

Kalp Yetmezliği Tahmininin Kategorik Olarak Farklı Tip Makine Öğrenmesi Yöntemleri ile Uygulanmasına Yönelik Bir Değerlendirme Çalışması

An Evaluation Study on the Application of Heart Failure Prediction with Categorically Different Type Machine Learning Methods

İsmail ATACAK * 

* Gazi Üniversitesi, Teknoloji
Fakültesi, Bilgisayar Mühendisliği
Bölümü, 06560, Ankara, Türkiye
iatacak@gazi.edu.tr; ORCID: 0000-0002-6357-0073

Öz

Kalp yetmezliği yaşam kalitesini olumsuz etkileyen ve tedavi edilmediğinde ölümcül sonuçlar doğurabilen ciddi bir sağlık problemidir. Erken teşhis ve doğru tedavinin uygulanması bu problemleri en aza indirebilir. Bu çalışmada farklı kategorilerde yer alan makine öğrenmesi (MÖ) yöntemlerinin kalp yetmezliği tahminindeki performanslarını ölçmek için bir model geliştirilerek, kategorik ve genel olarak performans analizleri gerçekleştirilmiştir. Kategori temelinde sınıflandırma problemlerinde başarılı sonuçlar ürettiği bilinen yöntemleri içeren ağaç, meta ve fonksiyon kategorileri tercih edilmiş ve her kategoriden beş yöntem kullanılmıştır. Deneysel çalışmalarda MÖ yöntemlerinin performansı Karışıklık matrisine dayanan temel metrikler ile sınıflandırma hata metrikleri üzerinden ölçülmüştür. Deneysel sonuçlar kategorik olarak değerlendirildiğinde en iyi performansların ağaç kategorisinde Duyarlılık ve Yanlış Negatif Oranı (False Negative Rate (FNR)) dışındaki metriklerde Alternatif Karar Ağacı (Alternating Decision Tree | ADTree (ADT)) yöntemi, meta kategorisinde ROC eğrisi altında kalan alan (Area under the curve (AUC)) dışındaki metriklerde Lojistik Artırma Regresyon (Logistic Boosting Regression | LogitBoost (LBST)) yöntemi ve fonksiyon kategorisinde Kesinlik ve Yanlış Pozitif Oranı (False Positive Rate (FPR)) dışındaki metriklerde Radyal Temelli Fonksiyon Sınıflandırıcı (Radial Bases Function Classifier (RBFC)) yöntemi ile elde edildiğini göstermektedir. Sonuçlara tüm yöntemlerin performansları açısından bakıldığında Doğruluk, Duyarlılık, F-skor, FNR ve Yanlış Sınıflandırma Oranı (Misclassification Rate (MCR)) metrikleri açısından 0.8725, 0.9173, 0.8885, 0.0827 ve 0.1275 değerleri ile RBFC yönteminin, Kesinlik, AUC ve FPR metrikleri açısından 0.8718, 0.9300 ve 0.1610 değerleri ile ADT yönteminin en iyi performansa sahip olduğu görülmüştür.

Anahtar kelimeler—Makine öğrenmesi, sınıflandırma, kalp yetmezliği

Abstract

Heart failure is a severe health problem that negatively impacts the quality of life and can lead to fatal consequences if left untreated. Early diagnosis and proper treatment can minimize these problems. In this study, a model was developed to measure the performances of machine learning (ML) methods in different categories for heart failure prediction, and performance analyses were performed on both categorical and general basis. Tree, meta and function categories, which include methods known to produce successful results in classification problems, were preferred as categories, and five methods from each category were used. In the experimental studies, the performance of the ML methods was measured using basic metrics and classification error metrics based on the confusion matrix. When the experimental results were evaluated categorically, the best performances were obtained with the Alternating Decision Tree (ADT) method in the tree category for the metrics except for Recall and False Negative Rate (FNR), the Logistic Boosting Regression (LBST) method in the meta category for the metrics except for Area under the curve (AUC), and the Radial Bases Function Classifier (RBFC) method in the function category for the metrics except for the Precision and False Positive Rate (FPR). When considering the results in terms of the performances of all methods, the RBFC method exhibited the best performance with values of 0.8725 for Accuracy, 0.9173 for Recall, 0.8885 for F-score, 0.0827 for FNR, and 0.1275 for Misclassification Rate (MCR). On the other hand, the ADT method showed the best performance in terms of Precision, AUC, and FPR metrics with values of 0.8718, 0.9300 and 0.1610, respectively.

Keywords— Machine learning, classification, heart failure

1. Giriş

Kalp damar bozukluklarını içeren kardiyovasküler hastalıklar (Cardiovascular Diseases (CVDs)) küresel bazda en önemli ölüm sebepleri arasında gösterilir. 2019 yılında tüm dünyada genelinde

gerçekleşen yaklaşık 17.9 milyon ölümün (bu oran tüm küresel ölümlerin %32' sine denk düşer) CVDs kaynaklı olduğu Dünya Sağlık Örgütü (World Health Organization (WHO)) verileri ile açıklanmıştır. WHO' nun bu hastalıkla ilgili aynı yıla ait başka bir analizinde bulaşıcı olmayan hastalıklara bağlı 17 milyon erken ölümün (70 yaş altı) %38'inin CVDs' e dayandığı ifade edilmiştir [1]. CVDs sınıfında yer alan kalp yetmezliği kalbin vücuda yeterince kan pompalayamadığı durumda oluşan bir hastalık olup, tedavi edilemediği taktirde hastaneye yatışa varacak derecede önemli sağlık sorunlarını beraberinde getirebilir. Hatta ileri safhaları ölümlere neden olabilir [2], [3]. CVDs' in istatistikleri kalp yetmezliği bazında değerlendirildiğinde, onların dünya genelinde ciddiye alınması gereken oldukça önemli bir hastalık olduğu anlaşılacaktır. Bu hastalıklar Avrupa Birliği ülkeleri ve Türkiye'de de yüksek riskli hastalıklar sınıfında yer alır. Avrupa birliği ülkelerinde hastalık kaynaklı ölüm oranları %12 ve Türkiye'de %14 ile oldukça yüksek bir seviyede seyredir [4]. Bu yüzden kalp yetmezliğinin hastalık ilerlemeden tespit edilmesi hem tedavinin amacına ulaşmasında hem de ölüm oranlarının en aza indirilmesinde hayati bir önem taşır. Bu hususta klinik çalışmalar ile birlikte uzman kararına katkı sağlayacak yapay zekâ temelli araçların kullanılması sürecin takibinde önemli bir role sahiptir.

Yapay zekâ temelli yöntemler ile tıbbi veri analizleri ve görüntüler üzerinden hastalığın seyri takip edilerek tedavi süreci düzenlenebilir, hastalık teşhis edebilirler, hastanın kişisel bilgilerini de dikkate alınarak hastalık riski tahmin edilebilir ve görüntü temelli anormallikler tanımlanabilirler. Bu anlamda bakıldığında MÖ yöntemleri tıbbi analiz ve değerlendirme süreçlerinde yaygın kullanılan bir yapay zekâ algoritması olarak karşımıza çıkar. Kalp yetmezliği tahmininde bu yöntemlerin kullanımı metodolojik olarak sırası ile veri toplama, veri temizleme ve düzenleme, veri bölme, sınıflandırma ve değerlendirme süreçlerinin yürürlüğe konmasını gerektirir. Şu ana kadar yapılan çalışmalar incelendiğinde Rasgele Orman (Random Forest (RF)), ekstremler Gradyan Artırma (eXtreme Gradient Boosting (XGBoost)), Destek Vektör Makineleri (Support Vector Machine (SVM)), Naive Bayes (NB), Lojistik Regresyon (Logistic Regression (LR)), K-En yakın komşuluk (K-Nearest Neighbors (KNN)), Karar Ağacı (Decision Tree (DT)), J48 Ağacı (J48) ve Adaptif Önyüklemeye (Adaptive Boosting (AdaBoost)) gibi MÖ yöntemlerinin kalp yetmezliği hatalığının tahmini ve teşhisi ile bu hastalığa yakalanma riskinin belirlenmesinde [5]–[8] başarılı bir şekilde kullanıldığı görülmektedir. Yapay zekâ temelli algoritmaları arasında günümüzün popüler yöntemlerinden derin öğrenme (DÖ) kalp yetmezliğinin teşhisi ve tahmininin farklı aşamalarında yaygın bir şekilde kullanılmaktadır. Bu yöntemler özellikle MRI, echokardiyografi ve CT taramaları gibi kalp yetmezliği teşhisinde kullanılan görüntüleri işleyerek kalp anormallikleri ve fonksiyonunun değerlendirmesinde etkin rol oynarlar [9]–[12]. DÖ yöntemleri elektrokardiyografi (EKG) verileri üzerinden kalp ritmi bozukluklarını da belirleyerek hastalığı tanımlayabilir. DÖ temelli yaklaşımlar olarak Evrişimli Sinir Ağı (Convolutional Neural Network (CNN)) [13], Yinelemeli sinir ağı (Recurrent Neural Network (RNN)) [14], Uzun Kısa Süreli Bellek (Long Short-Term Memory (LSTM)) [15] ve Kapı Özyinelemeli Geçit (Gated Recurrent Unit (GRU)) [16] yöntemleri bu hastalığın teşhisinde yaygın bir şekilde kullanılmıştır. Bazı derin öğrenme yapıları

hastalık teşhisi sürecinde özellik çıkarıcı gibi kullanılabilirler. Bununla çoğu zaman CNN-RNN, CNN-LSTM ve CNN-SVM gibi hibrid yaklaşımlarda karşılaşılmaktadır [17]–[19]. Bu yapıda ilk birimler özellik çıkarıcı gibi davranırken, diğerleri sınıflayıcı olarak işlev görürler. DÖ temelli yöntemler kalp yetmezliği teşhisinde yüksek performanslar sergilese de onlar yapısal olarak kompleksliği ve hesapsal olarak yüksek işlem yoğunluğunu içeren önemli dezavantajları da beraberinde getirirler.

Bu makalede kalp yetmezliği tahmininin kategorik olarak farklı tip MÖ yöntemleri ile uygulamasına yönelik bir çalışma sunulmuştur. Ağaç, meta ve fonksiyon olmak üzere üç kategoride toplam on beş MÖ yöntemini içeren bir kalp yetmezliği tahmin modeli oluşturularak deneysel çalışmalar yapılmış ve elde edilen sonuçlar hem kategorik temelde hem de genel performans olarak değerlendirilmiştir. Çalışmanın geri kalanı şu şekilde özetlenmiştir: Bölüm 2 çalışma ile ilgili geçmişte yapılmış güncel çalışmaları açıklar. Bölüm 3 çalışmada kullanılan veri setini, kalp yetmezliği teşhisi için oluşturulan modeli, modelde kullanılan MÖ yöntemlerini ve performans ölçüm metriklerini anlatır. Bölüm 4 deneysel sonuçları ve tartışmaları sunar. Son bölüm ise genel sonuçları değerlendirir.

2. Literatür Taraması

MÖ yöntemlerinin büyük miktardaki verileri ve örüntüleri hızlı bir şekilde işleyebilme kabiliyeti onu tıp alanındaki verilerin analizinde, hastalıkların teşhisinde ve tedavi yöntemlerinin geliştirilmesinde kullanılabilir bir araç haline getirmiştir. Bu bölüm başlığı altında ciddi tıbbi bir durum ve ölümcül bir hastalık olan kalp yetmezliğinin temel MÖ yöntemleri ve hibrid MÖ yöntemleri ile teşhisi ve tahminine yönelik yapılmış güncel literatür çalışmaları sunulmuştur.

Ali ve arkadaşları [20] etkili bir kalp yetmezliği tahmini için iki SVM' i istifleyen bir uzman sistem önerdiler. Bu bütünlük yapıda ilk SVM' i gereksiz özellikleri ortadan kaldırmak için, ikinci SVM' i ise tahmin sürecini gerçekleştirmek için kullandılar. Her iki SVM' i Hibrit IZgara Arama Algoritması (Hybrid Grid Search Algorithm (HGSA)) kullanarak optimize ettiler. Önerdikleri modelin etkinliğini Doğruluk, Duyarlılık, Kesinlik, Matthews korelasyon katsayısı (MCC) ve AUC metrikleri üzerinden test ettiler. Yaptıkları deneyler ile HGSA optimizasyonuna sahip uzman sistemin geleneksel SVM' den %3.3 daha yüksek bir performans ortaya koyduğunu ispatladılar. Ayrıca, önerdikleri yöntemi, %57,85-%91,83 aralığında doğruluk elde eden başka yöntemler ile kıyasladıklarında kendi yöntemlerinin daha iyi bir performans sergilediğini gördüler.

Chen ve arkadaşları [21], kalp yetmezliğinin olumsuz etkileri için Derin Piramit Evrişimli Sinir Ağları (Deep Pyramid Convolutional Neural Networks (DPCNN)) ile XGBoost' u birleştiren hibrid bir model önerdiler. Önerdikleri modelde DPCNN' i hastanın teşhis metinlerinden özellik çıkarmak için kullanırken, XGBoost' tan kalp yetmezliği tahmin modelini oluşturmada yararlandılar. 2014' ten 2018' e kadar kalp yetmezliği olan hastaların verilerine uygulanan modellerden elde edilen sonuçlar DPCNN- XGBoost hibrid modelinin, sade DPCNN ve XGBoost modellerine göre tahmin duyarlılığını %3 ve %31

oranında iyileştirdiğini gösterdiler.

Grgić ve arkadaşlarının yaptığı kalp hastalarının kardiyovasküler parametrelerini analiz ettiği çalışmalarında [22] RF ve LR algoritmaları kullanılarak 12 parametreye göre kardiyovasküler hastalığı olan hastalarda ölüm tahminini yaptılar. Her iki algoritmanın hiper parametrelerini Izgara Arama Algoritmasını (Grid Search Algorithm (GSA)) kullanarak optimize ettiler. Elde ettikleri sonuçların karşılaştırma analizlerinden %90 ile en yüksek doğruluğu RF algoritmasının elde ettiğini gördüler. Bu algoritma AUC performansında da %86 ile en yüksek performansı sergilediğini gözlemlədiler.

Mansur Huang ve arkadaşları [7] RF, SVM, NB ve LR' yi içeren dört popüler MÖ yöntemini Kaggle web sitesinden elde ettikleri halka açık veri setine uygulayarak deneysel çalışmalar yaptılar ve bu yöntemlerin kalp yetmezliği hastalığının tahminindeki performanslarını karşılaştırdılar. MÖ yöntemlerinin performansını Doğruluk, Kesinlik, Duyarlılık ve F-skör gibi Karışıklık matrisi metrikleri üzerinden değerlendirdiler. Deneysel sonuçları kullanarak yaptıkları analizler RF yönteminin 0.88' lik ortalama performans ile SVM, NB ve LR yöntemlerine göre daha yüksek olduğunu göstermiştir. Onlar aynı zamanda özellikler üzerine yaptıkları deneylerde 13 semptomun hepsinin sınıflandırma sürecinde önemli olduğunu gözlemlədiler.

Foziljonova ve Wasito [23] kalp yetmezliği olan hastalarda sağ kalım oranlarının tahmini için uygun MÖ tekniğinin belirlenmesi üzerine deneysel çalışmalar yaptılar ve bu çalışmalardan elde ettikleri sonuçları analiz ettiler. Bu çalışmalarında 2015 yılının Nisan-Aralık ayları arasında Pakistan'daki bir hastanede bulunan kalp yetmezliği olan hastalarından topladıkları verilere KNN, RF ve DT yöntemlerini uyguladılar. F1-skör ve AUC metrikleri ile performans değerlendirmesi yaptılar. Deneysel çalışmalarında her iki metrik açısından en iyi performansının 0.861 F1-skör ve 0.957 AUC değerleri ile RF yönteminin sağladığı görülmüştür.

Mudassar ve arkadaşları [24] 6' sınıflı topluluk öğrenmesi ve 3' ü geleneksel olmak üzere toplam 9 MÖ yöntemini UCI MÖ veri tabanından elde ettikleri 299 kalp yetmezliği hastasının kayıtlarını içeren veri setine uygulayarak bu yöntemlerin kalp yetmezliği teşhisindeki başarımlarını karşılaştırdılar. Kullandıkları veri seti veri etiketleri açısından dengesiz verileri içeren bir veri seti olduğu için aynı analizleri SMOTE işlemini uygulayarak dengeli hale getirdikleri veri seti üzerinden de gerçekleştirdiler. Deneysel çalışmalarında NB, Bagging, RF, XGBoost, LightBoost, DT, AdaBoost, LR ve ETC yöntemlerinin performanslarını Doğruluk, Kesinlik, Duyarlılık ve F1-skör metrikleri ile ölçtüler. 10 kat çapraz doğrulama ve 0.7' lik bölünme şartlarında yaptıkları deneylerde SMOTE' li ve SMOTE' siz durum için en iyi performansları sırası ile 0.926 ve 0.896' lık doğruluk değerleri ile LightBoost yönteminden elde ettiler.

Gold ve Lorshase [5] kalp yetmezliği tahmini için RF ve J48-AdaBoost olmak üzere iki tahmin modeli önerdiler. Veri seti olarak Cleveland Hastanesi veri tabanından toplanmış 13 özelliğe ve 303 örneğe sahip bir veri setini kullanarak yaptıkları deneylerden elde edilen sonuçlar J48-AdaBoost modelinin %92.3077 Doğruluk, %93.0 Kesinlik, %92.3 Duyarlılık, %92.4 F1-skör ve %97.6 AUC değerleri ile RF modeline göre daha yüksek bir performans sergilediğini ispatlamıştır.

Zeng [25] klinik özellikler ile kalp hastalığı arasındaki korelasyonları gözlemek için sıcaklık haritası çıkarmış ve DT, KNN, SVM ve XGBoost yöntemlerini kullanarak oluşturduğu MÖ temelli modeller aracılığı ile hastalık tahmini yapmıştır. Bu modeller arasından performans olarak en iyisini seçmek için Doğruluk, Duyarlılık ve F1-skör metriklerinden yararlanmıştır. Deneysel veri olarak 5 bağımsız veri setini birleştiren Kaggle web sitesinden elde edilmiş kalp yetmezliği tahmini (KYT) veri setini kullanmıştır. 0.7' lik bölünme şartlarında yapılan deneylerden elde edilen sonuçlar SVM yönteminin %88.8 Doğruluk, %89.3 Duyarlılık ve %90.7 F1-skör değerleri ile en yüksek performansa ulaştığını göstermiştir. Ayrıca ısı haritası üzerinden egzersizle indüklenen anjina (ExerciseAngina) ve ST depresyonu (Oldpeak) kalp hastalığı oluşumu ile yakın ilişkiye sahip olduğu ve bunun sonucunda kalp hastalığı riskinden kaçınmak için bu iki özelliğe çok dikkat edilmesi gerektiği de kanıtlanmıştır.

Coşkun ve Kuncan [26] NB, SVM, DT, RF ve Doğrusal Diskriminant Analizi (Linear Discriminant Analysis (LDA)) sınıflandırıcılarını kaggle web sitesinden elde ettikleri 918 örneğin oluşturduğu KYT veri setine uygulayarak MÖ temelli yöntemlerin kalp yetmezliği hastalığındaki etkinliğini araştırdılar. Veri setine etiket kodlama ve normalizasyon gibi birtakım ön işlem süreçlerini uyguladıktan sonra 0.8' lik bölünme oranında eğitim ve test verilerini oluşturdular. Bu verileri kullanarak yaptıkları deneyler RF yönteminin %90.76' lık başarımları ile en yüksek performansı elde ettiğini göstermiştir. Bu algoritmanın sınıflar düzeyindeki performansı ise kalp yetmezliği olan hastalarda %93.1 ve hastalığı olmayanlarda %87.8 olmuştur.

3. Materyal ve Metotlar

Bu bölümde, çalışmada kullanılan veri seti, kalp yetmezliği tahmini için oluşturulan model, modeli oluşturan birleşenler ve onun değerlendirilmesinde kullanılan metrikler sunulmaktadır. İlk olarak modelde kullanılan veri seti ve onu oluşturan özellikler anlatılmıştır. Sonra Kalp yetmezliği tahmini için kullanılan modelin oluşum aşamaları tüm detayları ile açıklanmıştır. Son olarak ise modelin değerlendirilmesinde kullanılan performans metriklerinden bahsedilmiştir.

3.1. Veri Seti

Çalışmamızda kalp yetmezliğini tahmin etme amaçlı kullanılan MÖ temelli yöntemlerin uygulaması Kaliforniya Üniversitesi-Irvine (University of California Irvine (UCI)) Makine Öğrenimi Veri tabanındaki "Cleveland", "Hungarian", "Switzerland", "Long Beach VA" ve "Stalog" veri setlerinin birleştirilmesi ile oluşturulmuş Kaggle web sitesinden alınan KYT veri seti üzerinden yapılmıştır [27]. 725 erkek ve 193 kadın katılımcıdan oluşan bu set 410' u normal ve 508' i kalp hastası toplam 918 veri örneği içerir. İlgili veri seti yaş, cinsiyet, kan basıncı, kolesterol seviyeleri, glikoz seviyeleri, elektrokardiyogram sonuçlarını, kalp hastalığı durumu gibi farklı nitelikleri beraberinde getiren 12 özelliğe sahiptir. Tablo 1' de KYT veri setinin özellikleri ile bu özelliklerin açıklamaları, türleri ve değer aralıklarına ait özet bir sunum gösterilmiştir.

Tablo 1. KYT veri setinin özellikleri ile onun açıklamaları, türleri ve değer aralıklarına ait özet bir sunum.

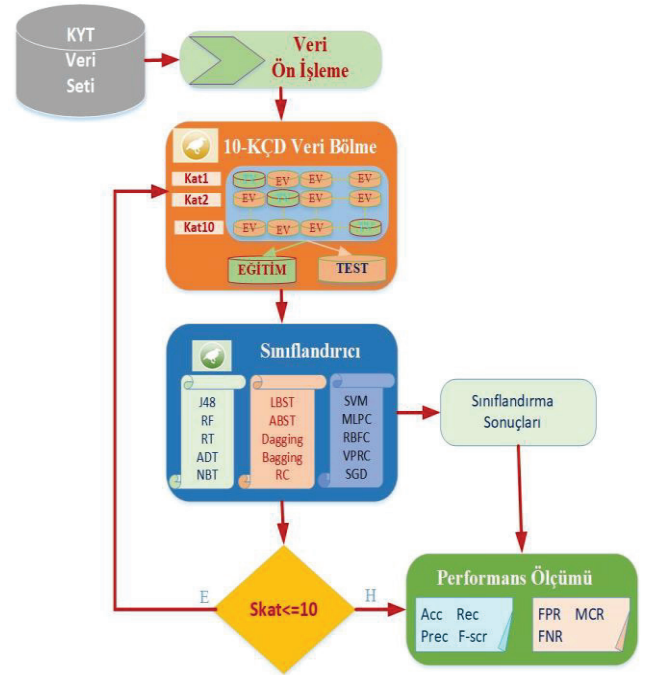
Özellik Adı	Tanımı	Türü	Değeri
Age	Hastanın yaşı	Tam sayı	28-77
Sex	Hastanın cinsiyeti	Kategorik	M: Bay F: Bayan
ChestPainType	Göğüs ağrısı tipi	Kategorik	TA: Tipik Anjin ATA: Atipik Anjin NAP: Anjinal Olmayan Ağrı ASY: Asemptomatik
RestingBP	Dinlenmiş kan basıncı	Tam sayı	80-200
Cholesterol	Serum kolesterolü	Tam sayı	0-603
FastingBS	Açlık kan şekeri	Tam sayı	1: Eğer FastingBS>120 mg/dl 0 Aksi durumda Normal: Normal ST: ST-T dalga anormalliğine sahip olmak
RestingECG	Dinlenmiş EKG sonuçları	Kategorik	LVH: Estes'in kriterlerine göre keskin sol ventrikül hipertrofisi göstermesi veya onun olasılığı
MaxHR	Maksimum kalp atış hızı	Tam sayı	60-202
ExerciseAngina	Egzersizle ilişkili anjin	Boolean	Y: Evet N: Hayır
Oldpeak	Depresyonda ölçülen Eskitepe	Float	-2.6-6.2
ST_Slope	En yüksek egzersiz ST segmentinin eğimi	Kategorik	Up: Yukarı eğimli Flat: Düz Down: Aşağı Eğimli
HeartDisease	Çıkış sınıfı	Tam sayı	1: Kalp hastalığı 0: Normal

Tabloda verilen 11 özellik, kalp hastalığı olup olmadığını sınıflandırmak için kullanılan "HeartDisease" bağımlı değişkeni ile ilişkilendirilir. Bağımlı değişkenin 1 değerini alması kalp hastası olma durumunu, 0 değerini alması ise bu hastalığa sahip olmama (normal olma) durumunu temsil eder.

3.2. Kalp Yetmezliği Tahmini için MÖ Temelli Model

Bu çalışmada ağaç, fonksiyon ve meta kategorisinde yer alan 15 MÖ yönteminin kalp yetmezliği tahminindeki performanslarını test edebilmek için oluşturulan modele ait şema Şekil 1' de gösterilmiştir. Hem bireysel hem de kategorik bazda MÖ yöntemlerinin performansını ölçümünde kullanılan model; ön işlem, 10-kat çapraz doğrulamalı (10-KÇD) veri bölme, sınıflandırma ve performans ölçümü olmak üzere dört temel prosedürü birleştirir. Bu model ilk olarak KYT veri setinden gelen ikili ya da kategorik formattaki verileri bir veri dönüşüm ön işlem prosedürü üzerinden sayısal formatlı veriler haline dönüştürür. Sonra bu veriyi MÖ yöntemlerinin eğitim ve test süreçlerini yürürlüğe konulmasında kullanılan eğitim ve test veri setlerini oluşturabilmek için 10-KÇD veri bölme

prosedürüne göndererek aynı veri setinden elde edilmiş 10 farklı bölünmüş eğitim ve test verisini elde eder. Daha sonra bu veriler sınıflandırma prosedürüne gönderilerek her bir eğitim ve test verisi için önce eğitim sonra test süreci yürürlüğe konarak sınıflandırma tahmini sonuçları elde edilir. Bölünmüş veri indeksi 10' a ulaştığında performans ölçüm prosedürü tüm bölünmüş test verisi için elde eden sınıflandırma sonuçlarına göre MÖ sınıflandırıcıların ortalama performansını hesaplar.



Şekil 1. Kalp yetmezliği tahmini için oluşturulan MÖ temelli model.

A. Ön İşlem Prosedürü

Kaggle web sitesinden alınan KYT veri seti farklı veri setlerinin birleştirilmesi sonucu oluştuğu için eksik özellik ve değerler içermemektedir. Bununla birlikte veri setinin içerdiği özelliklerin değerleri tür olarak sayısal, kategorik ya da ikili format olarak değişmektedir. Bu yüzden ön işlem prosedüründe ikili ve kategorik formattaki özellik değerleri tam sayı formatına dönüştürülmüştür. Bu anlamda veri seti üzerinde yapılan değişimler Tablo 2' de verilmiştir.

Tablo 2. Ön işlem süreci sonunda ikili ve kategorik formattaki özellik değerlerindeki değişimler

Özellik Adı	Türü(önceki)	Türü(yeni)	Değer Değişimi
Sex	Kategorik	Tam sayı	M: 1, F: 0
ChestPainType	Kategorik	Tam sayı	TA: 3, ATA: 2 NAP: 1, ASY: 0
RestingECG	Kategorik	Tam sayı	Normal:2, ST: 1 LVH: 0
ExerciseAngina	Boolean	Tam sayı	Y: 1, N: 0
ST_Slope	Kategorik	Tam sayı	Up: 2, Flat: 1 Down: 0

B. 10-KÇD' lı Veri Bölme Prosedürü

MÖ yöntemlerinin performansını doğru bir şekilde ölçebilmek için bu süreçte 10 kat çapraz doğrulama yaklaşımı kullanılmıştır. Bu yüzden sınıflandırıcıların işlevlerini yerine getirebilmelerinde ihtiyaç duydukları eğitim ve test verilerinin elde edilmesinde 10-KÇD' lı bir veri bölme prosedüründen yararlanılmıştır. Bu prosedür veri ön işlem sürecinden gelen veriyi kullanarak birbirinden farklı 10 bölünmüş eğitim ve test veri setini oluşturur.

C. Sınıflandırma Prosedürü

Çalışmada sınıflandırıcı olarak ağaç, meta ve fonksiyon kategorisinde yer alan ve birçok sınıflandırma probleminde başarılı sonuçlar ürettiği bilinen 15 MÖ temelli yöntemden yararlanılmıştır. Bu prosedürde ilk kategori olan ağaç kategorisinde J48, RF, Rasgele Ağaç (Random Tree (RT)), ADT ve NB Ağacı (NB Tree (NBT)) kullanılırken, Meta kategorisinde Önyüklemeye Toplaması (Bootstrap Aggregating (Bagging)), Ayrık Toplama (Disjoint Aggregating (Dagging)), ABST, LBST ve Rasgele Komite (Random Committee (RC)) kullanılır. Fonksiyon kategorisinde ise SVM, Çok Katmanlı Algılayıcı Sınıflandırıcısı (Multilayer Perceptron Classifier (MLPC)), RBFC, Oylanan Algılayıcı (Voted Perceptron (VPRC)) ve Stokastik Dik İniş (Stochastic Gradient Descent (SGD)) yer alır. Bu yöntemler aşağıda alt başlıkla halinde anlatılmıştır.

J48: C4.5 algoritmasının bir uzantısı olup, kök düğümünü kullanarak örneklerin sınıflandırılmasını sağlayan bir karar ağacı algoritmasıdır. Etkili bir yöntem olduğu için veri madenciliği ve sınıflandırma problemlerinde yaygın bir şekilde kullanılır. Yöntem içerisinde karar ağaçlarını oluşturulmasında veri kümesindeki örnekleri özellik değerlerine göre alt gruplara ayıran ve ismi "böl ve fetet" olan bir yaklaşımdan yararlanır [28]. Örneklerin değerlendirilmesinde veri setini en iyi şekilde bölen eşik değeri kullanılır. Bu değer genel olarak Eşitlik 1' de formülü verilen bilgi kazancı ölçümünden elde edilir.

$$IG(A) = H(S) - \sum_{t \in T} p(t)H(t) \quad (1)$$

Burada $H(S)$ S kümesinin entropisini, T S' nin A tahmincisine göre bölünmesiyle oluşturulan alt kümeyi, $p(t)$ t' deki eleman sayısının S kümesindeki eleman sayısına oranı ve $H(t)$ t alt kümesinin entropisini temsil eder.

RF: Eğitim sürecinde bir dizi karar ağacı oluşturduktan sonra onların tahminlerini ortalama ya da oylama teknikleri ile birleştirerek nihai çözüme ulaşan MÖ temelli bir topluluk algoritmasıdır. RF yönteminin ormanda oluşturduğu yeteri miktarda bir ağaç topluluğu ile aşırı uyum problemlerini giderebilmesi, eksik veriler üzerinden çalışabilmesi ve kategorik değerler için modellenebilmesi gibi bir takım önemli avantajları onu veri biliminde yaygın kullanılan bir yöntem haline getirmiştir [29]. Bu yöntemde karar ormanın oluşturulmasındaki ilk adım veri setinden rasgele örnekleme yöntemi ile alt kümelerin elde edilmesidir. Sonra bu alt kümelere CART ya da ID3 gibi algoritmalar uygulanarak karar ağaçları oluşturulur. Bu ağaçlar kendisi için seçilmiş eğitim örnekleri kullanılarak büyütülür. Sonuçta büyütülmüş ve üzerinde budama yapılmamış bu ağaç yapıları bir araya getirilerek karar ormanı oluşturulur. Bu yöntem bir gözleme ait tahmin edilen çıktıyı, ormandaki ağaçların çoğunluk oyuna göre

belirler.

RT: Rasgele bir şekilde oluşturulmuş ağaçları içeren ve danışmanlık öğrenmeyi kullanan bir karar ağacı yaklaşımıdır [30]. Bu yöntem, Gini indeksi ve bilgi kazancı gibi fonksiyonları kullanmadan rasgele seçtiği bir özellik üzerinden düğüm genişlemesi yaparak karar ağaçlarını oluşturur. Rasgele oluşturduğu karar ağaçlarının budamasını gereksiz ve eksik gördüğü düğümleri kaldırarak yapar. Bu yapıda her ağaç bir sınıf olasılık dağılımı üretir ve oluşturulan ağaçlardan elde edilen sınıf dağılım çıktılarının ortalaması, nihai çıktıyı verir.

ADT: Geniş bir uygulama alanına sahip olan yöntem karar ağaçlarının, oylanan karar ağaçlarının ve oylanan karar kütüklerinin bir genellemesidir [31], [32]. ADT, bir tahmin koşulunu ve tahmin düğümlerini belirtmek için tek bir sayı içeren karar düğümlerinin bir ürünü olarak ortaya çıkar. Bu yöntemde her karar düğümünün yerini ayrıncı düğüm ve tahmin düğümü olmak üzere iki düğüm alır. Karar ağaçlarında olduğu bir örnek ağaç ile birlikte kökten başlayarak yapraklara kadar tüm yolları takip eder ve üzerinden geçilen tüm düğümleri toplar.

NBT: Naive Bayes sınıflandırıcısı ve karar ağaçlarını bileştiren hibrid bir modeldir. Yapısal olarak bu algoritma bize geleneksel J48 algoritmasının her yaprak düğümüne NB sınıflandırıcısının uygulanmış bir formunu sunar. Bu yöntem ağaç yapısını oluştururken, tıpkı J48 algoritmasında olduğu gibi bilgi kazanımını dikkate alarak düğümleri bölme ve ağaçları budama sürecini yürütür. Süreç içerisinde ağaç büyüdükten sonra yapraklara ait veriler ile her bir yaprak için bir Naive Bayes oluşturur. NBT kendisine uygulanan örnekleri sınıflandırırken yapraklar üzerindeki bu Naive Bayes' leri kullanır [33]. Düğümlerdeki Naive Bayes' lere sınıflandırma sonuçlarına aşağıdaki denklem üzerinden karar verirler [34].

$$N(x) = \arg \max_{y=0 \text{ or } 1} P(y) \prod_{n=1}^N P(x_n | y) \quad (2)$$

Burada $N(x)$ yaprak düğümlerinin ürettiği sınıflandırma sonuçlarını, x giriş örneklerini, y sınıfları, $P(y)$ sınıfların önsel olasılığını, $P(x_n | y)$, verilen x örneğinin n ' inci özelliğinin koşullu olasılığını ve N x ' in toplam özellik sayısını temsil eder.

Bagging: Öğrenci topluluğu oluşturarak bunların sonuçlarının birleştirilmesini esas alan bir topluluk algoritmasıdır. Bu yöntem sonuca ulaşmak için ilk olarak orijinal veri setinden rasgele bir şekilde aldığı örnekleri %100 olacak şekilde çoklar. Sonra oluşturduğu m adet eğitim örneğini kullanarak eğittiği modellere test veri setini uygular. Daha sonra elde ettiği test sonuçlarını ortalama yaklaşımı ile birleştirerek nihai sonuca ulaşır [35]. Sonuçların birleştirilmesi Eşitlik 3' te verilen denklem kullanılarak yapılır.

$$f(E) = \frac{1}{M} \sum_{m=1}^M f_m(E) \quad (3)$$

Burada $f_m(E)$ E girdi kümesi için modellerin skor olarak bireysel çıktılarını, M model sayısını ve $f(E)$ E girdi kümesi için Bagging yönteminin nihai çıktısını gösterir.

Dagging: Aynı veri kümesinden farklı alt küme modelleri oluşturarak eğittikten sonra sonuçlarının birleştiği başarılı bir topluluk algoritmasıdır. Yapısal olarak Bagging algoritmasına benzemekle

birlikte farklılık temel sınıflandırıcıların oluşturulmasında yararlanılacak örnek kümesinin elde edilmesinde meydana gelir. Bir başka deyişle Dagging, temel sınıflandırıcıları elde etmek için yeniden örnekleme yerine çok tabakalı örnekleme kullanır [36], [37]. Nihai sonucu temel sınıflandırıcıların çıktılarını oylayarak elde eder. Bu yonteme ait modelin oluşturulması şu şekilde özetlenebilir: ilk aşamada eğitim veri seti her örnek sadece bir defa kullanılarak k örnek içerecek şekilde L alt kümeye bölünür. Sonra her alt kümeyi temsil eden bir regresyon modeli oluşturulur. Son aşamada ise modellerin sonuçları oylanarak en çok oyu alan modelin çıktısı Dagging yönteminin nihai sonucu olarak atanır. Bu yöntemin veri setini alt örnekleme kümelerine ayırarak işlemesi, onu veri madenciliği ve makine öğrenimi alanlarında özellikle büyük veri kümeleriyle yapılan çalışmalar için kullanışlı bir yöntem haline getirmiştir.

ABST: 1995 yılında Robert Schapire ve Yoav Freund' in geliştirdiği artırılmalı tip bir topluluk algoritmasıdır. Algoritmanın esası zayıf öğrenicileri bir araya getirerek daha güçlü bir öğrenici modeli oluşturmaya dayanır [38]. Algoritmada eğitim zayıf öğrenicilerle başlar. Yanlış tahmin edilen örneklerin ağırlıkları öğrenici çıktılarının ağırlıklı hata oranlarına göre güncellenir. Süreç belirlenen sayıda öğrenici eğitilene kadar sürer. Bu işlemlerin sonunda tüm öğrenicilerin ağırlığı onların performansına göre hesaplanır ve Eşitlik 4' te verilen formülde gösterildiği gibi bu ağırlıklar üzerinden bireysel sonuçlar birleştirilerek nihai sonuca ulaşılır [39].

$$H_k(x) = +/-(\sum_{k=1}^K a_k h_k(x)) \quad (4)$$

Burada K zayıf öğrenicilerin sayısını, h_k k ' nıncı zayıf öğrenicinin çıkışı ve a_k k ' nıncı zayıf öğrenicinin ağırlığını temsil eder.

LBST: Özellikle ikili sınıflandırma problemlerinde kullanılan başarılı bir MÖ yaklaşımıdır. Bu yöntem gürültülü ve aykırılıkları içeren verileri işlemede ABST' un sınırlamalarını gidermek için geliştirilmiştir. LBST' u ABST' tan ayıran en önemli özelliği onun aykırı verilere ve gürültüye karşı daha az duyarlı hale getiren kayıp fonksiyonunda vuku bulur [40]. ABST sınıflandırma sürecinde kayıp fonksiyonu olarak üssel bir fonksiyondan yararlanırken, buna karşı LBST kayıp işlevini doğrusal bir fonksiyon olan binom log olasılığını kullanarak yerine getirir. N örnek ve iki sınıfa ait bir eğitim verisi için ağırlıklar ve çalışma cevabı Eşitlik 5' de verilen formüller ile hesaplanabilir [41].

$$w_i = p(x_i)(1 - p(x_i)), \quad z_i = \frac{y_i - p(x_i)}{p(x_i)(1 - p(x_i))} \quad (5)$$

Burada x_i i ' ninci özellik vektörünü, y_i bu vektöre ait sınıf etiketini ve $p(x_i)$ x_i özelliğinin olasılığını temsil eder.

RC: Farklı temel sınıflayıcıların çıktılarını birleştirerek tahmin sonucunu üreten ve topluluk öğrenme yaklaşımını esas alan bir yaklaşımdır. Bu yöntemde her bir temel sınıflandırıcı, aynı veri kümesinden yararlanarak onların rasgele bir şekilde davranmasını sağlayan farklı tohumlar ile oluşturulur [42]. Temel sınıflandırıcıların bu şekildeki yapılandırılması bir yandan yöntem için güçlü bir tahmin kombinasyonu sağlarken, diğer taraftan sınıflayıcılarda oluşacak aşırı

uyum problemini azaltabilir. Nihai tahmin sonucuna, girdi verilerine göre bağımsız tahminler üreten temel sınıflandırıcıların çıktılarının genel ortalaması alınarak ulaşılır.

SVM: Sınıflandırma ve regresyon problemlerinin çözümünde yaygın kullanılan bir denetimli öğrenme algoritmasıdır. Bu yöntem sınıflara ait veri noktalarını en yüksek marj' a sahip bir hiper düzlem üzerinden ayırır. Hiper düzlemde marjın sınırları destek vektörleri adı verilen veri noktaları belirler. Marj sınırları arasındaki mesafe arttıkça yöntemin genelleme yeteneği artar ve bunun sonucunda daha iyi tahmin sonuçlarına ulaşır. SVM verilere ait sınıflara bölme bir başka deyişle hiper düzlem oluşturma görevini doğrusal ya da çekirdek (Kernel) fonksiyonları üzerinden yerine getirir [43]. Verilerin bir çizgi kullanılarak ikiye bölünebildiği durumlarda aşağıda verildiği gibi doğrusal tip bir fonksiyondan yararlanılabilir.

$$w_0 x + b_0 = 0 \quad (6)$$

Burada x giriş vektörünü, w_0 ağırlık vektörünü ve b_0 sapmayı gösterir. Verilerin doğrusal bir fonksiyon ile bölünemediği durumlarda ise Kernel fonksiyonları kullanılabilir [44].

$$K(x_i, x_j) = \{x_i^T x_j + c\}^d \quad (7)$$

Formülde x_i ve x_j giriş uzayı vektörü, d polinomun derecesi ve $x_i^T x_j$ ' nin transpozudur. c , en yüksek ve en düşük mertebeli polinom arasındaki denge için kullanılan bir parametredir.

MLPC: İleri beslemeli bir yapay sinir ağı modeli olan çok katmanlı algılayıcıyı (MLP) esas alan güçlü bir algoritmadır. Yapısal olarak giriş, gizli ve çıkış olmak üzere üç temel katmandan oluşur. Bu yapıda giriş katmanı dış dünyadan gelen verileri ağırlıklandırarak gizli katmana gönderir. MLP' nin hesaplama motoru olan gizli katman giriş katmanından gelen bilgileri işleyerek çıkış katmanına aktarır. Çıkış katmanı bu bilgileri eşik değerine göre işleyerek sınıflandırır [45]. Katmanlardaki işlemler işlem elemanı olarak isimlendirilen birimler üzerinden gerçekleştirilir. Onlar Eşitlik 8' de gösterildiği gibi kendilerine gelen bilgilerin ağırlıklı toplamlarını bularak net girdilerini hesaplar ve bu girdiyi bir eşik değeri ile birlikte aktivasyon fonksiyonundan geçirerek çıktılarını elde ederler.

$$net_k = \sum_{i=1}^n x_i * w_i + b_k, \quad O_k = f(net_k) \quad (8)$$

Burada net_k k ' nıncı işlem elemanı için net girdiyi, x_i işlem elemanına gelen i ' inci bilgi girişini, w_i işlem elemanına gelen i ' inci girişin bağlantı ağırlığını, $f(\bullet)$ işlem elemanı için aktivasyon fonksiyonunu ve O_k k ' nıncı işlem elemanının çıkışını gösterir.

RBFC: Sınıflandırma amaçlı kullanılan normalize edilmiş bir Gaussian tip radyal temel fonksiyon (RBF) ağıdır. Tipik bir RBF sınıflandırıcısı giriş, gizli ve çıkış olmak üzere üç katmandan oluşur. Eğitim süreci WEKA' nın optimizasyon sınıfı ile yürürlüğe konur ve bu süreçte hatayı en aza indirmede Broyden-Fletcher-Goldfarb-Shanno (BFGS) algoritması kullanılır. Başlangıç değerleri k-Means algoritması ile belirlenen merkezlerin herhangi birinin en yakın komşuna en uzak olan mesafe dikkate alınarak ayarlanır. Aşağıda Gaussian radyal temel fonksiyonu kullanılarak eğitim süreci sonunda öğrenilen modele ait çıkış temsili verilmiştir [46].

$$f(x_1, x_2, \dots, x_n) = h \left(w_{L,0} + \sum_{i=1}^m w_{L,i} \left(-\sum_{j=1}^n \frac{\alpha_j^2 (x_j - c_{L,j})^2}{2\sigma_{L,j}^2} \right) \right) \quad (9)$$

Burada x_1, x_2, \dots, x_n giriş vektörünü, $h(\bullet)$ logistik fonksiyonu, m temel fonksiyonların sayısını, $w_{L,i}$ her temel fonksiyonun ağırlığını, $c_{L,i}$ temel fonksiyon merkezlerini, α_j^2 j. özniteliğin ağırlığını ve $\sigma_{L,j}^2$ ise varyanslarını temsil eder.

VPRC: Doğrusal olmayan verilerin sınıflandırılmasında kullanılan etkili bir yaklaşımdır. Bu yöntem yapısal olarak orijinal perceptron ile Helmbold ve Warmuth' un tanıttığı çevrim içi öğrenme algoritmalarının yığın öğrenme algoritmasına uyarlanmasını esas alır [47]. O, nihai sonucu daha az hata yapan modelleri daha yüksek ağırlıklar ile etiketleyerek elde eder. Ağırlıklara, yanlış sınıflandırıldığında kaydedilen tahmin vektörleri ile doğru tahminde süreleri izlenerek karar verilir. Eğitim sürecinde de o, her hatadan sonra oluşan tahmin vektörleri kaydeder ve vektörlerin hata yapılmaya kadar geçen süresi hesaplanarak tahmin vektörlerine ağırlıklar atanır [48]. Bu süreç her veri akışında tekrar eder. Bu yüzden model deneyim kazandıkça daha iyi performans sergiler.

SGD: Dik iniş (GD) algoritmasının gelişmiş bir sürümüdür. Bu yöntem, SVM ve LR algoritmaları gibi doğrusal sınıflayıcılar altında konveks kayıp fonksiyonları ile ayırt edici öğrenmeyi destekleyebildiği için modellerin optimizasyonunda yaygın bir şekilde kullanılır. SGD' yi GD algoritmasından ayıran en önemli özelliği, onun veri kümesinin tamamının gradyanı almak yerine bu kümeden rasgele seçilmiş örnekler kullanılarak yaklaşık bir gradyan hesaplamasıdır. Bu özelliği onun hesapsal maliyetini düşürerek büyük veri kümelerinde hem kolay bir şekilde uygulanabilmesini sağlar hem de verimliliğini artırır [49]. Eşitlik 10' da verilen gradyan hesaplaması ile birlikte dikkate alınan tüm veri örneklerine ait ağırlık ve etiketler güncellenir. Bu güncelleme minimum maliyete ulaşana kadar her adımda yenilenir.

$$\nabla f(x) = \frac{1}{n} \sum_{i=1}^n \nabla f_i(x) \quad (10)$$

Formülde $f(x)$ amaç fonksiyonunu, $\nabla f(x)$ amaç fonksiyonunun gradyan inişini, n toplam veri örneği sayısını ve $\nabla f_i(x)$ ise rastgele sınırlı veri örnekleri için hesaplanan gradyan inişini gösterir. Seçilen veri örnekleri için hesaplanan gradyanlar ile ağırlıkların güncellenmesi Eşitlik 11' deki formül üzerinden yapılır [50].

$$x \leftarrow x - \alpha \nabla f_i(x) \quad (11)$$

Burada x veri örneğini, α ise öğrenme hızını açıklar.

D. Performans Ölçüm Prosedürü

Bu prosedür MÖ yöntemlerinin performansının ölçülmesinde gerçek değerler ile tahmin edilen değerler arasındaki ilişkiyi veren sayıların oluşturduğu bir matris olan Karışıklık matrisinden yararlanır. Bu matriste gerçek ve tahmin edilen değerler arasındaki sayısal ilişki; Doğru Pozitif (True Positive (TP)), Doğru Negatif (True Negative (TN)), Yanlış Pozitif (False Positive (FP)) ve Yanlış Negatif (False Negative (FN)) olmak üzere dört temel bileşen ile etiketlenir. Karışıklık matrisinde TP bileşeni gerçekte pozitif olan ve

sınıflandırma yönteminin de pozitif olarak tahmin ettiği örnek sayısını, TN bileşeni gerçekte negatif olan ve sınıflandırma yönteminin de negatif olarak tahmin ettiği örnek sayısını, FP bileşeni gerçekte negatif olmasına rağmen sınıflandırma yönteminin pozitif olarak tahmin ettiği örnek sayısı ve FN bileşeni gerçekte pozitif olmasına rağmen sınıflandırma yönteminin negatif olarak tahmin ettiği örnek sayısını temsil eder [51]. Performans ölçüm prosedüründe MÖ yöntemlerinin sınıflandırma sürecindeki performansının ölçülmesinde Doğruluk (Accuracy (Acc)), Kesinlik (Precision (Prec)), Duyarlılık (Recall (Rec)), F-skor (F-score (F-scr)), Yanlış Sınıflandırma Oranı (Misclassification Rate (MCR)), Yanlış Pozitif Oranı (False Positive Rate (FPR)), Yanlış Negatif Oranı (False Negative Rate (FNR)) ve ROC eğrisi altında kalan alan (Area Under Curve (AUC)) metriklerinden yararlanılmıştır.

Tablo 3. Performans ölçümlerinde kullanılan metriklerin formül ve tanımları

Metrik Adı	Formülü	Tanımı
Acc	$\frac{TP + TN}{TP + TN + FP + FN}$	Doğru tahmin edilen örnek sayısının toplam örnek sayısına oranı ile temsil edilen bu metrik sınıflandırma yönteminin genel doğruluğunu verir.
Prec	$\frac{TP}{TP + FP}$	Sınıflandırma yönteminin pozitif olarak tahmin ettiği örnek sayısının ne kadarının doğru tahmin edildiğini ölçer.
Rec	$\frac{TP}{TP + FN}$	Sınıflandırma yönteminin gerçek pozitif sınıfı içeren örneklerin ne kadarını doğru bir şekilde tahmin ettiğini gösteren ölçüdür.
F-scr	$2 \times \frac{Prec \times Rec}{Prec + Rec}$	Duyarlılık ve kesinlik metriklerinin harmonik ortalamalarını olarak tek bir değerde birleştiren bu metrik sınıflandırma yönteminin performansını daha geniş bir çerçeveden değerlendirme imkânı sağlar.
MCR	$\frac{FP + FN}{TP + TN + FP + FN}$	Bu metrik sınıflandırma yönteminin genel doğruluğunu veren ölçülerden biri olup, toplam bazda yanlış tahmin edilmiş örnek oranını gösterir.
FPR	$\frac{FP}{FP + TN}$	Sınıflandırma yönteminin negatif sınıfı tahmin ederken yanlış pozitif olarak sınıflandırdığı örneklerin oranını veren bir ölçüdür.
AUC	$\int TPR.d(FPR)$	ROC eğrisinin altında kalan alanı temsil eden bu metrik, sınıflandırma yönteminin tüm olası durumlarını değerlendirebildiği için onunla ilgili genel performansı en iyi şekilde özetleyen önemli bir ölçüdür.

Doğruluk metriği özellikle dengeli veri setlerinde yöntemin tüm sınıfları eşit derecede ne kadar doğru tahmin ettiğini veren bir ölçüdür. Kesinlik ve duyarlılık metriklerini ise sırası ile yanlış pozitiflerin ve

yanlış negatiflerin tespitinin kritik öneme sahip olduğu durumlarda tercih edilir. Bununla birlikte her iki metrik arasında ters bir ilişki vardır ve bu yüzden ölçümlerde metrikler arası bir denge durumu olmasına ihtiyaç duyulabilir. Bu bağlamda ölçümlerde kesinlik ve duyarlılık metrikleri açısından bir denge unsuru olan F-skor metriğinden yararlanılabilir [52]. Ölçümlerde sınıf ayırt etme yeteneğini göstergesi olan ROC alanı [53], sınıflandırma yönteminin tekbir sayısal değer üzerinden genel performansını yansıttığı için farklı yöntemlerin sınıflandırma sürecindeki performanslarını karşılaştırmada oldukça kullanışlı bir metriktir.

4. Deneysel Sonuçlar

Kalp yetmezliği tahmininde kullanılan MÖ yöntemlerinin

sınıflandırma sürecindeki performanslarını ölçme amaçlı Bölüm 3’ te oluşturduğumuz modelin uygulaması WEKA yazılımı üzerinden gerçekleştirilmiştir. Yazılımın bize hem sınıflandırma süreçlerinin analizinde MÖ yöntemleri olarak daha geniş bir algoritma yelpazesi sunması, hem de oluşturulmuş modeli daha hızlı bir şekilde yürürlüğe koyma imkânı sağlaması uygulamanın bu ortamda geliştirilmesindeki temel etkidir. Güvenilir bir performans değerlendirmesi yapabilmek için deneysel çalışmalarda 10 kat çapraz doğru tekniği kullanılmıştır. Ağaç, meta ve fonksiyon kategorisinde yer alan sınıflandırıcılar kategorik ve genel performans olarak değerlendirilmiştir. Deneysel çalışmalarda MÖ temelli yöntemlerin ayar parametrelerinde Tablo 4’ te verilen değerler ve WEKA yazılımının da ilgili yöntemlere ait default değerleri dikkate alınmıştır.

Tablo 4. Kalp yetmezliği tahmini için oluşturulan modelin deneysel çalışmalarında kullanılan MÖ temelli yöntemlerin ayarlanmış parametreleri

Yöntem	Parametre değeri	Yöntem	Parametre değeri	Yöntem	Parametre değeri
<i>J48</i>	BatchSize. 100 CoffidenceFactor. 0.45 MinNumObj. 2, NumFolds. 3 Seed. 1	<i>ABST</i>	BatchSize. 100 Classifier. DecisionStump NumDecimalPlaces. 2 WeightTreshold. 100 Seed. 1, NumIteration. 50	<i>SVM</i>	BatchSize. 100 SVMType. C-SVC CachSize. 40, cofe0. 0.0 Cost. 1.0, degree. 3, loss. 0.1 Eps. 0.001, gamma. 0.25 KernelType. RBF: exp
<i>RF</i>	BatchSize. 100 BagSizePercent. 100 Maxdepth. 2, Seed. 1 NumDecimalPlaces. 2 NumExecutionSlots. 1 NumIteration. 100	<i>LBST</i>	BatchSize. 100, Zmax. 3 Classifier. DecisionStump NumDecimalPlaces. 2 NumThreads. 1 PoolSize. 1, shrinkage. 1.0 WeightTreshold. 100 Seed. 1, NumIteration. 15	<i>MLPC</i>	BatchSize. 100 Act.Func. Appr.Sigmoid Lossfunc. Appr.AbsulateErr NumDecimalPlace. 2 PoolSize. 1, ridge. 1.0E-4 Tolerance. 1.0E-8 Seed. 1.
<i>RT</i>	BatchSize. 100, KValue.. 1 Maxdepth. 0, Seed. 1 NumDecimal Places. 2 MinVarianceProbe. 0.001	<i>Dagging</i>	BatchSize. 100 Classifier. DecisionStump NumDecimalPlaces. 2 NumFolds. 10, Seed. 1	<i>RBFC</i>	BatchSize. 100 Scale,Optm.opt. Use. Sc. Per. NumDecimalPlace. 2 PoolSize. 1, ridge. 0.1 Tolarence. 1.0E-8, Seed. 2
<i>NBT</i>	BatchSize. 100, Seed. 1 NumDecimal Places. 2	<i>Bagging</i>	BatchSize. 100 BagSizepercent. 100 Classifier. DecisionStump NumDecimalPlaces. 2 NumExecutionSlots. 1 Seed. 1, NumIteration. 10	<i>VPRC</i>	BatchSize. 100 Exponent. 2. NumDecimalPlace. 2 NumIteration. 1 Seed. 2
<i>ADT</i>	BatchSize. 100, Seed. 1 NumDecimal Places. 2 NumofBoostingIteration. 10	<i>RC</i>	BatchSize. 100 Classifier. RamdomTree NumDecimalPlaces. 2 Seed. 1, NumIteration. 10	<i>SGD</i>	BatchSize. 100, epochs. 100 Epsilon. 0, Lamda. 0.001 Learningrate. 0.0025 Lossfunc. Log loss(LR)

Tablo 5’ te 10 kat çapraz doğrulama şartlarında MÖ yöntemlerin temel metrikler açısından performansları gösterilmiştir. Deneysel sonuçlar kategorik temelde karşılaştırıldığında değerlendirmeler şu şekilde sunulabilir: Ağaç kategorisinde Duyarlılık metriği dışında 0.86383 Doğruluk, 0.8718 Kesinlik, 0.8778 F-skor ve 0.9300 AUC değeri ile en yüksek performansı ADT yöntemi elde etmiştir. Bu kategoride Duyarlılık metriği açısından en iyi performansa 0.8976 değeri ile NBT yöntemi ulaşmıştır. Ağaç kategorisinde RT; 0.70915 Doğruluk, 0.7444 Kesinlik, 0.7224 Duyarlılık, 0.7333 F-skor ve 0.7430 AUC değerleri ile en düşük performansa sahip yöntem olmuştur. Meta

kategorisinde AUC metriği dışında 0.8682 Doğruluk, 0.8617 Kesinlik, 0.9075 Duyarlılık ve 0.8840 F-skor değerleri ile LBST en yüksek performansı sergilemiştir. Bu kategoride 0.9220 ile en yüksek AUC değerine sahip ABST yöntemine göre LBST 0.0030 daha düşük bir AUC performansı ortaya koymuştur. Meta kategorisinde RC yöntemi Doğruluk, Kesinlik, Duyarlılık ve F-skor metrikleri açısından sırası ile 0.7756, 0.7786, 0.8307 ve 0.8038 değerleri ile en düşük performansa sahip olurken, AUC metriğinde bu sonucu 0.7900 değeri ile Bagging yöntemi ortaya koymuştur. Fonksiyon kategorisinde ise Kesinlik metriği dışında 0.8725 Doğruluk, 0.9173 Duyarlılık, 0.8885 F-skor ve 0.9240 AUC

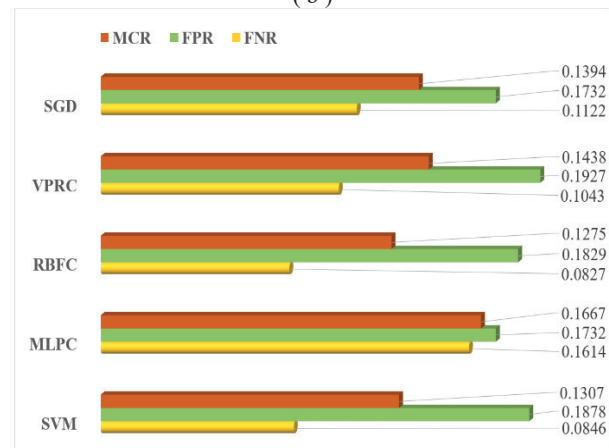
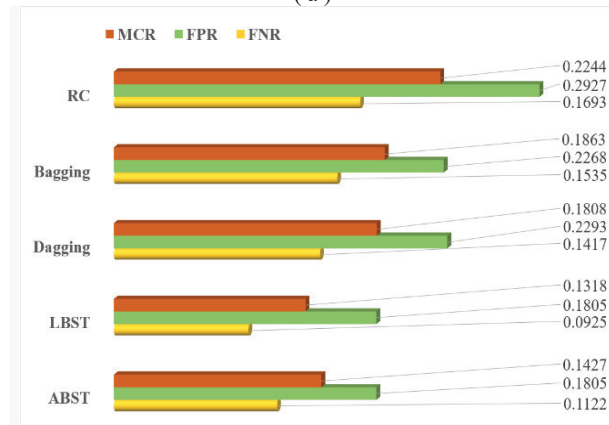
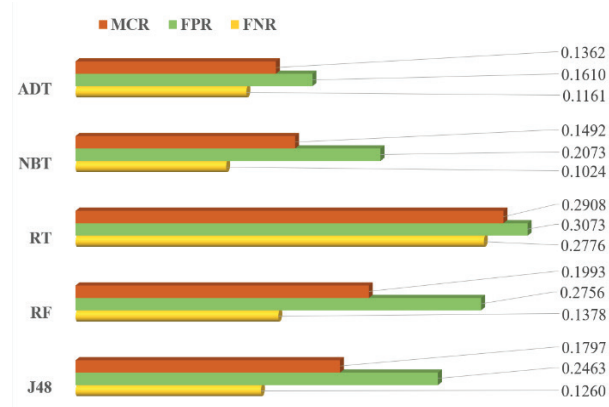
değerleri ile RBFC yöntemi en yüksek performansa ulaşmıştır. Kesinlik metriğinde ise en iyi sonuç 0.8640 değeri ile SGD yönteminden elde edilmiştir. Bu kategoride Kesinlik metriği dışında tüm metrikler açısından en düşük performansı MLPC yöntemi ortaya koymuştur. MLPC bu sonucu 0.8333 Doğruluk, 0.8386 Duyarlılık, 0.8478 F-skor ve 0.8520 AUC değerleri ile elde etmiştir. Kesinlik metriğinde en düşük performansa sahip yöntem 0.8521 değeri ile VPRC olmuştur. Sonuçlar genel performans bazında değerlendirildiğinde Doğruluk, Duyarlılık ve F-skor metrikleri açısından 0.8725, 0.9173 ve 0.8885 değerleri ile RBFC yönteminin, Kesinlik ve AUC metrikleri açısından ise 0.8718 ve 0.9300 değerleri ile ADT yönteminin en yüksek performansı ortaya koyduğu gözlenmiştir.

Tablo 5. 10 kat çapraz doğrulama şartlarında MÖ yöntemlerin temel metrikler açısından performansları

Kategori	Yöntem	Acc	Prec	Rec	F-scr	AUC
Ağaç	J48	0.8203	0.8147	0.8740	0.8433	0.8110
	RF	0.8007	0.7949	0.8622	0.8272	0.8620
	RT	0.7092	0.7444	0.7224	0.7333	0.7430
	NBT	0.8508	0.8429	0.8976	0.8694	0.9100
	ADT	0.8638	0.8718	0.8839	0.8778	0.9300
Meta	ABST	0.8573	0.8590	0.8878	0.8732	0.9220
	LBST	0.8682	0.8617	0.9075	0.8840	0.9190
	Dagging	0.8192	0.8226	0.8583	0.8401	0.8750
	RC	0.7756	0.7786	0.8307	0.8038	0.8410
Fonksiyon	SVM	0.8693	0.8579	0.9154	0.8857	0.8640
	MLPC	0.8333	0.8571	0.8386	0.8478	0.8520
	RBFC	0.8725	0.8614	0.9173	0.8885	0.9240
	VPRC	0.8562	0.8521	0.8957	0.8733	0.8680
	SGD	0.8606	0.8640	0.8878	0.8757	0.9160

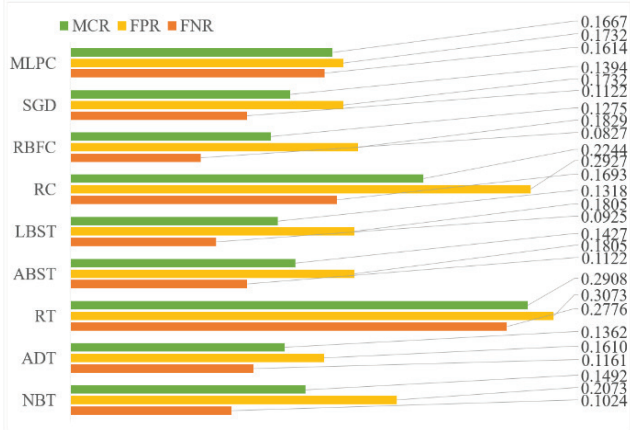
Şekil 2' de 10 kat çapraz doğrulama şartlarında ağaç, meta ve fonksiyon kategorisinde yer alan yöntemlerin hata ölçüm metrikleri açısından gösterdiği performanslar verilmiştir. Ağaç kategorisindeki yöntemlerin hata metrikleri açısından karşılaştırmasını gösteren Şekil 2.a' dan anlaşılacağı üzere en yüksek performansları 0.1024 değeri ile NBT yöntemi FNR metriğinden, 0.1610 ve 0.1362 değerleri ile ADT yöntemi FPR ve MCR metriklerinden elde etmişlerdir. Bu kategoride tüm metrikler açısından en düşük performansa 0.2776 FNR, 0.3073 FPR ve 2908 MCR değerleri ile RT yöntemi sahip olmuştur. Şekil 2.b' deki performans sonuçlarından LBST yönteminin FNR ve MCR metrikleri açısından 0.0925 ve 0.1318 değerleri en yüksek performansa sahip olduğu ve FPR metriğinde 0.1805 değeri ile ABST yöntemi ile birlikte en yüksek performansı sergilediği görülmektedir. Meta kategorideki yöntemler için sonuçlar en düşük performans açısından değerlendirildiğinde 0.1693 FNR, 0.2927 FPR ve 0.2244 MCR değerleri ile RC yönteminin bu sonuca ulaştığı anlaşılmaktadır. Şekil 2.c' deki performans sonuçları, fonksiyon kategorisindeki RBFC yönteminin FNR ve MCR metrikleri açısından en yüksek performansı sırası ile 0.0827 ve 0.1275 değerleri ile elde ettiğini göstermiştir. FPR metriği açısından en iyi performans ise 0.1732 değeri ile SGD ve MLPC yöntemleri arasında paylaşılmıştır. Bu kategoride en düşük performansa FNR ve MCR metriklerinde 0.1614 ve 0.1667 değerleri ile MLPC

yöntemi ve FPR metriğinde 0.1927 değeri ile VPRC yöntemi sahip olmuştur.



Şekil 2. 10 kat çapraz doğrulama şartlarında ağaç, meta ve fonksiyon kategorisinde yer alan yöntemlerin kategorik olarak hata ölçüm metrikleri açısından gösterdiği performanslar: (a) ağaç kategorisi, (b) meta kategorisi ve (c) fonksiyon kategorisi.

Genel performans olarak en iyi ve en kötü performansa sahip yöntem, her bir kategoriden iki en yüksek ve bir en düşük performansa sahip toplam dokuz MÖ yöntemi dikkate alınarak yapılan analizler sonucu belirlenmiştir. Bu kapsamda analiz edilen dokuz MÖ yönteminin hata ölçüm metrikleri açısından performanslarının karşılaştırılması Şekil 3’ te gösterilmiştir. Şekil üzerinde verilen performans sonuçları bize tüm yöntemler arasında FNR ve MCR metriklerinde 0.0827 ve 0.1275 değerleri ile RBFC yönteminin ve FPR metriğinde 0.1610 değeri ile ADT yönteminin en yüksek performansı elde ettiğini göstermiştir. Tüm metrikler açısından en düşük performansa ise 0.2776 FNR, 0.3073 FPR ve 0.2908 MCR değerleri ile RT yöntemi sahip olmuştur.



Şekil 3. 10 kat çapraz doğrulama şartlarında tüm kategoriler içerisinde en yüksek ve en düşük performansa sahip yöntemlerin hata ölçüm metrikleri açısından karşılaştırılması

5. Sonuçlar

Dünya genelinde ciddi sağlık sorunlarından biri olarak ifade edilen kalp yetmezliğinin erken teşhisi, onun seyrini değiştirerek etkin tedavi sürecini başlatabilir. Bu durum bir yandan hastanın yaşam kalitesini artırabilirken, diğer taraftan onun sağlık durumu sonuçlarına da olumlu yönde yansiyebilir. Literatürdeki çalışmalar bize büyük veri setlerinde derinlemesine analiz yapma kabiliyeti ve karmaşık ilişkileri saptama yeteneği sayesinde MÖ yöntemlerinin kalp yetmezliği hastalığı teşhisinde geleneksel yöntemlere göre ön planda olduğunu gösteriyor. Bu bağlamda bu çalışma; farklı kategoride yer alan farklı tip MÖ yöntemlerinin kalp yetmezliği teşhisindeki etkinliğini göstermek için bir modelin oluşturulması ve bu model üzerinden kategorik ve genel olarak MÖ yöntemlerinin analiz ve değerlendirmelerini içerir. MÖ yöntemleri olarak ağaç, meta ve fonksiyon kategorisinin her birinden 5 adet olmak üzere toplam 15 algoritma kullanılmıştır. Doğru bir performans ölçümü gerçekleştirebilmek için hem Doğruluk, Kesinlik, Duyarlılık, F-skör ve AUC ölçümlerini içeren sınıflandırma problemlerinin temel metrikleri hem de FNR, FPR ve MCR gibi hata ölçümünü esas alan metriklerden yararlanılmıştır. Deneysel çalışmalar kategorik temelde ele alındığında ağaç kategorisinde Duyarlılık ve FNR metrikleri dışındaki metriklerde ADT yönteminin, meta kategorisinde AUC metriği dışında metriklerde LBST yönteminin

ve fonksiyon kategorisinde Kesinlik ve FPR metrikleri dışındaki metriklerde RBFC yönteminin en yüksek performansı ortaya koyduğunu göstermiştir. Meta ve fonksiyon kategorilerinde LBST ve RBFC yöntemlerinin başarısız gözüktüğü metrikler AUC ve Kesinlik için ABST ve SGD yöntemleri daha başarılı gözükseler de bu metrikler açısından dört yöntemde de performans sonuçlarının birbirine oldukça yakın olduğu gözlenmiştir. Tüm yöntemler genelinde sonuçlar değerlendirildiğinde Doğruluk, Duyarlılık, F-skör, FNR ve MCR metriklerinde RBFC yöntemi, Kesinlik, AUC ve FPR metriklerinde ADT yönteminin en yüksek performansa sahip olduğu görülmüştür. Bu yöntemlerde Duyarlılık, F-skör, AUC ve FNR metrikleri açısından skorların birbirine oldukça yakın olduğu dikkate alındığında her iki metodunda kalp yetmezliği teşhisinde başarılı bir şekilde kullanılabileceği sonucuna varabiliriz. Gelecek çalışmamızda kalp hastalıklarının teşhisinde yüksek doğruluklu tahminler elde edebilmek için makine öğrenme ve derin öğrenme algoritmalarının avantajlarını birleştiren hibrid modellerin oluşturulması planlanmaktadır. Oluşturulan modeller ile büyük veri setlerinin özellikleri çıkartılarak etkili tahminler yapılması hedeflenmektedir.

6. Kaynaklar

- [1] WHO, “Cardiovascular Diseases (CVDs),” [Online]. Available: [https://www.who.int/news-room/factsheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/factsheets/detail/cardiovascular-diseases-(cvds)) (Accessed; Sep. 19, 2023).
- [2] S. Bozkurt Keser and K. Keskin, “Kalp Yetmezliği Hastalarının Sağ Kalım Tahmini: Sınıflandırmaya Dayalı Makine Öğrenmesi Algoritmalarının Bir Uygulaması,” Afyon Kocatepe University Journal of Sciences and Engineering, vol. 23, no. 2, pp. 362–369, May 2023, doi: 10.35414/AKUFEMUBID.1033377.
- [3] G. Gürgen and S. Serttaş, “Kalp Yetmezliği Hastalığının Erken Teşhisinde Makine Öğrenimi Algoritmalarının Performans Karşılaştırması,” Euroasia Journal of Mathematics, Engineering, Natural & Medical Sciences, vol. 165, no. 10, pp. 165–174, 2023, doi: 10.5281/zenodo.8238065.
- [4] SBSGM, “Sağlık İstatistikleri Yıllığı 2019,” Ankara, 2019. [Online]. Available: <https://sbsgm.saglik.gov.tr/Eklenti/40564/0/saglik-istatistikleri-yilligi-2019pdf.pdf> (Accessed: Nov. 07, 2023).
- [5] O. Gold and A. Iorshase, “Heart failure prediction framework using random forest and J48 with Adaboost algorithms,” Science World Journal, vol. 18, no. 2, pp. 165–175, Oct. 2023, doi: 10.4314/SWJ.V18I2.1.
- [6] S. Kaushik and R. Birok, “Heart Failure prediction using Xgboost algorithm and feature selection using feature permutation,” IEEE International Conference on Electrical, Computer and Communication Technologies, 2021, pp.1-6, doi: 10.1109/ICECCT52121.2021.9616626.
- [7] N. S. Mansur Huang, Z. Ibrahim, and N. Mat Diah, “Machine learning techniques for early heart failure prediction / Nur Shahellin Mansur Huang, Zaidah Ibrahim and Norizan Mat Diah,” 2021, [Online]. Available: <https://mjoc.uitm.edu.my/> (Accessed: Nov. 07, 2023).

- [8] Z. Masetic and A. Subasi, "Congestive heart failure detection using random forest classifier," *Comput Methods Programs Biomed*, vol. 130, pp. 54–64, Jul. 2016, doi: 10.1016/J.CMPB.2016.03.020.
- [9] S. Kermani, M. Ghelich Oghli, A. Mohammadzadeh, and R. Kafieh, "NF-RCNN: Heart localization and right ventricle wall motion abnormality detection in cardiac MRI," *Physica Medica*, vol. 70, pp. 65–74, Feb. 2020, doi: 10.1016/J.EJMP.2020.01.011.
- [10] K. A. Lara Hernandez, T. Rienmüller, D. Baumgartner, and C. Baumgartner, "Deep learning in spatiotemporal cardiac imaging: A review of methodologies and clinical usability," *Comput Biol Med*, vol. 130, Mar. 2021, doi: 10.1016/J.COMPBIOMED.2020.104200.
- [11] N. Lessmann et al., "Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest CT," <https://doi.org/10.1117/12.2216978>, vol. 9785, pp. 255–260, Mar. 2016, doi: 10.1117/12.2216978.
- [12] N. Zhang et al., "Deep Learning for Diagnosis of Chronic Myocardial Infarction on Nonenhanced Cardiac Cine MRI," *Radiology*, vol. 291, no. 3, pp. 606–607, 2019, doi: 10.1148/RADIOLOGY.2019182304.
- [13] M. Porumb, E. Iadanza, S. Massaro, and L. Pecchia, "A convolutional neural network approach to detect congestive heart failure," *Biomed Signal Process Control*, vol. 55, pp. 101597, Jan. 2020, doi: 10.1016/J.BSPC.2019.101597.
- [14] A. N. Shihab, M. J. Mokarrama, R. Karim, S. Khatun, and M. S. Arefin, "An iot-based heart disease detection system using rnn," *Advances in Intelligent Systems and Computing*, vol. 1200 AISC, pp. 535–545, 2021, doi: 10.1007/978-3-030-51859-2_49/COVER.
- [15] G. Maragatham and S. Devi, "LSTM Model for Prediction of Heart Failure in Big Data," *J Med Syst*, vol. 43, no. 5, pp. 1–13, May 2019, doi: 10.1007/S10916-019-1243-3.
- [16] S. Gao, Y. Zheng, and X. Guo, "Gated recurrent unit-based heart sound analysis for heart failure screening," *Biomed Eng Online*, vol. 19, no. 1, Jan. 2020, doi: 10.1186/S12938-020-0747-X.
- [17] A. Çınar and S. A. Tuncer, "Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks," *Comput Methods in Biomech and Biomed Engin*, vol. 24, no. 2, pp. 203–214, 2021, doi: 10.1080/10255842.2020.1821192.
- [18] D. Li, X. Li, J. Zhao, and X. Bai, "Automatic staging model of heart failure based on deep learning," *Biomed Signal Process Control*, vol. 52, pp. 77–83, Jul. 2019, doi: 10.1016/J.BSPC.2019.03.009.
- [19] V. K. Sudha and D. Kumar, "Hybrid CNN and LSTM Network for Heart Disease Prediction," *SN Comput Sci*, vol. 4, no. 2, pp. 172, Mar. 2023, doi: 10.1007/S42979-022-01598-9.
- [20] L. Ali et al., "An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure," *IEEE Access*, vol. 7, pp. 54007–54014, 2019, doi: 10.1109/ACCESS.2019.2909969.
- [21] Y. Chen, X. Qin, L. Zhang, and B. Yi, "A Novel Method of Heart Failure Prediction Based on DPCNN-XGBOOST Model," *Computers Materials & Continua*, vol. 65, no. 1, pp. 495–510, Jul. 2020, doi: 10.32604/CMC.2020.011278.
- [22] V. Grgić, D. Mušić, and E. Babović, "Model for predicting heart failure using Random Forest and Logistic Regression algorithms," *IOP Conf Ser Mater Sci Eng*, vol. 1208, no. 1, pp. 012039, Nov. 2021, doi: 10.1088/1757-899X/1208/1/012039.
- [23] N. Foziljonova and I. Wasito, "Prediction of survival rate of heart failure patients using machine learning techniques," *J Theor Appl Inf Technol*, vol. 100, no. 9, pp. 2703–2714, May 2022, [Online]. Available: <https://www.jatit.org/volumes/Vol100No9/2Vol100No9.pdf> (Accessed: Nov. 07, 2023).
- [24] M. Mudassar, M. Afzal, and T. Muhammad, "A Machine Learning Based Predictive Model to Diagnose Heart Failure Patients using Imbalanced Classification Problem," In *2023 4th International Conference on Advancements in Computational Sciences, ICACS 2023 - Proceedings, 2023*, pp. 1–8, doi: 10.1109/ICACS55311.2023.10089759.
- [25] M. Zeng, "The Prediction of Heart Failure based on Four Machine Learning Algorithms," *Highlights in Science Engineering and Technology*, vol. 39, pp. 1377–1382, Apr. 2023, doi: 10.54097/HSET.V39I.6771.
- [26] C. Coşkun and F. Kuncan, "Evaluation of Performance of Classification Algorithms in Prediction of Heart Failure Disease," *KSÜ Mühendislik Bilimleri Dergisi*, vol. 25, no. 4, pp. 622–632, Dec. 2022, doi: 10.17780/KSUJES.1144570.
- [27] fedesoriano., "Heart Failure Prediction Dataset," [Online]. Available: <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> (Accessed: Nov. 08, 2023).
- [28] C. N. Villavicencio, J. J. E. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "Covid-19 prediction applying supervised machine learning algorithms with comparative analysis using weka," *Algorithms*, vol. 14, no. 7, pp. 1–22, Jul. 2021, doi: 10.3390/A14070201.
- [29] Ç. Cem Berke, "Rastgele Orman Algoritması," [Online]. Available: <https://medium.com/@cemthecebi/rastgele-orman-algoritmas%C4%B1-1600ca4f4784> (Accessed: Sep. 08, 2023).
- [30] S. Gayathri, A. K. Krishna, V. P. Gopi, and P. Palanisamy, "Automated Binary and Multiclass Classification of Diabetic Retinopathy Using Haralick and Multiresolution Features," *IEEE Access*, vol. 8, pp. 57497–57504, 2020, doi: 10.1109/ACCESS.2020.2979753.
- [31] J. Li, H. Peng, H. Hu, Z. Luo, and C. Tang, "Multimodal Information Fusion for Automatic Aesthetics Evaluation of Robotic Dance Poses," *Int J Soc Robot*, vol. 12, no. 1, pp. 5–20, Jan. 2020, doi: 10.1007/S12369-019-00535-W/METRICS.
- [32] Y. Zhao and Y. Zhang, "Comparison of decision tree methods for finding active objects," *Advances in Space Research*, vol. 41, no. 12, pp. 1955–1959, 2008, doi: 10.1016/J.ASR.2007.07.020.

- [33] E. Doğru, "Makine Öğrenmesi Teknikleri Kullanarak Hizmet Aksatma Saldırıları Tespiti," Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2020.
- [34] H. Yan and W. Chen, "Landslide susceptibility modeling based on GIS and ensemble techniques," *Arabian Journal of Geosciences*, vol. 15, no. 8, pp. 1–22, Apr. 2022, doi: 10.1007/S12517-022-09974-8.
- [35] Y.-S. Chen et al., "Comparable Studies of Financial Bankruptcy Prediction Using Advanced Hybrid Intelligent Classification Models to Provide Early Warning in the Electronics Industry," *Mathematics*, vol. 9, pp. 2622, vol. 9, no. 20, pp. 2622, Oct. 2021, doi: 10.3390/MATH9202622.
- [36] K. Khosravi, Z. Sheikh Khozani, and J. Hatamiafkoueih, "Prediction of embankments dam break peak outflow: a comparison between empirical equations and ensemble-based machine learning algorithms," *Natural Hazards*, vol. 118, no. 3, pp. 1989–2018, Sep. 2023, doi: 10.1007/S11069-023-06060-4.
- [37] B. T. Pham et al., "GIS-based ensemble soft computing models for landslide susceptibility mapping," *Advances in Space Research*, vol. 66, no. 6, pp. 1303–1320, Sep. 2020, doi: 10.1016/J.ASR.2020.05.016.
- [38] F. Bulut, "Simflandırıcı Topluluklarının Dengesiz Veri Kümeleri Üzerindeki Performans Analizleri," *Bilişim Teknolojileri Dergisi*, vol. 9, no. 2, pp. 153–159, 2016, [Online]. Available: <http://search.yayin/detay/220680> (Accessed: Nov. 08, 2023).
- [39] P. Ghosh et al., "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and lasso feature selection techniques," *IEEE Access*, vol. 9, pp. 19304–19326, 2021, doi: 10.1109/ACCESS.2021.3053759.
- [40] B. T. Pham and I. Prakash, "Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping," *Geocarto Int.*, vol. 34, no. 3, pp. 316–333, Feb. 2019, doi: 10.1080/10106049.2017.1404141.
- [41] M. H. Kamarudin, C. Maple, T. Watson, and N. S. Safa, "A LogitBoost-Based Algorithm for Detecting Known and Unknown Web Attacks," *IEEE Access*, vol. 5, pp. 26190–26200, Nov. 2017, doi: 10.1109/ACCESS.2017.2766844.
- [42] M. Kaya Keleş, "Breast cancer prediction and detection using data mining classification algorithms: A comparative study," *Tehnicki Vjesnik*, vol. 26, no. 1, pp. 149–155, Feb. 2019, doi: 10.17559/TV-20180417102943.
- [43] Akca MF, "Nedir Bu Destek Vektör Makineleri? (Makine Öğrenmesi Serisi-2)." [Online]. Available: <https://medium.com/deep-learning-turkiye/nedir-bu-destek-vekt%C3%B6r-makinelerimakine%C3%B6%C4%9Frenmesi-serisi-2-94e576e4223e> (Accessed: Sep. 10, 2023).
- [44] M. P. Akhter, Z. Jiangbin, I. R. Naqvi, M. Abdelmajeed, A. Mehmood, and M. T. Sadiq, "Document-Level Text Classification Using Single-Layer Multisize Filters Convolutional Neural Network," *IEEE Access*, vol. 8, pp. 42689–42707, 2020, doi: 10.1109/ACCESS.2020.2976744.
- [45] M. A. Hussain and L. Gogoi, "Performance analyses of five neural network classifiers on nodule classification in lung CT images using WEKA: a comparative study.," *Phys Eng Sci Med*, vol. 45, no. 4, pp. 1193–1204, Dec. 2022, doi: 10.1007/S13246-022-01187-3.
- [46] W. Chen et al., "Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China," *Science of the Total Environment*, vol. 626, pp. 1121–1135, Jun. 2018, doi: 10.1016/J.SCITOTENV.2018.01.124.
- [47] C. Eyüpoğlu, "Büyük Veride Etkin Gizlilik Koruması İçin Yazılım Tasarımı," *Doktora Tezi*, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2018.
- [48] A. Firdaus, N. B. Anuar, M. F. A. Razak, and A. K. Sangaiah, "Bio-inspired computational paradigm for feature investigation and malware detection: interactive analytics," *Multimed Tools Appl*, vol. 77, no. 14, pp. 17519–17555, Jul. 2018, doi: 10.1007/S11042-017-4586-0/METRICS.
- [49] J. Huang, S. Ling, X. Wu, and R. Deng, "GIS-Based Comparative Study of the Bayesian Network, Decision Table, Radial Basis Function Network and Stochastic Gradient Descent for the Spatial Prediction of Landslide Susceptibility," *Land (Basel)*, vol. 11, no. 3, pp. 436, Mar. 2022, doi: 10.3390/LAND11030436.
- [50] T. Srinivas et al., "Novel Based Ensemble Machine Learning Classifiers for Detecting Breast Cancer," *Math. Probl. in Eng.*, vol. 2022, pp. 1–15, May 2022, doi: 10.1155/2022/9619102.
- [51] H. Torun, "Karışıklık Matrisi (Confusion Matrix)," [Online]. Available: <https://hakan.io/karisiklik-matrisi-confusion-matrix/> (Accessed: Sep. 15, 2023).
- [52] Y. Tang, "Python Algorithms: Accuracy, Precision, Recall, and F Score." [Online]. Available: <https://pythonalgorithms.com/2022/05/16/accuracy-precision-recall-and-f-score/> (Accessed: Sep. 17, 2023).
- [53] Ş. Ay, "Model Performansını Değerlendirmek — Metrikler," [Online]. Available: <https://medium.com/deep-learningturkiye/modelperformans%C4%B1n%C4%B1de%C4%9Ferlendirmek-metrikler-cb6568705b1> (Accessed: Sep. 17, 2023).

Özgeçmiş



İsmail Atacak, 1972 yılında Konya’da doğdu, 1994 yılında Gazi Üniversitesi Elektronik ve Bilgisayar Eğitimi Bölümünden mezun oldu. 1998 ve 2005 yıllarında Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik-Bilgisayar Eğitimi Anabilim dalında sırası ile Yüksek Lisans ve Doktora eğitimlerini tamamladı. 2007-2012 yılları arasında Teknik Eğitim Fakültesi, Elektronik ve Bilgisayar Eğitimi Bölümünde Yardımcı Doçent olarak görev yaptıktan sonra, aynı üniversitenin Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümünde çalışmaya başladı ve halen bu bölümde Dr. Öğr. Üyesi olarak çalışmaya devam etmektedir. Bulanık Mantık, Makine Öğrenmesi, Derin Öğrenme, Optimizasyon Algoritmaları ve Veri Bilmi ilgi duyduğu araştırma alanları arasında yer alır.