



## Classification of Arrhythmia and Cardiac Anomalies from ECG Signals: A Machine Learning-Based Comparison on MIT-BIH and PTB-DB

Sercan Tohma <sup>1\*</sup>

<sup>1</sup> Department of Computer Engineering, Iskenderun Technical University, Hatay, Türkiye.

### ABSTRACT

Abstract In this study, two publicly available electrocardiogram (ECG) datasets were used to address (i) five-class heartbeat classification on MIT-BIH and (ii) binary (normal/abnormal) cardiac anomaly detection on PTB-DB. Raw time series samples (187-point segments) were used directly as feature vectors; after z-score standardization, Random Forest, Histogram-Based Gradient Boosting, Decision Tree, Naive Bayes, and Logistic Regression models were trained. Experimental evaluation was performed using accuracy (Acc), precision (Prec), recall (Rec), F1, and AUC metrics. The results indicated that the best performance in both datasets was achieved with the Histogram-Based Gradient Boosting model (MIT-BIH: Acc = 0.9801, AUC = 0.9930; PTB-DB: Acc = 0.9825, AUC = 0.9945). The findings show that classical machine learning models, with appropriate preprocessing and parameterization, can offer competitive performance in ECG-based arrhythmia and cardiac anomaly classification problems.

### ARTICLE INFO

**Received** 21.11.2025,  
**Accepted** 20.12.2025,  
**Publication Date** 25.12.2025

### Keywords:

ECG, Arrhythmia  
Classification, MIT-BIH,  
PTB-DB, Machine Learning,  
Histogram-Based Gradient  
Boosting.

**Distributed Under** CC-BY 4.0



## ECG Sinyallerinden Aritmi ve Kardiyak Anomali Sınıflandırması: MIT-BIH ve PTB-DB Üzerinde Makine Öğrenmesi Tabanlı Bir Karşılaştırma

### ÖZET

Özet Bu çalışmada, halka açık iki elektrokardiyogram (ECG) veri seti kullanılarak (i) MIT-BIH üzerinde beş sınıflı atım (heartbeat) sınıflandırması ve (ii) PTB-DB üzerinde ikili (normal/anormal) kardiyak anomali tespiti problemleri ele alınmıştır. Ham zaman serisi örnekleri (187 noktalı segmentler) doğrudan özellik vektörü olarak kullanılmış; z-skor standardizasyonu sonrasında Random Forest, Histogram Tabanlı Gradient Boosting, Decision Tree, Naive Bayes ve Logistic Regression modelleri eğitilmiştir. Deneysel değerlendirme; doğruluk (Acc), kesinlik (Prec), duyarlılık (Rec), F1 ve AUC ölçütleri ile gerçekleştirilmiştir. Sonuçlarda, her iki veri setinde de en iyi performansın Histogram Tabanlı Gradient Boosting modeli ile elde edildiği gözlenmiştir (MIT-BIH: Acc = 0.9801, AUC = 0.9930; PTB-DB: Acc = 0.9825, AUC = 0.9945). Elde edilen bulgular, klasik makine öğrenmesi modellerinin, uygun ön işleme ve parametrisasyon ile ECG tabanlı aritmi ve kardiyak anomali sınıflandırma problemlerinde rekabetçi performans sunabildiğini göstermektedir.

### MAKALE BİLGİSİ

Received 21.11.2025,

Accepted 20.12.2025,

Publication Date 25.12.2025

### Keywords:

ECG, aritmi sınıflandırma, MIT-BIH, PTB-DB, makine öğrenmesi, Histogram Tabanlı Gradient Boosting.

Distributed Under CC-BY 4.0



### GİRİŞ

Kardiyovasküler hastalıklar dünya genelinde yüksek morbidite ve mortalite oranlarına sahip olup, erken tanı ve risk sınıflandırması klinik açıdan kritik önemdedir. Elektrokardiyogram (ECG), non-invaziv ve görece düşük maliyetli bir ölçüm yöntemi olarak kalp ritmi ve iletim bozukluklarının değerlendirilmesinde yaygın şekilde kullanılmaktadır. Bununla birlikte, uzun süreli ECG kayıtlarının uzmanlar tarafından manuel incelenmesi zaman alıcı olup, özellikle yüksek hasta yükü olan klinik ortamlarda otomatik karar destek sistemlerine duyulan ihtiyacı artırmaktadır.

Bu bağlamda çalışma kapsamındaki temel problem; ECG atımlarından aritmi sınıflarının belirlenmesi ve/veya normal–anormal kardiyak durum ayrımının yapılmasıdır. Aritmiler (örneğin supraventriküler ve ventriküler kökenli atımlar) klinik risk açısından önemli farklılıklar gösterir ve doğru sınıflandırma tedavi planlamasını doğrudan etkileyebilir. Benzer şekilde miyokard

enfarktüsü (MI) ve diğer kardiyak anomalilerin erken tespiti, olası komplikasyonların önlenmesi ve prognozun iyileştirilmesi açısından önem taşır.

Son yıllarda yapay zekâ ve makine öğrenmesi yaklaşımları, ECG sınıflandırma problemlerinde yoğun olarak kullanılmaya başlanmıştır. Derin öğrenme tabanlı yaklaşımlar özellikle geniş ölçekli veri setlerinde yüksek başarılar raporlamaktadır (Hannun vd., 2019; Ribeiro vd., 2020; Wagner vd., 2020; Strodthoff vd., 2021). Örneğin Hannun ve ark. (2019) ambulatuvar ECG’de derin sinir ağları ile kardiyolog seviyesinde aritmi tespiti gerçekleştirmiştir. Ribeiro ve ark. (2020) 12 derivasyonlu ECG üzerinde derin öğrenme ile çoklu tanı etiketlerinin otomatik tahminini hedefleyen uçtan uca bir modelleme yaklaşımı sunmuştur. PTB-XL veri seti ve bu veri setine dayalı benchmark çalışmalar, model karşılaştırmalarının standartlaştırılmasına katkı sağlamıştır (Wagner vd., 2020; Strodthoff vd., 2021).

MIT-BIH veri kümesi üzerinde farklı CNN tabanlı mimariler kullanılarak yapılan aritmi sınıflandırma çalışmaları yaygın olarak raporlanmıştır (Mahmud vd., 2020; Essa ve Xie, 2021; Kanani ve Padole, 2020). Veri temsili ve transfer öğrenme gibi yaklaşımlar da kalp atışı sınıflandırma problemine uyarlanmıştır (Kachuee vd., 2018). Öte yandan, PTB-DB’ye benzer veri kümelerinde kardiyak anomalilerin (özellikle MI gibi patolojilerin) tespiti için otoenkoder ve diğer derin öğrenme tabanlı yöntemler önerilmiştir (Airlangga, 2024). PTB-XL üzerinde yapılan çalışmalarda, farklı ön işleme stratejilerinin (örneğin, R-tepe tespiti) ve ağ tasarımlarının sınıflandırma başarısı üzerindeki etkisi araştırılmıştır (Śmigiel vd., 2021a; Śmigiel vd., 2021b).

Literatürdeki öne çıkan çalışmalar kısaca şu şekilde özetlenebilir:

Hannun ve ark. (2019), derin sinir ağı tabanlı bir yaklaşım kullanarak ambulatuvar EKG kayıtlarında birden fazla aritmi sınıfını otomatik olarak tespit ederek klinik düzeyde sonuçlar bildirdi. Ribeiro ve ark. (2020), derin öğrenme kullanarak 12-kurşunlu EKG’de birden fazla tanı etiketinin otomatik tahminini hedefleyen uçtan uca bir modelleme yaklaşımı sundu. Wagner ve ark. (2020), PTB-XL veri setini tanımlayarak büyük ölçekli ve standartlaştırılmış bir EKG veri kaynağı sağladı. Strodthoff ve ark. (2021), PTB-XL üzerinde farklı derin öğrenme mimarileri için kıyaslama sonuçları sunarak karşılaştırmalı bir değerlendirme çerçevesi önerdi. Mahmud ve ark. (2020), verimli bir 1D-CNN mimarisi (DeepArrNet) kullanarak MIT-BIH benzeri kalp atışı verilerinde aritmi tespiti/sınıflandırması gerçekleştirdi. Essa ve Xie (2021), birden fazla derin öğrenme modelini birleştiren bir topluluk yaklaşımıyla kalp atışı aritmisi sınıflandırma performansını iyileştirmeyi amaçlamıştır. Kanani ve Padole (2020), zaman serisi artırma stratejileriyle derin öğrenmeye dayalı bir kalp atışı sınıflandırma yaklaşımı önermiştir. Śmigiel ve ark. (2021a), PTB-XL üzerinde farklı derin öğrenme tekniklerini karşılaştırarak EKG sinyali sınıflandırma performansını raporlamıştır. Śmigiel ve ark. (2021b), R-tepe tespiti gibi ön işleme adımlarının PTB-XL sınıflandırma performansı üzerindeki etkisini araştırmıştır. Airlangga (2024), otoenkoder tabanlı bir derin öğrenme yaklaşımıyla PTB Tanısal EKG Veritabanında kardiyak

anomali tespiti gerçekleştirmiştir. Kachuee ve ark. (2018), EKG Kalp Atışı Kategorizasyon Veri Kümesi üzerinde derin öğrenmeye dayalı, aktarılabılır bir gösterim önermiştir.

Bu çalışmada; MIT-BIH veri setinde beş sınıflı heartbeat sınıflandırması ve PTB-DB veri setinde ikili (normal/anormal) sınıflandırma problemleri, klasik makine öğrenmesi modelleri altında karşılaştırmalı olarak değerlendirilmiştir. Amaç; aynı deney düzeni altında farklı modeller ile performansını raporlamak ve elde edilen sonuçları Acc/Prec/Rec/F1/AUC ölçütleri ile özetleyerek literatürdeki eğilimlerle birlikte tartışmaktır.

## MATERIAL and METHODS

### Veri Seti

Bu çalışmada kullanılan veri setleri, PhysioNet altyapısı üzerinden paylaşılan ve literatürde yaygın biçimde kullanılan ECG kaynaklarına dayanmaktadır (Goldberger ve ark., 2000). MIT-BIH Arrhythmia Database, aritmi analiz çalışmalarında temel bir referans veri seti olarak kabul edilmektedir (Moody ve ark., 2001 ). PTB Diagnostic ECG Database ise farklı kardiyak durumları içeren tanısız ECG kayıtlarını barındıran bir kaynaktır (Bousseljot ve ark., 1995).

Çalışmada kullanılan dosyalar, her örneğin 187 noktalı ECG segmenti + 1 etiket olacak şekilde düzenlendiği (önceden segmentlenmiş) bir formattadır. Veri seti istatistikleri aşağıdaki gibidir:

MIT-BIH (5 sınıf):Eğitim: 87 554 örnek, Test: 21 892 örnek, Toplam: 109 446 örnek, Özellik sayısı: 187, Sınıflar: {0, 1, 2, 3, 4}

PTB-DB (2 sınıf):Toplam: 14 552 örnek, Özellik sayısı: 187, Sınıflar: {0: normal, 1: anormal}

Bu çalışmada, veri kümesi sınıf oranları korunacak şekilde (stratified) %80 eğitim ve %20 test olacak biçimde bölünmüştür.

Analiz ve ön işlemler

Standartlaştırma (z-skor normalizasyonu)

ECG segmentlerindeki ölçek farklılıklarının model eğitimini olumsuz etkilemesini önlemek amacıyla her özellik boyutu z-skor standardizasyonu ile dönüştürülmüştür. Bu adımda, eğitim verisi üzerinden ortalama ( $\mu$ ) ve standart sapma ( $\sigma$ ) hesaplanmış; her bir özellik değeri için aşağıdaki dönüşüm uygulanmıştır:

$$z=(x-\mu)/\sigma$$

Bu sayede tüm özellikler benzer ölçek aralığına taşınarak optimizasyon süreci daha kararlı hale getirilmiş, özellikle doğrusal modellerde görülebilecek yakınsama problemleri azaltılmıştır. Literatürde PTB-XL gibi veri setleri üzerinde yapılan derin öğrenme çalışmalarında da benzer normalizasyon/ölçekleme adımlarının kullanıldığı görülmektedir (Strodthoff ve Ark, 2021; Śmigiel ve Ark., 2021a). Ayrıca ROC-AUC gibi olasılık temelli ölçütlerin daha tutarlı hesaplanabilmesi için modelin iyi kalibre olmasına katkı sağlayabilmektedir (Fawcett , 2005).

### Eğitim-test ayrımı

MIT-BIH veri setinde eğitim ve test ayrımı, veri sağlayıcılar tarafından hazır olarak sunulmaktadır (ayrı dosyalar). PTB-DB için ise veri, sınıf oranlarını koruyacak şekilde stratified olarak %80 eğitim ve %20 test biçiminde ayrılmıştır. Stratified bölme, sınıf dengesizliğinin değerlendirme sırasında yanlılığa yol açmasını azaltır ve test kümesinin temsil gücünü artırır. Bu sayede farklı model ailelerinin karşılaştırılması daha adil bir zeminde yapılabilir ve AUC gibi eşik bağımsız ölçütler daha güvenilir biçimde yorumlanabilir (Fawcett, 2006). Benzer şekilde benchmark çalışmalarında da standart bir test protokolü, yöntemlerin kıyaslanabilirliği için temel bir gereklilik olarak vurgulanmaktadır (Strodthoff ve Ark, 2021).

## Makine öğrenmesi modelleri

### Random Forest (RF)

Random Forest, birden fazla karar ağacının bootstrap örneklemeleri üzerinde eğitilmesi ve nihai kararın çoğunluk oyu ile verilmesi prensibine dayanan topluluk (ensemble) yöntemidir (Breiman , 2001). Ağaçlar arası çeşitlilik, varyansı düşürerek genelleme performansını artırır. RF, doğrusal olmayan karar sınırlarını modelleyebilir ve gürültüye karşı görece dayanıklıdır. Bu çalışmada RF, temel hiperparametrelerle (200 ağaç) kullanılmıştır.

### Histogram Tabanlı Gradient Boosting (HGB)

Gradient Boosting yaklaşımı, zayıf öğrencilerin (çoğunlukla karar ağaçları) ardışık olarak eğitilip hataların kademeli biçimde düzeltilmesi fikrine dayanmaktadır (Friedman , 2001). Histogram tabanlı varyant, sürekli özellikleri kovalar (bins) üzerinden ayrıklaştırarak eğitim hızını artırır ve büyük veri kümeleri üzerinde verimli çalışır. Bu çalışmada HGB, ağaç derinliği ve iterasyon sayısı gibi temel hiperparametreler ayarlanarak kullanılmıştır.

### Decision Tree (DT)

Karar ağaçları, özellik uzayını ardışık bölmelere ayırarak karar kuralları üreten, yorumlanabilir yapıda modellerdir. Doğrusal olmayan ilişkileri yakalayabilirler; ancak tek ağaç yapısı, özellikle

yüksek boyutlu ve gürültülü veride aşırı uyum (overfitting) riskine açıktır. Bu çalışmada DT modeli, aşırı karmaşıklığı sınırlamak amacıyla maksimum derinlik kısıtı ile eğitilmiştir.

#### Naive Bayes (NB)

Naive Bayes sınıflandırıcıları, sınıf koşullu bağımsızlık varsayımı altında özelliklerin dağılımına dayalı olasılıksal sınıflandırma yapar. Yüksek boyutlu veride hızlı eğitim avantajına sahip olmakla birlikte, ECG gibi zamansal korelasyon içeren sinyallerde bağımsızlık varsayımı modelin ayırt ediciliğini sınırlayabilmektedir. Bu çalışmada GaussianNB türevi kullanılmıştır.

#### Logistic Regression (LR)

Lojistik regresyon, doğrusal karar sınırı üzerinden olasılık tahmini yapan temel bir sınıflandırma yöntemidir. Çok sınıflı senaryolarda multinomial yaklaşım ile genelleştirilebilir. Basitliği ve hesaplama verimliliği avantaj sağlasa da, karmaşık ve doğrusal olmayan örüntüler içeren ECG sinyallerinde ağaç tabanlı yöntemlerin gerisinde kalması beklenmektedir.

#### Performans ölçütleri

Modeller, doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall), F1 ve ROC-AUC ölçütleri ile değerlendirilmiştir. ROC analizi ve AUC için standart yaklaşım (Fawcett, 2006), dengesiz veri durumlarında Precision–Recall eğrilerinin önemi ise (Saito and Rehmsmeier, 2015) ile desteklenmektedir.

$$\text{Accuracy (Acc): } \text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision (Prec): } \text{Prec} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall (Rec): } \text{Rec} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1 skoru: } \text{F1} = \frac{2 \cdot (\text{Prec} \cdot \text{Rec})}{\text{Prec} + \text{Rec}}$$

AUC: ROC eğrisi altında kalan alanı ifade eder. Çok sınıflı MIT-BIH senaryosunda one-vs-rest yaklaşımı ile ağırlıklı AUC değeri hesaplanmıştır.

MIT-BIH çok sınıflı probleminde Precision/Recall/F1 değerleri ağırlıklı ortalama (weighted average) olarak raporlanmıştır. PTB-DB ikili probleminde ise Precision/Recall/F1, pozitif sınıf (anormal = 1) üzerinden hesaplanmıştır.

## SONUÇLAR ve TARTIŞMA

Bu çalışmada, MIT-BIH veri seti üzerinde beş sınıflı heartbeat sınıflandırması ve PTB-DB veri seti üzerinde normal/anormal kardiyak durum sınıflandırması gerçekleştirilmiştir. Tüm modeller aynı ön işleme adımlarına tabi tutulmuş, test kümeleri üzerinde Acc, Prec, Rec, F1 ve AUC metrikleri hesaplanmıştır. Sonuçlar Tablo 1 ve Tablo 2’de özetlenmektedir.

### MIT-BIH (5 sınıf) sonuçları

Tablo 1. MIT-BIH veri seti üzerinde elde edilen sınıflandırma sonuçları..

Model	Acc	Prec	Rec	F1	AUC
<b>Random Forest</b>	0.9748	0.9749	0.9748	0.9732	0.9946
<b>HistGradientBoosting</b>	0.9801	0.9796	0.9801	0.9792	0.9930
<b>Decision Tree</b>	0.9607	0.9588	0.9607	0.9593	0.9141
<b>Naive Bayes</b>	0.1906	0.7995	0.1906	0.2024	0.7458
<b>Logistic Regression</b>	0.9148	0.9042	0.9148	0.9030	0.9182

MIT-BIH veri setinde ağaç tabanlı ensemble yöntemlerin (HGB ve RF), doğrusal (LR) ve olasılıksal (NB) modellere göre belirgin üstünlük sağladığı görülmektedir. Histogram Tabanlı Gradient Boosting modeli, hem doğruluk hem de F1 metriğinde en yüksek değere ulaşmış; Random Forest ise özellikle AUC açısından HGB ile benzer ve yüksek performans sergilemiştir. Naive Bayes modelinin doğruluğunun düşük kalması, ECG örüntülerinde özellikler arası bağımlılığın yüksek olmasına karşılık sınıf koşullu bağımsızlık varsayımının yetersiz oluşu ile ilişkilendirilebilir.

### 3.2. PTB-DB (2 sınıf) sonuçları

Tablo 2. PTB-DB veri seti üzerinde elde edilen sınıflandırma sonuçları.

Model	Acc	Prec	Rec	F1	AUC
-------	-----	------	-----	----	-----

<b>Random Forest</b>	0.9711	0.9711	0.9895	0.9802	0.9936
<b>HistGradientBoosting</b>	0.9825	0.9835	0.9924	0.9879	0.9945
<b>Decision Tree</b>	0.9148	0.9440	0.9377	0.9408	0.8930
<b>Naive Bayes</b>	0.6379	0.8994	0.5614	0.6913	0.7785
<b>Logistic Regression</b>	0.8282	0.8477	0.9291	0.8865	0.8693

PTB-DB veri setinde HGB modeli, hem doğruluk hem de F1 ve AUC metriklerinde en yüksek performansı sağlamıştır. Random Forest modeli de oldukça yüksek bir AUC değeriyle (0.9936) rekabetçi sonuçlar üretmiştir. Logistic Regression, pozitif sınıf için yüksek duyarlılık (Rec = 0.9291) üretmesine rağmen genel doğruluk ve F1 skorları açısından ağaç tabanlı ensemble yöntemlerin gerisinde kalmıştır. Naive Bayes modelinin doğruluğunun görece düşük olması, MIT-BIH senaryosunda olduğu gibi bağımsızlık varsayımının ECG verisinin yapısıyla tam uyumlu olmamasına bağlanabilir.

### 3.3. Tartışma ve kısıtlar

Elde edilen bulgular, ECG segmentlerinin ham zaman örnekleriyle temsil edildiği durumda dahi ağaç tabanlı ensemble yöntemlerin güçlü performans sergileyebildiğini göstermektedir. HistGradientBoosting modelinin her iki veri setinde de en yüksek doğruluk ve F1 skorlarını üretmesi, gradient boosting tabanlı ağaç yöntemlerinin ECG gibi karmaşık, doğrusal olmayan ve yüksek boyutlu zaman serisi verisine iyi uyum sağlayabildiğini desteklemektedir.

Bununla birlikte, kullanılan veri setlerinin önceden segmentlenmiş olması, gerçek klinik senaryolara doğrudan genellenebilirlik açısından bir sınırlılık oluşturmaktadır. Uzun süreli ECG kayıtlarında gürültü, artefaktlar, ritim değişimleri ve hasta bazlı farklılıklar, model performansını olumsuz etkileyebilir. Ayrıca MIT-BIH veri setinde sınıflar dengesizdir ve nadir görülen aritmi türlerinde başarı tekil sınıf bazında daha düşük olabilir. Bu çalışmada ağırlıklı ortalama metrikler raporlanmış olmakla birlikte, gelecekte makro ortalama metrikler, sınıf ağırlıklandırma, odak kaybı (focal loss) ve yeniden örnekleme stratejileri ile nadir sınıfların performansını iyileştirmeye odaklanan çalışmalar yapılabilir.

PTB-DB senaryosunda ikili sınıflandırma problemi, MI gibi patolojilerin erken tespiti açısından önemli bir adım olmakla birlikte, kardiyak durumların klinik spektrumu göz önüne alındığında daha ince tanı etiketleri (örneğin alt gruplara ayrılmış MI, iskemi vb.) ile çok sınıflı yaklaşımların

değerlendirilmesi gelecekteki çalışmalar için anlamlı bir yöndür. Bunun yanı sıra, sinyal tabanlı özellik çıkarımı (örneğin dalgacık tabanlı öznelikler, frekans alanı özellikleri) veya derin özellikler (temsil öğrenme) ile klasik makine öğrenmesi yöntemlerinin hibrit kullanımı da potansiyel bir araştırma alanı olarak öne çıkmaktadır.

Genel olarak bu çalışma, ham ECG segmentlerinin basit bir z-skor normalizasyonu sonrasında dahi klasik makine öğrenmesi modelleriyle yüksek sınıflandırma performanslarına ulaşabildiğini göstermekte; aynı zamanda derin öğrenme tabanlı literatür sonuçları ile karşılaştırılabilir bir referans zemin sunmaktadır.

#### 4. Sonuç

Bu çalışmada, MIT-BIH Arrhythmia Database ve PTB Diagnostic ECG Database veri setleri kullanılarak ECG tabanlı aritmi ve kardiyak anomali sınıflandırma problemleri klasik makine öğrenmesi yöntemleri ile kapsamlı biçimde incelenmiştir. Ham ECG segmentlerinin (187 örnek nokta) yalnızca z-skor standardizasyonu ile ön işleme tabi tutulduğu bir deney düzeninde, ağaç tabanlı topluluk (ensemble) yöntemlerinin hem çok sınıflı (MIT-BIH, 5 sınıf) hem de ikili (PTB-DB, normal/anormal) senaryolarda yüksek performans sergileyebildiği gösterilmiştir.

Deneysel sonuçlar, Histogram Tabanlı Gradient Boosting (HGB) modelinin her iki veri setinde de en yüksek doğruluk ve F1 skorlarını ürettiğini ortaya koymuştur (MIT-BIH: Acc = 0.9801, AUC = 0.9930; PTB-DB: Acc = 0.9825, AUC = 0.9945). Random Forest modeli de özellikle AUC açısından HGB'ye yakın ve klinik açıdan anlamlı kabul edilebilecek seviyede sonuçlar üretmiştir. Buna karşın, Naive Bayes modeli her iki veri setinde de belirgin biçimde daha düşük doğruluk değerleri üretmiş; bu durum, ECG sinyallerinde özellikler arası korelasyonun yüksek olması ve sınıf koşullu bağımsızlık varsayımının bu bağlamda yetersiz kalması ile açıklanabilir. Lojistik regresyon ise hesaplama açısından verimli olmakla birlikte, doğrusal karar sınırı varsayımı nedeniyle karmaşık ve doğrusal olmayan örüntüleri temsil etmede ensemble yöntemlerinin gerisinde kalmıştır.

Elde edilen bulgular, derin öğrenme tabanlı daha karmaşık mimarilerin literatürde raporlanan başarımına rağmen, uygun ön işleme ve model seçimi ile klasik makine öğrenmesi yöntemlerinin de ECG sınıflandırma problemlerinde rekabetçi bir alternatif olabileceğini göstermektedir. Özellikle veri boyutunun, hesaplama kaynaklarının veya açıklanabilirlik gereksinimlerinin sınırlı olduğu uygulamalarda, HGB ve RF gibi ağaç tabanlı modeller pratik ve güçlü bir çözüm sunmaktadır.

Bununla birlikte, çalışmanın bazı kısıtları bulunmaktadır. Kullanılan veri setleri önceden segmentlenmiş ve görece kontrollü koşullarda elde edilmiş kayıtlardan oluşmaktadır. Gerçek klinik uygulamalarda uzun süreli kayıtlar, gürültü, artefakt, hasta bazlı varyasyon ve sensör kaynaklı

problemler model performansını azaltabilir. Ayrıca MIT-BIH veri setindeki sınıf dengesizliği, nadir görülen aritmi türlerinde sınıf bazlı performansın daha ayrıntılı incelenmesini gerektirmektedir.

Gelecek çalışmalarda, sınıf dengesizliğini hedefleyen yeniden örnekleme ve kayıp fonksiyonu tasarımları, dalgacık tabanlı veya frekans alanı özellikleri ile zenginleştirilmiş hibrit temsil yaklaşımları, segmentlenmemiş uzun ECG kayıtları üzerinde uçtan uca işleyen karar destek sistemleri ve klasik makine öğrenmesi modellerinin derin özelliklerle (deep features) birleştirildiği hibrit mimariler araştırılabilir. Ayrıca model açıklanabilirliği ve klinik yorumlanabilirlik odaklı yöntemlerin (özellikle ağaç tabanlı modellerle ilişkilendirilebilen kural tabanlı açıklama mekanizmaları) geliştirilmesi, bu tür sistemlerin klinik entegrasyonunu kolaylaştıracak önemli bir adım olacaktır.

## KAYNAKÇA

- Airlangga, G. (2024). Enhancing cardiac anomaly detection through deep learning autoencoder: An in-depth analysis using the PTB diagnostic ECG database. *G-Tech: Jurnal Teknologi Terapan*, 8(1), 584–592. <https://doi.org/10.33379/gtech.v8i1.3921>
- Bousseljot, R., Kreiseler, D., & Schnabel, A. (1995). Nutzung der EKG-signal-datenbank CARDIODAT der PTB über das internet. *Biomedizinische Technik/Biomedical Engineering*, 317–318. <https://doi.org/10.1515/BMTE.1995.40.s1.317>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Essa, E., & Xie, X. (2021). An ensemble of deep learning-based multi-model for ECG heartbeats arrhythmia classification. *IEEE Access*, 9, 103452–103464. <https://doi.org/10.1109/ACCESS.2021.3098986>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5). <https://doi.org/10.1214/AOS/1013203451>
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220. <https://doi.org/10.1161/01.CIR.101.23.e215>
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1), 65–69. <https://doi.org/10.1038/s41591-018-0268-3>

- Kachuee, M., Fazeli, S., & Sarrafzadeh, M. (2018). ECG heartbeat classification: A deep transferable representation. 2018 IEEE International Conference on Healthcare Informatics (ICHI), 443–444. <https://doi.org/10.1109/ICHI.2018.00092>
- Kanani, P., & Padole, M. (2020). ECG heartbeat arrhythmia classification using time-series augmented signals and deep learning approach. *Procedia Computer Science*, 171, 524–531. <https://doi.org/10.1016/j.procs.2020.04.056>
- Mahmud, T., Fattah, S. A., & Saquib, M. (2020). DeepArrNet: An efficient deep CNN architecture for automatic arrhythmia detection and classification from denoised ECG beats. *IEEE Access*, 8, 104788–104800. <https://doi.org/10.1109/ACCESS.2020.2998788>
- Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), 45–50. <https://doi.org/10.1109/51.932724>
- Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., Ferreira, M. P. S., Andersson, C. R., Macfarlane, P. W., Meira Jr., W., Schön, T. B., & Ribeiro, A. L. P. (2020). Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*, 11(1). <https://doi.org/10.1038/s41467-020-15432-4>
- Saito, T., & M. Rehmsmeier (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Śmigiel, S., Pałczyński, K., & Ledziński, D. (2021a). ECG signal classification using deep learning techniques based on the PTB-XL dataset. *Entropy*, 23(9), 1121. <https://doi.org/10.3390/e23091121>
- Śmigiel, S., Pałczyński, K., & Ledziński, D. (2021b). Deep learning techniques in the classification of ECG signals using r-peak detection based on the PTB-XL dataset. *Sensors*, 21(24), 8174. <https://doi.org/10.3390/s21248174>
- Strodthoff, N., Wagner, P., Schaeffter, T., & Samek, W. (2021). Deep learning for ECG analysis: Benchmarks and insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1519–1528. <https://doi.org/10.1109/JBHI.2020.3022989>
- Wagner, P., Strodthoff, N., Bousseljot, R. D., Samek, W., & Schaeffter, T. (2020). PTB-XL, a large publicly available electrocardiography dataset. *Scientific Data*, 7(1). <https://doi.org/10.1038/s41597-020-0495-6>