi sayısı genellikle 5 ile sınırlıdır. Bu çalışmada ise kısa süreli (10 saniye) EKG sinyallerinden zaman, frekans, zaman-frekans, entropi ve karmaşıklık öznitelikleri çıkarılarak öznitelik seçimi gerçekleştirilmiş ve çeşitli sınıflandırıcılarla az sayıda veriden aritmilerin sınıflandırılması yüksek doğrulukla sağlanmıştır.

Bu çalışmada kullanılan veri setinin ilk kullanıldığı ve 17 sınıflı sınıflandırmanın gerçekleştirdiği çalışma Plawiak (2018) tarafından sinyallerin GSY'dan frekans değerlerinin öznitelik olarak çıkarıldığı ve Genetik Algoritma ile seçilen özniteliklerin Olasılıksal Sinir Ağı, DVM, K-NN, Radyal Baz Fonksiyonlu Sinir Ağı yöntemleri kullanılarak sınıflandırılmasını içeren yeni bir yaklaşımdır. Çalışmada 4-fold ve 10-fold çapraz doğrulama sonuçlarına yer verilmiş ve 10-fold ile SVM algoritması %90,2 doğruluk elde edilmiştir. Tuncer vd. (2019) tarafından yapılan çalışmada, ADD ve yeni bir yöntem olan 1-boyutlu hegzadesimal yerel desen (1D-HLP) tekniği birleştirilmiştir. 10 saniyelik EKG sinyalleri, ADD ile beş seviyeye kadar ayrıştırılmış ve her seviyeden 1D-HLP yöntemiyle 3072 boyutlu özellikler çıkarılmıştır. Bu özellikler birleştirilerek büyük bir özellik seti oluşturulmuş ve ardından mahalle bileşeni analizi (NCA) ile 64, 128 ve 256 özellik olacak şekilde boyut indirgeme uygulanmıştır. Elde edilen bu özellikler, 1 en yakın komşu (1NN) sınıflandırıcısına dört farklı uzaklık metriği (city block, Euclidean, Spearman, ve cosine) ile uygulanmış ve 17 aritmi sınıfını %95 doğrulukla sınıflandırmayı başarmışlardır. Bu çalışmada öncelikle 3072 öznitelik elde edilmiş ve 256 adeti seçilmiştir. Çalışılan öznitelik sayısı gerçekleştirilen bu çalışma ile karşılaştırıldığında oldukça fazladır bu durum sınıflandırıcının eğitim süresini arttıracaktır. Plawiak ve Pławiak ve Acharya, (2020), 17 aritmi sınıfının tespiti için yenilikçi bir üç katmanlı derin genetik sınıflandırıcılar topluluğu (DGEC) geliştirilmiştir. Bu yöntem, topluluk öğrenimi, derin öğrenme ve evrimsel hesaplamayı bir araya getiren hibrit bir yaklaşımdır.

Çalışmada, Welch yöntemi ve ayrık Fourier dönüşümü ile spektral güç yoğunluğu tahmin edilerek öznelilikler güçlendirilmiştir. Genetik algoritma ile öznelilik seçimi ve sınıflandırıcı parametrelerinin genetik optimizasyonu ile geliştirilen DGEC sistemi, %94.62 doğruluk ile 17 sınıflı sınıflandırmıştır. Fakat çalışmalarının daha uzun sistem tasarımı, daha uzun eğitim ve optimizasyon gerektiren karmaşık yapıya sahip olmasının dezavantaj olduğunu belirtmişlerdir. Alqudah ve Alqudah (2022) çalışmasında, EKG sinyalleri üzerinde vuruş bazlı analiz yapılarak 17 sınıfa kadar kalp aritmilerini tespit etmek için yeni bir derin öğrenme metodolojisi sunulmuştur. Yöntem, her bir EKG vuruşunu analiz ederek iris spektrogramı hesaplamakta ve bu spektrogramı bir evrişimli sınır ağına (CNN) beslemektedir. Çalışma, 17 aritmi sınıfında sırasıyla 97.494 ± 1.26 doğruluk oranına ulaşmıştır. Demiroğlu vd., (2024)'nin çalışmasında, Parçacık Sürü Optimizasyonu (PSO) ile Tunable Q-Faktör Dalgacık Dönüşümü (TQWT) parametreleri optimize edilmiştir ve Hamsi Hash Fonksiyonu (Hamsi-Pat) kullanılarak sinyalin özellikleri çıkarılmıştır. 10 saniyelik alt sinyaller 25 alt banda bölünmüş ve her alt bant için 536 öznelilik elde edilerek $536 \times 25 = 13400$ özellik boyutuna ulaşılmıştır. İteratif Komşuluk Bileşeni Analizi (INCA) ile en iyi özellikler seçildikten sonra K-NN algoritması ile sınıflandırılmıştır. Bu yöntem, 17 sınıflı sınıflandırmada %98,5 doğruluk oranına ulaşmış ve literatürdeki diğer

çalışmalara kıyasla üstün performans göstermiştir. Fakat farklı yöntemlerin birleştirilmesi ve 'füzyon' yaklaşımının kullanılması, teorik olarak karmaşıklığı artırarak bir dezavantaj oluşturmıştır. Yöntemler arasındaki uyumun sağlanması ve her birinin optimal performans göstermesi zaman alıcı ve hesaplama açısından maliyetli olacaktır. Tablo 7'de verilen çalışmalarla karşılaştırıldığında gerçekleştirilen bu çalışma EKG sinyallerinden daha az sayıda öznelilikle ve daha az hesaplama karmaşıklığı ile yüksek doğrulukla 17 farklı aritminin sınıflandırılabilceğini göstermektedir.

Sınıflandırma modellerinin performansı, en anlamlı özelliklerin dikkatle seçilmesine bağlıdır. Gereksiz özelliklerin kaldırılması, veri kümesinin boyutunu azaltarak modelleme sürecine önemli avantajlar sağlar. Özellik sayısının azaltılmasının faydaları araştırmacılar tarafından kapsamlı bir şekilde incelenmiştir (Samb vd., 2012; Maldonado vd., 2014). Bu avantajlar şunlardır: Aşırı uyum riskinin azaltılması, bu sayede modelin genelleme yeteneğinin artması ve daha isabetli tahminler elde edilmesi; ayrıca, daha az özellik kullanımı sayesinde hesaplama maliyetinin düşürülmesi ve modelin daha hızlı çalışması. Bu nedenle bu çalışmada basit, hızlı ve hesaplama karmaşıklığı az olan ReliefF yöntemi ile öznelilik seçimi yapılarak çeşitli makine öğrenimi algoritmaları ile sınıflandırma gerçekleştirilmiştir.

Tablo 7. Aynı veri seti ile gerçekleştirilen çalışmaların karşılaştırılması

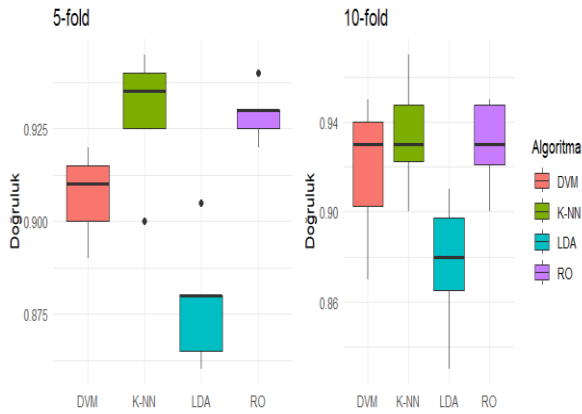
Yazar, Çalışma yılı	Sınıf Sayısı	Öznelilik Çıkarımı	Öznelilik Seçimi	Sınıflandırma yöntemi	Sonuç (Doğruluk)
Plawiak, (2018)	17	GSY'den frekans içeriği	Genetik Algoritma	Evrimsel-Nöral Sistem (DVM tabanlı)	%90.20
Tuncer vd., (2019)	17	ADD- 3072 öznelilik	Mahalle bileşeni analizi	1D-HLP	%95
Yıldırım vd., (2018)	17	Ham verilerin ADD ile yeniden ölçeklenmesi	Yok	1B-Evrimsel Sınır Ağı (1D-CNN)	%91.33
Plawiak ve Acharya, (2020)	17	GSY'den frekans içeriği	Genetik Algoritma	Sınıflandırıcıların derin genetik topluluğu, 3 katmanlı sistem	%94.62
Alqudah ve Alqudah, (2022)	17	irisgram'dan derin özellikler	-	2B-CNN	%97.49
Demiroğlu vd., (2024)	17	Hamsi-Pat	INCA	K-NN	%98.5
Önerilen çalışma, Altıntop, (2024)	17	zaman, frekans, zaman-frekans, entropi ve karmaşıklık öznelilikleri	ReliefF	RO, DVM, K-NN, LDA	%93.4 K-NN ile, %99.23 RO ile (SMOTE)

Şekil 11'de hem 5-fold hem de 10-fold çapraz doğrulama sonuçlarına göre algoritmaların doğruluk performansları karşılaştırılmaktadır. K-NN algoritması her iki çapraz doğrulama yönteminde de en yüksek doğruluk oranlarını göstermekte olup, diğer algoritmalara kıyasla daha tutarlı ve istikrarlı sonuçlar vermektedir. RO algoritması

da oldukça yüksek doğruluk oranlarıyla dikkat çekmekte ve K-NN'ye benzer bir başarı sergilemektedir. DVM algoritması ise ortalama doğruluk açısından makul bir performans sergilemekte, ancak K-NN ve RO'ya kıyasla biraz daha düşük doğruluk göstermektedir. LDA ise her iki çapraz doğrulama yöntemi için de en düşük doğruluk

oranlarına sahip algoritma olarak öne çıkmaktadır. Özellikle LDA'nın performansı diğer algoritmalarla kıyasla hem düşük hem de daha fazla değişkenlik göstermektedir. Genel olarak, K-NN ve RO'nun bu veri seti üzerinde daha üstün performans gösterdiği, DVM'nin de rekabetçi olduğu, ancak LDA'nın nispeten zayıf kaldığı bilgisine ulaşılmaktadır.

Sınıflandırma sonuçları açısından K-NN, her iki çapraz doğrulama yönteminde de %93'ün üzerinde doğruluk gösterirken, RO algoritması da benzer bir başarı sergileyerek %93 doğruluğu aşmıştır. DVM algoritması %91 civarında bir doğrulukla orta seviyede performans göstermiştir, LDA ise en düşük doğruluğa sahip algoritma olarak dikkat çekmektedir. SMOTE yöntemi ile veri sayıları dengelendiğinde ise RO algoritmasının %99.23 ile en yüksek doğruluğu sağlayan algoritma olduğu görülmektedir.



Şekil 11. ReliefF ile seçilen öznelikler ile elde edilen doğruluk değerleri

Bunu K-NN algoritması %98.48 doğrulukla takip ederken, DVM ve LDA daha düşük doğruluk oranları sunmaktadır. Sonuç olarak, RO ve K-NN algoritmalarının genel olarak en yüksek doğruluk oranlarına sahip olduğu ve tutarlı sonuçlar verdiği, LDA'nın ise daha düşük doğruluk ve daha yüksek varyans gösterdiği söylenebilir. Bu nedenle, EKG sinyallerinden aritmi sınıflandırılması problemlerinde yüksek doğruluk ve tutarlılık arayan çalışmalar için RO ve K-NN algoritmalarının tercih edilebileceği sonucuna varılabilir.

5. Sonuç

Bu çalışmada, EKG sinyallerinden çıkarılan özneliklerle 17 farklı kalp aritmi sınıfının yüksek doğrulukla sınıflandırılması hedeflenmiştir. Literatürdeki yöntemler genellikle yüksek veri sayısına ve karmaşık algoritmalarla dayanırken, bu çalışma, daha az sayıda öznelikle ve daha basit algoritmalarla başarılı sonuçlar elde edilmiştir. Elde edilen sonuçlar, K-NN ve RO algoritmalarının, en yüksek doğruluk oranlarını sağladığını ortaya koymuştur. K-NN algoritması %93'ün üzerinde doğruluk gösterirken, RO algoritması SMOTE yöntemi ile uygulandığında %99.23 gibi olağanüstü bir başarı elde etmiştir. DVM algoritması ise rekabetçi bir

performans sergilemekle birlikte, K-NN ve RO'nun gerisinde kalmıştır. LDA algoritması ise en başarısız sınıflandırma algoritması olmuştur. Bu çalışma ile bilgisayar destekli sistemlerin, klinik uygulamalarda EKG dalgalarının doğru ve tutarlı bir şekilde analiz edilmesine önemli ölçüde katkıda bulunabileceğini ortaya koymuştur. Elde edilen sonuçlar, koroner arter hastalığı, kalp yetmezliği, miyokard enfarktüsü gibi kardiyovasküler hastalıkların teşhisinde ve yönetiminde bilgisayar destekli sistemlerin kullanılmasının gerekliliğini göstermektedir. Bu çalışmanın dezavantajı veri sayısının az olmasıdır. Gelecek çalışmalarda, daha karmaşık aritmi sınıflandırma problemleri üzerinde odaklanarak bu bulguları genişletmesi ve farklı veri setleri ile örnek sayısının artırılarak doğruluk oranının artırılması planlanmaktadır. Bu sayede, EKG sinyallerinin analizi ve aritmi teşhisi için daha etkili ve güvenilir yöntemlerin geliştirilmesi sağlanabilecektir.

Katkı Oranı Beyanı

Yazarın katkı yüzdeleri aşağıda verilmiştir. Yazar makaleyi incelemiş ve onaylamıştır.

	Ç.G.A.
K	100
T	100
Y	100
VTI	100
VAY	100
KT	100
YZ	100
KI	100
GR	100
PY	100
FA	100

K= kavram, T= tasarım, Y= yönetim, VTI= veri toplama ve/veya işleme, VAY= veri analizi ve/veya yorumlama, KT= kaynak tarama, YZ= Yazım, KI= kritik inceleme, GR= gönderim ve revizyon, PY= proje yönetimi, FA= fon alımı.

Çatışma Beyanı

Yazar bu çalışmada hiçbir çıkar ilişkisi olmadığını beyan etmektedir.

Etik Onay Beyanı

Bu çalışmada hayvanlar ve insanlar üzerinde herhangi bir çalışma yapılmadığı için etik kurul onayı alınmamıştır.

Kaynaklar

- Afaq Y, Manocha A. 2024. Blockchain and deep learning integration for various application: a review. JCIS, 64: 92-105.
- Allabun S. 2024. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. IJACSA, 15: 4.
- Alqudah AM, Alqudah A. 2022. Deep learning for single-lead ECG beat arrhythmia-type detection using novel iris spectrogram representation. Soft Comput, 26: 1123-1139.
- Altıntop ÇG, Latifoğlu F, Akın AK. 2022. Can patients in deep coma hear us? Examination of coma depth using physiological

- signals. *Biomed Signal Process Control*, 77: 103756.
- Bishop CM, Nasrabadi NM. 2006. *Pattern recognition and machine learning*. Springer, New York, USA, pp: 738.
- Breiman L. 2001. Random forests. *Machine learning*, 45: 5-32.
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. 2002. SMOTE: Synthetic minority over-sampling technique. *JAIR*, 16: 321-357.
- Cortes C, Vapnik V. 1995. Support-vector networks. *Mach Learn*, 20(3): 273-297.
- Das MK, Ari S. 2014. Electrocardiogram beat classification using S-transform based feature set. *JMMB*, 14:1450066.
- Daydulo YD, Thamineni B, Dawud AA. 2023. Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals. *BMC Med Inform Decis Mak*, 23: 232.
- Demiroğlu U, Şenol B, Matuš R. 2024. A fused electrocardiography arrhythmia detection method. *Multim Tools Appl*, 83: 49057-49089.
- Devadas RM. 2021. Cardiac arrhythmia classification using svm, knn and naive bayes algorithms. *IRJET*, 8(5): 3937-3941.
- Dhyani S, Kumar A, Choudhury S. 2023. Analysis of ECG-based arrhythmia detection system using machine learning. *MethodsX*, 10: 102195.
- Dinakarrao SMP, Jantsch A, Shafique M. 2019. Computer-aided arrhythmia diagnosis with bio-signal processing: A survey of trends and techniques. *ACM CSUR*, 52(2): 1-37.
- Ebrahimi Z, Loni M, Daneshmandi M, Gharehbaghi A. 2020. A review on deep learning methods for ECG arrhythmia classification. *Expert Syst Appl*, 7: 100033.
- Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Stanley HE. 2000. *PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals*. *Circulation*, 101(23): e215-e220.
- Goldberger AL, Goldberger ZD, Shvilkin A. 2017. *Clinical electrocardiography: a simplified approach: clinical electrocardiography: a simplified approach e-book*. Elsevier Health Sciences, London, UK, pp: 187.
- Gopinathannair R, Etheridge SP, Marchlinski FE, Spinale FG, Lakkireddy D, Olshansky B. 2015. Arrhythmia-induced cardiomyopathies: mechanisms, recognition, and management. *JACC*, 66(15): 1714-1728.
- Gou J, Ma H, Ou W, Zeng S, Rao Y, Yang H. 2019. A generalized mean distance-based k-nearest neighbor classifier. *Expert Syst Appl*, 115: 356-372.
- Gupta V, Mittal M. 2020. A novel method of cardiac arrhythmia detection in electrocardiogram signal. *IJMEI*, 12(5): 489-499.
- Huang Z, Yang C, Zhou X, Huang T. 2018. A hybrid feature selection method based on binary state transition algorithm and ReliefF. *IEEE J-BHI*, 23(5): 1888-1898.
- Huikuri HV, Castellanos A, Myerburg RJ. 2001. Sudden death due to cardiac arrhythmias. *NEJM*, 345(20): 1473-1482.
- Inan OT, Giovannardi L, Kovacs GTA. 2006. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE TBME*, 53(12): 2507-2515.
- Kira K, Rendell L. A. 1992. A practical approach to feature selection. *Machine learning proceedings*. Elsevier, Amsterdam, Holland, pp: 249-256.
- Kohli N, Verma NK. 2011. Arrhythmia classification using SVM with selected features. *IJEST*, 3(8): 122-131.
- Kononenko I. 1994. Estimating attributes: Analysis and extensions of RELIEF. *European conference on machine learning proceedings*. Springer, Berlin, Germany, pp: 171-182.
- Kossmann CE. 1953. The normal electrocardiogram. *Circulation*, 8(6): 920-936.
- Liou JW, Wang PS, Wu YT, Lee SK, Chang SD, Liou M. 2022. ECG approximate entropy in the elderly during cycling exercise. *Sensors*, 22(14): 5255.
- Maldonado S, Weber R, Famili F. 2014. Feature selection for high-dimensional class-imbalanced data sets using support vector machines. *Inf Sci*, 286: 228-246.
- Marinho LB, Nascimento NMM, Souza JWM, Gurgel MV, Rebouças Filho PP, de Albuquerque VHC. 2019. A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification. *FGCS*, 97: 564-577.
- Martis RJ, Acharya UR, Adeli H. 2014. Current methods in electrocardiogram characterization. *Comput Biol Med*, 48: 133-149.
- Martis RJ, Acharya UR, Lim CM, Mandana KM, Ray AK, Chakraborty C. 2013. Application of higher order cumulant features for cardiac health diagnosis using ECG signals. *Int J Neural Syst*, 23(04): 1350014.
- Mazaheri V, Khodadadi H. 2020. Heart arrhythmia diagnosis based on the combination of morphological, frequency and nonlinear features of ECG signals and metaheuristic feature selection algorithm. *Expert Syst Appl*, 161: 113697.
- Mian Qaisar S, Hussain SF. 2023. An effective arrhythmia classification via ECG signal subsampling and mutual information based subbands statistical features selection. *J Ambient Intell Humaniz Comput*, 14(3): 1473-1487.
- Mishra A, Sahu SS, Sharma R, Mishra SK. 2021. Denoising of electrocardiogram signal using S-transform based time-frequency filtering approach. *Arab J Sci Eng*, 2021: 1-11.
- Mitra M, Samanta RK. 2013. Cardiac arrhythmia classification using neural networks with selected features. *Proc Technol*, 10: 76-84.
- Mizobuchi A, Osawa K, Tanaka M, Yumoto A, Saito H, Fuke S. 2021. Detrended fluctuation analysis can detect the impairment of heart rate regulation in patients with heart failure with preserved ejection fraction. *Cardiol J*, 77(1): 72-78.
- Mondéjar-Guerra V, Novo J, Rouco J, Penedo MG, Ortega M. 2019. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. *Biomed. Signal Process Control*, 47: 41-48.
- Moody GB, Mark RG. 2001. The impact of the MIT-BIH arrhythmia database. *IEEE Eng Med Biol*, 20(3): 45-50.
- Nascimento NMM, Marinho LB, Peixoto SA, do Vale Madeiro JP, de Albuquerque VHC, Filho PPR. 2020. Heart arrhythmia classification based on statistical moments and structural co-occurrence. *CSSP*, 39(2): 631-650.
- Oh SL, Ng EYK, San Tan R, Acharya UR. 2019. Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types. *Comput Biol Med*, 105: 92-101.
- Ojha MK, Wadhwani S, Wadhwani AK, Shukla A. 2022. Automatic detection of arrhythmias from an ECG signal using an auto-encoder and SVM classifier. *Phys Eng Sci Med*, 45(2): 665-674.
- Organization WH. 2021. *Cardiovascular Diseases*. URL: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 (accessed date: October 2, 2024)
- Osowski S, Linh TH. 2001. ECG beat recognition using fuzzy hybrid neural network. *IEEE Trans Biomed Eng*, 48(11): 1265-1271.
- Peng C, Havlin S, Stanley HE, Goldberger AL. 1995. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5(1): 82-87.
- Plawiak P, Acharya UR. 2020. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. *Neural Comput Appl*, 32(15):11137-11161.

- Plawiak P. 2017. ECG signals (1000 fragments). Mendeley Data, v3, London, UK.
- Plawiak P. 2018. Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system. *Expert Syst Appl*, 92: 334-349.
- Rodriguez E, Lerma C, Echeverria JC, Alvarez-Ramirez J. 2008. ECG scaling properties of cardiac arrhythmias using detrended fluctuation analysis. *Physiol Meas*, 29(11): 1255.
- Sahoo S, Subudhi A, Dash M, Sabut S. 2020. Automatic classification of cardiac arrhythmias based on hybrid features and decision tree algorithm. *IJAC*, 17(4): 551-561.
- Samb ML, Camara F, Ndiaye S, Slimani Y, Esseghir MA. 2012. A novel RFE-SVM-based feature selection approach for classification. *IJAST*, 43(1): 27-36.
- Sangaiah AK, Arumugam M, Bian GB. 2020. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. *Artif Intell Med*, 103:101788.
- Sharma LD, Rahul J, Aggarwal A, Bohat VK. 2023. An improved cardiac arrhythmia classification using stationary wavelet transform decomposed short duration QRS segment and Bi-LSTM network. *Multidimens Syst Signal Process*, 34(2): 503-520.
- Shi H, Wang H, Huang Y, Zhao L, Qin C, Liu C. 2019. A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Comput Methods Programs Biomed*, 171: 1-10.
- Shilaskar S, Ghatol A. 2013. Feature selection for medical diagnosis: Evaluation for cardiovascular diseases. *Expert Syst Appl*, 40(10):4146-4153.
- Singhal S, Kumar M. 2023. A systematic review on artificial intelligence-based techniques for diagnosis of cardiovascular arrhythmia diseases: challenges and opportunities. *Arch Comput Methods Eng*, 30(2): 865-888.
- Tuncer T, Dogan S, Plawiak P, Acharya UR. 2019. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. *Knowl-Based Syst*, 186: 104923.
- Vásquez-Iturralde F, Flores-Calero M, Grijalva-Arévalo F, Rosales-Acosta A. 2024. Automatic Classification of Cardiac Arrhythmias using Deep Learning Techniques: A Systematic Review. *IEEE Access*, 12: 118467-118492.
- Wan X, Liu Y, Mei X, Ye J, Zeng C, Chen Y. 2024. A novel atrial fibrillation automatic detection algorithm based on ensemble learning and multi-feature discrimination. *MBEC*, 62(6): 1809-1820.
- Wang D, Meng Q, Chen D, Zhang H, Xu L. 2020. Automatic detection of arrhythmia based on multi-resolution representation of ECG signal. *Sensors*, 20(6): 1579.
- Yang H, Wei Z. 2020. Arrhythmia recognition and classification using combined parametric and visual pattern features of ECG morphology. *IEEE Access*, 8:47103-47117.
- Yang J, Yan R. 2020. A multidimensional feature extraction and selection method for ECG arrhythmias classification. *IEEE Sens J*, 21(13): 14180-14190.
- Yeh YC, Chiou CW, Lin HJ. 2012. Analyzing ECG for cardiac arrhythmia using cluster analysis. *Expert Syst Appl*, 39(1): 1000-1010.
- Yıldırım Ö, Plawiak P, Tan RS, Acharya UR. 2018. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput Biol Med*, 102: 411-420.
- Zimetbaum P, Goldman A. 2010. Ambulatory arrhythmia monitoring: choosing the right device. *Circulation*, 122(16): 1629-1636.